

MIT Open Access Articles

The role of the CD[subscript 5]8 locus in multiple sclerosis

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

Citation: De Jager, Philip L et al. "The role of the CD58 locus in multiple sclerosis." Proceedings of the National Academy of Sciences 106.13 (2009): 5264-5269.

As Published: <http://dx.doi.org/10.1073/pnas.0813310106>

Publisher: National Academy of Sciences

Persistent URL: <http://hdl.handle.net/1721.1/50257>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



The role of the *CD58* locus in multiple sclerosis

Philip L. De Jager^{a,b,c,1}, Clare Baecher-Allan^a, Lisa M. Maier^{a,c}, Ariel T. Arthur^d, Linda Ottoboni^{a,c}, Lisa Barcellos^e, Jacob L. McCauley^f, Stephen Sawcer^g, An Goris^h, Janna Saarelaⁱ, Roman Yelensky^{c,j}, Alkes Price^{c,k}, Virpi Leppäⁱ, Nick Patterson^l, Paul I. W. de Bakker^{b,c}, Dong Tran^{a,c}, Cristin Aubin^{a,c}, Susan Pobywajlo^a, Elizabeth Rossin^{a,c}, Xinli Hu^c, Charles W. Ashley^a, Edwin Choy^{c,i}, John D. Rioux^{c,i}, Margaret A. Pericak-Vance^f, Adrian Ivinson^m, David R. Boothⁿ, Graeme J. Stewartⁿ, Aarno Palotie^{c,h}, Leena Peltonen^{c,h}, Bénédicte Dubois^g, Jonathan L. Haines^o, Howard L. Weiner^a, Alastair Compston^g, Stephen L. Hauser^{p,q}, Mark J. Daly^{c,i}, David Reich^{c,i}, Jorge R. Oksenberg^{p,q}, and David A. Hafler^{a,c}.

^aDivision of Molecular Immunology, Center for Neurologic Diseases, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115; ^bPartners Center for Personalized Genetic Medicine, Boston, MA 02115; ^cProgram in Medical and Population Genetics, Broad Institute of Harvard University and Massachusetts Institute of Technology, Cambridge, MA 02139; ^dDepartment of Medicine and the Nerve Research Foundation, University of Sydney, Sydney NSW 2145, Australia; ^eDivision of Epidemiology, School of Public Health, University of California, Berkeley, CA 94720-7360; ^fMiami Institute for Human Genetics, Miller School of Medicine, University of Miami, Miami, FL 33136; ^gDepartment of Clinical Neurosciences, Addenbrooke's Hospital, University of Cambridge, Box 165, Hills Road, Cambridge CB2 2QQ, United Kingdom; ^hSection of Experimental Neurology, University of Leuven, B-3000 Leuven, Belgium; ⁱDepartment of Molecular Medicine, National Public Health Institute and the Program of Molecular Medicine of the University of Finland, Biomedicum, Haartmaninkatu 8, 00290, Helsinki, Finland; ^jCenter for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114; ^kDepartment of Genetics, Harvard Medical School, Boston, MA 02115; ^lMontréal Heart Institute and Université de Montréal, Montréal, Québec H3C 3J7, Canada; ^mHarvard Neurodiscovery Center, Harvard Medical School, Boston, MA 02115; ⁿInstitute for Immunology and Allergy Research, Westmead Millennium Institute, Sydney, Australia; ^oCenter for Human Genetics Research, Vanderbilt University Medical Center, Nashville, TN 37232-0700; ^pDepartment of Neurology, School of Medicine, University of California, San Francisco, CA 94143-0435; and ^qInstitute for Human Genetics, School of Medicine, University of California, San Francisco, CA 94143-0435

Communicated by Elliott D. Kieff, Harvard University, Boston, MA, December 29, 2008 (received for review July 23, 2008)

