Clinical Immunology 180 (2017) 100–105

Contents lists available at ScienceDirect

Clinical Immunology

journal homepage: www.elsevier.com/locate/yclim

Genetic modifiers of multiple sclerosis progression, severity and onset

A. Dessa Sadovnick [a,b], Anthony L. Traboulsee [b], Yinshan Zhao [b], Cecily Q. Bernales [a], Mary Encarnacion [a], Jay P. Ross [a], Irene M. Yee [a], Maria G. Criscuoli [a], Carles Vilariño-Güell [a,*]

[a] Department of Medical Genetics, University of British Columbia, Vancouver, Canada
[b] Division of Neurology, Faculty of Medicine, University of British Columbia, Vancouver, Canada

ARTICLE INFO

Article history:
Received 18 January 2017
accepted with revision 9 May 2017
Available online 10 May 2017

Keywords:
Multiple sclerosis
Progression
Severity
Association
Genetic

ABSTRACT

The genetic contribution to clinical outcomes for multiple sclerosis (MS) has yet to be defined. We performed exome sequencing analysis in 100 MS patients presenting opposite extremes of clinical phenotype (discovery cohort), and genotyped variants of interest in 2016 MS patients (replication cohort). Linear and logistic regression analyses were used to identify significant associations with disease course, severity and onset. Our analysis of the discovery cohort nominated 38 variants in 21 genes. Replication analysis identified PSMG4 p.W99R and NLRP5 p.M459I to be associated with disease severity ($p = 0.002$ and 0.008). CACNA1H p.R1871Q was found associated with patients presenting relapsing remitting MS at clinical onset ($p = 0.028$) whereas NLRP5 p.M459I and EIF2AK1 p.K558R were associated with primary progressive disease ($p = 0.031$ and 0.023). In addition, PSMG4 p.W99R and NLRP5 p.R761L were found to correlate with an earlier age at MS clinical onset, and MC1R p.R160W with delayed onset of clinical symptoms ($p = 0.010$–0.041).

1. Introduction

Multiple sclerosis (MS) is one of the most prevalent diseases of the central nervous system. It is characterized by myelin loss, varying degrees of axonal pathology, and progressive neurological dysfunction resulting in a myriad of clinical symptoms [1–3]. The majority of patients (~85%) present relapsing-remitting MS (RRMS) characterized by a sudden onset of symptoms followed by partial or complete recovery. Approximately 15% of patients initially present primary progressive MS (PPMS) and exhibit an accumulation of irreversible neurological symptoms from the clinical onset of disease [4].

Case-control studies have identified many genes containing variants which may influence an individual's risk of developing MS [5], but very little is known about genetic factors implicated in determining disease onset, clinical course and severity [6–9]. Here, we used a genome-wide sequencing approach to genetically distinguish two distinct sets of MS patients presenting opposite extremes of clinical course ("discovery cohort"). We then genotyped variants of interest in a large series of MS patients ("replication cohort") for the identification and validation of genes and biological processes which may affect MS susceptibility, age at onset of clinical symptoms, and protect against or predispose to a more progressive and debilitating disease course.

2. Methods

2.1. Participants

A total of 2016 MS patients ascertained through the Canadian Collaborative Project on the Genetic Susceptibility to Multiple Sclerosis (CCPGSMS) were included in this study [10]. The Research Ethic Board (REB) at the University of British Columbia approved the study, and all participants provided informed consent. Blood samples and detailed clinical information including patient demographics, family history of MS, disease course, the first clinical onset symptom(s), and disease severity as measured by the expanded disability status scale (EDSS) [11] were collected. All samples were of Caucasian ancestry (based on four grandparents), and patients were diagnosed with MS according to Poser criteria prior to 2001 and McDonald thereafter [12,13]. Basic demographics and clinical details are provided in Supplemental Table 1.

