The Genetic Association of Variants in CD6, TNFRSF1A and IRF8 to Multiple Sclerosis: A Multicenter Case-Control Study

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1371/journal.pone.0018813">http://dx.doi.org/10.1371/journal.pone.0018813</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Public Library of Science</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/69041">http://hdl.handle.net/1721.1/69041</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by/2.5/">http://creativecommons.org/licenses/by/2.5/</a></td>
</tr>
</tbody>
</table>
The Genetic Association of Variants in CD6, TNFRSF1A and IRF8 to Multiple Sclerosis: A Multicenter Case-Control Study

The International Multiple Sclerosis Genetics Consortium*†

Abstract

Background: In the recently published meta-analysis of multiple sclerosis genome-wide association studies De Jager et al. identified three single nucleotide polymorphisms associated to MS: rs17824933 (CD6), rs1800693 (TNFRSF1A) and rs17445836 (61.5 kb from IRF8). To refine our understanding of these associations we sought to replicate these findings in a large more extensive independent sample set of 11 populations of European origin.

Principal Findings: We calculated individual and combined associations using a meta-analysis method by Kazeem and Farral (2005). We confirmed the association of rs1800693 in TNFRSF1A (p = 1.9 × 10−7, OR 1.1, 7,665 cases, 8,051 controls) and rs17445836 near IRF8 (p = 5.3 × 10−10, OR 0.8, 6,895 cases, 7,580 controls and 596 case-parent trios) The SNP rs17824933 in CD6 also showed nominally significant evidence for association (p = 2.19 × 10−5, OR 1.1, 8,047 cases, 9,174 controls, 604 case-parent trios).

Conclusions: Variants in TNFRSF1A and in the vicinity of IRF8 were confirmed to be associated in these independent cohorts, which supports the role of these loci in etiology of multiple sclerosis. The variant in CD6 reached genome-wide significance after combining the data with the original meta-analysis. Fine mapping is required to identify the predisposing variants in the loci and future functional studies will refine their molecular role in MS pathogenesis.

Introduction

Multiple sclerosis (MS) is a complex neurological autoimmune disease with few known predisposing factors. Both genetic and environmental components have been predicted to play a role in MS etiology and the role of the HLA-locus, HLA-DRB1 in particular, is well recognized [1,2]. Recently, genome-wide association and candidate gene studies have revealed significant associations to MS outside the HLA-locus in IL2RA [2], IL7R [2], CD58 [3], CLEC16A [4], TYK2 [5], STAT3 [6], IL12A, MPHOSPH9/CDJ2AP1, ETB [2], KIF21B [2,7], TMEFI [2,7], CYP27B1 [8], CD276 [4], CD40 [8], CBLB [9] and RGS1 [10], but with modest odds ratios suggesting the involvement of other loci.

In a recently published meta-analysis of six genome-wide analysis (GWA) study sets of 2,624 MS cases and 7,220 controls from four populations of European origin (United States, United Kingdom, Netherlands and Switzerland), De Jager et al. identified three single nucleotide polymorphisms (SNPs) associated with MS with significance exceeding the genome-wide significance level of p < 5×10−8: rs1800693 in TNFRSF1A, rs17445836 61.5 kb from IRF8 and rs17824933 in CD6 [11]. De Jager et al. replicated these findings in 2,215 cases and 2,116 controls from UK and US. Recently, there have been reports showing significant genetic differences in allele frequencies between populations even within Europe [12,13,14] which has led to speculation of allelic heterogeneity. We set out to replicate the association of these SNPs to MS in a more extensive sample set with varying European origins.

Results

We investigated the top three SNP associations by De Jager et al. (rs1800693 in TNFRSF1A, rs17445836 61.5 kb from IRF8 and

* E-mail: virpi.leppa@thl.fi

† Members of the International Multiple Sclerosis Genetics Consortium is provided in the Acknowledgments.

Citation: The International Multiple Sclerosis Genetics Consortium (2011) The Genetic Association of Variants in CD6, TNFRSF1A and IRF8 to Multiple Sclerosis: A Multicenter Case-Control Study. PLoS ONE 6(4): e18813. doi:10.1371/journal.pone.0018813

Editor: Lucienne Chatenoud, Université Paris Descartes, France

Received October 7, 2010; Accepted March 16, 2011; Published April 28, 2011

Copyright: © 2011 Leppä et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by National Institutes of Health (grants RO1 NS 43559 and RO1 NS049477), Center for Excellence for Complex Disease Genet...
and rs17824933 in CD6 in an independent sample set of 11 populations of varying European origins, comprising a total of 8,493 cases, 9,280 controls and 608 case-parent trios (Table 1). Cases and controls were selected from the same populations to minimize population stratification. We performed meta-analysis using a method by Kazeem and Farrall (2005) [13] and observed nominal association (p<0.05) with multiple sclerosis for rs17824933 in CD6 in four of the eleven cohorts (Figure 1a), for rs1800693 in TNFRSF1A in four out of nine available cohorts (Figure 1b) and for rs17445836 near IRF8 in five out of nine available cohorts (Figure 1c) (see materials and methods for details).