Multiple sclerosis (MS) is an inflammatory disease of the central nervous system associated with demyelination and axonal loss. A whole genome association scan suggested that allelic variants in the *CD58* gene region, encoding the costimulatory molecule LFA-3, are associated with risk of developing MS. We now report additional genetic evidence, as well as resequencing and fine mapping of the *CD58* locus in patients with MS and control subjects. These efforts identify a *CD58* variant that provides further evidence of association with MS ($P = 1.1 \times 10^{-6}$, OR 0.82) and the single protective effect within the *CD58* locus is captured by the rs2300747^g allele. This protective rs2300747^g allele is associated with a dose-dependent increase in *CD58* mRNA expression in lymphoblastic cell lines ($P = 1.1 \times 10^{-10}$) and in peripheral blood mononuclear cells from MS subjects ($P = 0.0037$). This protective effect of enhanced *CD58* expression on circulating mononuclear cells in patients with MS is supported by finding that *CD58* mRNA expression is higher in MS subjects during clinical remission. Functional investigations suggest a potential mechanism whereby increases in *CD58* expression, mediated by the protective allele, up-regulate the expression of transcription factor FoxP3 through engagement of the *CD58* receptor, CD2, leading to the enhanced function of CD4⁺CD25^{high} regulatory T cells that are defective in subjects with MS.

genetic | human | RNA | quantitative trait | inflammation

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system associated with demyelination, axonal loss, and brain atrophy; susceptibility to this disease is affected by both genetic variation and environmental risk factors (1, 2). The initial episode of neurologic dysfunction results in a clinical diagnosis of a clinically isolated demyelinating syndrome (CIS), and a second episode leads to a diagnosis of MS (1). Increasing evidence suggests that activated, autoreactive T cells play a central role in MS pathophysiology, as evidenced by the efficacy of treatments such as Natalizumab (anti-VLA-4 monoclonal antibody) that block lymphocyte egress from the vascular compartment into the CNS (3). Furthermore, the control of activated T cells by natural regulatory CD4⁺ T cells is impaired in subjects with MS (4). This population of regulatory CD4⁺ T cells expresses high levels of the IL-2 receptor (CD25) and FoxP3, an important transcription factor for regulatory T cells (4). We have now begun to integrate these immunologic observations with results of our genetic studies in patients with MS.

Two novel MS susceptibility loci have recently been identified using a genome-wide association scan approach, and these 2 loci, *IL2RA* and *IL7R*, have now been validated in independent subject collections (5–9). In the genome scan, several other loci, including the *CD58* locus, displayed suggestive evidence of association with MS susceptibility. Since *CD58* (LFA-3) costimulates and enhances T cell receptor signaling by engaging CD2 (10), the *CD58* locus is an attractive target for understanding the role of genetic variation in immune system dysfunction associated with MS. Here, we first refine and then enhance the association between a polymorphism in *CD58* and MS susceptibility. Then, using data from Epstein-Barr virus (EBV) transformed B cells (lymphoblastic cell lines), we find the putative *CD58* protective allele to be significantly associated with higher *CD58* RNA levels, and we validate this observation by measuring mRNA expression in circulating mononuclear cells isolated ex vivo from subjects with MS. Moreover, we present evidence that a higher level of *CD58* mRNA expression is seen during the clinically quiescent phase of MS and, finally, that higher *CD58* expression may function in part by enhancing FoxP3 expression in regulatory T cells.

