Interferon regulatory factor 5 (IRF5) gene variants are associated with multiple sclerosis in three distinct populations

G Kristjansdottir,1 J K Sandling,1 A Bonetti,2 I M Roos,3 L Milani,1 C Wang,1 S M Gustafsdottir,4 S Sigurdsdottir,1 A Lundmark,1 P J Tienari,2 K Koivisto,5 T Ilola,6 M Reunen,8 P Peltonen,9 J Saarela,9 J Hillert,7 T Olsson,10 U Landegren,4 A Alcina,11 O Fernández,12 L Leyva,12 M Guerrero,13 M Lucas,14 G Izquierdo,15 F Mateus,11 A-C Syvänen1

For numbered affiliations see end of article

ABSTRACT

Background: IRF5 is a transcription factor involved both in the type I interferon and the toll-like receptor signalling pathways. Previously, IRF5 has been found to be associated with systemic lupus erythematosus, rheumatoid arthritis and inflammatory bowel diseases. Here we investigated whether polymorphisms in the IRF5 gene would be associated with yet another disease with features of autoimmunity, multiple sclerosis (MS).

Methods: We genotyped nine single nucleotide polymorphisms and one insertion-deletion polymorphism in the IRF5 gene in a collection of 2337 patients with MS and 2813 controls from three populations: two case-control cohorts from Spain and Sweden, and a set of MS trio families from Finland.

Results: Two single nucleotide polymorphism (SNPs) (rs4728142, rs3807306), and a 5 bp insertion-deletion polymorphism located in the promoter and first intron of the IRF5 gene, showed association signals with values of p<0.001 when the data from all cohorts were combined. The predisposing alleles were present on the same common haplotype in all populations. Using electrophoretic mobility shift assays we observed allele specific differences in protein binding for the SNP rs4728142 and the 5 bp indel, and by a proximity ligation assay we demonstrated increased binding of the transcription factor SP1 to the risk allele of the 5 bp indel.

Conclusion: These findings add IRF5 to the short list of genes shown to be associated with MS in more than one population. Our study adds to the evidence that there might be genes or pathways that are common in multiple autoimmune diseases, and that the type I interferon system is likely to be involved in the development of these diseases.

Multiple sclerosis (MS, OMIM 126200) is an inflammatory disease estimated to affect over two million individuals worldwide. MS is not well recognised as an autoimmune disease, but exhibits features of autoimmunity—that is, activation of the immune system in the absence of apparent ongoing infection.1 In MS the presumed target for the autoimmune process is the central nervous system, and the disease manifests itself by immune mediated demyelination and damage to axons. A spectrum of neurological symptoms are found among MS patients, including sensory or motor pareses, visual disturbances, ataxia, pain, cognitive dysfunction and fatigue. MS is a complex disease caused by interaction between environmental and inherited factors. The disease shows familial clustering, and twin studies have revealed that a large portion of this clustering can be attributed to shared genes.2

A remarkably low number of susceptibility genes for MS have been identified so far. Genome-wide linkage studies have indicated several possible susceptibility loci, but the only locus to be identified across most studies is the major histocompatibility complex (MHC) on chromosome 6p21, where the HLA-DRB1*1501 allele is a well established genetic risk factor for MS.3 4 This locus does not, however, account for the whole genetic component of MS, and multiple loci with smaller contributions to disease susceptibility are likely to exist. Numerous candidate gene studies have also been performed in MS, but findings from one population have been difficult to replicate in other populations. The protein kinase C α (PRKCA) gene is one of few genes reported to be associated with MS in more than one population.5 6 Recently, the interleukin 7 receptor a chain gene (IL7RA) was found to be associated with MS in two independent candidate gene studies and in a genome-wide association study.7 8 The genome-wide association study also identified variants in the interleukin 2 receptor a chain gene (IL2RA) as risk factors for MS,7 which is in accordance with findings in a previous candidate gene study.9

It is relatively common that patients affected by an autoimmune disease suffer from another autoimmune disease, and that members of the same family suffer from different autoimmune diseases. For example, in families with systemic lupus erythematosus (SLE, OMIM 152700), MS and rheumatoid arthritis (RA, OMIM 180800) occur more frequently than in the general population.11 Such observations suggest shared genes or involvement of common cellular pathways in these diseases. This hypothesis is supported by reports on genes found to be associated with more than one autoimmune disease in experimental models of RA and MS.12 Shared genes in autoimmune diseases are becoming apparent also in humans, such as PTPN22 in RA and SLE,13 MHC2TA in RA and MS,14 and recently the suggested involvement of CD24 in MS and SLE.15
The type I interferon (IFN) system has been postulated to play a key role in autoimmune diseases. Increased expression of IFN induced genes has been detected in autoimmune diseases like SLE, RA, Sjögren’s syndrome, and in a subgroup of MS patients. The interferon regulatory factors (IRFs) are major regulators of genes activated by the type I IFNs, and a role in the regulation of the immune system is well established for the majority of the members of the IRF family of nine genes. The role of IRF5 in the immune response is not as well established as for other IRFs, but IRF5 has recently received attention in studies of autoimmunity. The IRF5 gene displays a complex transcription pattern with three alternative non-coding 5’ exons and at least nine alternatively spliced mRNAs. IRF5 is expressed in dendritic cells, monocytes and B cells, but its expression can be induced in other cell types by the type I IFNs. IRF5 regulates the toll-like receptor (TLR) dependent activation of inflammatory cytokines and functions downstream of the TLR-MyD88 pathway where it is activated by MyD88 and TNF receptor associated factor 6 (TRAF6). Our original finding of an association between the IRF5 gene and SLE, which has been replicated in multiple populations, as well as our recent findings of association between IRF5 and RA and inflammatory bowel diseases (IBD), provide additional support for the important role of IRF5 and the type I IFN system in autoimmune diseases. Inspired by these findings, and by the role of recombinant IFN β as a standard treatment of MS, we investigated whether polymorphisms in the IRF5 gene would also be associated with MS. We found that polymorphisms in the IRF5 gene displayed associations with MS in three independent patient cohorts from Spain, Sweden and Finland. The three most strongly associated polymorphisms are located in the promoter region and first intron of IRF5. A functional role is suggested by increased protein binding to the risk alleles of two of these polymorphisms.