A subset of 100 MS patients representing clinically opposite extremes of disease phenotype, 50 "severe-progressive" and 50 "mild-relapsing", were identified and formed the "discovery cohort". Severe-progressive MS was defined as PPMS course with an EDSS score ≥ 6 within 10 years of the onset of disease. Mild-relapsing MS included MS patients with RRMS at clinical onset and an EDSS ≤ 3 after 14 or

* Corresponding author at: 2215 Wesbrook Mall, University of British Columbia, Vancouver, V6T 1Z3, BC, Canada.
E-mail address: carles@can.ubc.ca (C. Vilariño-Güell).

http://dx.doi.org/10.1016/j.clim.2017.05.009
1521-6616/

more years of disease duration. The demographics for these groups are provided in Supplemental Table 2.

2.2. Exome sequencing and genotyping

Exonic regions were enriched using an Ion AmpliSeq exome kit (57.7Mb) and sequenced in an Ion Torrent Proton sequencer (Life Technologies, Carlsbad, CA, USA) with a minimum average coverage of 70 reads per base and an average read length of 150 bases. Sequence mapping and variant identification parameters were as previously described [14]. All variants of interest and rs3135388 were genotyped using a MassArray iPLEX platform (Sequenom, San Diego, CA, USA) according to manufacturer instructions.

2.3. Statistical analysis

Deviations from the Hardy-Weinberg equilibrium (HWE) were assessed using chi-square tests. A combined statistical approach implementing the Sequence Kernel Association Test (SKAT) and PLINK was used to analyze the discovery cohort [15–17]. For SKAT, missense and nonsense variants were collapsed within genes to define "regions of interest". Genes with a p-value below 0.003 were considered to differ between mild-relapsing and severe-progressive MS in the discovery cohort. Within genes of interest, variants observed in at least four samples, and with an allelic frequency difference >20% between groups were selected for replication. PLINK was implemented for the analysis of individual variants; those presenting an allelic p-values < 0.001 were selected for further genotyping in all MS patients from the CCPGSMS. Although these post hoc p-values are not considered significant, they were chosen as thresholds to select variants for replication in the complete series given the small size of the discovery cohort.

Variants of interest were genotyped in the whole cohort and analyzed with (combined) and without (replication) the inclusion of the discovery cohort. Association analyses for disease course at onset of disease were assessed by logistic regression models. Associations with EDSS and MS severity score (MSSS) [18] were assessed with linear regression models. In these analyses, genotypes were dichotomised as presence versus absence of the minor allele (dominant model). Odds ratios (OR) and 95% confidence intervals (CI) for logistic regression or effect coefficients and standard error (SE) for linear regressions were estimated. All models were adjusted for gender, age at MS onset, clinical course and/or disease duration as indicated. p-Values < 0.0015 (0.05/33) were considered significant after Bonferroni correction for multiple testing.

3. Results

Exome sequencing analysis was performed in 100 MS patients representing opposite extremes of clinical disease spectrum, defined as mild-relapsing and severe-progressive (Supplemental Table 2). This analysis identified an average of 50,114 variants per individual, and a total of 284,727 different variants in the whole dataset. Of those, 67,055 were found to cause an amino acid substitution or protein truncation and were evaluated for differences in frequency between patients with mild-relapsing and severe-progressive disease. Statistical analyses of the discovery cohort data identified 38 variants in 21 genes below the p-value thresholds selected for replication (Supplemental Table 3). These variants were genotyped in 2016 MS patients; this includes the discovery cohort, which was used to confirm exome calls and genotyping accuracy. Genotyping counts and frequencies for the whole series are provided in Supplemental Table 4. Average genotyping call rate was 99.1%, and all variants were in HWE with the exception of CACNA1H p.V664A ($p < 0.05$), which was excluded from

Table 1
Linear regression analysis with MSSS. Significant associations are indicated in bold. SE, standard error.