Results

Fine Mapping, Resequencing, and Validation of the *CD58* Susceptibility Locus in MS. We recently performed a whole genome association screen for MS susceptibility genes and identified a suggestive association at SNP rs12044852 both in the screening and replication phase ($P = 1.9 \times 10^{-5}$ in the combined analysis) (Fig. 1; ref. 5). We therefore initiated a fine mapping effort in the chromosomal region that contains rs12044852 to better characterize this association to MS. Using our collection of subjects with MS from the Brigham and Women's Hospital in Boston, MA [supporting information (SI) Table S1], we initially surveyed

Author contributions: P.L.D.J., C.B.-A., and D.A.H. designed research; P.L.D.J., C.B.-A., A.T.A., L.O., S.S., A.G., J.S., V.L., C.A., S.P., E.R., X.H., C.W.A., D.R.B., B.D., and D.R. performed research; P.L.D.J., C.B.-A., A.T.A., L.O., R.Y., A. Price, N.P., D.T., D.R.B., M.J.D., and D.A.H. analyzed data; and P.L.D.J., L.M.M., L.B., J.L.M., S.S., A.G., J.S., R.Y., A. Price, V.L., N.P., P.I.W.d.B., E.C., J.D.R., M.A.P.-V., A.I., D.R.B., G.J.S., A. Palotie, L.P., B.D., J.L.H., H.L.W., A.C., S.L.H., M.J.D., D.R., J.R.O., and D.A.H. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: pdejager@rics.bwh.harvard.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0813310106/DCSupplemental.