**PATIENTS AND METHODS**

**Clinical samples**

Cohorts of MS patients collected in Spain, Sweden and Finland were included in the study. All patients had clinically or laboratory supported definite MS according to the Poser criteria or fulfilled the criteria of McDonald for MS. In total 660 Spanish MS patients were recruited at three public hospitals in the south of Spain: Clínico in Granada (n = 130), Carlos Haya in Málaga (n = 563) and Virgen de la Macarena in Seville (n = 167). The Spanish controls were 833 blood donors without history of inflammatory disease attending the blood banks of Granada (n = 441), Seville (n = 211) and Málaga (n = 181); 67% of the patients and 50% of the controls were female. The Swedish cohort consisted of 1166 MS patients (67% females) and 1255 controls (65% females) of Nordic ethnicity recruited at Danderyd’s Hospital or Karolinska University Hospital in Huddinge or in Solna, all located in the Stockholm County of Sweden. Controls were consecutive blood donors of Nordic origin that visited three blood donor facilities in the Stockholm area in 2001 and in 2004/2005. The Finnish cohort consisted of 511 MS trio-families recruited at five centres in Finland: the University Central Hospitals in Helsinki (n = 148 families), Tampere (n = 102), Kuopio (n = 79), Oulu (n = 73), and the Central Hospital of Seinäjoki (n = 109); 71% of the Finnish patients were female. The study was approved by the respective local ethic committees and informed consent was obtained from all study participants.

**Genotyping**

Nine single nucleotide polymorphisms (SNPs) and one 5 bp biallelic insertion-deletion polymorphism (CGGGG indel) in the IRF5 gene were genotyped. The SNPs were genotyped using multiplex fluorescent minisequencing (single base extension) with the SNPstream system (Beckman Coulter). The SNP rs4728142 was also genotyped using a homogeneous minisequencing assay with fluorescence detection (FP-TDI) ( Analyst AD, Molecular Probes) in 1440 Swedish samples. The genotype call rate in the samples was on average 96.6% and the genotyping accuracy was 99.8% as estimated from 13,800 genotype comparisons (20% of the genotypes) between repeated assays. The CCGGGG indel was amplified as a 100/105 bp polymerase chain reaction (PCR) fragment using fluorescent primers with subsequent fragment analysis on an ABI PRISM 3730 DNA Analyser (Applied Biosystems, Foster City, California, USA). The GeneMapper v.3.7. software was used for genotype calling. Alternatively, the PCR amplified fragments were separated on 4% MetaPhor high resolution agarose gels (Cambrex Bio Science Rockland Inc, Maine, USA) and visualised using ethidium bromide staining. The call rate for genotyping the CCGGGG indel was on average 96.0% in the three cohorts and the genotyping accuracy was 99.1% as estimated from genotype comparison between repeated assays for 15% of the subjects. Samples with a genotype call rate <50% for the 10 markers and four trios with inheritance errors were excluded from the analysis. The PCR and extension primers are provided.

---

**Figure 1** Schematic illustration of the IRF5 gene with the positions of the analysed polymorphisms. The exons and the 3’-UTR are shown as boxes with the exons labelled 1-9 and the untranslated alternative exons in the 5’-end of IRF5 labelled as 1a, 1b and 1c. The translation initiation site is indicated by an arrow above the gene. The single nucleotide polymorphism (SNP) rs2004640 is located at the splice junction of alternative exon 1b, where it alters a consensus splice donor site that allows expression of IRF5 mRNA containing exon 1b. The SNP rs10954213 located in the 3’-untranslated region (UTR) of the gene correlates with an altered length of the IRF5 3’ UTR and thereby affects the stability of the IRF5 transcript. The SNPs rs2280714, located —6 kb downstream of the IRF5 gene, has been reported to be correlated with variation in IRF5 mRNA expression levels. A 5 bp insertion-deletion polymorphism (CGGGG indel) in the promoter region of IRF5, identified by sequencing the IRF5 gene in Swedish individuals, was included in our study because it is predicted to alter a binding site for the transcription factor SP1, which could affect the expression of IRF5. The SNP rs12539741 was included because this SNP, and several other SNPs located 3’ of IRF5 that are in almost full linkage disequilibrium (LD) with it, show very strong association signals with systemic lupus erythematosus.
in the supplementary table S1. All genotyped polymorphisms fulfilled the criteria of Hardy–Weinberg equilibrium in the control samples.

