nature genetics

Variation in interleukin 7 receptor α chain (*IL7R*) influences risk of multiple sclerosis

Frida Lundmark[1], Kristina Duvefelt[2], Ellen Iacobaeus[3], Ingrid Kockum[1,3], Erik Wallström[3], Mohsen Khademi[3], Annette Oturai[4], Lars P Ryder[5], Janna Saarela[6], Hanne F Harbo[7,8], Elisabeth G Celius[8], Hugh Salter[9], Tomas Olsson[3] & Jan Hillert[1]

Multiple sclerosis is a chronic, often disabling, disease of the central nervous system affecting more than 1 in 1,000 people in most western countries. The inflammatory lesions typical of multiple sclerosis show autoimmune features and depend partly on genetic factors. Of these genetic factors, only the *HLA* gene complex has been repeatedly confirmed to be associated with multiple sclerosis, despite considerable efforts. Polymorphisms in a number of non-*HLA* genes have been reported to be associated with multiple sclerosis, but so far confirmation has been difficult. Here, we report compelling evidence that polymorphisms in *IL7R*, which encodes the interleukin 7 receptor α chain (IL7Rα), indeed contribute to the non-HLA genetic risk in multiple sclerosis, demonstrating a role for this pathway in the pathophysiology of this disease. In addition, we report altered expression of the genes encoding IL7Rα and its ligand, IL7, in the cerebrospinal fluid compartment of individuals with multiple sclerosis.

IL7Rα (also known as CD127), encoded by *IL7R*, is a member of the type I cytokine receptor family and forms a receptor complex with the common cytokine receptor gamma chain (CD132) in which IL7 is the ligand[1]. The IL-7–IL7Rα ligand-receptor pair is crucial for proliferation and survival of T and B lymphocytes in a nonredundant fashion, as shown in human and animal models, in which genetic aberrations lead to immune deficiency syndromes[2]. *IL7R* is located on chromosome 5p13, a region occasionally suggested to be linked with multiple sclerosis[3].

We have considered *IL7R* a promising candidate gene in multiple sclerosis, and we have recently reported genetic associations with three *IL7R* SNP markers in up to 672 Swedish individuals with multiple sclerosis and 672 controls, as well as two associated haplotypes spanning these markers[4]. To confirm these associations, we assessed an independent case-control group consisting of 1,820 individuals with multiple sclerosis and 2,634 healthy controls from the Nordic countries (Denmark, Finland, Norway and Sweden), independent from the data set analyzed in ref. 4 (**Supplementary Table 1** online). Of these, 91% of the affected individuals had experienced an initially relapsing-remitting course of multiple sclerosis (RRMS), whereas 9% had a primary progressive course (PPMS). In addition, we analyzed the expression of *IL7R* and *IL7* in the peripheral blood as well as in cells from the cerebrospinal fluid (CSF).

We genotyped the three previously associated SNPs, located in intron 6 (rs987106 and rs987107) and exon 8 (rs3194051). rs987106 and rs3194051 were in high linkage disequilibrium (LD) ($r^2 = 0.99$, $|D' = 1.00|$), and rs987107 was in partial LD ($r^2 = 0.29$, $|D' = 0.99|$) with rs987106 and rs3194051, as they were located in the same haplotype block. The size of the study allowed full power (100%) to detect an odds ratio (OR) of 1.3. All SNPs were genotyped using the Sequenom hME assays. The observed control genotypes conformed to Hardy-Weinberg equilibrium. All three SNPs confirmed significant association with multiple sclerosis in this nonoverlapping case-control group, with very similar ORs as in the previous study (rs987107, $P = 0.002$; rs987106, $P = 0.001$; rs3194051, $P = 0.002$). A test for heterogeneity between the data sets from Norway, Denmark, Finland and Sweden showed no evidence of stratification, thus permitting a combined analysis (Mantel-Haenszel–corrected and crude ORs are shown in **Table 1**). We estimated the three-marker haplotype frequencies using the EM algorithm in Haploview[5]. The estimated distribution of haplotypes differed significantly between affected individuals and controls ($P = 0.001$), with two haplotypes associating with multiple sclerosis, one conferring an increased risk of disease ($P = 0.0004$) and the other conferring a decreased risk of disease ($P = 0.003$) (**Supplementary Table 2** online), in accordance with previous results[4].