© 2009 by The National Academy of Sciences of the USA

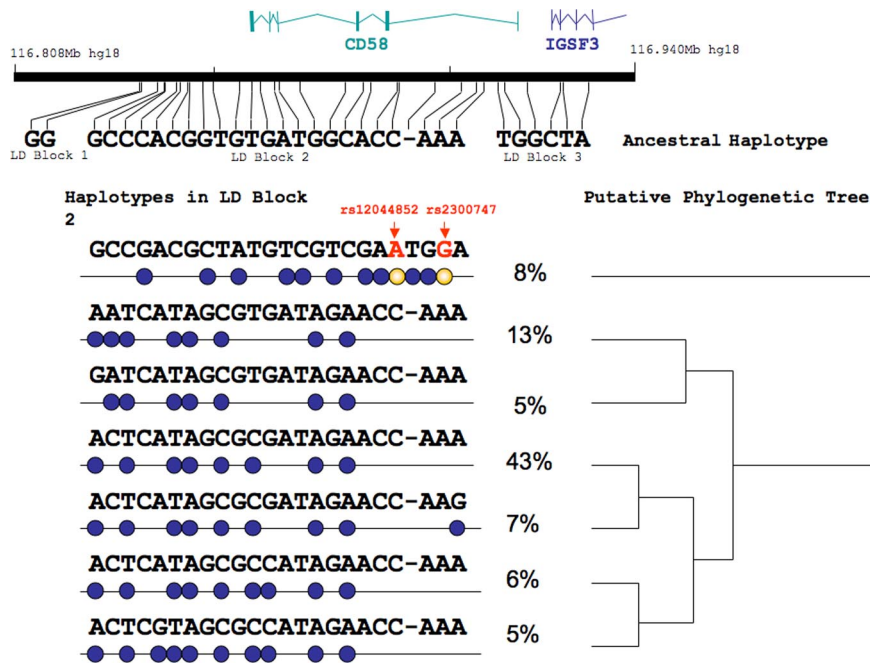


Fig. 1. The minor alleles of the 2 SNPs with the lowest association P -values are found on the same haplotype that spans most of the *CD58* gene. The segment of chromosomal DNA under study is shown in black, with its physical boundaries noted based on human genome assembly 18; the location of the *CD58* gene is shown in light blue. The flanking genes are *IGSF3* (centromeric, navy blue) and *ATP1A1* (telomeric, outside of the segment shown here). Below, we show the 31 SNPs and 1 insertion/deletion polymorphism (indel: $-T$) found within this chromosomal segment that we have genotyped. These polymorphisms are divided into 3 distinct blocks of linkage disequilibrium based on genotyping data from both cases and healthy control individuals from the Brigham and Women's Hospital collection (see Fig. S1 for the detailed linkage disequilibrium structure and list of haplotypes found in this region). All haplotypes in the second block of linkage disequilibrium with a frequency >0.05 in the Brigham and Women's Hospital samples are shown with their unique sequences of alleles. The ancestral alleles of each polymorphisms, as recorded in the HapMap (27) or dbSNP resources (<http://www.ncbi.nlm.nih.gov/projects/SNP/>), are used to determine the ancestral haplotype, which serves as the reference haplotype, at the top of the figure. The derived (non-ancestral) allele of a SNP is highlighted by a navy blue circle on each haplotype. The minor alleles of the 2 SNPs (rs12044852 and rs2300747; Table S2) with a P value < 0.0005 in our primary association analysis to MS are highlighted in yellow; both minor alleles are on the same haplotype that is found at a frequency of 8% in the Brigham and Women's Hospital samples. Their location is highlighted in red. It is interesting to note that this 8% haplotype, which is underrepresented in subjects with MS, is the most divergent of the haplotypes with a frequency $>5\%$ in subjects of European ancestry studied here. The frequency of each haplotype is noted to the right, and the dendritic tree illustrates a possible evolutionary relationship of the different haplotypes.

24 SNPs that capture common variation within this region of the genome (Table S2). These SNP data allowed us to define groups of markers that are correlated and define chromosomal segments that tend to be inherited as a block (Fig. 1). Further analysis in the context of this linkage disequilibrium structure shows that the association to MS is located within the central chromosomal segment (Fig. S1). More specifically, one version of this chromosomal segment, a haplotype found at 8% frequency in our subjects, contains the minor "A" allele of rs12044852 (rs12044852^A) and is under-represented in subjects with MS ($P = 0.0015$). None of the haplotypes containing the more common "G" major allele at rs12044852 has significant evidence of association (Fig. S1). These results suggest that an allele protecting subjects from MS exists somewhere within a 76 kb segment of DNA that only contains the *CD58* gene (Fig. 1).

We then sequenced 16 selected individuals over this 76,048 bp DNA segment. Seventeen putative new SNPs were identified and underwent validation (Table S3). To identify a better marker for the *CD58* association to MS, we then genotyped those SNPs that had been validated and demonstrated some level of correlation with rs12044852. We also genotyped additional SNPs that provided information regarding genetic variation within the *CD58* gene region but which had not yet been assessed by the initial panel of 24 SNPs (Table S2). Using these data, we assessed the possibility of other independent associations to MS susceptibility within the *CD58* locus (allelic heterogeneity) by performing a conditional analysis in our fine mapping data to account for the effect of the most associated SNP (Table S4). Once the effect

of the associated *CD58* protective allele is accounted for, as estimated by either of the 2 best markers (rs12044852^A or rs2300747^G), there is no residual evidence of association to MS susceptibility within the *CD58* locus. This result suggests that a single allele or a group of alleles that are strongly correlated explains the association of the *CD58* locus to MS susceptibility.

We then extended our mapping effort by genotyping the 15 most associated SNPs from our original fine mapping screen in an additional 1,278 trio families with MS (Table S5a). The 2 best SNPs from this trio analysis [rs2300747 and rs12044852, which are strongly correlated ($r^2 = 0.929$) in HapMap CEU samples; ref. 11] were then genotyped in an additional 3,341 MS cases and their controls. Once all data are pooled, rs2300747 is the most associated marker ($P = 1.1 \times 10^{-6}$, odds ratio 0.82, 95% confidence interval 0.75–0.89) (Fig. S2 and Table S5b). The minor allele rs2300747^G is found in the protective haplotype that contains the *CD58* gene (Fig. 1), and we therefore consider this allele to be a marker for a protective effect in MS susceptibility.