**Electrophoretic mobility shift assay (EMSA)**

For each allele of the polymorphism, pairs of single stranded 5' biotinylated and unlabelled 30–37 bp oligonucleotides (obtained from IDT Inc, Coralville, Iowa, USA) were allowed to anneal to generate double stranded probes (supplementary table S2). Twenty fmoles of labelled probes was incubated for 20 min with 2 mL of nuclear extract prepared from blood cells in a freshly made binding buffer supplemented with poly(dI-dC)-poly(dI-dC) and protease inhibitors. Competition experiments were performed using a 100-fold molar excess of unlabelled probe. The binding reactions were analysed using electrophoresis on 6% polyacrylamide gels and transferred to nylon membranes (Bio-Rad Laboratories, Hercules, California, USA). The biotinylated fragments were detected by a chemiluminescent procedure (LightShift Chemiluminescent EMSA kit, Pierce Biotechnology, Rockford, Illinois, USA).

**Proximity ligation assay (PLA)**

Polyclonal antibody against SP1 was purchased from Santa Cruz Biotechnology (Cat. no. sc-14027, Santa Cruz, California, USA). The antibody was biotinylated using D-biotin-N-hydroxysuccinimide (Nordic Biosite, Täby, Sweden) according to the recommendation by the manufacturer. The biotinylated antibody was diluted in phosphate buffered saline (PBS) containing 1% bovine serum albumin (BSA) to a final

## Table 1

| Polymorphism | Spanish c/c* | Swedish c/c* | Finnish trios | Combined p value |
|--------------|--------------|--------------|---------------|-----------------|
| rs729302     | 0.22         | 0.29         | 0.047         | 0.07            |
| rs4728142    | 0.0029       | 0.020        | 0.035         | 0.0002          |
| rs3757305    | 0.0005       | 0.43         | 0.53          | 0.04            |
| CGGGG indel  | 0.011        | 0.0093       | 0.056         | 0.0006          |
| rs2004640    | 0.0011       | 0.082        | 0.54          | 0.003           |
| rs3807306    | 0.0037       | 0.049        | 0.012         | 0.0002          |
| rs10954213   | 0.031        | 0.15         | 0.010         | 0.003           |
| rs11770589   | 0.29         | 0.49         | 0.070         | 0.16            |
| rs2280714    | 0.0030       | 0.43         | 0.30          | 0.02            |
| rs12539741   | 0.074        | 0.10         | 0.74          | 0.11            |

*p/c denotes case–control; p values below 0.05 are indicated in bold.
concentration of 100 nM. The anti-SP1 antibody was then combined in a 1:1 ratio with a streptavidin–oligonucleotide conjugate (Solulink, San Diego, California, USA) in a volume of 20 µl, incubated at room temperature for 1 h, and then stored at 4°C until use. High pressure liquid chromatography (HPLC) purified DNA probes (Biomers.net, Ulm, Germany) were made partially double stranded according to Gustafsdottir et al.31 The partially double stranded DNA probes (25 pM) were incubated with 200 ng of Jurkat nuclear extract (Active Motif, Carlsbad, California, USA) in PBS containing 1% BSA, 16 µg/ml of sheared polyA bulk nucleic acid (Sigma-Aldrich, Stockholm, Sweden), 1 mM D-biotin (Molecular Probes, Eugene, Oregon, USA) in a total volume of 9 µl for 30 min at room temperature; 6 µl of 25 pM anti-SP1-DNA conjugate was added to the mixture, and the incubation was continued for 2 h at 20°C. After the incubations, 55 µl of a combined mixture for ligation and real-time PCR amplification/detection was added. Ligation and PCR amplification/detection was performed according to Gustafsdottir et al.31 Sequences of the primers, probes and the streptavidin–oligonucleotides used in PLA are provided in supplementary table S2.

Statistical analysis
A χ² test (p = 0.05) was used to assess that the genotype distributions of the polymorphisms fulfilled the criteria of Hardy–Weinberg equilibrium. The PLINK software (http://pngu.mgh.harvard.edu/~purcell/plink/) was used to compare the allele counts in cases and controls by Fisher’s exact test and to calculate odds ratios (OR) with 95% confidence intervals. PLINK was also used to perform the sliding window haplotype association analysis. The Haplovew v.3.3. software was used to determine linkage disequilibrium (LD) between the polymorphisms. In the family cohort, the genotype data were analysed using the TRANSMIT 2.5.2. computer program package.32 Alleles that were included in the analysis were required to be transmitted from a least 20 informative meioses, corresponding to the Var (O-E) value of 5 or more in the TRANSMIT analysis. A combined p value from case–control cohorts and trios were calculated using the analytical expression by Jost, k(1-<ln(k)>)+(<ln(k))/2, where k is the product of the p values from the different cohorts (http://www.loujost.com). This formula is an analytical solution of Fischer’s original formula for combining p values.33 The ssSNPer tool was used to determine the pair-wise correlation between SNPs in the HapMap data.34

