

Article

Plasma Protein Levels Analysis in Multiple Sclerosis Sardinian Families Identified C9 and CYP24A1 as Candidate Biomarkers

Andrea Nova ^{1,*},[†] , Teresa Fazia ¹,[†] , Ashley Beecham ^{2,3}, Valeria Saggi ⁴, Marialuisa Piras ⁴, Jacob L. McCauley ^{2,3} , Carlo Berzuini ⁵  and Luisa Bernardinelli ¹

¹ Department of Brain and Behavioral Sciences, University of Pavia, 27100 Pavia, Italy; teresa.fazia01@ateneopv.it (T.F.); luisa.bernardinelli@unipv.it (L.B.)

² John P. Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL 33146, USA; ABeecham@med.miami.edu (A.B.); jmccauley@med.miami.edu (J.L.M.)

³ Dr. John T. Macdonald Foundation Department of Human Genetics, Miller School of Medicine, Miami, FL 33136, USA

⁴ Divisione di Neurologia, Presidio Ospedaliero S. Francesco, ASL Numero 3 Nuoro, 08100 Nuoro, Italy; valeria.saggi@tiscali.it (V.S.); marialpiras@tiscali.it (M.P.)

⁵ Centre for Biostatistics, The University of Manchester, Manchester M13 9PL, UK; carlo.berzuini@manchester.ac.uk

* Correspondence: andrea.nova01@universitadipavia.it

† These authors contributed equally to this work.

Abstract: Here we investigate protein levels in 69 multiple sclerosis (MS) cases and 143 healthy controls (HC) from twenty Sardinian families to search for promising biomarkers in plasma. Using antibody suspension bead array technology, the plasma levels of 56 MS-related proteins were obtained. Differences between MS cases and HC were estimated using Linear Mixed Models or Linear Quantile Mixed Models. The proportion of proteins level variability, explained by a set of 119 MS-risk SNPs as to the literature, was also quantified. Higher plasma C9 and CYP24A1 levels were found in MS cases compared to HC ($p < 0.05$ after Holm multiple testing correction), with protein level differences estimated as, respectively, 0.53 (95% CI: 0.25, 0.81) and 0.42 (95% CI: 0.19, 0.65) times plasma level standard deviation measured in HC. Furthermore, C9 resulted in both statistically significantly higher relapsing-remitting MS (RRMS) and secondary-progressive MS (SPMS) compared to HC, with SPMS showing the highest differences. Instead, CYP24A1 was statistically significantly higher only in RRMS as compared to HC. Respectively, 26% (95% CI: 10%, 44%) and 16% (95% CI: 9%, 39%) of CYP24A1 and C9 plasma level variability was explained by known MS-risk SNPs. Our results highlight C9 and CYP24A1 as potential biomarkers in plasma for MS and allow us to gain insight into molecular disease mechanisms.

Keywords: biomarkers; plasma proteins; family data; suspension bead array technology; explained variability



Citation: Nova, A.; Fazia, T.; Beecham, A.; Saggi, V.; Piras, M.; McCauley, J.L.; Berzuini, C.; Bernardinelli, L. Plasma Protein Levels Analysis in Multiple Sclerosis Sardinian Families Identified C9 and CYP24A1 as Candidate Biomarkers. *Life* **2022**, *12*, 151. <https://doi.org/10.3390/life12020151>

Academic Editor: Sonia Longhi

Received: 9 December 2021

Accepted: 13 January 2022

Published: 20 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Multiple sclerosis (MS) (OMIM 126200) is an inflammatory demyelinating disorder of the central nervous system (CNS) and the most frequent cause of chronic neurological disability [1–3]. Epidemiological studies have provided evidence on the multifactorial nature of MS susceptibility involving both genetic and environmental factors [4,5]. In the disease process several pathophysiological mechanisms, including inflammation, demyelination, axonal damage, and repair mechanisms are involved. These mechanisms are not uniformly represented across patient populations, thus contributing to the heterogeneity in phenotypic expression of the disease, its prognosis, and response to therapies [6]. Thus, the need for suitable biomarkers to assist in the treatment decisions and in the choice of effective therapy strategy is compelling.

As defined in [7], “[a] biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biological process, pathogenic processes, or pharmacologic responses to a therapeutic intervention”.

For neurological diseases such MS, cerebrospinal fluid (CSF), given its proximity to the CNS, would be the preferred body fluid in which to search for candidate biomarkers rather than plasma or serum. Clearly, however, the CSF sample collection is a more invasive procedure with potential risks compared to plasma sampling. For this reason, CSF samples from healthy individuals are seldom taken especially for discovery-driven approaches. On the other hand, the presence in blood plasma of leakage or secretion products from various tissues or blood cells can reflect the pathological and physiological state of certain tissues [8] so that plasma, easily obtained through a non-invasive procedure, can be considered a good body fluid on which to search for promising diagnostic biomarkers.

Proteome in plasma has started to be studied in a high throughput manner, thanks to the technological development of the Human Protein Atlas (HPA) project (<https://www.proteinatlas.org/>, accessed on 12 July 2021) which has allowed the production of a large resource of polyclonal antibodies [9].

In the past years, taking advantage of new proteomic technologies [10], the field of biomarker discovery has gradually shifted from the aim of finding a unique perfect surrogate marker to the construction of a composite panel of candidate biomarkers with higher performance, particularly suitable in the context of complex multifactorial diseases, such as MS, where many factors are involved in the development and progression of the disease.

Despite many efforts, the identification of a suitable set of biomarkers for MS based on peripheral blood is just beginning due to the complexity of MS [11]. Affinity proteomic methods have been used to identify MS-related proteins both in plasma and in CSF, leading to the identification of IRF8, IL7, SLC30A7, METTL14, and GAP43 [12–18]. A recent work [19], using antibody suspension bead array, has led to the identification of nine proteins to be regarded as potentially useful biomarkers for MS treatment, where plasma levels decrease after treatment with natalizumab, with two of these, PEBP1 and RTN3, showing the most significant changes.

However, very few of the many proposed candidate molecular MS biomarkers, specifically antibodies measuring humoral immune response (i.e., anti-natalizumab antibodies, neutralizing antibodies to interferon beta, IgG, anti-aquaporin-4 antibodies, and antibodies against varicella zoster and John Cunningham viruses) [20,21], have been validated and used in clinical practice, thus affecting clinical decision making [22].

Our study is framed within the field of biomarker discovery. More specifically, it was focused on the identification of new putative candidate biomarkers (proteins) for MS in plasma with the aim of gaining more insight into the pathogenic aspect of MS. To this aim we used protein profile data obtained via antibody suspension bead affinity-based array technology [23] from MS subjects and healthy related controls belonging to twenty multiplex families from the founder population of the Nuoro province, Sardinia (Italy).

The Mediterranean island of Sardinia is characterized by an age- and sex-adjusted prevalence of MS of 330 per 100,000 inhabitants, among the highest reported worldwide [24]. Furthermore, being a founder population isolated in a geographical, cultural, or religious sense, evolved from a small number of ancestors, subjected to limited immigration and hence characterized by high genetic homogeneity, it is the ideal population for the identification of disease determinants [25,26].

Our study was mainly structured into two phases. In the first phase we compared, in affected and non-affected individuals, 56 plasma protein profiles, a priori selected for their suggested association with MS from previous studies, to explore the possibility of identifying a set of reliable disease-specific diagnostic biomarkers. The highlighted proteins were further investigated by evaluating protein level differences between the MS disease-course classes.

In the second phase, we selected a panel of MS-risk genetic variants highlighted by the last study of the International Multiple Sclerosis Genetic Consortium (IMSGC) [27] to quantify the plasma level variability explained by these genetic variants in a colocalization perspective. If a consistent portion of variability can be explained by these MS-risk variants, this would add more evidence to the validity of our discovered biomarkers, and at the same time a potential biological function could be attributed to these variants. The rationale behind this second phase was that signals highlighted by Genome-Wide Association Studies (GWAS) are enriched with variants that could influence the levels of quantitative traits acting as quantitative trait loci (QTL), often in a tissue-specific manner. In fact, over three quarters of GWAS hits appear not to be associated with potentially deleterious protein-coding variants [28–30]. Therefore, the identified MS-risk variants can act as regulators of plasma protein levels and altogether explain proteins' variability.

In summary, our analysis allows us to gain more insights into the understanding of molecular mechanisms underlying MS, to identify potential biomarkers for the disease, and to quantify proteins' plasma level variability as explained by a set of well-established MS-risk SNPs to gain insight on their biological function and molecular disease mechanisms.

2. Results

2.1. Sample Description

Twenty Sardinian pedigrees, containing from 6 to 66 subjects (median = 22 subjects) and from 2 to 15 MS patients (median = 3 cases), were selected for the analysis, for a total of 533 subjects (96 affected and 437 healthy controls). For this study, families were further divided into 69 sub-families containing from 3 to 25 subjects (median = 7), to better capture the shared environment effect. Moreover, 56 antibodies targeting unique plasma protein levels (listed in Table S1), made available from the HPA project, and 131,497 genotyped variants from ImmunoChip, were available as follows:

- 212 subjects (58 sub-families, 69 MS cases, 143 healthy controls) had protein levels measured.
- 272 subjects (53 sub-families, 58 MS cases, 214 healthy controls) had ImmunoChip data.
- 135 subjects (43 sub-families, 43 MS cases, 92 healthy controls) had both protein levels and ImmunoChip data
- 77 subjects (29 sub-families, 26 MS cases, 51 healthy controls) had protein levels only
- 137 subjects (47 sub-families, 15 MS cases, 122 healthy controls) had ImmunoChip data only.
- 184 subjects (62 sub-families, 12 MS cases, 172 healthy controls) had neither protein levels nor ImmunoChip data.

Table 1 reports descriptive statistics for sex, age at the day of blood sampling, age at MS onset, and MS disease course for the 212 subjects having protein levels.

Table 1. Sample descriptive statistics for 212 subjects having protein levels.

	<i>n</i>	% Females	Age at Blood Sampling		Age at MS Onset	
			Median	IQR ¹	Median	IQR ¹
Number of subjects	212	54.7%	49.3	36.0–64.5		
MS cases	69 (32.5%)	66.7%	39.3 ³	32.5–48.2	27.0 ⁴	21.0–34.0
RRMS ⁵	41 (59.4%) ²	73.2%	34.9	29.4–42.4	26.0	20.0–34.0
SPMS ⁶	18 (26.1%) ²	61.1%	45.7	42.2–50.8	29.0	24.5–37.0
Not reported	10 (14.5%) ²	50.0%	41.2 ³	36.1–45.9	23.0 ⁴	22.8–29.0
Healthy controls	143 (67.4%)	49.0%	57.2 ³	40.5–69.0		

¹ Interquartile Range; ² Percentages are referred to the total of MS cases; ³ 3 subjects had missing age at sampling; ⁴ 6 subjects had missing age at onset; ⁵ RRMS = Relapse-Remitting MS; ⁶ SPMS = Secondary Progressive MS.

2.2. Differences in Protein Levels

2.2.1. Power Evaluation

The primary endpoint was to detect significant protein level differences between MS cases and healthy controls. The statistical power to detect effect sizes ≥ 0.3 and ≥ 0.5 (in terms of healthy controls standard deviation (SD) scale) was evaluated using a simulation approach for both the linear mixed model (LMM) and linear quantile mixed model (LQMM) (this latter to be used in case of non-normality of LMM residuals). An effect size of 0.5 was chosen in the spirit of the distributional methods to define the minimum clinically important difference (MCID) as 0.5 times the SD of the endpoint of interest [31], even if in this case it is simply a measure of a potentially meaningful physiological change. The sample size was fixed to 212, the proportion of MS cases to 0.32 (as in our sample), and the pedigree structure as that of our Sardinian families. In the end 206 subjects were analyzed, since 6 subjects had a missing age at blood sampling day. Simulating protein levels 2000 times, the power to detect an effect size ≥ 0.5 at significance level $\alpha = 0.05$ was $>90\%$ for both LMM and LQMM, while the power to detect an effect size ≥ 0.3 at $\alpha = 0.05$ was, still, $>90\%$ for LMM but $\approx 60\%$ for LQMM.