According to data from the HapMap CEU population, *IL7R* is located within a tight LD block containing no additional genes. To

[1]Division of Neurology, Department of Clinical Neuroscience, Karolinska Institutet at Karolinska University Hospital–Huddinge, SE-141 86 Stockholm, Sweden. [2]Clinical Research Centre, Mutation Analysis Facility, Karolinska University Hospital, SE-141 86 Huddinge, Sweden. [3]Neuroimmunology Unit, Department of Clinical Neuroscience, Karolinska Institutet at Karolinska University Hospital–Solna, SE-171 76 Stockholm, Sweden. [4]Danish Multiple Sclerosis Research Centre and [5]Department of Clinical Immunology, Rigshospitalet, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark. [6]Department of Molecular Medicine, National Public Health Institute, FI-00290 Helsinki, Finland. [7]Institute of Immunology, University of Oslo, N-0027 Oslo, Norway. [8]Department of Neurology, Ullevål University Hospital, N-0407 Oslo, Norway. [9]Pharmacogenomics Section, Department of Disease Biology, AstraZeneca R&D, SE-151 85 Södertälje, Sweden. Correspondence should be addressed to J.H. (jan.hillert@ki.se).

Received 22 February; accepted 20 June; published online 29 July 2007; doi:10.1038/ng2106

Table 1 Summary of data confirming association of *IL7R* with multiple sclerosis in the Nordic case-control group

SNP	Allele 1	Allele 2	Affected individuals 11	Affected individuals 12	Affected individuals 22	Controls 11	Controls 12	Controls 22	*P* value	OR (95% c.i.) carriership	OR (95% c.i.) homozygosity	OR (95% c.i.) Mantel-Haenszel
rs987107	C	T	976	651	152	1,417	991	157	0.002	0.70 (0.55–0.88)[a]	1.43 (1.14–1.82)[b]	1.42 (1.12–1.80)
rs987106	A	T	485	848	424	784	1,278	519	0.001	0.79 (0.69–0.91)[a]	1.26 (1.09–1.46)[b]	1.17 (1.07–1.42)
rs3194051	A	G	979	657	149	1,430	982	152	0.002	0.70 (0.55–0.89)[a]	1.43 (1.13–1.80)[b]	1.32 (1.04–1.68)

The Nordic case-control group consisted of 1,820 individuals with multiple sclerosis and 2,634 healthy controls. The samples are independent of the material analyzed in ref. 4. All three SNPs confirmed association with multiple sclerosis; single-point values were calculated with two-sided *P* values using a χ^2 test as a confirmation of an initial association. ORs for single SNPs were calculated using the approximation of Woolf for the 95% confidence interval (c.i.). Mantel-Haenszel combined ORs were calculated to adjust for the contribution of the different countries.
[a]Carrier of allele 1. [b]Homozygous for allele 2.

more precisely map the risk-modifying variants within this LD block, we investigated 12 additional SNPs (ten tag SNPs and two nonsynonymous SNPs); one of these, rs6897932, was selected based on preliminary data now included in a related study in this issue[6]. We studied these 12 SNPs in 1,210 Swedish individuals with multiple sclerosis and 1,234 blood donors—a combination of the Swedish 1 and Swedish 2 groups, where Swedish 1 refers to individuals that were included in ref. 4, and Swedish 2 is a group of affected individuals and controls included in the confirmatory group and not overlapping with Swedish 1 (**Supplementary Table 1**). In this group, 93% of affected individuals had relapsing-remitting multiple sclerosis (RRMS), and 7% had primary-progressive multiple sclerosis (PPMS). The tag SNPs were selected based on their ability to tag (that is, to provide indirect information for a larger number of SNPs within the LD block) (**Supplementary Table 3** online). We selected the tag SNPs using Tagger software based on their r^2 values, with $r^2 = 0.9$ as a cut-off in order to capture genotype information[7]. The ten selected tag SNPs provided inferred genotype information for a total of 69 SNPs in the LD block, 35 of which were located within *IL7R*. In addition, we included the three originally associated markers (rs987106, rs987107 and rs3194051) to increase the comparability of haplotypes between the two studies.