RESULTS
Ten polymorphisms in the IRF5 gene on chromosome 7q32 were genotyped in MS patient samples collected in three European countries. The selected polymorphisms include five SNPs and one insertion-deletion polymorphism (indel) in the promoter region or first intron of IRF5, two SNPs in the 5’UTR and two SNPs downstream of IRF5 (fig 1). The polymorphisms were selected because they have previously been shown to be associated with SLE,25 35 36 RA26 and IBD,27 or have been suggested to modulate the expression of IRF5.37

The polymorphisms were first genotyped in a Spanish cohort of MS patients (n = 660) and controls (n = 553). Seven of the 10 polymorphisms showed nominally significant signals of association with MS (p<0.05) (table 1). To replicate this finding, we genotyped the same set of polymorphisms in an independent case–control cohort with 1166 MS patients and 1285 matched controls from Sweden. Two of the SNPs, rs4728142 and rs3807306, and the CGGGG indel polymorphism, showed a nominally significant association (p<0.05) with MS. Each of these three polymorphisms was also associated with MS in the Spanish samples (table 1). In a further attempt to verify these findings from two case–control cohorts of MS patients, we genotyped the same set of polymorphisms in 511 Finnish MS trio families. Using a transmission disequilibrium test, four polymorphisms showed nominally significant association with MS (table 2). Two SNPs, rs4728142 and rs3807306, were nominally significantly associated with MS in all three cohorts (tables 1 and 2).

Information on HLA-DRB1*15 genotype was available from the Swedish cases and controls, and from the Finnish patients (supplementary table S3). In the Swedish dataset most of the association signal was observed in the DR15-negative stratum, whereas in the Finnish dataset the DR2 stratum provided most of the association signal (supplementary table S4). When combining the p values from all cohorts, seven of the 10 polymorphisms exhibited association with MS with p values ranging from 0.0002 to 0.04. The strongest association signals in this combined analysis were observed for the SNP rs4728142 (p = 0.0002) located ~5 kb upstream of IRF5, the CGGGG indel located 64 bp upstream of exon 1a of IRF5 (p = 0.0005), and the SNP rs3807306 in the first intron of IRF5 (p = 0.0002) (table 3, fig 1).

The linkage disequilibrium (LD) pattern of the polymorphisms was similar between the three populations (figure 2), and displayed relatively high LD with pair-wise r² values of 0.61–0.85 for the three associated polymorphisms in the Spanish, Swedish and Finnish unaffected subjects. We performed haplotype association tests in the three populations to investigate whether haplotypes could capture the association signal at a higher significance than the individual polymorphisms, and whether all three populations harbour the same disease associated haplotype(s). The haplotype analysis was performed using a five marker sliding window approach. The associated risk alleles of the polymorphisms were all present on
### Table 4  Sliding window haplotype association analysis:

|   | Spain | Sweden | Finland |
|---|-------|--------|---------|
|   | F_{aff} | F_{corr} | p Value | F_{aff} | F_{corr} | p Value | F_{obs} | Obs-Exp | p Value | Combined p value | Combined corrected p value |
| A | A | G | in | T | 0.49 | 0.44 | 0.0036 | 0.44 | 0.41 | 0.019 | 0.41 | + | 0.010 | 0.000081 | 0.0032 |
| C | G | T | del | G | 0.17 | 0.18 | 0.55 | 0.17 | 0.18 | 0.61 | 0.20 | – | 0.42 | 0.68 | NS |
| A | G | T | del | G | 0.13 | 0.17 | 0.0021 | 0.15 | 0.15 | 0.60 | 0.14 | + | 0.47 | 0.021 | NS |
| A | G | in | T | T | 0.51 | 0.46 | 0.0034 | 0.47 | 0.44 | 0.019 | 0.43 | + | 0.028 | 0.00018 | 0.0070 |
| G | T | del | G | G | 0.28 | 0.34 | 0.0016 | 0.28 | 0.29 | 0.49 | 0.26 | – | 0.51 | 0.016 | NS |
| G | G | del | G | G | 0.10 | 0.11 | 0.28 | 0.12 | 0.13 | 0.20 | 0.10 | – | 0.29 | 0.22 | NS |
| G | in | T | T | A | 0.53 | 0.48 | 0.0114 | 0.49 | 0.45 | 0.0061 | 0.45 | + | 0.012 | 0.00012 | 0.0046 |
| T | del | G | G | G | 0.27 | 0.32 | 0.0018 | 0.28 | 0.29 | 0.53 | 0.25 | – | 0.54 | 0.019 | NS |
| G | del | G | G | A | 0.10 | 0.11 | 0.29 | 0.12 | 0.13 | 0.28 | 0.11 | – | 0.79 | 0.48 | NS |
| in | T | T | A | A | 0.42 | 0.39 | 0.16 | 0.34 | 0.32 | 0.18 | 0.30 | + | 0.068 | 0.052 | NS |
| del | G | G | G | G | 0.27 | 0.32 | 0.0018 | 0.28 | 0.30 | 0.26 | 0.25 | – | 0.52 | 0.011 | NS |
| ins | T | T | A | G | 0.11 | 0.09 | 0.068 | 0.15 | 0.13 | 0.088 | 0.15 | + | 0.24 | 0.034 | NS |
| T | T | A | A | T | 0.44 | 0.41 | 0.13 | 0.35 | 0.33 | 0.11 | 0.31 | + | 0.044 | 0.022 | NS |
| G | G | G | G | C | 0.27 | 0.32 | 0.0026 | 0.28 | 0.29 | 0.35 | 0.25 | – | 0.42 | 0.015 | NS |
| G | A | A | A | T | 0.11 | 0.13 | 0.22 | 0.13 | 0.14 | 0.35 | 0.13 | – | 0.56 | 0.39 | NS |
| T | A | A | T | C | 0.46 | 0.42 | 0.0376 | 0.39 | 0.37 | 0.19 | 0.39 | + | 0.016 | 0.010 | NS |
| G | G | G | C | C | 0.27 | 0.32 | 0.0028 | 0.28 | 0.29 | 0.38 | 0.25 | – | 0.44 | 0.018 | NS |
| G | A | A | T | C | 0.11 | 0.13 | 0.25 | 0.13 | 0.14 | 0.40 | 0.13 | – | 0.59 | 0.46 | NS |