In all except three cohorts (Denmark, Italy and Norway for the CD6 rs17824933 C allele) allele frequency differences between cases and controls had a trend towards the same direction as seen in the original meta-analysis [11] (Figure 1). Most of the individual cohorts had limited estimated power (varying between 25–82%; alpha 0.05) to observe the association by themselves (Table S1). Nevertheless, the estimated power for a combined analysis was >99% (alpha 0.05) to detect association to variants with the same effect sizes as observed in the original meta-analysis (rs1800693 OR 1.2, rs17445836 OR 0.80, rs17824933 OR 1.18).

The combined analysis confirmed independent associations with two of the SNPs with odds ratios comparable to those observed in the original meta-analysis: rs1800693 in TNFRSF1A (p = 4.19 × 10⁻⁶, OR 1.12, 95% CI 1.07–1.18) and rs17445836 near IRF8 (p = 5.34 × 10⁻⁹, OR 0.84, 95% CI 0.80–0.89) (Figure 1b and c, respectively). Non-significant association for rs17824933 in CD6 was also observed (p = 2.19 × 10⁻³, OR 1.11, 95% CI 1.06–1.17) (Figure 1a). Combining the replication data with the original meta-analysis data from Dr. Jager et al. did not significantly change the observed odds ratios (Figure 1). We noticed an unequal distribution of minor allele frequencies across European populations as might be expected [12,13,14] in the rs17445836 and rs17824933 SNPs (Figure 1). However, the Breslow-Day test confirmed that there was no major heterogeneity in the odds ratios, although the allele frequency differences were significant between several populations when controls from different populations were compared in a pairwise manner with a standard association tests (Table S2).

### Discussion

We conclude that the SNPs rs1800693 (TNFRSF1A) and rs17445836 (IRF8) are convincingly associated to MS in this independent replication set. This supports the role of these genes in MS etiology. The rs17824933 (CD6) showed nominally significant association in the analysis combining the replication cohorts, although the association in most of the individual cohorts was not significant. It is possible that the lack of association in some cohorts is due to true population heterogeneity, but the individual cohorts in our study do not have enough power to draw any definite conclusions. Especially, since the cohorts showing an opposite trend have little power by themselves. None of these three genes (CD6, TNFRSF1A or IRF8) had shown association above the replication inclusion threshold in the IMMSGC [2] or Gene MSA [16] original publications (p<10⁻⁴), but by combining the data in a meta-analysis the full advantage of these cohorts could be used to mine more MS susceptibility affecting genes [11].

Rare mutations in previously validated MS susceptibility genes have been implicated in rare monogenic disorders. For example, mutations in IL2R4 [17] and IL7R [18] cause immunodeficiency and mutation in TTK [19] and STAT3 [20] have been reported to cause hyper-IgE syndrome. Similarly, mutations in TNFRSF1A can cause TRAPS, a disease of the immune system characterized by periodic fevers [21]. It is interesting, that both TRAPS and relapsing-remitting form of multiple sclerosis are characterized by periodic activations of autoimmunity. A recent study in a small German cohort reported that 24% (6/25) of patients with clinically isolated syndrome (CIS) or MS with TRAPS-like symptoms were carrying an amino-acid changing allele R92Q of the SNP rs4149504 in TNFRSF1A [22]. In addition, they reported that the frequency of the R92Q allele was 4.66% in a general MS