Three of the tag SNPs showed significant association in the single-point analysis that survived after Bonferroni correction for multiple comparisons: rs2303137 ($P_{\text{uncorrected}} = 0.004$, $P_{\text{Bonferroni}} = 0.05$), rs6871748 ($P_{\text{uncorrected}} = 0.003$ and $P_{\text{Bonferroni}} = 0.04$) and the exon 6 nonsynonymous SNP rs6897932 ($P_{\text{uncorrected}} = 0.001$, $P_{\text{Bonferroni}} = 0.02$) (**Table 2**). Logistic regression analysis did not identify any interaction with *HLA-DRB1*1501*, the most strongly associated allele of *HLA-DRB1* (data not shown).

We performed logistic regression analysis to identify the key associated SNPs. For rs6897932, rs2303137 and rs987106, the codominant coding resulted in a better model than either dominant or recessive coding ($P = 0.0007$, $P = 0.002$ and $P = 0.02$, respectively). For the other markers, the dominant model performed slightly better than the codominant model. For all SNPs, the recessive model showed the lowest degree of fit (**Supplementary Table 4** online).

We next tested whether SNPs were independently associated with multiple sclerosis. Using the codominant coding of markers and including all typed markers, the best model with forward or stepwise

Table 2 Fine-mapping of *IL7R* by genotyping 15 SNPs

SNP	Allele 1	Allele 2	Affected individuals 11	Affected individuals 12	Affected individuals 22	Controls 11	Controls 12	Controls 22	*P* value, uncorrected	*P* value, Bonferroni-corrected	OR (95% c.i.) carriership
rs11567698	G	T	14	273	864	11	247	950	0.04	NS	1.22 (1.01–1.49)[a]
rs11567701	G	T	636	427	97	651	483	68	0.01	NS	0.65 (0.47–0.90)[a]
rs10461959	C	T	52	356	737	45	374	759	0.40	NS	0.82 (0.54–1.24)[b]
rs3777090	A	T	519	508	125	575	524	104	0.08	NS	0.78 (0.59–1.02)[a]
rs11567773	G	T	1,119	9	0	1,205	7	0	–		–
rs1494555	C	T	123	499	521	108	524	581	0.15	NS	0.81 (0.62–1.06)[b]
rs6897932	C	T	642	423	82	595	500	115	0.001	0.02	0.76 (0.65–0.90)[b]
rs987107	C	T	630	430	100	657	485	72	0.01	NS	0.81 (0.73–0.94)[a]
rs987106	A	T	301	566	256	379	591	241	0.02	NS	0.80 (0.67–0.96)[a]
rs3194051	A	G	637	430	94	660	481	66	0.02	NS	0.83 (0.72–0.96)[a]
rs10058453	C	T	124	515	512	105	525	568	0.11	NS	0.79 (0.60–1.04)[b]
rs6871748	C	T	84	433	638	113	510	579	0.003	0.04	0.75 (0.64–0.88)[a]
rs13169780	A	C	853	286	23	900	290	18	0.47	NS	0.76 (0.40–1.41)[a]
rs6870944	A	G	820	308	24	857	315	27	0.99	NS	1.04 (0.60–1.82)[a]
rs2303137	A	T	188	575	402	245	610	348	0.004	0.05	0.77 (0.65–0.92)[a]

The above data are from 1,210 individuals with multiple sclerosis and 1,234 healthy controls. Single-point values were calculated with two-sided *P* values, using a χ^2 test. ORs for single SNPs were calculated using the approximation of Woolf for the 95% confidence interval (c.i.) and are presented here as ORs for carriership. Bonferroni correction was used to correct for multiple testing. NS, nonsignificant.
[a]Carrier of allele 1. [b]Carrier of allele 2.

Table 3 Results of logistic regression analysis showing model fit

	Model		
	rs6897932 only	rs2303137 only	rs6897932 and rs2303137
Model fit (–2logL)	3,196.2	3,197.8	3,194.9
P value	0.0006	0.002	0.002
rs6897932			
P value	0.0006		0.09
Estimate (95% c.i.)	0.80 (0.70–0.91)		0.86 (0.72–1.02)
rs2303137			
P value		0.002	0.05
Estimate (95% c.i.)		0.83 (0.73–0.93)	0.91 (0.77–1.07)

Results reflect data from 1,131 affected individuals and 1,184 controls. The best fit was obtained for rs6897932 alone ($P = 0.0006$), closely followed by rs2303137 alone or both markers together ($P = 0.002$). c.i., confidence interval.