*Haplotype frequency in Finnish trios. †Difference between observed and expected transmissions of the associated allele; + denotes increased and – denotes decreased transmission to affected offspring. ‡39 haplotypes were tested and a conservative Bonferroni correction factor of 39 was applied. The three highest frequency haplotypes in each window are shown. p Values are indicated in bold.
DISCUSSION

Here we describe the association of three polymorphisms in the IRF5 gene with MS in three independent patient cohorts from Spain, Sweden and Finland. The SNPs rs4728142 and rs3807306 reached nominal significance for association in all cohorts and revealed strong signals of association with MS (p = 0.0002) when the data from all three cohorts was combined. The SNP rs4728142 was located ~5 kb upstream of the alternative exon 1a of IRF5, and rs3807306 is located in the first intron of the gene, 96 bp upstream of alternative exon 1c (fig 1). The SNP rs3807306 was the only IRF5 polymorphism reported in the recently performed genome-wide association study on MS.39 In that study the SNP rs3807306 exhibited association with MS with a TDT p value of 0.014 (https://imsgc.org/), but it did not pass the threshold for inclusion in the confirmatory phase of the study. This study also identified the T allele as the risk allele. It is notable that the initial TDT screening phase of this genome-wide study only had about 6% power to detect each locus at the chosen cut off and with the observed modest risk ratios.7 The effect of this low power becomes apparent when looking at the SNPs in IL2RA and IL7RA, which did not pass the initial p value cut off for the TDT, but turned out to be the most strongly associated markers after the replication phase. In our study the CGGGG indel polymorphisms, located 64 nucleotides upstream of the alternative exon 1a of IRF5, also showed evidence of association with MS. Although the CGGGG indel did not reach statistical significance independently in all three cohorts, the combined analysis revealed a clear association signal between the CGGGG indel and MS (p = 0.0005), with the longer allele (4 x CGGGG) as the risk allele.

The risk alleles of the SNPs rs4728142 and rs3807306 and of the CGGGG indel are present on the same common haplotype in each of the three populations (table 4). In a recent study on IBD we analysed the same set of polymorphisms as in the current study on MS, and found that the same three polymorphisms were associated with IBDs, with the strongest signal of association for the CGGGG indel.27 We have recently performed an association study of a comprehensive set of polymorphisms in IRF5 that were identified by sequencing the introns and exons of IRF5 in SLE patients.41 This analysis identified a set of correlated polymorphisms in IRF5 that gave strong signals of association with SLE (p<10^-5), including the SNPs rs4728142, and rs3807306 and the CGGGG indel. In SLE logistic regression analysis conditional only on the CGGGG indel abolished all the other association signals from this set of correlated polymorphisms. Because the effect size of the polymorphisms are lower in MS than in SLE we cannot distinguish which one of them would be the most likely causal variant in MS. In an earlier association study on RA we analysed five SNPs in IRF5, including the SNPs rs729302, rs375385, rs2004640 and 3807306,

the most common haplotype, which had a frequency of 0.43–0.51 and was the same in all three cohorts (table 4). The association signals were comparable to those in the analysis of individual polymorphisms.

We used electrophoretic mobility shift assays (EMSA) to test for differential protein binding to the alleles of the three polymorphisms rs4728142, rs3807306 and the CGGGG indel, which displayed association signals with values of p<0.001 in the combined analysis of the three MS cohorts. This analysis revealed a stronger binding of protein to the risk alleles of the SNP rs4728142 (the A allele) and of the CGGGG indel polymorphism (the 4 x CGGGG allele) (fig 3). For the CGGGG indel polymorphisms the insertion of one CGGGG repeat in the longer (4 x CGGGG) allele is predicted to create an additional and third binding site for the transcription factor SP1, while the shorter (3 x CGGGG) allele has two SP1 binding sites (TFSEARCH: http://www.cbrc.jp/research/db/TFSEARCH.html).39 By using the proximity ligation assay41 42 we confirmed that SP1 protein binds to the alleles of the CGGGG indel, and that an increased amount of SP1 is bound to the 4 x CGGGG allele than to the 3 x CGGGG allele of the indel polymorphism (fig 4).
but not the CGGGG indel, and found the strongest signal of association with MS for the SNP rs3807306. Together, the association results from the current study on MS and previous studies on IBD, SLE and RA indicate that one or more of these correlated polymorphisms in the promoter and first intron of the IRF5 gene could be a universal risk factor for chronic inflammatory disorders, but further studies are required to dissect if the mechanism is the same in all of these disorders.