2.2.2. Protein Levels Significantly Different between MS Cases and Healthy Controls

Differences between MS cases and healthy controls were evaluated for the 56 protein levels in 69 MS cases and 143 unaffected controls, belonging to 58 Sardinian sub-families, by fitting LMM as in Equation (1). Six subjects with missing age at blood sampling were excluded from the analysis. The normality of residuals was checked following Kim guidelines [32] and only in 14 models was the normality assumption respected. Thus, for the remaining 42 protein levels, LQMM was performed. Statistically significant results, after family-wise error (FWER) = 0.05 control using Holm procedure, obtained both fitting LMM or LQMM are reported in Table 2. Estimates represent the protein level difference between MS cases and healthy controls, adjusted for sex, age at blood sampling, kinship effect and shared environment effect, expressed as the number of SDs of the protein levels in healthy controls; thus, all estimates are to be considered on the healthy control SD scale (HC SD). As shown in Table 2, two proteins, complement C9 and cytochrome P450 family 24 subfamily A member 1 (CYP24A1), were significantly different between MS cases and healthy controls, after multiple testing Holm correction, and thus they were considered as potential biomarkers. Among the results not statistically significant, after multiple testing correction, it is worth mentioning three additional proteins considered suggestive, i.e., somatostatin (SST), ceruloplasmin (CP), and tissue type plasminogen activator (PLAT), having a sex- and age-adjusted absolute difference between MS case and healthy control levels of at least 0.30 HC SD with an uncorrected p -value < 0.005 . Complete results for all 56 protein levels are provided in Supplementary Table S2.

Table 2. Significant and suggestive protein level differences between MS cases and healthy controls.

Protein	Chromosome	HPA ¹	Model ²	Estimate ³	SE ⁴	95% CI ⁵	Raw p -Value ⁶	Corrected p -Value ⁷
C9	5	029577	LMM	0.529	0.143	[0.248, 0.810]	<0.001	0.012
CYP24A1	20	022261	LQMM	0.418	0.116	[0.188, 0.647]	<0.001	0.023
SST	3	019472	LQMM	-0.371	0.126	[-0.619, -0.123]	0.004	0.190
CP	3	001834	LQMM	-0.381	0.130	[-0.638, -0.125]	0.004	0.195
PLAT	8	003412	LMM	-0.322	0.113	[-0.543, -0.101]	0.004	0.217

¹ Human Protein Atlas antibody product name (Human Platelet Antigen); ² LMM = linear mixed model, LQMM = linear quantile mixed model; ³ difference between MS cases and healthy controls expressed as number of standard deviations in healthy controls. Estimate is based on the mean for LMM models and on the median for LQMM models; ⁴ standard error; ⁵ confidence interval; ⁶ adjusted for sex, age at blood sampling day, kinship effect, and shared environment effect; ⁷ corrected, due to multiple testing, using Holm procedure.

Specifically, C9 protein levels in MS cases were estimated, using LMM, on average 0.53 HC SD higher as compared to healthy control (95% CI: 0.25, 0.81) with a corrected p -value < 0.012 . A CYP24A1 estimate was obtained through LQMM and, thus, is referred to the median (50th quantile) rather than the mean of protein levels. Therefore, the median of CYP24A1 protein levels in MS cases was 0.42 HC SD higher compared to healthy controls (95% CI: 0.19, 0.65) with a corrected p -value < 0.023 . Box plots, for these two plasma protein levels, are provided in Figure 1.

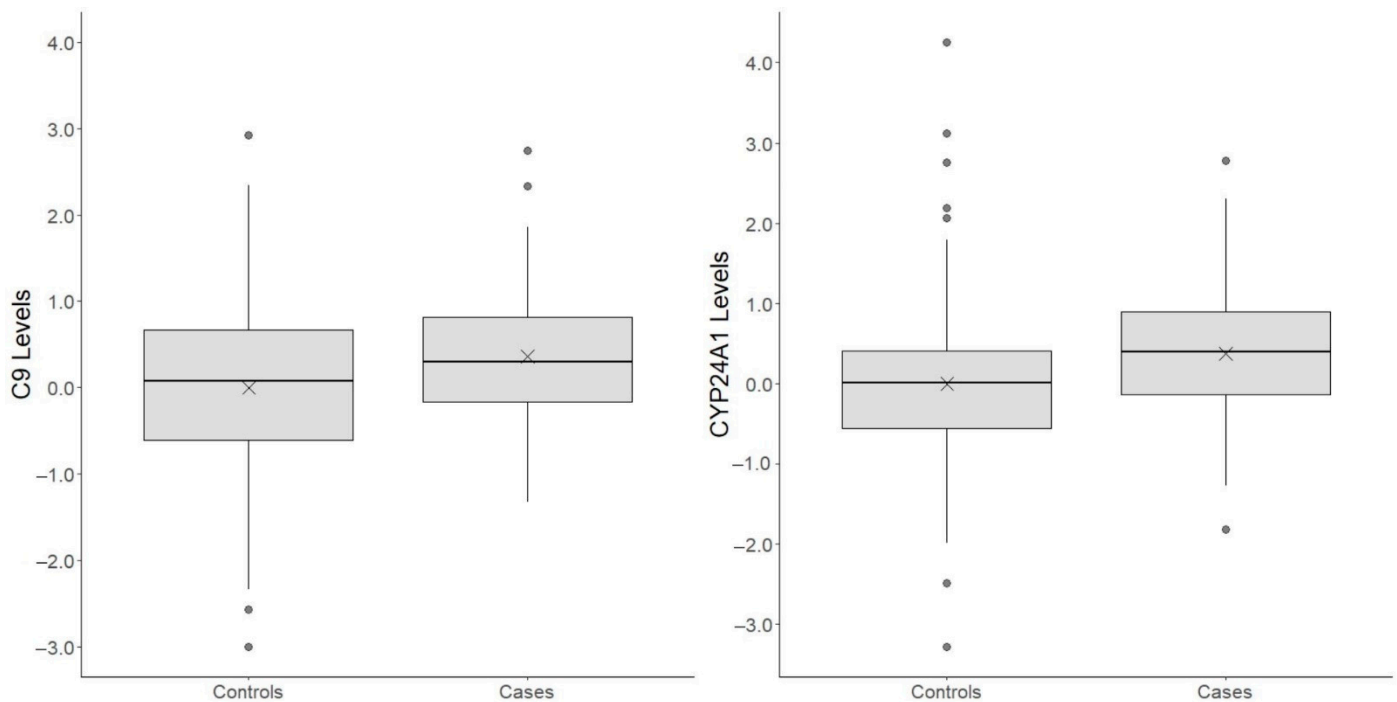


Figure 1. Box plots for statistically significant protein level differences between MS cases and healthy controls. Plasma protein levels were standardized using mean and standard deviation from healthy control protein levels; thus, 1 unit on the y-scale represents 1 standard deviation in healthy control protein levels. “X” represents the mean protein level within a group.

The sex- and age-adjusted estimated median of SST and CP plasma protein levels was lower in MS cases, i.e., -0.37 HC SD (95% CI: -0.62 , -0.12 , uncorrected $p < 0.004$) and -0.38 HC SD (95% CI: -0.64 , -0.13 , uncorrected $p < 0.004$), respectively. Additionally, PLAT plasma protein levels were lower, on average, in MS cases, i.e., -0.32 HC SD (95% CI: -0.54 , -0.10 , uncorrected $p < 0.004$). Box plots, for these three plasma protein levels, are provided in Figure 2. Pearson’s coefficient correlation matrix for C9, CYP24A1, SST, CP, and PLAT plasma protein levels is depicted in Figure 3. SST and CP were highly correlated, having Pearson’s coefficient $r = 0.71$, while the other proteins were weakly correlated, i.e., $r < 0.15$.

CYP24A1, SST, and CP protein level differences between MS cases and healthy controls have also been explored at the 25th and 75th quantiles of the protein level distribution (Table 3). The results obtained at these quantiles were almost identical to those of the 50th quantile, meaning that the non-normality of residuals obtained using linear regression to infer plasma protein levels’ entire distribution did not meaningfully impact estimates at different quantiles. Therefore, CYP24A1, SST, and CP protein level differences between MS cases and healthy controls at the median could be generalized to higher and lower levels of the distribution.

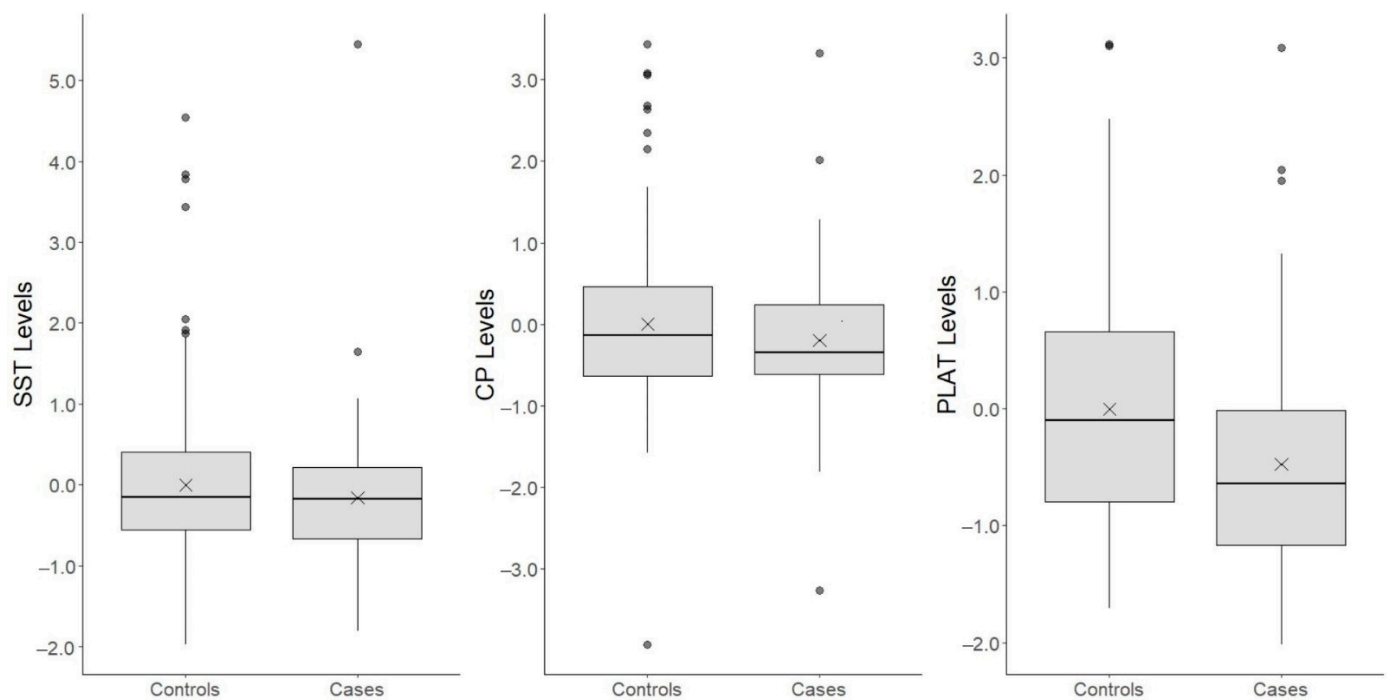


Figure 2. Box plots for SST, CP, and PLAT protein level differences between MS cases and healthy controls. Plasma protein levels were standardized using mean and standard deviation from healthy control protein levels; thus, 1 unit on the y-scale represents 1 standard deviation in healthy controls' protein levels. "X" represents the mean protein level within a group.

Table 3. 25th and 75th quantile estimates for protein level differences between MS cases and healthy controls in proteins resulted statistically significant at 50th quantile after multiple testing correction.

Protein	Quantile	Estimate ¹	SE ²	95% CI ³	p-Value ⁴
CYP24A1	25th	0.413	0.117	[0.182, 0.643]	<0.001
CYP24A1	75th	0.418	0.116	[0.188, 0.647]	<0.001
SST	25th	−0.391	0.126	[−0.639, −0.143]	0.002
SST	75th	−0.371	0.126	[−0.621, −0.122]	0.004
CP	25th	−0.386	0.130	[−0.643, −0.129]	0.003
CP	75th	−0.379	0.130	[−0.635, −0.122]	0.004

¹ Difference between MS cases and healthy controls expressed as number of standard deviations in healthy controls; ² standard error; ³ confidence interval; ⁴ adjusted for sex, age at blood sampling day, kinship effect, and shared environment effect.