selection was one containing only rs6897932 ($P = 0.004$). With backward selection, the best model contained only rs2303137 ($P = 0.007$). Using dominant coding and forward selection, the best model contained rs6897932, rs3194051 and rs987106 ($P = 0.0007$), and with forward selection, only rs6897932 was included ($P = 0.002$). Thus, we conclude that the most associated markers are rs6897932 and rs2303137. We tested the model fit for the models containing rs6897932 and rs2303137 alone and the two markers together. The best fit was obtained for rs6897932 alone ($P = 0.0006$), closely followed by rs2303137 alone or both markers together ($P = 0.002$) (**Table 3**).

We estimated haplotype frequencies based on the algorithm from ref. 8. We included in the haplotype analysis 12 SNPs that were located in the same LD block and defined the boundaries of the haplotype block. We tested the robustness of the haplotype block using the EM algorithm as well as Bayesian-based PHASE software; both identified the same LD block structure with virtually identical frequencies (data not shown). The analysis uncovered a significant association with the most common haplotype, conferring a protective effect ($P_{uncorrected} = 0.0004$, $P_{permutated} = 0.0007$) (**Supplementary Table 5** online). We did not identify any significant risk haplotypes using the 12 SNP markers, because the positively associated three-marker haplotype observed in the confirmation analysis (as shown in **Supplementary Table 2**) was split into four different haplotypes, each insignificantly more common in affected individuals than controls (**Supplementary Table 5**). The T allele of rs6897932 in exon 6 was unique for the protective haplotype. To assess the haplotype association further, we performed a sliding-window analysis of the haplotype block, including four markers in the window. We observed the strongest effect when we included markers rs6897932, rs987107, rs987106 and rs3194051 together ($P = 0.0002$; **Table 4**). In summary, all analyses point toward marker rs6897932, or haplotypes including this marker, as being most strongly associated with risk of multiple sclerosis.

From these data, we conclude that the genetic evidence for involvement of *IL7R* in multiple sclerosis susceptibility is solid, based on the fact that the initial associations have been replicated and confirmed in an independent case-control group twice the size of the original sample. We estimated an etiologic fraction of 12% for *IL7R* in multiple sclerosis based on the risk genotype (CC) of rs6897932, compared with a reported etiologic fraction of 40% for the previously established risk factor HLA-DR15 (ref. 9).

To explore the biological relevance of the IL7Rα-IL7 signaling pathway in multiple sclerosis, we designed two mRNA assays and assessed expression patterns in parallel samples of peripheral blood mononuclear cells (PBMC) and cerebrospinal fluid (CSF). We analyzed expression in 75 individuals with multiple sclerosis, including 65 with RRMS, of which 53 were in remission and 12 were in relapse. Ten individuals had chronic progressive multiple sclerosis (PPMS or secondary progressive multiple sclerosis). None had received immunomodulatory treatments during the 3 months before sampling. For comparison, we collected PBMCs and CSF from 48 controls with other noninflammatory neurological diseases (ONDs), all without CSF IgG oligoclonal bands, as well as PBMCs from 20 healthy controls. In general, expression of *IL7R* was higher in the CSF compared with PBMCs both in individuals with multiple sclerosis and in individuals with ONDs ($P = 0.0001$). There was no difference in the expression of *IL7R* in PBMCs between individuals with multiple sclerosis, individuals with ONDs or healthy controls. In contrast, we observed increased expression of both *IL7R* ($P = 0.0001$) and *IL7* ($P = 0.0001$) mRNA in the CSF of individuals with multiple sclerosis in comparison with individuals with ONDs, suggesting a pathophysiological importance of IL7 and IL7Rα in multiple sclerosis within the central nervous system (**Fig. 1**). Analysis of *IL7R* mRNA expression in CSF mononuclear cells in relation to the genotype of the multiple sclerosis risk–modifying SNP, rs6897932, did not uncover any correlation ($P = 0.81$).