Using EMSA, as a preliminary functional test for the three polymorphisms in IRF5 that gave the strongest association signals with MS, we observed stronger protein binding to the risk alleles of the SNP rs47278142 (the A allele) and of the CGGGG indel polymorphisms (the 4×CGGGG allele), whereas both alleles of the SNP rs3807306 appear to bind an equal amount of protein. We confirmed experimentally using an antibody against SP1 in the proximity ligation assay that an increased amount of transcription factor SP1 binds to the risk allele (the 4×CGGGG allele) of the CGGGG indel polymorphism. This result is also supported indirectly by a study in which they demonstrated that SP1 binds to a similar sequence motif in the IRF4 and IRF5 genes.

It is notable that the SNP rs12559741 located in the 3′-end of the IRF5 gene does not show any association with MS in our study. In SLE this SNP and two other linked SNPs, the SNPs rs2070197 in the 3′-UTR of IRF5 and rs10488681 located ~5 kb downstream of IRF5, give particularly strong association signals. Two recent genome-wide association studies failed to detect an association with RA for the SNP rs10488681 or its proxies and in our study on IBD we did not detect an association with the SNP rs10488681 either. Thus it appears that in SLE there are two groups of independently associated polymorphisms in the IRF5 gene region, whereas in MS, RA and IBD association from only one of these groups is observed. According to data from the HapMap project (www.hapmap.org), the SNP rs10488681 is in complete LD with multiple SNPs located in a 100 kb region downstream of IRF5, which also contains the transportin 3 (TNPO3) gene. The polymorphisms in the IRF5 gene that we found to be associated with MS in our study are not strongly correlated (r² 0.1−0.2) with the SNP rs10488681 or its proxies in the TNPO3 gene, indicating that it actually is IRF5, and not TNPO3, that is primarily responsible for the association with MS that we observe.

Three different European populations were included in the present study, and the observed allele and haplotype frequencies did not show major inter-population differences. The disease predisposing alleles are the same and occur on the same major haplotype in the Spanish, Swedish and Finnish populations. These results indicate that the predisposing alleles are widely distributed in the Caucasian population. The effect sizes of the risk alleles are relatively small, with odds ratios of about 1.2, indicating that large datasets are needed to replicate these findings. The strength of our study is that we used both case−controls and family based association testing. Population stratification artefacts are common in case−control settings, but unlikely in family based studies. On the other hand, the TDT may be biased by erroneous detection of association to alleles with high frequency in the analysed populations.

MS is a common disease, and most likely caused by interaction between multiple common allelic variants of genes. The association of IRF5 polymorphisms with MS in the cohorts studied here suggests that IRF5 is one of these genes that contribute to the disease. Interestingly, in the animal models for MS, experimental autoimmune encephalomyelitis, deletion of the IFNβ gene leads to more severe disease, suggesting that the inherent type I IFN function contributes to the autoimmune disease. Currently, the most common therapy in MS is IFNβ, which has been shown to reduce the magnetic resonance imaging activity and relapse rate in MS. This therapeutic effect is consistent with an immunoregulatory role of the type I IFN pathway in MS. The findings from our study add IRF5 to the short list of genes with confirmed association with MS. Our study also contributes to the evidence that there might be genes or pathways that are common between multiple autoimmune diseases, and that the type I IFN signalling system, to which the IRF5 gene belongs, is likely to be one of these pathways.

Author affiliations: 1Molecular Medicine, Department of Medical Sciences, Uppsala University, Uppsala, Sweden; 2Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Research Program, University of Helsinki, Helsinki, Finland; 3Department of Clinical Neuroscience, Division of Neurology, Karolinska Institutet, Karolinska University Hospital Huddinge, Stockholm, Sweden; 4Rudbeck Laboratory, Department of Genetics and Pathology, Uppsala University, Uppsala, Sweden; 5Central Hospital of Seinäjoki, Seinäjoki, Finland; 6Department of Neurology, University of Tampere and Tampere University Hospital, Tampere, Finland; 7Department of Neurology and Neuroscience, University of Kuopio and Kuopio University Hospital, Kuopio, Finland; 8Department of Neurology, University of Oulu and Oulu University Hospital, Oulu, Finland; 9Department of Molecular Medicine, Bioremedicums, National Public Health Institute, Helsinki, Finland; 10Department of Clinical Neuroscience, Neuroimmunology Unit, Karolinska Institutet, Stockholm, Sweden; 11Instituto de Parasitología y Biomedicina López Neyra, Consejo Superior de Investigaciones Científicas, Granada, Spain; 12Servicio de Neurología, Instituto de Neurociencias Clínicas, Hospital Regional, Universitario Carlos Haya, Málaga, Spain; 13Servicio de Neurología, Hospital Clínico San Cecilio, Granada, Spain; 14Servicio de Biología Molecular, Hospital Universitario Virgen Macarena, Sevilla, Spain; 15Unidad de Esclerosis Múltiple, Hospital Universitario Virgen Macarena, Sevilla, Spain

Acknowledgements: The SNPs were genotyped using equipment at the SNP technology platform in Uppsala (www.genotyng.se).