The five proteins were further explored evaluating protein level differences (sex- and age-adjusted) between the MS course classifications. Relapse-remitting MS (RRMS) and secondary progressive MS (SPMS) were compared to healthy controls (HC), and RRMS and SPMS cases were compared to each other. Results are reported in Table 4 and the number of subjects included in these analyses are summarized as follows (omitting subjects with missing age at blood sampling day):

- RRMS vs. HC comparisons comprised 181 subjects (41 RRMS cases and 140 healthy controls)
- SPMS vs. HC comparisons comprised 158 subjects (18 SPMS cases and 140 healthy controls)
- SPMS vs. RRMS comparisons comprised 59 subjects (18 SPMS cases and 41 RRMS cases)

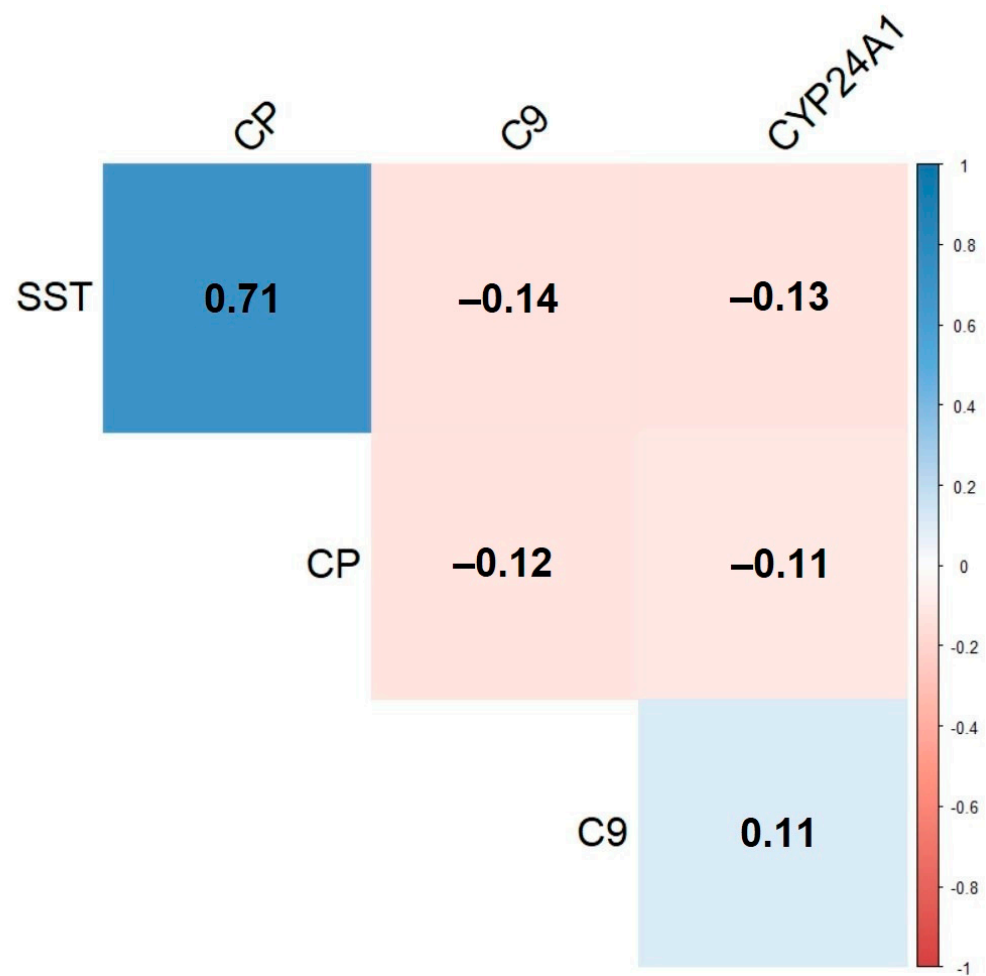


Figure 3. Pearson pairwise correlations between C9, CYP24A1, SST, CP, and PLAT plasma protein levels.

C9 protein levels were significantly higher in both RRMS and SPMS cases compared to healthy controls ($p < 0.05$). On average, plasma C9 levels were, respectively, 0.44 HC SD and 0.62 HC SD higher in RRMS cases and SPMS cases compared to healthy controls. Both estimates were close to the estimate obtained considering overall MS cases (0.53 HC SD, see Table 3), with SPMS cases showing the highest differences.

Estimated median CYP24A1 protein levels resulted 0.46 HC SD higher ($p < 0.05$) in RRMS cases compared to healthy controls, while in SPMS cases the difference was lower, i.e., 0.32 HC SD ($p > 0.05$). As for C9, estimates were close to the estimate obtained considering overall MS cases (0.42 HC SD, see Table 3), with RRMS cases showing the highest difference. Boxplots are reported in Figure 4 for RRMS vs. HC and SPMS vs. HC comparisons in C9 and CYP24A1 plasma proteins. Regarding suggestive proteins SST and CP protein levels, SPMS cases' protein levels in comparison with healthy controls showed a higher difference (in the median) compared to RRMS cases vs. healthy control comparison (-0.51 vs. -0.28 HC SD for SST, -0.40 vs. -0.28 HC SD for CP), with the former being statistically significant ($p < 0.05$). On the contrary, plasma PLAT levels' median was significantly lower in RRMS cases compared to healthy controls (-0.44 HC SD, $p < 0.05$), with SPMS vs. HC PLAT difference being slightly lower in magnitude (-0.35 , $p > 0.05$). Due to small sample size, i.e., $n = 56$, all comparisons between RRMS cases and SPMS cases resulted in estimated protein level differences being close to zero and in estimated large standard errors, making these comparisons not interpretable.

Table 4. Two-group comparison within MS course classifications for statistically significant and suggestive plasma biomarkers.

Protein	Chromosome	HPA ¹	Model ²	Comparison	Estimate ³	SE ⁴	95% CI ⁵	p-Value ⁶
C9	5	029577	LMM	RRMS vs. HC	0.441	0.174	[0.100, 0.782]	0.011
			LMM	SPMS vs. HC	0.624	0.229	[0.176, 1.073]	0.006
			LQMM	SPMS vs. RRMS	0.187	0.308	[−0.431, 0.807]	0.545
CYP24A1	20	022261	LQMM	RRMS vs. HC	0.457	0.156	[0.151, 0.766]	0.004
			LQMM	SPMS vs. HC	0.317	0.217	[−0.111, 0.746]	0.146
			LMM	RRMS vs. SPMS	−0.090	0.319	[−0.716, 0.536]	0.778
SST	3	019472	LQMM	RRMS vs. HC	−0.284	0.152	[−0.583, 0.015]	0.063
			LQMM	SPMS vs. HC	−0.507	0.173	[−0.849, −0.166]	0.004
			LQMM	RRMS vs. SPMS	−0.065	0.201	[−0.468, 0.338]	0.748
CP	3	001834	LQMM	RRMS vs. HC	−0.281	0.177	[−0.630, 0.068]	0.114
			LQMM	SPMS vs. HC	−0.395	0.141	[−0.673, −0.117]	0.006
			LQMM	RRMS vs. SPMS	0.011	0.255	[−0.501, 0.523]	0.967
PLAT	8	003412	LMM	RRMS vs. HC	−0.441	0.144	[−0.723, −0.159]	0.002
			LMM	SPMS vs. HC	−0.346	0.182	[−0.704, 0.012]	0.058
			LMM	RRMS vs. SPMS	0.221	0.321	[−0.407, 0.850]	0.491

¹ Human Protein Atlas antibody product name (human platelet antigen); ² LMM = linear mixed model, LQMM = linear quantile mixed model; ³ For RRMS vs. HC and SPMS vs. HC comparison: difference between MS cases and healthy controls expressed as number of standard deviations in healthy controls. For SPMS vs. RRMS: difference between SPMS cases and RRMS expressed as number of standard deviations in RRMS cases. Estimate is based on the mean for LMM models and on the median for LQMM models; ⁴ standard error; ⁵ confidence interval; ⁶ adjusted for sex, age at blood sampling day, kinship effect, and shared environment effect; RRMS = relapse-remitting MS; SPMS = secondary progressive MS; HC = healthy controls.

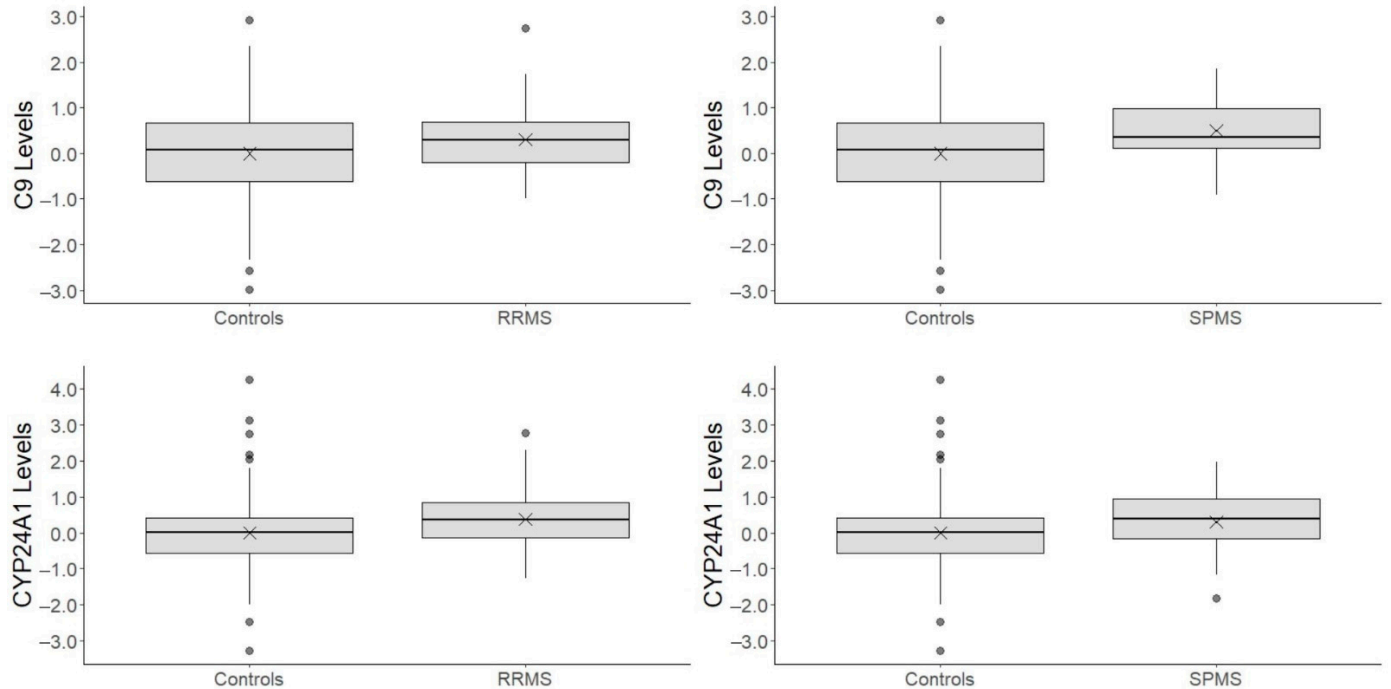


Figure 4. Box plots for C9 and CYP24A1 plasma protein level differences between relapse-remitting MS (RRMS) and secondary progressive MS (SPSM) compared to healthy controls. Plasma protein levels were standardized using mean and standard deviation from healthy controls’ protein levels; thus, 1 unit on the y-scale represents 1 standard deviation in healthy controls’ protein levels. “X” represents the mean protein level within a group.

2.3. Results of Plasma Protein Levels Variability Explained by MS-Risk Variants

A total of 92 unaffected Sardinian controls (48 female and 44 males) from 39 sub-families were included in the subsequent analysis but one subject was omitted due to missing age at the day of blood sampling. The aim of this second analysis was to quantify the variability in protein levels in plasma, explained by well-known MS-risk genetic variants.