The identification of non-HLA genes influencing the risk of autoimmune disorders has been difficult. In multiple sclerosis, linkage analysis in nuclear families has failed to identify additional loci with genome-wide significance[10]. Although a large number of genetic associations have been described in initial reports, larger follow-up studies have invariably failed to give support for the gene in question[11,12]. Reasons may include the weak influence of each genetic factor relative to the often limited numbers of cases and controls in each study, making both type I and type II errors likely. Clearly, *IL7R* represents an exception. Others studying PPMS have reported a suggestive but nonsignificant association with *IL7R* that later improved with the addition of markers and more affected individuals to become significant and was supported by differential expression[13,14]. In conjunction with our earlier results[4], the evidence presented here and in the related paper in this issue[6] (which also reports differences in expression of *IL7R* mRNA and IL7Rα protein) convincingly corroborate the case

Table 4 Four-marker sliding window haplotype analysis over the entire *IL7R* haplotype block

	rs6897932	rs987107	rs987106	rs3194051	Affected individuals	Controls	Haplotype *P*	Haplotype *P* permutation
Haplotype 1	C	C	A	A	0.475	0.441	0.021	NS
Haplotype 2	T	C	T	A	0.253	0.301	0.0002	0.0002
Haplotype 3	C	T	T	G	0.265	0.253	0.385	NS

Using a four-marker window, we identified one haplotype window that presented a slightly stronger protective association than the full haplotype. A permutation test based on 10,000 permutations was performed in order to correct for multiple testing. NS, nonsignificant.

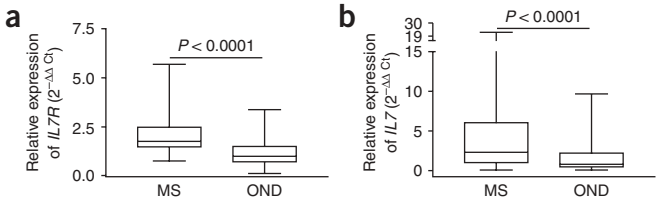

Figure 1 Expression analysis of *IL7R* and *IL7* in individuals with multiple sclerosis. (**a**) *IL7R* mRNA expression is higher in the CSF of individuals with multiple sclerosis (MS) than in the CSF of individuals with other noninflammatory neurological disorders (OND). (**b**) *IL7* mRNA expression is higher in the CSF of individuals with multiple sclerosis than in the CSF of individuals with other noninflammatory neurological disorders. Horizontal bars represent minimum and maximum measurements.

for *IL7R* in multiple sclerosis. In addition, the same chromosomal region, 5p13, has been identified in two other inflammatory diseases, asthma and type 1 diabetes, suggesting a more general role for *IL7R* in autoimmune disease mechanisms[15,16].

We have yet to explain the mechanism behind the importance of the IL7-IL7Rα ligand-receptor pair in multiple sclerosis. Increased expression in the CSF compartment would support a simplistic interpretation of higher signaling inducing immune cell proliferation and survival. Alternatively, the mechanism could involve an altered balance between soluble and membrane-bound IL7Rα, which has been suggested to be under genetic control (D. Booth & G.J. Stewart, personal communication). Data presented in ref. 6 also support the inference that polymorphisms studied here influence differential splicing, with resulting altered expression of the soluble receptor. Our observed lack of correlation between mRNA expression and genotype also needs to be explored further but could be explained by an altered balance between soluble and membrane-bound receptor or by an altered expression in a subpopulation of T cells only. Notably, low expression of IL7Rα by $CD4^+$ $CD25^+$ T cells correlates with high expression of FoxP3 and has recently been identified as an important characteristic of regulatory T cells, suggesting that impaired suppression of autoimmune mechanism may be important in multiple sclerosis[17].

In summary, evidence is mounting for a role for *IL7R* in the genetics of multiple sclerosis, making it, to our knowledge, the only non-HLA gene thus far convincingly associated with this disease, and putting it in a small but growing family of genes with confirmed importance for autoimmune diseases, including *NOD2* in Crohn's disease[18–20], *IRF5* in lupus[21] and *PTPN22* in rheumatoid arthritis[22] and in type 1 diabetes[23]. Further characterization of the mechanisms behind its role in multiple sclerosis may uncover new options for therapeutic intervention.