The *CD58* Protective Allele Affects RNA Expression. The rs2300747 polymorphism is found within the first intron of *CD58* and does not have a known functional consequence. Thus, we investigated the effect of the MS associated allele on expression of *CD58* RNA using data generated by the Sanger Institute from EBV-transformed lymphoblastic cell lines (LCL) used in the HapMap project (11, 12). Using the quantitative trait analysis module implemented in the PLINK toolkit (13), we find that the protective rs2300747^G allele is associated with increased expres-

sion of *CD58* RNA in 60 unrelated LCLs of European ancestry [Centre d'Etude du Polymorphisme Humain (CEPH), Utah residents with ancestry from northern and western Europe (CEU) LCLs] ($P = 0.038$) and in 89 unrelated LCLs of East Asian ancestry [Han Chinese in Beijing (CHB) and Japanese in Tokyo (JPT) LCLs] ($P = 1.1 \times 10^{-10}$). The higher frequency of rs2300747^G in the larger sample of East Asian LCLs (frequency = 0.66 vs. 0.13 in the CEU LCLs) explains in part the more extreme association of rs2300747^G with higher levels of *CD58* expression in the East Asian LCLs. This association of rs2300747^G with higher expression of *CD58* RNA is best illustrated by plotting the RNA expression values of individual LCLs and organizing them by genotype class (Fig. 2A): the rs2300747^{GG} homozygote class has a higher level of *CD58* RNA expression than does the rs2300747^{AA} homozygote class, and the rs2300747^{AG} heterozygote class has an intermediate level of expression. This effect on *CD58* RNA expression is also observed in an independent set of 400 independent LCLs from subjects of British ancestry for which similar data have been collected (14). In these samples, rs2300747 has not been genotyped, but the minor allele of rs6677309, a SNP which is strongly correlated with rs2300747 ($r^2 = 0.87$ in CEU HapMap samples; ref. 11), is seen to be associated with higher levels of *CD58* RNA expression ($P = 2.1 \times 10^{-5}$). In addition, the correlation of rs2300747^G with RNA expression is specific to *CD58*: it is not seen with the 2 genes flanking *CD58*. Using the more informative East Asian HapMap LCLs, we repeated the quantitative trait analysis and found no evidence for association of rs2300747^G with RNA expression of the flanking *ATPIA1* gene ($P = 0.96$) on the telomeric side or of the *IGSF3* gene ($P = 0.50$) on the centromeric side. The LCLs from individuals of European ancestry showed similar results (data not shown). Thus, this putative “protective” *CD58* allele for MS may exert its effect on disease risk by specifically increasing the expression of *CD58* RNA in a dose-dependent manner.

We next validated this in vitro observation using ex vivo data: we examined a data set derived by extracting mRNA from circulating mononuclear cells from 239 subjects with remitting-relapsing MS (RR MS) or CIS for evidence of correlation between rs2300747 and the expression of *CD58* RNA. Since some of these subjects with MS were treated at the time of sampling, we first established that there was no significant difference in *CD58* RNA expression among subjects that are untreated ($n = 81$), treated with glatiramer acetate ($n = 64$), or treated with an IFN beta formulation ($n = 94$) (Fig. S3). We then pooled all of these subjects into a genotypic analysis. There is only one rs2300747^{GG} homozygote, so we compared the level of *CD58* RNA expression in rs2300747^{AA} homozygotes to the expression level observed in subjects bearing at least one rs2300747^G allele. This analysis validates our in vitro LCL results by demonstrating that the protective rs2300747^G allele is associated with higher expression of *CD58* RNA ($P = 0.0037$) in mononuclear cells of RRMS and CIS subjects (Fig. 2B).