To this aim we firstly identified among a set of 139 MS-risk SNPs, as from the IMSCG study [27], those associated with the plasma levels of the two highlighted proteins (i.e., C9 and CYP24A1) in our sample. Of these 139 MS genetic risk variants, 119 were included in the analysis, 4 SNPs were removed due to linkage disequilibrium ($r^2 > 0.2$) and 16 SNPs were removed due to minor allele frequency (MAF), < 0.10 in our sample of 91 healthy controls. Firstly, 119 univariate models as in Equation (3) were performed for C9 and CYP24A1, for a total of 238 statistical tests. Secondly, among the statistically significant SNP–protein associations at $\alpha = 0.10$, the best set of SNPs were selected using a stepwise regression procedure. Subsequently, a multivariable LMM model, which includes the best set of SNPs and which adjusts for age at blood sampling day, sex, kinship effect, and shared environment effect, was fitted to estimate the marginal proportion of protein level variability jointly explained by significant SNPs at $\alpha = 0.01$ (R^2_{SNPs}); a 95% confidence interval was estimated as (bias-corrected accelerated) BCa interval, at $\alpha = 0.05$, simulating 1000 bootstrap replications as explained in the Materials and Methods section. In Table 5 the results for R^2_{SNPs} , along with a 95% confidence interval and significant SNPs at $\alpha = 0.01$, are reported. A total of 16% (95% CI: 9%, 39%) of plasma C9 level variability was jointly explained by two MS-risk SNPs, i.e., rs2104286 and rs1610555; their effect sizes have been estimated, respectively, as 0.61 (95% CI: 0.32, 0.90) and 0.32 (95% CI: 0.08, 0.56) HC SD, meaning that a single effect allele (the minor allele) was associated with an average increase of 0.61/0.32 HC SD, controlling for sex, age, kinship effect, shared environment effect, and the other SNPs included in the multivariable model. Since the model is additive and linear, a recessive homozygote genotype would then lead to doubled increases in C9 plasma levels, i.e., 1.22 and 0.64 HC SD.

As to the plasma CYP24A1 level, 26% (95% CI: 10%, 44%) of its variability was jointly explained by three MS-risk SNPs, i.e., rs1870071, rs11567694, and rs7665090, with the first explaining 12% of the total variability; their estimated effect sizes were, respectively, 0.64 (95% CI: 0.33, 0.96), 0.55 (95% CI: 0.23, 0.88) and -0.35 (95% CI: -0.60 , -0.11) HC SD. Thus, a portion of both C9 and CYP24A1 plasma level variability was significantly explained by MS-risk SNPs.

We have also explored R^2_{SNPs} in the suggestive, but not significant after multiple testing correction, proteins found in the first analysis (see Section 2.2.2.), i.e., SST, CP, and PLAT (see Table 5). Plasma SST levels were not associated significantly with any MS-risk SNPs in the multivariable model. As to CP plasma levels, 16% (95% CI: 6%, 33%) of their variability was explained by three MS-risk SNPs, each of them explaining $\approx 5\%$ of the total variability. Interestingly, for PLAT plasma levels, R^2_{SNPs} were the highest among the five proteins analyzed, with almost one-third (32%, 95% CI: 13%, 49%) of plasma level variability explained by five significant MS-risk SNPs, among which rs11567694 explained the largest part (12%), with its effect allele resulting in an estimated PLAT level increase of 0.63 HC SD (95% CI: 0.24, 0.91). For all the five proteins, the 119 univariate and respective multivariate SNP–protein level associations are reported in Supplementary Table S3.

Table 5. Marginal proportion of protein level variability explained by significant MS-risk SNP allele additive effect.

Protein	Chr ¹	HPA ²	N° SNPs ³	SNP ⁴	SNP Position (chr:bp) ⁵	Effect Allele ⁶	MAF ⁷	Additive Effect ⁸	SE ⁹	p-Value ¹⁰	R ² _{SNP} ¹¹	R ² _{SNPs} ¹²	95% CI ¹³
C9	5	029577	5	rs2104286	10:6099045	C (T)	0.19	0.61	0.15	<0.001	0.10	0.16	[0.09–0.39]
				rs1610555	18:67543147	T (G)	0.32	0.32	0.12	0.009			
				rs1870071	19:16505106	C (T)	0.21	0.64	0.16	<0.001	0.12		
CYP24A1	20	022261	4	rs11567694	5:35857704	G (A)	0.23	0.55	0.17	<0.001	0.09	0.26	[0.10–0.44]
				rs7665090	4:103551603	G (A)	0.46	−0.35	0.12	0.004	0.05		
SST	3	019472	2	-	-	-	-	-	-	-	-	-	-
CP	3	001834	5	rs2082881	2:25038268	G (A)	0.24	0.43	0.15	0.004	0.06	0.16	[0.06–0.33]
				rs756699	5:133446575	C (T)	0.13	0.54	0.20	0.006	0.05		
				rs17886724	17:40496163	G (A)	0.35	−0.32	0.12	0.009	0.05		
PLAT	8	003412	8	rs11567694	5:35857704	G (A)	0.23	0.63	0.14	<0.001	0.11	0.32	[0.13–0.49]
				rs3766374	1:160720554	A (G)	0.21	−0.49	0.14	<0.001	0.08		
				rs12086448	1:160393905	G (A)	0.43	0.39	0.11	<0.001	0.06		
				rs212397	6:159474624	C (A)	0.27	−0.32	0.12	0.008	0.05		
				rs9947399	18:56271544	G (A)	0.47	0.29	0.11	0.009	0.04		

¹ Chromosome position of coding gene; ² Human Protein Atlas antibody product name (human platelet antigen); ³ number of SNPs, selected in the stepwise procedure, included in the multivariable SNP–protein level model; ⁴ SNPs significantly associated with protein levels at $\alpha = 0.01$ in the multivariable model; ⁵ SNP position based on human genome 19; ⁶ the effect allele is represented by the minor allele in our ImmunoChip data. The allele between brackets is the reference allele; ⁷ minor allele frequency; ⁸ additive effect due to one effect allele on protein levels resulting from the multivariable SNP–protein level model (controlling for sex, age, kinship effect, shared environment effect, and the other SNPs included in the multivariable model); ⁹ standard error; ¹⁰ p-value for null hypothesis of additive effect equal to 0; ¹¹ proportion of protein levels variability explained by the additive effect of the specific significant SNP; ¹² proportion of protein levels variability jointly explained by additive effects of significant SNPs at $\alpha = 0.01$; ¹³ 95% confidence interval.

3. Discussion

MS is a heterogeneous disease, therefore reliable biomarkers that can capture the different aspects of this heterogeneity are required to understand its mechanism and pathophysiology, especially to diagnose, predict prognosis and response to therapies, and not least to develop new therapeutic strategies [22].

Our study tried to contribute to MS biomarkers research, specifically in blood plasma, where a set of well-established biomarkers has yet to be found and the research is still ongoing. In this work we have highlighted two proteins in plasma, i.e., C9 and CYP24A1, that could be regarded as putative diagnostic biomarkers for MS as their level in plasma was increased in affected subjects as compared to controls, which may serve as a good basis to extend the current understanding of the disease and the literature regarding biomarkers for MS in plasma. These findings could help to establish a set of biomarkers which could jointly discriminate between MS-affected and unaffected subjects with considerable accuracy [12]. C9 and CYP24A1 plasma levels were approximately 0.5 HC SD higher compared to healthy controls, therefore showing a clinically meaningful but “moderate” change [31], in line with results from previous studies searching for MS biomarkers from plasma protein levels [12–18].

Particularly, C9 plasma protein levels in MS cases were 0.53 HC SD higher (95% CI: 0.25, 0.81) compared to healthy controls, with a slightly higher difference found in SPMS cases compared with healthy controls (i.e., 0.62 HC SD, 95% CI: 0.18, 1.07). C9 is a constituent of the membrane attack complex (MAC); it plays a key role in the innate and adaptive immune response promoting phagocytic/endocytic elimination of non-self-antigens and immune complexes by inducing apoptosis [33]. In demyelinating disorders, such as MS, MAC could play a dual role, neurodegenerative or neuroprotective. Specifically, complement components, which constitute the MAC, can originate in plasma after the disruption of the blood brain barrier (BBB) or they can be synthesized locally in CNS cells, where they promote neuroinflammation and demyelination during the acute phase; on the other hand, during the chronic phase of the disease, they provide neuroprotection by promoting axon preservation, oligodendrocyte survival, and remyelination [34]. Therefore, the biological background supports our finding of increased C9 plasma levels in MS cases due to a systemic production of complement components consequent to a neuroinflammation state. Previous studies have investigated C9 level differences between MS cases and healthy controls in both CSF and plasma.

While decreased C9 levels in CSF were found in different studies [35–37], Ingram et al. [35] was the only study to evaluate C9 plasma levels with a consistent sample size. The study [35] focused on the analysis of plasma complement components’ levels differences between MS cases and healthy controls, driven by a previous investigation which highlighted increased levels of complement components in MS cases’ plasma blood [38–40]. Their results showed a significant decrease for C9 plasma levels in both RRMS, SPMS cases, and overall MS cases, but at the same time significantly increased plasma levels of all the other complement proteins analyzed, i.e., C3 (where a difference between MS cases and healthy controls did not result significantly in our sample), C4, C4a, C1 inhibitor, and factor H were highlighted. Therefore, given that complement protein levels have been suggested as potential plasma biomarkers due to their increase in MS cases [35,38–40], a hint that is also supported by the biological background as described before [34], the finding in [35] does not necessarily contradict our result, but leads to the need to pinpoint plasma complement components levels, and therefore also plasma C9 levels, as candidate plasma MS biomarkers which should be further investigated in confirmatory studies. CYP24A1 plasma levels were higher, 0.42 HC SD (95% CI: 0.19, 0.65), as compared to healthy controls, with similar differences found in RRMS cases and SPMS cases comparisons with healthy controls. This mitochondrial protein initiates the degradation of calcitriol, the physiologically active form of vitamin D3, by hydroxylation of the side chain. In regulating the level of vitamin D3, this enzyme has a vital role in vitamin D metabolism [41]. Vitamin D deficiency has been highlighted as a risk factor for developing MS, and it was proposed primarily to explain

the observation of a latitude gradient in MS prevalence. To support vitamin D deficiency as an MS risk-factor, MS-risk genetic polymorphisms have been discovered near to genes encoding key enzymes in vitamin D metabolism, and a low vitamin D status have been associated with a high risk of developing MS, high risk of MS relapses, and high level of MS-related disability [42]. Ramasamy et al. [43] found MS-risk SNP rs2248359, listed among the 200 prioritized MS-risk SNPs in the IMSCG meta-analysis study, to be strongly associated with increased gene expression of CYP24A1 in the frontal cortex, while two additional studies linked MS-risk SNPs rs2248137, located in CYP24A1 gene, to decreased serum vitamin D levels [41,44]. These findings led to the speculation that MS-risk genetic variants in the CYP24A1 gene could result in an increased expression or activity of this enzyme, leading to low circulating vitamin D levels. This claim is consistent with our finding of higher plasma CYP24A1 levels in MS cases compared to healthy controls, which could represent a risk factor, due to genetics or other unknown causes.

SST, CP, and PLAT have been selected among the suggestive, but not statistically significant after multiple correction, plasma protein levels different in MS cases compared to healthy controls. Therefore, it is important to underline that these discoveries are not controlled at fixed I type error (0.05) after multiple testing correction but could still give meaningful validations, about previous findings, and new unexplored insights.

SST plasma levels were lower, 0.37 HC SD (95% CI: -0.62 , -0.12), in MS cases compared to healthy controls, with a higher decrease found in SPMS cases compared to healthy controls (i.e., -0.51 HC SD, 95% CI: -0.85 , -0.17). SST is a growth hormone inhibitory peptide which plays a prominent role in several pathological conditions, including inflammation, diabetes mellitus, Alzheimer's, Parkinson's, and acquired immunodeficiency syndrome [45]. In previous studies, SST levels in MS patients' CSF were significantly decreased compared to healthy controls, and this decrease has been linked to diminished cognitive function and increased cytokine levels [46,47]. Moreover, SST has been linked to control cytokines and cytokine-related pathways, including toll-like receptor4 and nuclear factor-kappa B, which led to the hypothesis that SST might play a role in inflammation inducing damage to the BBB [45]. Considering the results from our sample, plasma SST levels were also decreased in MS cases compared to healthy controls besides CSF findings mentioned previously, leading to the possibility of investigating this potential biomarker in a less invasive way.