### Table 1. Summary of all independent replication sample sets.

<table>
<thead>
<tr>
<th>Sets</th>
<th>N trios</th>
<th>N ctrl</th>
<th>N MS</th>
<th>% PMMS</th>
<th>Sex ratios F:M MS, ctrl</th>
<th>EDSS</th>
<th>disease duration</th>
<th>Genotyping platform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>0</td>
<td>1,021</td>
<td>776</td>
<td>13.7</td>
<td>1.8:1, 1.1:1</td>
<td>4.8</td>
<td>14</td>
<td>TaqMan® (Applied Biosystems)</td>
</tr>
<tr>
<td>Denmark</td>
<td>0</td>
<td>1,090</td>
<td>634</td>
<td>7.6</td>
<td>2.0:1, 1.6:1</td>
<td>4.1</td>
<td>12</td>
<td>Sequenom® iPLEX® Gold</td>
</tr>
<tr>
<td>Finland</td>
<td>0</td>
<td>1,077</td>
<td>792</td>
<td>9.4</td>
<td>2.4:1, 1.4:1</td>
<td>4.5</td>
<td>21</td>
<td>Sequenom®, TaqMan® Gold</td>
</tr>
<tr>
<td>France</td>
<td>608</td>
<td>0</td>
<td>0</td>
<td>12.0</td>
<td>2.4:1, 1.0:1</td>
<td>3.4</td>
<td>9.1</td>
<td>TaqMan® (Applied Biosystems)</td>
</tr>
<tr>
<td>Germany</td>
<td>0</td>
<td>911</td>
<td>930</td>
<td>&lt;1%</td>
<td>n.a.</td>
<td>n.a.</td>
<td>7</td>
<td>Sequenom® iPLEX® Gold</td>
</tr>
<tr>
<td>Italy</td>
<td>0</td>
<td>629</td>
<td>828</td>
<td>11.1</td>
<td>2.0:1, 1.0:1</td>
<td>3.2</td>
<td>32</td>
<td>TaqMan® (Applied Biosystems)</td>
</tr>
<tr>
<td>Norway</td>
<td>0</td>
<td>1,027</td>
<td>662</td>
<td>17.7</td>
<td>2.6:1, 2.0:1</td>
<td>4.6</td>
<td>16</td>
<td>Sequenom®, TaqMan® Gold</td>
</tr>
<tr>
<td>Spain</td>
<td>0</td>
<td>501</td>
<td>501</td>
<td>19.9</td>
<td>1.8:1, 1.1:1</td>
<td>4.2</td>
<td>14</td>
<td>TaqMan® (Applied Biosystems)</td>
</tr>
<tr>
<td>Sweden</td>
<td>0</td>
<td>1,723</td>
<td>2,016</td>
<td>5.8</td>
<td>2.5:1, 2.0:1</td>
<td>3.3</td>
<td>n.a.</td>
<td>Sequenom® iPLEX® Gold</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0</td>
<td>714</td>
<td>656</td>
<td>14.4</td>
<td>2.8:1, 2.8:1</td>
<td>4.8</td>
<td>18</td>
<td>Sequenom® iPLEX® Gold</td>
</tr>
<tr>
<td>United States</td>
<td>0</td>
<td>587</td>
<td>644</td>
<td>12.0</td>
<td>1.1:1, 1.1:1</td>
<td>4.1</td>
<td>15</td>
<td>Sequenom® iPLEX® Gold</td>
</tr>
<tr>
<td>Total</td>
<td>608</td>
<td>9,280</td>
<td>8,439</td>
<td></td>
<td>2.1:1, 1.4:1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All sample sets for the replication are independent, cases had clinically definite MS by either the Poser or McDonald criteria and anonymous population samples from respective populations were used as controls. The clinical parameters for MS patients describe the percentage of primary progressive MS (PPMS) of all cases, the mean EDSS score and the mean disease duration. The original GWA meta-analysis sample sets by De Jager et al. that were used in the combined analysis of the original GWA results and our independent replication have been described elsewhere [11,16].

*The Norwegian and Finnish samples were genotyped with the Applied Biosystems TaqMan® platform for rs1800693 and Sequenom® iPLEX® Gold for rs17624933 and rs17445836.

doi:10.1371/journal.pone.0018813.t001
The association analyses were performed according to Kazeem and Farrall [15]. The reported HWE p value is reported for cases and controls combined, but no significant deviation was observed within cases or controls when analyzed separately (data not shown). Figure 1a represents the results for rs17824933 in CD6. The replication -line is the combined result of all independent sample sets in the replication (8,047 cases, 9,174 controls, 604 case-parent trios) and “Combined with De Jager et al. GWA” set includes the De Jager et al. [11] GWA data set (2,624 cases, 7,220 controls). Figure 1b summarizes the results for rs1800693 in TNFRSF1A. Genotyping was unsuccessful in two sample sets (Danish case – control set and French case-parent trios) for rs1800693. Independent replication data set (“Replication”) included total of 7,665 cases and 8,051 controls and the “Combined with De Jager et al. GWA” set includes available genotypes from De Jager et al. [11] (1,829 cases, 2,591 controls). Figure 1c is a summary of results for rs17445836 (61.5 kb from IRF8). The genotyping was unsuccessful in two sample sets (Spanish and German case – control sets). The independent replication set (Replication) includes in total 6,895 cases, 7,580 controls and 596 case-parent trios and the “Combined with De Jager et al. GWA” set includes available genotypes from De Jager et al. [11] (2,624 cases, 7,220 controls).