Funding: Financial support for the study was provided by the Swedish Research Council and the Knut and Alice Wallenberg Foundation (to A-CS), the Academy of Finland, the Sigrid Juselius Foundation and Helsinki University Central Hospital (to PJT and LP), the Center of Excellence for Disease Genetics of the Academy of Finland (to LP), the Paula Foundation (to PJT), the Finnish Cultural Foundation (to PJT and AB), the Helsinki Biomedical Graduate School (to AB), the Nils and Bibi Jensen’s Foundation (to JH) and by the grants PN-SAF 2006−202023 (to AA) and FIS-P1141298 (to FM).

Competing interests: None.

Ethics approval: The study was approved by the respective local ethics committees in Spain, Sweden and Finland.

Patient consent: Informed consent was obtained from the patients for publication of their details in this report.

REFERENCES

1. Davidson A, Diamond B. Autoimmune diseases. N Engl J Med 2001;345:340−50.
2. Kalman B, Lublin FD. The genetics of multiple sclerosis. A review. Biomed Pharmacother 1999;53:356−70.
3. Sawcer S, Ban M, Maranian M, Yeo TW, Compton A, Kirby A, Daly MJ, De Jager PL, Walsh E, Landers ES, Rioux JD, Hafler DA, Ivinson A, Rimmer J, Gregory SG, Schmidt P, Pencak−Vance MA, Akesson E, Hillert J, Datta P, Oturai A, Ryder LP, Harbo HF, Sparklund A, Myhr KM, Laaksonen M, Booth D, Heard R, Stewart G, Lincoln R, Barcellos LF, Hauser SL, Oksenberg JR, Kerealy SJ, Haines JL. A high-density screen for linkage in multiple sclerosis. Am J Hum Genet 2005;77:654−67.
4. Lincoln MR, Montpetit A, Cader MZ, Saarela J, Dymtent DA, Titular M, Ferretti V, Tienari PJ, Sadovnick AD, Petlonen L, Ebers GC, Hudson TJ. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. Nat Genet 2005;37:1108−12.
5. Saarela J, Kallo SP, Chen D, Montpetit A, Kojola K, Choi E, Asselta R, Bronnikov D, Lincoln MR, Sadovnick AD, Tienari PJ, Kovisto K, Palotie A, Ebers GC, Hudson TJ, Petlonen L. PRKCA and multiple sclerosis: association in two independent populations. PLoS Genet 2006;2:e42.
6. Barton A, Woolmore JA, Ward D, Eyre S, Hinks A, Oller WE, Strange WC, Fryer AA, John S, Hawkins CP, Worthington J. Association of protein kinase C alpha (PRKCA) gene with multiple sclerosis in a UK population. Brain 2004;127(Pt 8):1711−22.
7. Hafler DA, Compton A, Sawcer S, Lander ES, Daly MJ, De Jager PL, de Bakker PI, Gabriel SB, Minel DB, Ivinson AJ, Pericak−Vance MA, Gregory SG, Rioux JD, McCauley JL, Haines JL, Barcellos LF, Cree B, Oksenberg JR, Hauser SL. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007;357:851−62.
Sigurdsson S, van Baarsen LG, Takaoka A, Sigurdsson S, Bave U, van der Pouw Kraan TC, Wang L, Swanberg M, Olsson T, Corporaal S, Lundmark F, Gregory SG, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, Johnson KP, Kristjansdottir G, Milani L, Libioulle C, Sigurdsson S, Louis E, Wiman AC, Arthritis Rheum: gene with rheumatoid arthritis. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene confers risk of inflammatory bowel diseases. Hum Mol Genet 2008;17:369–76.

van Baarsen LG, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg NA, on October 2, 2023 by guest. Protected by copyright.http://jmg.bmj.com/ J Med Genet: first published as 10.1136/jmg.2007.055012 on 19 February 2008. Downloaded from

---

8. Gregory SG, Schmidt S, Seth P, Oksenberg JR, Hart J, Prokop A, Cailler SJ, Ban M, Goris A, Barcellos LF, Lincoln R, McAuley J, Sawcer SJ, Compton DA, Dubois B, Hamsten A, Garcia-Marcos P, Reichenberg MA, Haines JL. Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. Nat Genet 2007;39:1083–108.

9. Lundmark F, Duvefelt K, Jacobsson E, Kockum I, Wallstrom E, Khademi M, Dutta A, Ryder LP, Saarela J, Harbo HF, Celli SG, Saltner H, Olsson T, Hilt J. Variation in interleukin 7 receptor alpha chain (IL7R) influences risk of multiple sclerosis. Nat Genet 2007;39:1108–13.

10. Matsen F, Caro-Maldonado A, Fedetz M, Fernandez O, Milne RL, Guerrero M, Delgado C, Alcaín A. IL2RA/CD25 polymorphisms contribute to multiple sclerosis susceptibility. J Neuroimmunol 2007;184:682–4.

11. Corporaal S, Bijl M, Kallenborg CS. Familial occurrence of autoimmune diseases and autoantibodies in a Caucasian population of patients with systemic lupus erythematosus. Clin Rheumatol 2002;21:108–13.