CP was lower, 0.38 HC SD (95% CI: -0.64 , -0.13), compared to healthy controls, with similar differences found in RRMS cases and SPMS case comparisons with healthy controls. The plasma CP level was highly negatively correlated with SST ($r = -0.71$). CP is an acute phase protein which plays a fundamental role in copper (Cu) and Fe metabolism; it has been shown, in in vivo studies, to be an effective antioxidant in the CNS protecting neural cells against oxidative stress [48]. Our result is in conflict with previous findings which highlighted increased CP serum levels in MS cases compared to healthy controls, attributing this increase to the inability to incorporate CP due to the highly oxidative environment found in the serum of MS patients [49,50]. Nevertheless, decreased serum CP levels have been detected in Parkinson and Alzheimer's disease cases compared to healthy controls, indicating that this deficiency may contribute to neuronal cell death due to an iron-mediated increase in free radicals [51–53]. Therefore, further studies are needed to clarify this unexpected finding.

Lastly, PLAT, or tPa, protein levels in MS cases were estimated on average lower, 0.32 HC SD (95% CI: -0.54 , -0.10), as compared to healthy controls, with a higher decrease found in RRMS cases compared to healthy controls (i.e., -0.44 HC SD, 95% CI: -0.72 , -0.16). PLAT is a secreted serine protease that converts the proenzyme plasminogen to plasmin, a fibrinolytic enzyme [54]. This protease, expressed in neurons throughout the developing and the adult nervous systems [55], has pleiotropic roles in the CNS including neuronal migration, axonal growth, and synaptic plasticity, as well as a contribution to neurodegeneration in pathological states and the regulation of neurovascular and neuronal response [56–58]. PLAT levels in CSF were found to be highly increased in MS cases

compared to healthy controls, while, in plasma, PLAT levels were lower in MS cases (but not statistically significantly) [59]. In another study [60] PLAT levels in CSF were instead found to be significantly decreased in SPMS cases compared to healthy controls. Interestingly, and consistently, plasma plasminogen activator inhibitor-1 (PAI-1) levels were found to be significantly higher in MS case controls, specifically RRMS cases, consistently with our finding of a higher plasma PLAT decrease found in RRMS cases compared to healthy controls (as described previously) [61]. An imbalance in the PLAT–PAI-1 ratio has been previously highlighted to play a role in compromising axonal integrity [62]. In fact, decreased PLAT activity, due to increased PAI-1 levels, impairs the capacity of PLAT to clear the fibrinogen deposits (fibrinolysis), at sites of BBB breakdown, demyelinated axons, and inflammatory MS lesions, therefore contributing to axonal damage [63,64]. These results are therefore consistent with the estimated lower plasma PLAT protein levels in MS cases found in our sample, and further highlight the potential risk role played by an imbalanced PLAT–PAI-1 ratio in the neurodegeneration.

In the subsequent analysis, for the identified candidate biomarkers, the variability explained by a set of well-established MS-risk SNPs was quantified, with the aim of improving the functional understanding of GWASs results, helping understanding disease etiology, and gaining more evidence in favor of the identified biomarkers. The analysis was performed on healthy controls only, to avoid MS status exerting potential reverse causation on plasma protein levels. This analysis revealed a significant portion of C9 and CYP24A1 variability explained by, respectively, two and three MS-risk SNPs, i.e., 16% (95% CI: 9%, 39%) for C9 plasma levels and 26% (95% CI: 10%, 44%) for CYP24A1 plasma levels.

Suggestive additional biomarkers SST, CP, and PLAT plasma level variability were also explored. While plasma SST levels were not significantly associated with any MS-risk SNPs, 16% (95% CI: 6%, 33%) of CP plasma protein level variability was jointly explained by three MS-risk SNPs. Interestingly, almost one-third of PLAT plasma level variability was jointly explained by five MS-risk SNPs, i.e., 32% (95% CI: 13%, 49%).

In summary, we found four different proteins (C9, CYP24A1, CP, and PLAT), with a potential role in MS susceptibility as highlighted previously from different studies, whose plasma level variability was significantly explained ($\geq 16\%$) by well-established MS-risk SNPs. Therefore, these results could help explain the possible role of these genetic variants, for which, to date, the causal pathway of disease susceptibility is largely unknown.

The study presented strengths and limitations. The strength for the first analysis, where plasma protein level differences were evaluated between MS cases and healthy controls, was the high power achieved with our sample size, which made the results more reliable, while potential limitations were represented by: (1) the lack of concomitant treatment data, taken (if any) by MS cases at the time of blood sampling, which could have potentially affected plasma protein levels [19,65], even though the literature on this topic is still not widely explored; (2) the lack of validation of plasma protein level measurements with other methodologies (such as Western blot). The second analysis, where biomarker level variability, explained by a set of well-known MS-risk SNPs, was quantified, was limited by a lower sample size ($N = 91$) compared to the first analysis, which led to a lower power to detect significant signals and potentially leading to inflated effect sizes in significant discoveries (winner's curse). Nevertheless, unlike GWASs and large-scale QTL analyses, where millions of genetic variants are tested, only a small subset of SNPs were considered and tested. Therefore, this second analysis represented a first exploratory, in a colocalization perspective, attempt to link MS-risk SNPs to MS-related plasma protein levels which could serve as a starting point for further investigations with larger samples.

Together our results represent an important step toward the identification of candidate diagnostic biomarkers in plasma for MS. Even if a future dedicated assay is required to analytically validate our results and to extend them in clinical practice, our study represents the first step in defining the role of circulating proteins towards a rational therapeutic strategy choice, other than increasing the comprehension of the mechanisms underlying the disease. Furthermore, the identification of reliable peripheral, easy to obtain, candidate

biomarkers for MS, to be further investigated in forthcoming studies, bears the potential to partially substitute the current invasive lumbar puncture procedure for obtaining CSF and for an improved diagnosis of MS, monitoring the disease activity and progression, and evaluation of treatment responses.

4. Materials and Methods

4.1. Sample Collection and Genotyping

A register of MS cases diagnosed according to Poser's criteria [66] was established in Sardinia's Nuoro province, Italy, in 1995. From this register, data from 20 Sardinian families with 96 MS cases and 437 healthy controls for a total of 533 subjects were available. A genealogical questionnaire was used to reconstruct the pedigrees from grandparents to first-degree cousins. Subjects affected with MS were classified as RRMS or SPMS disease course.

Genotyping data were obtained using Immunochip [67] from a previous study [68], where the quality control-filtered dataset included 131,497 SNPs.

R Studio, version 1.4.1106, has been used for all analyses.

4.2. Measuring Plasma Protein Levels

Proteins in plasma were studied in a high throughput manner by using the large resource of polyclonal antibodies made available by the HPA project (www.proteinatlas.org/, accessed on 12 July 2021). A total of 56 plasma protein levels were obtained using a bead-based antibody array format, consisting of human platelet antigen antibodies immobilized onto microspheres in suspension. Through a setup that profiles proteins via direct labelling of whole and unfractionated samples with biotin, targets could be analyzed in a multiplexed manner using antibody-conjugated bead arrays of a desired composition with instant data acquisition from the flow cytometer-like Flexmap 3D instrument. To determine relative signal intensities from the binding of antibodies to their target antigens, median fluorescent intensities are displayed when counting at least 50 events per bead ID. The obtained data were processed using probabilistic quotient normalization over the entire dataset to account for possible differences in sample dilution. Finally, 56 antibodies, targeting 56 unique proteins (listed in Supplementary Table S1), selected from suggestions by the clinical collaborators and from proteomic, mRNA, and cDNA expression data found in the literature for their suggested association with MS, both functionally and as a biomarker, were included for the analysis.

4.3. Differences in Protein Levels

The first step of our analysis was to search for significant and potentially clinically important differences in protein levels between MS cases and healthy controls.

Since subjects were related, differences in protein levels between MS cases and healthy controls in the 212 subjects (69 MS cases, 143 healthy controls), having measured protein levels, were evaluated using a linear mixed model (LMM) formulated as:

$$Y_{ij} = \beta_0 + MS_{ij} * \beta_1 + SEX_{ij} * \beta_2 + AGE_{ij} * \beta_3 + Z_{1i} * kinship_i + Z_{2i} * family_i + e_{ij} \quad (1)$$

where j denotes the individual and i the corresponding family, Y_{ij} is the standardized protein level (using mean and standard deviation from healthy control, as explained below), β_0 is the intercept term, MS_{ij} is the MS status (reference = controls) with β_1 the corresponding fixed effect, SEX_{ij} is the sex of the individual (reference = males) with β_2 the corresponding fixed effect, AGE_{ij} is the age of the individual at the day of the blood sampling (obtained as the difference of the day of blood sampling and date of birth) with β_3 the corresponding fixed effect. $kinship_i$ was the random effect accounting for a familiar relationship distributed as $N(0, \sigma^2_{GA})$, where A was the kinship matrix multiplied by 2 (or the relationship matrix). $family_i$ was the random effect accounting for the shared environmental effect with other members of the sub-family effect distributed as $N(0, \sigma^2_{CH})$, where H is the matrix with value "1" for the individuals belonging to

the same family, and e_{ij} was the residual error assumed to be distributed as $N(0, \sigma^2_I)$. Z_{1i} and Z_{2i} denote the random effect model matrices for kinship_{*i*} and family_{*i*}. σ^2_G and σ^2_C were assumed to be independent. Sex and age at blood sampling were included in the model as fixed effects to avoid potential confounding. The female-to-male MS prevalence ratio in the province of Nuoro (Sardinia, Italy) was reported to be 2:1 [24], a result in line with the worldwide estimate of 2.3–3.5:1 [69]. The association between sex and immune response level was explained in depth in the review by Klein SL and Flanagan [70], where they noted: “sex is one variable that influences innate and adaptive immune responses, resulting in sex-specific outcomes from infectious and autoimmune diseases, malignancies, and vaccines;” this claim was further supported by other subsequent studies [71,72]. Age is also considered a potential confounder for immune protein levels and MS association, as highlighted in a review of neuronal and glial CSF biomarkers in MS [73]; moreover, Lind et al. highlighted the magnitude of plasma proteins changes in adults during a 10-year follow-up, with 61 out of 84 changing significantly [74]. Thus, these insights justified the inclusion of sex and age at blood sampling in the model. `relmatLmer` function, from `lme4qtl` R package [75], was used to fit LMM using the maximum likelihood (ML) method. The inference for MS fixed effect β_1 was based on Wald test statistic [76].

Since protein levels were rarely normal distributed, but rather right-skewed, using LMM could have led to incorrect inference in the presence of the non-normality of residual distribution. Moreover, an average protein level difference may have masked stronger or weaker differences that may have existed at other points of the distribution. The normality of residuals obtained from LMM was evaluated following guidelines from Kim [32], i.e., in our sample, we rejected the null hypothesis of normality of residuals for an absolute *z*-value over 3.29 for skewness or excess kurtosis statistics. If the null hypothesis of the normality of residuals assumption was rejected, an LQMM, as formulated by Geraci and Bottai, was used instead [77]. The model is based on quantile regression (QR), a methodology which extends regression for the mean to the analysis of the entire conditional distribution of the outcome variable and does not make assumptions about the model residuals, thus providing a complete picture of the distributional effects using maximum likelihood methods. The model follows the same equation as Equation (1) with the exception that the protein level differences in MS cases and healthy controls were evaluated at the 50th quantile τ (i.e., the median). The 25th and 75th quantiles were also explored for proteins where statistical significance for the median was achieved. The β_1 standard error $se(\hat{\beta}_1)$ estimate was based on block-bootstrap; $B = 1000$ bootstrap samples were obtained by resampling the $i = 1, \dots, M$ sub-families with replacement, and the standard deviation of B distribution $sd(B)$ was used as standard error estimate $se(\hat{\beta}_1)$. Parametric 95% confidence intervals are reported based on the *t*-value statistic. [78]. The `lqmm` function from `lqmm` R package [79] was used to fit LQMM. Additional statistical details on LQMM can be found in Supplementary Materials S1.