doi:10.1371/journal.pone.0018813.g001

Samples and genotyping
All samples had clinically definite MS by either the Poser criteria or McDonald criteria and anonymous population samples from respective populations were used as controls. (Table 1) All cohorts used in this independent replication were genotyped in local centers using either Taqman (Applied Biosystems, CA, USA) or Sequenom® iPLEX® Gold platform (SEQUENOM, CA, US) and manufacturer protocols, except for the Danish and Norwegian samples that were genotyped in Finland for rs17445836 and rs17824933 (Sequenom®iPLEX® Gold) (Table 1). The original meta-analysis sample sets from De Jager et al., that we used in the combined analysis of the original GWA and our replication results (Figure 1, last line), and their genotyping have been described elsewhere [11,16].

Statistical analyses
We excluded from the analysis all samples with >1 missing genotype and SNPs with <90% success rate or Hardy-Weinberg disequilibrium (HWE) p<0.001 per population. Using these criteria we excluded rs17445836 (IRF8) from the Spanish and German cohorts and rs1800693 (TNFRSF1A) from the Danish and French cohorts.

We performed both an independent replication analysis and a combined analysis using the original De Jager et al. GWA sample set. The analyses were performed according to Kazeem and Farrall [15] and the calculations were done using R 2.9.0 (www.r-project.org). The Hardy-Weinberg (dis)equilibrium analysis p values were calculated using PLINK v1.06 (http://pngu.mgh.harvard.edu/~purcell/plink/) . The T (Transmitted alleles) and U (Under-transmitted alleles) for the case-parent trios have been obtained from PLINK v1.06 transmission disequilibrium test (TDT) analysis.

Supporting Information
Table S1 Power calculations for all study sets. All calculations were done using Researcher’s toolkit’s Statistical Power Calculator’s two-tailed test with percentages by DSS (http://www.dssresearch.com/toolkit/spcalc/power_p2.asp) alpha = 5% for false positive probability, fixed MAFs calculated from the ORs of the combined effects and allele frequencies from the original study by De Jager et al. 2009. These results show that most of the individual sample sets have only moderate power to detect the association by themselves, but together have over 99% power to detect...
detect these variants with these effect sizes. The power for trios was not estimated. (DOC)

Table S2 Differences in rs17824933, rs1800693 and rs17445836 minor allele frequencies between population based controls. This table shows results for pair-wise associations between controls from different populations. We used the controls from populations on the left as cases and controls from the population above as controls. For French samples, healthy parents from case-parent trio samples were used as population controls. Uncorrected p-values are shown, but all values below p 0.000303 are significant (\( \alpha = 0.05 \)) after Bonferroni correction. Table S2a has the results for rs17624933 in CD6, Table S2b describes the results for rs1800693 in TNFRSF1A and Table S2c describes results for 17445836 61.5 kb from IRF8. (DOC)

Table S3 Ethics committee approvals for all cohorts. This study has been approved by appropriate local ethics committees as listed in this table by sample set. For each cohort we report the ethics committee or equivalent authority and the approval number. (DOC)

Acknowledgments

We wish to thank all participating MS patients and their families. We also wish to thank Liisa Arala and Anne Vikman for their invaluable assistance and technical support. Concerning the statistical analyses, we sincerely thank Dr. Samuli Ripatti who has provided valuable assistance and advice. The DIMG and Gene MSA consortia are acknowledged for the data from the original meta-analysis. Danish Multiple Sclerosis Society is acknowledged for supporting the Danish sample collection and The Norwegian Bone Marrow Donor Registry is acknowledged for collaboration in establishment of the Norwegian control material. Dr. Mauri Reunanen (Oulu University Hospital and University of Oulu), Dr Tuula Pirttilä (Kuopio University Hospital and University of Kuopio) and Dr Keijo Koivisto (Seinäjoki Central Hospital) are thanked for their efforts in recruiting Finnish MS patients. We also would like to acknowledge the Institute for Molecular Medicine Finland FIMM Technology Center for genotyping assistance. The French network REFGENSEP acknowledges the collaboration of CIC Pitie-salpetriere (Centre d’Investigation Clinique) and Génétion.