METHODS

Subjects. For the genetic association studies, two case-control groups were used: Nordic and Swedish. The Nordic case-control group consisted of 1,820 individuals with multiple sclerosis from Sweden, Norway, Denmark and Finland (72% female and 28% male; mean age at sampling, 46.4 years) and fulfilled either Poser's criteria for definite multiple sclerosis or McDonald's criteria for multiple sclerosis, to confirm three previously associated markers. The identity of the Swedish individuals was assessed using the unique personal identification numbers provided by the Swedish tax administration to exclude any overlap with the data set used in ref. 4. In the Nordic group, the disease course was primary progressive (PPMS) in 11.2% of the affected individuals. We collected DNA from 2,634 healthy controls (53% females and 47% males, with a mean age at sampling of 38.6 years) in this group, matched with the same ancestry. The Swedish case-control material consisted of DNA from 1,210 individuals with multiple sclerosis, including the Swedish samples genotyped in the confirmation step (Swedish 2) (73% females and 27% males) fulfilling either Poser's criteria for definite multiple sclerosis or McDonald's criteria for multiple sclerosis, for fine mapping of *IL7R* after the confirmation. The disease course was primary progressive in 6.8% of the affected individuals. We collected DNA from 1,234 blood donors (63% female and 37% male; mean age at sampling, 45.6 years) as control samples, with ancestry matched to the affected individuals. Four individuals with multiple sclerosis from the Swedish 1 group had a relative in the Swedish 2 group. This did not influence the analysis (data not shown). Twenty-two other affected individuals from both groups reported in their medical records that they had a relative with multiple sclerosis; these 22 individuals were not related to one another.

For the mRNA expression studies, we collected PBMCs and CSF from 65 individuals with multiple sclerosis (68% female and 32% male; mean age, 37.4 years) with RRMS, including 53 individuals in remission and 12 individuals in relapse. In addition, we collected samples from ten individuals (60% female and 40% male, with a mean age at sampling of 53.5 years) with SPMS or PPMS. We also collected PBMCs and CSF from 48 individuals with noninflammatory ONDs (73% female and 23% male; mean age, 42.1 years) and PBMCs from 20 healthy subjects (75% female and 25% male; mean age, 31.5 years) (CSF from healthy subjects was not available). We obtained informed consent from all subjects, and the study was approved by the ethics committees of all participating institutions.

SNP selection. SNPs for genotyping were chosen from dbSNP and the HapMap. SNPs genotyped in the Nordic material were selected for confirmatory analysis based on associations reported in our previous publication[4]. We selected ten additional SNPs to tag the LD block harboring *IL7R* in order to acquire genotype information for a larger number of SNPs within and surrounding *IL7R*. Tagging of the complete LD block was based on Tagger software, with a minimum r^2 of 0.9 in order to ensure nearly perfect LD to infer information on all SNPs captured by the tag set. We also selected two additional nonsynonymous SNPs for genotyping (**Supplementary Table 3**).

Genotyping. All SNPs were genotyped using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (Sequenom). Assays for all SNPs were designed using SpectroDESIGNER software (Sequenom). Primer sequences are given in **Supplementary Table 6** online.

All samples from Norway, Denmark and Finland were genotyped using the hME assays. The amplification was performed in a total volume of 5 μl, with 2.5 ng of genomic DNA, 1 pmol of each amplification primer, 0.2 mM of each dNTP, 2.5 mM $MgCl_2$ and 0.2 units of HotStartTaq DNA Polymerase (Qiagen). The reaction was subjected to the following PCR conditions: a single cycle of denaturation at 95 °C for 15 min, followed by 45 cycles at 94 °C for 20 s, 60 °C for 30 s and 72 °C for 30 s, with a final extension at 72 °C for 7 min. The allele-specific extension step was performed in a total volume of 9 μl using 5 pmol of extension primer and the MassEXTEND Reagent Kit, and products were cleaned using SpectroCleaner (Sequenom). Products from primer-extension reactions were loaded on a 384-element chip nanoliter pipetting system (Sequenom) and were analyzed on a MassARRAY mass spectrometer (Bruker Daltonik). For genotype calling, we used SpectroTYPER RT 2.0 software (Sequenom).