Both LMM and LQMM were fitted on centered and standardized protein levels, using the mean and SD of protein levels in healthy controls, to get a better interpretability of the estimated coefficients. Thus, for a specific protein, an estimated coefficient β_1 would translate as an increase/decrease in MS case protein level (mean or median, depending on the model used) equal to β_1 times the SD of the protein levels in healthy controls. This interpretation gave direct and more understandable evidence of the magnitude of protein level differences between MS cases and healthy controls.

Once *p*-values were obtained both for LMM or LQMM, these were corrected to avoid type I error inflation due to multiple testing (i.e., 56 statistical tests, one for each protein). The Holm procedure was used to provide strong control on FWER at the 0.05 level since it does not require the independence of the test statistics [80], which we could not assume since proteins were correlated both positively and negatively. Plasma protein differences between MS cases and healthy controls, which did not reach statistical significance but showed a *p*-value < 0.005 and an absolute difference of at least 0.3 HC SD, were still

considered potentially interesting proteins to be investigated. For these proteins, both significant and “suggestive,” Pearson’s correlation coefficients were calculated.

Finally, comparisons for plasma protein levels significant after multiple testing corrections were explored within MS course classifications, i.e., protein level differences were tested between RRMS cases and SPMS cases compared to healthy controls, as well as between SPMS cases and RRMS cases, using the same model in Equation (1).

4.4. Power Calculation

Prior to the evaluation of protein level differences on our sample, we estimated the accuracy of the methods described in Section 4.3. applied on our sample size and its pedigree structure. Therefore, the power to detect a fixed β_1 effect = 0.3 and 0.5 at significance level $\alpha = 0.05$ was evaluated using both LMM and LQMM. The same pedigree from our data was used to simulate protein levels. For LMM, the parameters were fixed as follows (all parameters are explained in Equation (1)):

$$\begin{aligned}\beta_0 &= 0 \\ \beta_1 &= 0.3, 0.5 \\ \beta_2 &= 0.05 \\ \beta_3 &= 0.01 \\ \sigma^2_G &= 0.15 \\ \sigma^2_C &= 0.10 \\ \sigma^2_I &= 1 - (\sigma^2_G + \sigma^2_C + \text{var}(\beta_1 * \text{MS} + \beta_2 * \text{SEX} + \beta_3 * \text{AGE}))\end{aligned}\tag{2}$$

β effects were in the scale of plasma protein levels SD as measured in healthy controls (HC SD), e.g., for $\beta = 0.5$ the plasma protein levels in MS cases were higher by a half SD as measured in healthy controls. The power to detect a plasma protein level difference of 0.5 HC SD is of interest because this value has been arbitrarily selected to represent the MCID, which is defined as the smallest the change in score of a patient-reported outcome (either beneficial or harmful) that is important from the patient or clinician’s perspective [31]. Even if this measure is used to evaluate treatment efficacy, we used it as a reference value to highlight a potentially meaningful physiological change. The power to detect a plasma level difference of 0.3 HC SD has been evaluated to verify our power to detect lower plasma protein levels differences.

Sex and age at blood sampling were already defined based on our pedigree, and their fixed effects β_2 and β_3 fixed to 0.05 and 0.01, i.e., female/male protein level differences have been simulated as 0.05 HC SD and each age unit increase (in years) has been associated with a protein level increase by 0.01 HC SD (or, equally, an increase in 10 years in age has been associated with a protein level increase of 0.10 HC SD). Since an estimate of age and sex effects on each plasma protein level could not be found in the literature, and the effects could differ between proteins, we have fixed, a priori, small β effects. Regardless, sex and age effects are expected to have a minor influence on power evaluation compared to the impact of the sample size alone.

MS was simulated as a random Bernoulli variable, i.e., $\text{MS} \sim \text{Bernoulli}(0.325)$, where the parameter $p = 0.325$ is the actual proportion of MS cases in our data.

kinship and family random effects (kinship and shared environment effects) were simulated as a random multivariate normal variable, i.e., $\text{kinship} \sim \text{MVN}(0, \sigma^2_G A)$, $\text{family} \sim \text{MVN}(0, \sigma^2_C H)$. We chose a 0.15 value for σ^2_G and 0.10 for σ^2_C following the results from Liu et al. [81], which estimated the mean proportion of heritability and shared environment effect for 372 plasma protein levels.

Residuals e were generated from a random normal distribution, i.e., $e \sim N(0, \sigma^2_I)$, where σ^2_I was simply calculated as the variance component that sums to 1 after considering the sum of σ^2_G , σ^2_C and the variance of the fixed-effects component (i.e., $\text{var}(\beta_1 * \text{MS} + \beta_2 * \text{SEX} + \beta_3 * \text{AGE})$). Protein levels were then generated 2000 times using the formula in Equation (1). The power to detect effect sizes $\beta_1 = 0.3$ and 0.5 was evaluated at $\alpha = 0.05$ using LMM.

Simulations to evaluate power using the LQMM model were conducted fixing the same parameters as the previous LMM simulation except for residuals e ; these were generated from a random asymmetric Laplace distribution $ALD(0, \sigma, \tau)$ with scale parameter σ fixed to 0.2 and τ , which stands for the respective quantile to be estimated, fixed to 0.5. In each LQMM simulation, $se(\beta_1)$ was estimated using 50 block-bootstrap replications (as defined in Section 4.3.).

4.5. MS-Risk SNPs-Protein Levels Associations

The second step of our analysis was to quantify plasma protein level variability explained by a set of well-known MS-risk SNPs. This analysis, performed on the protein levels, was significantly different between MS cases and healthy controls in the previous step of the analysis only. The list of MS-risk SNPs was obtained from [27], where 200 autosomal SNPs outside the major histocompatibility complex (MHC) region were prioritized as significantly and strongly suggestive of being associated with MS risk. From these 200 prioritized signals only 139 SNPs could be selected for our analysis as included in our ImmunoChip data.

This “naïve” approach did not have any intention to establish causality between protein levels and MS but represents solely an attempt to investigate the potential biological function of well-established MS-risk SNPs for which, to date, the causal pathway is still unknown.

The analysis was conducted on the 92 healthy controls having both proteins and ImmunoChip data. We excluded MS cases from the analysis since the aim was to obtain SNP–protein associations in regular healthy conditions, avoiding potential reverse causation due to a different physiological status caused by the disease. To avoid a lack of precision and/or type I error inflation due to a reduced sample size, we kept in our sample only variants having $MAF > 0.10$. Moreover, variants in linkage disequilibrium, considering a maximum threshold for the r^2 statistic equal to 0.2, were also removed. This caused the removal of 16 and 4 variants, respectively, leading to 119 SNPs being included in the analysis.

Since all 119 MS-risk SNPs could not be included in a single model, as parameters would outnumber observations, we first refined the search for MS-risk SNPs potentially associated with protein levels following the approach in [82,83]. First, each SNP was included as a covariate in a univariate LMM formulated as:

$$Y_{ij} = \beta_0 + SEX_{ij} * \beta_1 + AGE_{ij} * \beta_2 + SNP_{ij} * \beta_3 + Z_{1i} * kinship_i + Z_{2i} * family_i + e_{ij} \quad (3)$$

where all the variables were already defined in Equation (1) except for SNP_{ij} , which denotes the number of effect alleles (minor allele), with β_1 the respective additive linear effect on protein levels. Among the SNP–protein level associations significant at $\alpha = 0.10$, the best set of SNP markers were selected using the stepwise regression procedure [84], where the inclusion and exclusion of each SNP of the model was determined at the 0.05 level. The best set of SNPs was then included in a multivariable LMM model, formulated as in Equation (3), and SNPs significantly associated with protein levels at $\alpha = 0.01$ were selected to estimate the marginal proportion of protein level variability explained by significant SNPs. This measure has been calculated using the marginal R^2 statistic as defined by Nakagawa and Schielzeth [85]:

$$R^2_{SNPs} = \frac{\sigma^2_{SNPs}}{\sigma^2_F + \sigma^2_G + \sigma^2_C + \sigma^2_I} \quad (4)$$

where σ^2_I , σ^2_C , σ^2_G were defined as in Equation (1), σ^2_F was the variance for the fixed effects components (i.e., sex, age at blood sampling, and the set of SNPs included in the multivariate model) and σ^2_{SNPs} was the variance for the significant SNPs, at $\alpha = 0.01$, the fixed effects component. A marginal R^2_{SNP} statistic was also calculated separately for each significant SNP (the sum of each R^2_{SNP} giving R^2_{SNPs}). More details are provided in

Supplementary Materials S1. A 95% confidence interval for R^2_{SNPs} was calculated making use of BCa interval, at $\alpha = 0.05$, calculated on $B = 1000$ block-bootstrap replications (as defined in Section 4.3) [78].

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/life12020151/s1>, Supplementary Table S1: List of plasma proteins measured; Supplementary Table S2: Plasma proteins differences between MS cases and healthy controls from LMM/LQMM models; Supplementary Table S3: SNPs-protein levels associations from univariate and multivariate LMM models; data_sardinian: Dataset used for the analyses, comprising pedigree, age, disease course, protein levels and genotyping information; R Codes: Folder with R codes used for all analyses; Details of statistical analysis are reported in Supplementary Materials S1.

Author Contributions: Conceptualization, A.N., T.F., C.B. and L.B.; methodology, A.N., T.F., C.B. and L.B.; formal analysis, A.N. and T.F.; investigation, V.S. and M.P.; data curation, A.B. and J.L.M.; writing—original draft preparation, A.N., T.F. and L.B.; writing—review and editing, all authors; supervision, L.B.; project administration, L.B.; funding acquisition, L.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FISM—Fondazione Italiana Sclerosi Multipla (Cod. 2009//R//2) and by Fondazione Cariplo (grant no. 2009–2528).

Institutional Review Board Statement: The study was approved by the ethics committee of the Azienda Sanitaria of Nuoro and was conducted in conformity with the 1954 Declaration of Helsinki in its currently applicable version and applicable Italian laws.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement: Raw data was uploaded as Supplementary Materials.