12. Olsson T, Jagodic M, Piehl F, Wallstrom E. Genetics of autoimmune neurinoma. Curr Opin Immunol 2000;12:643–9.

13. Lee YH, Ryo CH, Choi J, Jo DS, Song GS, Noh SK, Hanley JY. The PTFN22 C185ST functional polymorphism and autoimmune diseases—a meta-analysis. Rheumatology (Oxford) 2007;46:49–56.

14. Swanberg M, Lidman O, Padyukov L, Eriksson P, Akesson E, Jagodic M, Lobell A, Khademi M, Borjesson O, Lindgren CM, Lundman P, Brookes AJ, Keje J, Luthman H, Alfredsson L, Hilt J, Klæreskog L, Hamsten A, Piehl F, Olsson T. MHCT2A is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. Nat Genet 2005;37:485–92.

15. Wang L, Lin S, Rammohan KW, Liu Z, Liu QJ, Liu RH, Gunther N, Lima J, Zhou Q, Wang T, Zheng X, Birmingham DJ, Roxin BH, Hebert LA, Wu Y, Lund FY, Cooke G, Yu CY, Zheng P, Liu Y. A d nucleotide deletion in CD44 confirms protection against autoimmune diseases. PLoS Genet 2007;3:e49.

16. Ronnblom LE, Alg MV, Øberg KE. Autoimmunity after alpha-interferon therapy for malignant carcinoid tumors. Ann Intern Med 1991;115:178–83.

17. Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, Shank KB, Grande WJ, Hughes KM, Kapur V, Gregersen PK, Behrens TW. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A 2005;102:2682–4.

18. van der Pouw Kraan TC, Wijkström NA, van Baarsen LG, Voskuyl AE, Rustenburg NA, on October 2, 2023 by guest. Protected by copyright.http://jmg.bmj.com/ J Med Genet: first published as 10.1136/jmg.2007.055012 on 19 February 2008. Downloaded from

19. Bave U, Nordmark G, Lovgren T, Ronnelid J, Cajander S, Verweij CL. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. Ann Rheum Dis 2006;65:1006–12.

20. Bave U, Nordmark G, Lovgren T, Ronnelid J, Cajander S, Eloranta ML, Alm GV. Autoimmunity after alpha-interferon therapy for multiple sclerosis. Clin Exp Immunol 2005;140:477–84.

21. Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ, Shark KB, Grande WJ, Hughes KM, Kapur V, Gregersen PK, Behrens TW. Activation of the type I interferon system in primary Sjögren's syndrome: A generalization of the transmission/disequilibrium test for uncertain haplotype transmission. A generalization of the transmission/disequilibrium test for uncertainty haplotype transmission. Proc Natl Acad Sci U S A 2001;98:12182–7.

22. Cheung VG, Spelman RS, Edwards KV, TB, Morley T, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. Nature 2005;437:1365–9.

23. Graham RR, Kozryev SV, Baechler EC, Reddy MV, Ramsey JW, Ortmann WA, Bench T, Gonzalez-Monzano MF, Pena-Estel E, Peterson D, Daly M, Gregersen PK, Martin J, Altschuler D, Behrens TW, Alarcon-Riquelme ME. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet 2006;38:550–8.

24. Heiney SN, Wingender E, Reuter J, Hermjakob H, Kel AE, Kel OV, Ignevskaya EV, Anakea E, Pechokodyda K, Kola, Papasteriades C, Gomansky SN, Gregersen PK, Bohjanen PO. Association of IRF5 haplotypes with susceptibility to systemic lupus erythematosus. Hum Mol Genet 2008;17:872–81.

25. Sharma S, Grandvaux N, Mamane Y, Genn P, Azimi N, Waldmann T, Hiscott J. Regulation of IFN regulatory factor 4 expression in human T cell leukemia virus-I-transformed T cells. J Immunol 2002;169:3120–30.

26. Book Alexander HM, Yi-Lei LY, SP1 is required for prolactin activation of the interferon regulatory factor-1 gene. Mol Cell Endocrinol 2001;184:135–41.

27. Pempe RJ, Copatcs J, Davies L, Price AL, de Bakker P, Miller J, Peier M, Buffa NA, Blumenstiel B, Defelece M, Parkin M, Barry R, Winslow W, Healy C, Graham RR, Neale BM, Izmailova E, Rosenbom R, Parker AN, Glass R, Karlock EW, Maher N, Halder GA, Lee DM, Seldin MF, Remmers EF, Lee AT, Padyukov L, Alfredsson L, Coblyn J, Weinblatt ME, Gabriel SP, Pocock S, Kolesnikov NA, Eloranta ML, Alm GV, Oberg KE. Autoimmunity after alpha-interferon therapy for multiple sclerosis. Clin Exp Immunol 2005;140:477–84.

28. Mitchell AA, Guthrie DJ, Chakravarti A. Undetected genotyping errors cause apparent overtransmission of common alleles in the transmission/disequilibrium test. Am J Hum Genet 2003;72:598–610.

29. Teige I, Troschow A, Teige A, Mattsson R, Navikas V, Leanderd N, Holmran H, Issazadeh-Navikas S. IFN-beta gene deletion leads to augmented and chronic demyelinating experimental autoimmune encephalomyelitis. J Immunol 2003;170:677–84.

---

**Original article**

J Med Genet 2008;45:362–369. doi:10.1136/jmg.2007.055012

369