Samples from the Swedish material were genotyped using iPLEX (Sequenom). The amplification was performed in a total volume of 5 μl, with 10 ng of genomic DNA, 100 nM of each amplification primer, 500 mM of dNTP mix, 1.625 mM $MgCl_2$ and 0.5 units of HotStartTaq DNA Polymerase (Qiagen). The reaction was subjected to the following PCR conditions: a single cycle of denaturation at 95 °C for 15 min, followed by 45 cycles at 94 °C for 20 s, 56 °C for 30 s, 72 °C for 60 s and a final extension at 72 °C for 3 min. The allele-specific extension step was performed in a total volume of 9 μl using 5 pmol of extension primer and the MassEXTEND Reagent Kit; reaction products were cleaned using SpectroCleaner (Sequenom). The later steps were performed as for the hME assays. Data were analyzed independently by two persons using SpectroTyper software (Sequenom).

Quality control of genotyping. The genotyping of each assay was validated using a set of 14 trios of 42 individuals (individuals with multiple sclerosis and their parents) available through the HapMap consortium. Concordance analyses with the HapMap data, as well as analysis of the parent-offspring compatibility with the produced genotypes, showed a perfect match. The

success rate for the different SNP assays ranged from 96.5 to 98%. Genotype calling was performed independently by two persons using SpectroTyper software (Sequenom). For each 384-well plate, half the wells included material from affected individuals, and half contained material from controls. Internal concordance analysis was performed by genotyping 140 samples twice for each SNP and resulted in 100% concordance. Three different DNA samples were used in duplicate in each 384-well plate. The success of each SNP genotyping assay was calculated as the number of genotypes retrieved over the possible number of genotypes, including genotypes from quality control samples.

Preparation of PBMCs and cells from CSF. Peripheral blood was collected in cell preparation tubes containing sodium citrate (Vacutainer CPT, Becton Dickinson), and CSF was collected in siliconized glass tubes. CSF samples were immediately centrifuged, and the pellet was recovered and stored at –70 °C until use. PBMCs were separated by density gradient centrifugation. Cells from interphase were collected and washed twice with Dulbecco's PBS. The proportion of viable cells was assessed by trypan blue exclusion. More than 95% of the cells were viable. Finally, the cells were pelleted, frozen on dry ice and stored at –70 °C until use.

Relative quantification of mRNA by real-time quantitative RT-PCR. Cell pellets were lysed, and total RNA was extracted (PicoPure RNA isolation kit, Arcturus Bioscience). Samples were incubated with 27,000 units of DNase (Qiagen RNase-free DNase set) for 15 min at room temperature (20 °C) in order to avoid amplification or detection of contaminating genomic DNA. The RNA concentration was measured spectrophotometrically, and the quality of the RNA was determined using the Agilent 2100 Bioanalyzer. Only samples containing completely pure RNA were used in the RT-PCR analysis. Reverse transcription was performed with 10 µl total RNA (1–5 ng), random hexamer primers (0.1 µg; Gibco BRL) and Superscript Reverse Transcriptase (200 units; Gibco BRL). Real-time PCR was performed using a BioRad iQ5 iCycler Detection System with a three-step PCR protocol (95 °C for 10 min, followed by 50 cycles of 95 °C for 15 s, 60 °C for 30 s and 72 °C for 30 s) with SYBR green fluorophore. Reactions were performed in a total volume of 25 µl, including 12.5 µl 2× SYBR Green PCR Master Mix (BioRad Laboratories), 1.5 µl of a 5 µM solution of each primer, 2 µl of the previously reverse-transcribed cDNA template and 7.5 µl double-distilled H_2O. Primers were designed with Primer Express Software (Perkin Elmer) in our laboratory. The primer sequences for *IL7*, *IL7R* and *GAPDH* are given in **Supplementary Table 6**.