Acknowledgments: We would like to thank Burcu Ayoglu and Peter Nillson from SciLifeLab, Dept. of Protein Science, KTH Royal Institute of Technology, Stockholm, for providing us protein data.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Gourraud, P.A.; Harbo, H.F.; Hauser, S.L.; Baranzini, S.E. The genetics of multiple sclerosis: An up-to-date review. *Immunol. Rev.* **2012**, *248*, 87–103. [CrossRef]
2. Nylander, A.; Hafler, D.A. Multiple sclerosis. *J. Clin. Investig.* **2012**, *122*, 1180–1188. [CrossRef] [PubMed]
3. Fazio, T.; Pastorino, R.; Notartomaso, S.; Busceti, C.; Imbriglio, T.; Cannella, M.; Gentilini, D.; Morani, G.; Ticca, A.; Bitti, P.; et al. Acid sensing ion channel 2: A new potential player in the pathophysiology of multiple sclerosis. *Eur J Neurosci.* **2019**, *49*, 1233–1243. [CrossRef] [PubMed]
4. Milo, R.; Kahana, E. Multiple sclerosis: Geoepidemiology, genetics and the environment. *Autoimmun. Rev.* **2010**, *9*, A387–A394. [CrossRef]
5. Compston, A.; Coles, A. Multiple sclerosis. *Lancet* **2008**, *372*, 1502–1517. [CrossRef]
6. Bielekova, B.; Martin, R. Development of biomarkers in multiple sclerosis. *Brain* **2004**, *127*, 1463–1478. [CrossRef]
7. Atkinson, A.J.; Colburn, W.A.; DeGruttola, V.G.; DeMets, D.L.; Downing, G.J.; Hoth, D.F.; Oates, J.A.; Peck, C.C.; Schooley, R.T.; Spilker, B.A.; et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin. Pharmacol. Ther.* **2001**, *69*, 89–95. [CrossRef]
8. Hu, S.; Loo, J.A.; Wong, D.T. Human body fluid proteome analysis. *Proteomics* **2006**, *6*, 6326–6353. [CrossRef] [PubMed]
9. Uhlén, M.; Björling, E.; Agaton, C.; Szgyarto, C.A.K.; Amini, B.; Andersen, E.; Andersson, A.C.; Angelidou, P.; Asplund, A.; Asplund, C.; et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteom.* **2005**, *4*, 1920–1932. [CrossRef]
10. Templin, M.F.; Stoll, D.; Schrenk, M.; Traub, P.C.; Vöhringer, C.F.; Joos, T.O. Protein microarray technology. *Trends Biotechnol.* **2002**, *20*, 160–166. [CrossRef]
11. Jafari, A.; Babajani, A.; Rezaei-Tavirani, M. Multiple Sclerosis Biomarker Discoveries by Proteomics and Metabolomics Approaches. *Biomark. Insights* **2021**, *16*, 11772719211013352. [CrossRef] [PubMed]

12. Byström, S.; Ayoglu, B.; Häggmark, A.; Mitsios, N.; Hong, M.G.; Drobin, K.; Forsström, B.; Fredolini, C.; Khademi, M.; Amor, S.; et al. Affinity proteomic profiling of plasma, cerebrospinal fluid, and brain tissue within multiple sclerosis. *J. Proteome Res.* **2014**, *13*, 4607–4619. [[CrossRef](#)] [[PubMed](#)]
13. Rithidech, K.N.; Honikel, L.; Milazzo, M.; Madigan, D.; Troxell, R.; Krupp, L.B. Protein expression profiles in pediatric multiple sclerosis: Potential biomarkers. *Mult. Scler.* **2009**, *15*, 455–464. [[CrossRef](#)]
14. Ismail, S.; Engin, K.; Ozan, K.; Mustafa, C.; Banu, A. Analysis of plasma protein biomarkers in childhood onset multiple sclerosis. *J. Neuroimmunol.* **2020**, *348*, 577359. [[CrossRef](#)]
15. Malekzadeh, A.; Leurs, C.; van Wieringen, W.; Steenwijk, M.D.; Schoonheim, M.M.; Amann, M.; Naegelin, Y.; Kuhle, J.; Killestein, J.; Teunissen, C.E. Plasma proteome in multiple sclerosis disease progression. *Ann. Clin. Transl. Neurol.* **2019**, *6*, 582–1594. [[CrossRef](#)] [[PubMed](#)]
16. Huang, J.; Khademi, M.; Fugger, L.; Lindhe, Ö.; Novakova, L.; Axelsson, M.; Malmeström, C.; Constantinescu, C.; Lycke, J.; Piehl, F.; et al. Inflammation-related plasma and CSF biomarkers for multiple sclerosis. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 12952–12960. [[CrossRef](#)]
17. Tejera-Alhambra, M.; Casrouge, A.; De Andrés, C.; Seyfferth, A.; Ramos-Medina, R.; Alonso, B.; Vega, J.; Fernández-Paredes, L.; Albert, M.L.; Sánchez-Ramón, S. Plasma biomarkers discriminate clinical forms of multiple sclerosis. *PLoS ONE* **2015**, *10*, e0128952. [[CrossRef](#)]
18. D'Ambrosio, A.; Pontecorvo, S.; Colasanti, T.; Zamboni, S.; Francia, A.; Margutti, P. Peripheral blood biomarkers in multiple sclerosis. *Autoimmun. Rev.* **2015**, *14*, 1097–1110. [[CrossRef](#)]
19. Bedri, S.K.; Nilsson, O.B.; Fink, K.; Månberg, A.; Hamsten, C.; Ayoglu, B.; Manouchehrinia, A.; Nilsson, P.; Olsson, T.; Hillert, J.; et al. Plasma protein profiling reveals candidate biomarkers for multiple sclerosis treatment. *PLoS ONE* **2019**, *14*, e0217208. [[CrossRef](#)] [[PubMed](#)]
20. Polman, C.H.; Bertolotto, A.; Deisenhammer, F.; Giovannoni, G.; Hartung, H.P.; Hemmer, B.; Killestein, J.; McFarland, H.F.; Oger, J.; Pachner, A.R.; et al. Recommendations for clinical use of data on neutralising antibodies to interferon-beta therapy in multiple sclerosis. *Lancet Neurol.* **2010**, *9*, 740–750. [[CrossRef](#)]
21. Calabresi, P.A.; Giovannoni, G.; Confavreux, C.; Galetta, S.L.; Havrdova, E.; Hutchinson, M.; Kappos, L.; Miller, D.H.; O'Connor, P.W.; Phillips, J.T.; et al. The incidence and significance of anti-natalizumab antibodies: Results from AFFIRM and SENTINEL. *Neurology* **2007**, *69*, 1391–1403. [[CrossRef](#)] [[PubMed](#)]
22. Comabella, M.; Montalban, X. Body fluid biomarkers in multiple sclerosis. *Lancet Neurol.* **2014**, *13*, 113–126. [[CrossRef](#)]
23. Schwenk, J.M.; Gry, M.; Rimini, R.; Uhlén, M.; Nilsson, P. Antibody Suspension Bead Arrays within Serum Proteomics. *J. Proteome Res.* **2008**, *7*, 3168–3179. [[CrossRef](#)]
24. Urru, S.A.M.; Antonelli, A.; Sechi, G.M. Prevalence of multiple sclerosis in Sardinia: A systematic cross-sectional multi-source survey. *Mult. Scler. J.* **2020**, *26*, 372–380. [[CrossRef](#)]
25. Peltonen, L.; Palotie, A.; Lange, K. Use of population isolates for mapping complex traits. *Nat. Rev. Genet.* **2000**, *1*, 182–190. [[CrossRef](#)]
26. Varilo, T.; Peltonen, L. Isolates and their potential use in complex gene mapping efforts. *Curr. Opin. Genet. Dev.* **2004**, *14*, 316–323. [[CrossRef](#)]
27. Consortium, I.M.S.G. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **2019**, *365*, eaav7188. [[CrossRef](#)]
28. Gibson, G.; Powell, J.E.; Marigorta, U.M. Expression quantitative trait locus analysis for translational medicine. *Genome Med.* **2015**, *7*, 60. [[CrossRef](#)]
29. Nicolae, D.L.; Gamazon, E.; Zhang, W.; Duan, S.; Eileen Dolan, M.; Cox, N.J. Trait-associated SNPs are more likely to be eQTLs: Annotation to enhance discovery from GWAS. *PLoS Genet.* **2010**, *6*, e1000888. [[CrossRef](#)]
30. Fazia, T.; Nova, A.; Gentilini, D.; Beecham, A.; Piras, M.; Saggi, V.; Ticca, A.; Bitti, P.; McCauley, J.L.; Berzuini, C.; et al. Investigating the Causal Effect of Brain Expression of CCL2, NFKB1, MAPK14, TNFRSF1A, CXCL10 Genes on Multiple Sclerosis: A Two-Sample Mendelian Randomization Approach. *Front Bioeng Biotechnol.* **2020**, *8*, 397. [[CrossRef](#)]
31. Mouelhi, Y.; Jouve, E.; Castelli, C.; Gentile, S. How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. *Health Qual. Life Outcomes* **2020**, *18*, 136. [[CrossRef](#)]
32. Kim, H.-Y. Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restor. Dent. Endod.* **2013**, *38*, 52–54. [[CrossRef](#)]
33. Witzel-Schlömp, K.; Hobart, M.J.; Fernie, B.A.; Orren, A.; Würzner, R.; Rittner, C.; Kaufmann, T.; Schneider, P.M. Heterogeneity in the genetic basis of human complement C9 deficiency. *Immunogenetics* **1998**, *48*, 144–147. [[CrossRef](#)]
34. Rus, H.; Cudrici, C.; Niculescu, F. C5b-9 complement complex in autoimmune demyelination and multiple sclerosis: Dual role in neuroinflammation and neuroprotection. *Ann. Med.* **2005**, *37*, 97–104. [[CrossRef](#)] [[PubMed](#)]
35. Ingram, G.; Hakobyan, S.; Hirst, C.L.; Harris, C.L.; Loveless, S.; Mitchell, J.P.; Pickersgill, T.P.; Robertson, N.P.; Morgan, B.P. Systemic complement profiling in multiple sclerosis as a biomarker of disease state. *Mult. Scler. J.* **2012**, *18*, 1401–1411. [[CrossRef](#)]
36. Compston, D.; Morgan, B.P.; Olesky, D.; Fifield, R.; Campbell, A.K. Cerebrospinal fluid C9 in demyelinating disease. *Neurology* **1986**, *36*, 1503–1506. [[CrossRef](#)] [[PubMed](#)]
37. Morgan, B.P.; Campbell, A.K.; Compston, D.A.S. Terminal component of complement (C9) In cerebrospinal fluid of patients with multiple sclerosis. *Lancet* **1984**, *2*, 251–254. [[CrossRef](#)]