The *CD58* Locus and Clinical Manifestations of MS. To assess whether *CD58* mRNA levels correlated with clinical disease activity, we analyzed RNA data captured from a different set of subjects with MS who were experiencing either a clinical relapse or remission. These data had been generated as part of an independent project that analyzed changes in RNA expression of 9,381 genes to discover relapse- and remission-specific patterns of gene expression in whole blood of untreated subjects with MS (15). Of the 38 putative MS susceptibility loci with evidence of replication in the recent whole genome association scan for MS (5), *CD58* is the only one whose RNA expression is enhanced in subjects in clinical remission (Fig. S4). The expression of *CD58* RNA in whole blood during a remission is, on average, 1.7-fold greater than baseline expression in healthy control subjects, and this

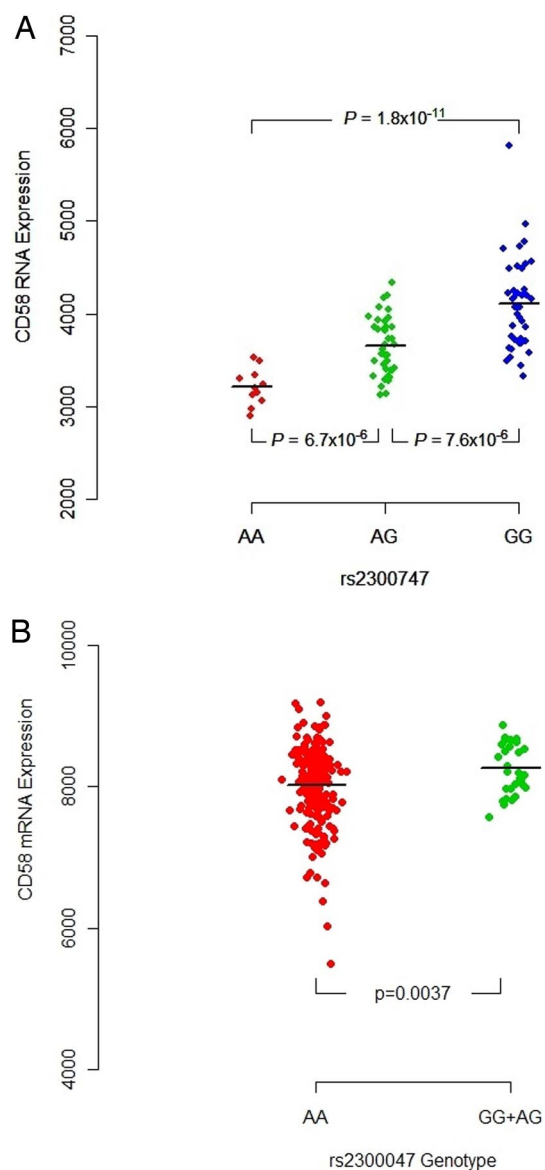


Fig. 2. The *CD58* protective allele increases *CD58* RNA expression. (A) To illustrate the association of the rs2300747^G marker to higher *CD58* expression, we have plotted the distribution of RNA expression values by genotype and show the results of pairwise comparisons between genotype categories. The location of the mean value for each category is denoted by a black line. We present the results of the genotypic comparisons among the 89 LCLs of East Asian ancestry that offer a much more robust estimate of the allele's effect given the higher frequency (0.66) of rs2300747^G in this population. The allele frequency of rs2300747^G is much lower (0.13) in the smaller collection of 60 LCLs of European ancestry; this precludes a robust pairwise assessment of genotype categories. The cell surface expression of *CD58* in these cell lines shows the same pattern of association to genotype—higher surface expression in rs2300747^{GG} subjects—but does not reach statistical significance given the limited number of cell lines that are available (data not shown). (B) To validate the initial observation obtained from the LCL data, we present evidence that the correlation of higher *CD58* RNA expression with rs2300747^G is also seen ex vivo in RNA data obtained from mononuclear cells of subjects with remitting-relapsing MS or CIS. Since a single MS subject had the rare rs2300747^{GG} genotype, this subject was pooled with the 29 rs2300747^{AG} heterozygotes (green) and compared to the 209 rs2300747^{AA} homozygotes (red) with MS or CIS. The location of the mean value for each category is denoted by a black line.

remission-associated increased expression of *CD58* is significantly greater than the levels of *CD58* expression seen in subjects with MS that are sampled during a relapse ($P = 0.011$) (Fig. 3).