			Replication (n = 1367)		Combined (n = 1467)	
Gene	Amino acid	dbSNP	p-Value	Effect ± SE	p-Value	Effect ± SE
FNDC7	I146V	rs17553619	0.549	−0.14 ± 0.23	0.149	−0.32 ± 0.22
VILL	L740F	rs9816693	**0.023**	**−0.39 ± 0.17**	0.173	−0.23 ± 0.17
FAT1	V482I	rs3733413	0.717	−0.06 ± 0.17	0.161	0.24 ± 0.17
PSMG4	W99R	rs4959786	**0.009**	**0.45 ± 0.17**	**0.002**	**0.53 ± 0.17**
SLC22A1	P341L	rs2282143	0.271	−0.54 ± 0.49	0.712	−0.18 ± 0.48
SLC22A1	M440I	rs35956182	0.801	−0.17 ± 0.68	0.517	0.39 ± 0.60
EIF2AK1	K558R	rs2640	0.791	0.07 ± 0.26	0.516	0.17 ± 0.26
EIF2AK1	L319H	rs34909691	0.107	−0.69 ± 0.43	0.321	−0.41 ± 0.42
ZNF107	E40D	rs73138709	0.468	0.19 ± 0.26	0.185	0.34 ± 0.26
FAM185A	L15F	rs141352868	0.866	−0.04 ± 0.25	0.794	−0.06 ± 0.24
LRRC17	K119E	rs3800939	0.584	−0.14 ± 0.25	0.323	−0.24 ± 0.24
CNGB3	I307V	rs13265557	0.980	−0.01 ± 0.24	0.568	−0.14 ± 0.24
TG	D1838N	rs2069561	0.593	−0.10 ± 0.18	0.667	−0.08 ± 0.18
ARMS2	A69S	rs10490924	0.405	−0.14 ± 0.16	**0.042**	**−0.33 ± 0.16**
TRIM66	L630V	rs7935453	0.364	0.15 ± 0.16	0.241	0.19 ± 0.16
TRIM66	R145K	rs61741649	0.827	−0.05 ± 0.23	0.557	0.13 ± 0.23
ANAPC7	S33N	rs141147170	0.538	0.21 ± 0.34	0.203	0.42 ± 0.33
CACNA1H	R1871Q	rs58124832	0.877	0.04 ± 0.24	0.791	−0.06 ± 0.24
CACNA1H	R2060H	rs1054644	0.152	−0.27 ± 0.19	0.132	−0.28 ± 0.18
MC1R	R160W	rs1805008	0.978	−0.01 ± 0.26	0.517	0.17 ± 0.25
ZNF830	S154T	rs3744355	0.208	−0.26 ± 0.20	0.125	−0.31 ± 0.20
GP6	R576K	rs10418074	0.228	−0.33 ± 0.28	0.772	−0.08 ± 0.27
GP6	K323T	rs1671152	0.316	−0.18 ± 0.17	0.163	−0.24 ± 0.17
GP6	E237K	rs1654416	0.515	−0.11 ± 0.17	0.345	−0.16 ± 0.17
GP6	P219S	rs1613662	0.210	−0.22 ± 0.17	0.126	−0.26 ± 0.17
RDH13	N21K	rs12609004	0.890	0.03 ± 0.23	0.277	0.25 ± 0.23
NLRP5	M459I	rs471979	**0.030**	**0.43 ± 0.20**	**0.008**	**0.50 ± 0.19**
NLRP5	R761L	rs17713875	0.673	−0.14 ± 0.33	0.736	0.11 ± 0.33
NLRP5	M912T	rs16986899	0.398	0.15 ± 0.17	0.227	0.21 ± 0.17
NLRP5	V1181I	rs10409555	0.610	0.08 ± 0.16	0.427	0.13 ± 0.16
NLRP5	R1195Q	rs36118060	0.406	0.15 ± 0.18	0.168	0.24 ± 0.18
FAM83C	R621Q	rs2425049	0.404	−0.14 ± 0.17	0.743	−0.05 ± 0.16

Bold numbers indicate significance at $p < 0.05$.