Consortium Authors

Virpi Leppä1,2, Ida Surakka1,2, Petti J. Tienari4, Irina Elouzaev5, Alastair Compton6, Stephen Sowers6, Neill Robertson6, Philip L. De Jager8,9, Cristin Aubin9, David A. Hafler9,10, Annette Bang Oturai11, Helle Cathrine Smestad22, Elisabeth G. Celius 22, Hanne F. Harbo 22, Sandra CD6

1 Institute for Molecular Medicine Finland FIMM, University of Helsinki, Helsinki, Finland
2 Public Health Genomics Unit, National Institute for Health and Welfare, Helsinki, Finland
3 Helsinki Biomedical Graduate School, University of Helsinki, Helsinki, Finland
4 Helsinki University Central Hospital, Department of Neurology and University of Helsinki, Biomedicum, Molecular Neurology Programme, Helsinki, Finland
5 University of Tampere and Tampere University Hospital, Department of Neurology, Tampere, Finland
6 Department of Clinical Neurosciences, University of Cambridge, Addenbrooke’s Hospital, Cambridge, UK
7 Department of Neurology, Ophthalmology and Audiological Medicine, School of Medicine, Cardiff University, Cardiff, CF14 4XN, UK
8 Program in Translational NeuroPsychiatric Genomics, Department of Neurology, Brigham & Women’s Hospital and Harvard Medical School, Boston, MA, USA
9 The Broad Institute of MIT and Harvard, Cambridge, MA, USA
10 Departments of Neurology and Immunobiology, Yale School of Medicine, New Haven, CT 06520-8018, USA
11 Danish Multiple Sclerosis Center, University of Copenhagen and Department of Neurology, Rigshospitalet, Copenhagen, Denmark
12 Klinik für Neurologie, Klinikum rechts der Isar, Technische Universität, München, Germany
13 Institut für Humangenetik, Technische Universität München, München, Germany
14 Institut für Humangenetik, Helmholtz Zentrum München, München, Germany
15 Institute of Epidemiology, Helmholtz Zentrum München-German Research Center for Environmental Health, Munich, Germany
16 Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany
17 Centre d’Esclerosi Múltiple de Catalunya, CEM-Cat, Unitat de Neuroimmunologia Clínica, Hospital Universitari Vall d’Hebron (HUH), Barcelona, Spain
18 Neuroimmunology Unit, Center for Molecular Medicine, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
19 The Multiple Sclerosis Research Group, Center for Molecular Medicine, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
20 Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
21 Laboratory for Neuroimmunology, Section for Experimental Neurology, Katholieke Universiteit Leuven, Herestraat 49 bus 1022, 3000 Leuven, Belgium
22 Department of Neurology, Oslo University Hospital, 0407 Oslo, Norway
23 Institute of Inimmunology, Oslo University Hospital, 0027 Oslo, Norway
24 University of Oslo, Oslo, Norway
25 Department of Medical Sciences and Interdisciplinary Research Center of Autoimmune Diseases (IRCAD), University of Eastern Piedmont, Novara, Italy
26 Department of Neurology, Ospedale Maggiore and IRCAD, Novara, Italy
27 Department of Neurology and Center for Experimental Neurological Therapy (CENTERS), Università La Sapienza, Roma, Italy
28 Departments of Neurology and Biomedicine University Hospital Basel, University of Basel, Switzerland
29 Department of Neurology, University of California at San Francisco, US
30 on behalf of the French Genetics MS network REFGENSEP, INSERM, UMR_S975, Paris, France, UPMC Univ Paris 06, UMR_S975, Centre de Recherche Institut du Cerveau et de la Moelle, CNRS 7225, Department of Neurology, Hôpital Pitie –Salpêtrière, AP-HP, Paris, France
31 Leuven University Center for Metabolic Bone Diseases and Division of Geriatric Medicine, University of Leuven, 3000 Leuven, Belgium
32 Department of Neurology, Vrije Universiteit Medical Centre, Amsterdam, The Netherlands
33 Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, United Kingdom
34 Program in Medical and Population Genetics and Genetic Analysis Platform, The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
35 Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, Finland
36Department of Gynecology and Pediatrics, Department of Child Psychiatry, Helsinki University Central Hospital, Helsinki, Finland
37Posthumously
Author Contributions
Conceived and designed the experiments: AG BD SB ABO HBS FS PSS VL JS AP LP BF BH SC JW HEW SD ML GR CP ILM HHF MC XM TO IK JH LA LK AC SS PLD DAH SLH. Performed the experiments: AG VL ICR SC LB ILM MFB MC IK LK CA. Analyzed the data: VL IS.

References