38. Ingram, G.; Hakobyan, S.; Hirst, C.L.; Harris, C.L.; Pickersgill, T.P.; Cossburn, M.D.; Lovless, S.; Robertson, N.P.; Morgan, B.P. Complement regulator factor H as a serum biomarker of multiple sclerosis disease state. *Brain* **2010**, *133*, 1602–161137. [[CrossRef](#)] [[PubMed](#)]
39. Ingram, G.; Hakobyan, S.; Robertson, N.P.; Morgan, B.P. Elevated plasma C4a levels in multiple sclerosis correlate with disease activity. *J. Neuroimmunol.* **2010**, *223*, 124–127. [[CrossRef](#)] [[PubMed](#)]
40. Sellebjerg, F.; Jaliashvili, I.; Christiansen, M.; Garred, P. Intrathecal activation of the complement system and disability in multiple sclerosis. *J. Neurol. Sci.* **1998**, *157*, 168–174. [[CrossRef](#)]
41. Agnello, L.; Scazzone, C.; Lo Sasso, B.; Ragonese, P.; Milano, S.; Salemi, G.; Ciaccio, M. CYP27A1, CYP24A1, and RXR- α Polymorphisms, Vitamin D, and Multiple Sclerosis: A Pilot Study. *J. Mol. Neurosci.* **2018**, *66*, 77–84. [[CrossRef](#)] [[PubMed](#)]
42. Simon, K.C.; Munger, K.L.; Ascherio, A. Vitamin D and multiple sclerosis: Epidemiology, immunology, and genetics. *Curr. Opin. Neurol.* **2012**, *25*, 246–251. [[CrossRef](#)]
43. Ramasamy, A.; Trabzuni, D.; Forabosco, P.; Smith, C.; Walker, R.; Dillman, A.; Sveinbjornsdottir, S.; Hardy, J.; Weale, M.E.; Ryten, M. Genetic evidence for a pathogenic role for the vitamin D3 metabolizing enzyme CYP24A1 in multiple sclerosis. *Mult. Scler. Relat. Disord.* **2014**, *3*, 211–219. [[CrossRef](#)] [[PubMed](#)]
44. Gil-Varea, E.; Matesanz, F.; Ferrer, R.; Spataro, N.; Fernández, O.; Sáiz, A.; Ramió-Torrentá, L.; Izquierdo, G.; Malhotra, S.; Navarro, A.; et al. The MS risk variant rs2762943 of the CYP24A1 gene is associated with decreased serum levels of the active form of vitamin D. In *Proceedings of the 34th Congress of the European-Committee-for-Treatment-and-Research-in-Multiple-Sclerosis (ECTRIMS)*; Sage Publications: Berlin, Germany, 2018; p. 831.
45. Basivireddy, J.; Somvanshi, R.K.; Romero, I.A.; Weksler, B.B.; Couraud, P.O.; Oger, J.; Kumar, U. Somatostatin preserved blood brain barrier against cytokine induced alterations: Possible role in multiple sclerosis. *Biochem. Pharmacol.* **2013**, *86*, 497–507. [[CrossRef](#)]
46. Roca, C.A.; Su, T.P.; Elpern, S.; McFarland, H.; Rubinow, D.R. Cerebrospinal fluid somatostatin, mood, and cognition in multiple sclerosis. *Biol. Psychiatry* **1999**, *46*, 551–556. [[CrossRef](#)]
47. Sarensen, K.V.; Alslev, T.; Christensen, S.E.; Jensen, N.B.; Orskov, H. Csf somatostatin in multiple sclerosis: Reversible loss of diurnal oscillation in relapses. *Neurology* **1987**, *37*, 1050–1053. [[CrossRef](#)] [[PubMed](#)]
48. Gutteridge, J.M.C. Lipid peroxidation and antioxidants as biomarkers of tissue damage. *Clin. Chem.* **1995**, *41*, 1819–1828. [[CrossRef](#)]
49. Siotto, M.; Filippi, M.M.; Simonelli, I.; Landi, D.; Ghazaryan, A.; Vollaro, S.; Ventriglia, M.; Pasqualetti, P.; Rongioletti, M.C.A.; Squitti, R.; et al. Oxidative stress related to iron metabolism in relapsing remitting multiple sclerosis patients with low disability. *Front. Neurosci.* **2019**, *11*, 13–86. [[CrossRef](#)] [[PubMed](#)]
50. De Riccardis, L.; Buccolieri, A.; Muci, M.; Pitotti, E.; De Robertis, F.; Trianni, G.; Manno, D.; Maffia, M. Copper and ceruloplasmin dyshomeostasis in serum and cerebrospinal fluid of multiple sclerosis subjects. *Biochim. Biophys. Acta Mol. Basis Dis.* **2018**, *1864*, 1828–1838. [[CrossRef](#)]
51. Gürer, R.; Aydın, Ş.; Gökçe, D.; Ayaz, T.; Işık, N. The Role of Ceruloplasmin in Neurodegeneration in Parkinson's Disease. *Turk. J. Neurol.* **2016**, *22*, 19–25. [[CrossRef](#)]
52. Jin, L.; Wang, J.; Zhao, L.; Jin, H.; Fei, G.; Zhang, Y.; Zeng, M.; Zhong, C. Decreased serum ceruloplasmin levels characteristically aggravate nigral iron deposition in Parkinson's disease. *Brain* **2011**, *134*, 50–58. [[CrossRef](#)] [[PubMed](#)]
53. Kessler, H.; Pajonk, F.G.; Meisser, P.; Schneider-Axmann, T.; Hoffmann, K.H.; Supprian, T.; Herrmann, W.; Obeid, R.; Multhaup, G.; Falkai, P.; et al. Cerebrospinal fluid diagnostic markers correlate with lower plasma copper and ceruloplasmin in patients with Alzheimer's disease. *J. Neural Transm.* **2006**, *113*, 1763–1769. [[CrossRef](#)] [[PubMed](#)]
54. Kruihof, E.K.O.; Dunoyer-Geindre, S. Human tissue-type plasminogen activator. *Thromb. Haemost.* **2014**, *112*, 243–254. [[CrossRef](#)]
55. Teesalu, T.; Kulla, A.; Simisker, A.; Sirén, V.; Lawrence, D.A.; Asser, T.; Vaheri, A. Tissue plasminogen activator and neuroserpin are widely expressed in the human central nervous system. *Thromb. Haemost.* **2004**, *92*, 358–368. [[CrossRef](#)] [[PubMed](#)]
56. Lee, T.W.; Tsang, V.W.K.; Birch, N.P. Physiological and pathological roles of tissue plasminogen activator and its inhibitor neuroserpin in the nervous system. *Front. Cell. Neurosci.* **2015**, *13*, 9–396. [[CrossRef](#)]
57. Moonen, G.; Grau-Wagemans, M.P.; Selak, I. Plasminogen activator-plasmin system and neuronal migration. *Nature* **1982**, *298*, 753–755. [[CrossRef](#)]
58. Yaamada, M.; Takaehashi, K.; Ukai, W.; Hashimoto, E.; Saito, T.; Yamada, M. Neuroserpin is expressed in early stage of neurogenesis in adult rat hippocampus. *Neuroreport* **2010**, *21*, 138–142. [[CrossRef](#)]
59. Akenami, F.O.T.; Koskiniemi, M.; Mustjoki, S.; Siren, V.; Färkkilä, M.; Vaheri, A. Plasma and cerebrospinal fluid activities of tissue plasminogen activator, urokinase and plasminogen activator inhibitor-1 in multiple sclerosis. *Fibrinolysis Proteolysis* **1997**, *11*, 109–113. [[CrossRef](#)]
60. Harris, V.K.; Diamanduros, A.; Good, P.; Zakin, E.; Chalivendra, V.; Sadiq, S.A. Bri2-23 is a potential cerebrospinal fluid biomarker in multiple sclerosis. *Neurobiol. Dis.* **2010**, *40*, 331–339. [[CrossRef](#)]
61. Onodera, H.; Nakashima, I.; Fujihara, K.; Nagata, T.; Itoyama, Y. Elevated plasma level of plasminogen activator inhibitor-1 (PAI-1) in patients with relapsing-remitting multiple sclerosis. *Tohoku J. Exp. Med.* **1999**, *189*, 259–265. [[CrossRef](#)]
62. Gverić, D.; Herrera, B.; Petzold, A.; Lawrence, D.A.; Cuzner, M.L. Impaired fibrinolysis in multiple sclerosis: A role for tissue plasminogen activator inhibitors. *Brain* **2003**, *126*, 1590–1598. [[CrossRef](#)]

63. Shaw, M.A.; Gao, Z.; McElhinney, K.E.; Thornton, S.; Flick, M.J.; Lane, A.; Degen, J.L.; Ryu, J.K.; Akassoglou, K.; Mullins, E.S. Plasminogen deficiency delays the onset and protects from demyelination and paralysis in autoimmune neuroinflammatory disease. *J. Neurosci.* **2017**, *37*, 3776–3788. [[CrossRef](#)] [[PubMed](#)]
64. Sillen, M.; Declerck, P.J. A narrative review on plasminogen activator inhibitor-1 and its (Patho)physiological role: To target or not to target? *Int. J. Mol. Sci.* **2021**, *22*, 2721. [[CrossRef](#)] [[PubMed](#)]
65. Adamczyk-Sowa, M.; Galiniak, S.; Zyracka, E.; Grzesik, M.; Naparło, K.; Sowa, P.; Bartosz, G.; Sadowska-Bartosz, I. Oxidative modification of blood serum proteins in multiple sclerosis after interferon beta and melatonin treatment. *Oxid. Med. Cell. Longev.* **2017**, *2017*, 7905148. [[CrossRef](#)]
66. Poser, C.M.; Paty, D.W.; Scheinberg, L.; McDonald, W.I.; Davis, F.A.; Ebers, G.C.; Johnson, K.P.; Sibley, W.A.; Silberberg, D.H.; Tourtellotte, W.W. New diagnostic criteria for multiple sclerosis: Guidelines for research protocols. *Ann. Neurol.* **1983**, *13*, 227–231. [[CrossRef](#)] [[PubMed](#)]
67. Consortium, I.M.S.G.; Beecham, A.H.; Patsopoulos, N.A.; Xifara, D.K.; Davis, M.F.; Kempainen, A.; Cotsapas, C.; Shah, T.S.; Spencer, C.; Booth, D.; et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat. Genet.* **2013**, *45*, 1353–1360.
68. Fazia, T.; Pastorino, R.; Foco, L.; Han, L.; Abney, M.; Beecham, A.; Hadjixenofontos, A.; Guo, H.; Gentilini, D.; Papachristou, C.; et al. Investigating multiple sclerosis genetic susceptibility on the founder population of east-central Sardinia via association and linkage analysis of immune-related loci. *Mult. Scler. J.* **2018**, *24*, 1815–1824. [[CrossRef](#)]
69. Harbo, H.F.; Gold, R.; Tintora, M. Sex and gender issues in multiple sclerosis. *Ther. Adv. Neurol. Disord.* **2013**, *6*, 237–248. [[CrossRef](#)]
70. Klein, S.L.; Flanagan, K.L. Sex differences in immune responses. *Nat. Rev. Immunol.* **2016**, *16*, 626–638. [[CrossRef](#)]
71. Ortona, E.; Pierdominici, M.; Rider, V. Editorial: Sex hormones and gender differences in immune responses. *Front. Immunol.* **2019**, *10*, 1076. [[CrossRef](#)]
72. Morris, G.P. Understanding sex-related differences in immune responses. *Sci. Transl. Med.* **2020**, *12*, eabd3631. [[CrossRef](#)]
73. Momtazmanesh, S.; Shobeiri, P.; Saghadzadeh, A.; Teunissen, C.E.; Burman, J.; Szalardy, L.; Klivenyi, P.; Bartos, A.; Fernandes, A.; Rezaei, N. Neuronal and glial CSF biomarkers in multiple sclerosis: A systematic review and meta-analysis. *Rev. Neurosci.* **2021**, *32*, 573–595. [[CrossRef](#)]
74. Lind, L.; Sundström, J.; Larsson, A.; Lampa, E.; Ärnlov, J.; Ingelsson, E. Longitudinal effects of aging on plasma proteins levels in older adults—Associations with kidney function and hemoglobin levels. *PLoS ONE* **2019**, *14*, e0212060. [[CrossRef](#)] [[PubMed](#)]
75. Ziyatdinov, A.; Vázquez-Santiago, M.; Brunel, H.; Martinez-Perez, A.; Aschard, H.; Soria, J.M. lme4qtl: Linear mixed models with flexible covariance structure for genetic studies of related individuals. *BMC Bioinform.* **2018**, *19*, 68. [[CrossRef](#)]
76. Wasserman, L. *All of Statistics*; Springer Texts in Statistics; Springer: New York, NY, USA, 2004; ISBN 978-1-4419-2322-6.
77. Geraci, M.; Bottai, M. Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics* **2007**, *8*, 140–154. [[CrossRef](#)] [[PubMed](#)]
78. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*; Chapman and Hall/CRC: Boca Raton, FL, USA, 1994; ISBN 9780429246593.
79. Geraci, M. Linear Quantile Mixed Models: The lqmm Package for Laplace Quantile Regression. *J. Stat. Softw.* **2014**, *57*, 1–29. [[CrossRef](#)]
80. Emmert-Streib, F.; Dehmer, M. Large-Scale Simultaneous Inference with Hypothesis Testing: Multiple Testing Procedures in Practice. *Mach. Learn. Knowl. Extr.* **2019**, *1*, 653–683. [[CrossRef](#)]
81. Liu, Y.; Buil, A.; Collins, B.C.; Gillet, L.C.; Blum, L.C.; Cheng, L.; Vitek, O.; Mouritsen, J.; Lachance, G.; Spector, T.D.; et al. Quantitative variability of 342 plasma proteins in a human twin population. *Mol. Syst. Biol.* **2015**, *11*, 786. [[CrossRef](#)]
82. Seral-Cortes, M.; Sabroso-Lasa, S.; De Miguel-Etayo, P.; Gonzalez-Gross, M.; Gesteiro, E.; Molina-Hidalgo, C.; De Henauw, S.; Gottrand, F.; Mavrogianni, C.; Manios, Y.; et al. Development of a Genetic Risk Score to predict the risk of overweight and obesity in European adolescents from the HELENA study. *Sci. Rep.* **2021**, *11*, 3067. [[CrossRef](#)]
83. Iqbal, A.; Kim, Y.S.; Kang, J.M.; Lee, Y.M.; Rai, R.; Jung, J.H.; Oh, D.Y.; Nam, K.C.; Lee, H.K.; Kim, J.J. Genome-wide Association Study to Identify Quantitative Trait Loci for Meat and Carcass Quality Traits in Berkshire. *Asian-Australas. J. Anim. Sci.* **2015**, *28*, 1537–1544. [[CrossRef](#)]
84. Neter, J.; Kutner, M.H.; Nachtsheim, C.J.; Wasserman, W. *Applied Linear Statistical Models*, 4th ed.; WCB McGraw Hill/Irwin: Columbus, OH, USA, 1996; Volume 1, ISBN 0256117365.
85. Nakagawa, S.; Schielzeth, H. A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods Ecol. Evol.* **2013**, *4*, 133–142. [[CrossRef](#)]