In preliminary experiments, the primer pairs were tested by a temperature gradient program (55 °C to 65 °C for the annealing step), and the PCR products were separated on agarose gels; in all cases, products were confined to a single band of the expected size for each specific target. Sequencing of the different bands by Cybergene confirmed homology with the reported sequences for human *IL7R* and *GAPDH*, respectively. Relative quantification of mRNA was performed using both the standard curve method and the comparative threshold cycle method using *GAPDH* as endogenous control. The standard curves were created using five serial dilutions (1:10, $1:10^2$, $1:10^3$, $1:10^4$ and $1:10^5$) of either the tested amplicon for each target or cDNA from human blood cells stimulated with lectin concanavalin A (ConA). The samples were processed in duplicate with primers against *GAPDH* and the target mRNA in different wells. Samples without added cDNA served as negative controls. The PCR efficiency was >90% in all PCR runs. A comparative threshold cycle (CT) was used to determine gene expression relative to the control (calibrator). Hence, steady-state mRNA levels were expressed as a multiple of the expression level of the calibrator. For each sample, the *IL7R* or *IL7* CT value was normalized using the formula $\Delta C = CT_{IL\text{-}7R\ (or\ IL\text{-}7)} - CT_{GAPDH}$. To determine relative expression levels, the following formula was used: $\Delta\Delta_{CT} = \Delta_{CT(sample)} - \Delta_{CT(calibrator)}$. Relative *IL7R* or *IL7* expression was calculated using the expression formula $2^{-\Delta\Delta ct}$. For PBMCs, we used the Δ_{CT} from healthy controls, and for CSF, we used Δ_{CT} from individuals with other noninflammatory neurological diseases as calibrators.

Primer sequences are given in **Supplementary Table 6**.

Statistical analysis. Hardy-Weinberg equilibrium (HWE) was tested to ensure that control samples were within allelic population equilibrium. To test for single-point association, two-sided *P* values for carriership were calculated using a χ^2 test with no correction for multiple testing (GraphPad Instat). Multiple testing for single-point values was performed using a Bonferroni correction, where *P* values were corrected based on ten independent observations. Single-point values for carriership in the confirmatory study were calculated using two-sided *P* values because of the prior associations. ORs were calculated using the approximation of Woolf for the 95% confidence interval (c.i.). A test of heterogeneity was performed in the Nordic group to allow for a combined analysis, and Mantel-Haenszel combined ORs were calculated (GraphPad Instat). Power calculations were performed using a 5% significance level for detecting an OR of 1.3. LD between pairs of SNPs was calculated using the measures $|D'|$ and r^2, implemented in Haploview version 3.32. Frequencies of haplotypes were estimated using the EM algorithm in Haploview and Haplotype Blocker, and to test for association, a χ^2 test was performed in Haploview, with 10,000 permutations in order to correct for multiple testing.

Logistic regression analysis was performed using SAS (SAS Institute) where the proc logistic command was used. This resulted in a standard logistic regression model in which the regression coefficients were logarithms of the ORs. Both forward, backward and stepwise selection were used. Cut-offs for the different models were $P = 0.3$ for forward selection and $P = 0.05$ for stepwise and backward selection. We first tested which coding of the SNP variables resulted in the best logistic regression model when we included only one marker in each model. We next tested which markers were independently associated with multiple sclerosis, using the codominant coding of markers and including all typed markers.

Differences in relative mRNA levels of *IL7R* and *IL7* in CSF or PBMCs were tested for significance with the nonparametric Wilcoxon signed-rank test. Correlation between mRNA levels of *IL7R* and genotype was analyzed using the Kruskal-Wallis rank sum test.

URLs. dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/; HapMap: http://www.hapmap.org/; Haploview: http://www.broad.mit.edu/mpg/haploview/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank all the patients and volunteers who participated in this study. This study was supported by grants from the Swedish Research Council (project number 11220), the Norwegian Research Council (project numbers 1154888/40 and 148351/300), the Eastern Norwegian Regional Health Authority (project number 29300604), the Danish Multiple Sclerosis Society, the Warwara Larsen Foundation, NeuroproMiSe LSMH-CT-2005-018637, the Swedish Society of the Neurologically Disabled, the Bibbi and Nils Jensens Foundation and the Swedish Foundation for Strategic Research and Karolinska Institutet. We are grateful to the Knut and Alice Wallenberg Foundation.

AUTHOR CONTRIBUTIONS

F.L., K.D., H.S. and J.H. planned, designed and managed the study. K.D. performed the genotyping; F.L. and I.K. performed the statistical analysis. J.H. and T.O. collected the Swedish patient and control groups. E.I., E.W., M.K., and T.O. collected material for and performed the expression analyses of IL7R and IL7. A.O., L.P.R., J.S., H.F.H. and E.G.C. took part in the planning of the study and contributed the non-Swedish groups of affected individuals and controls. J.H. and F.L. wrote the manuscript with assistance from all coauthors.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Noguchi, M. *et al.* Interleukin-2 receptor gamma chain: a functional component of the interleukin-7 receptor. *Science* **262**, 1877–1880 (1993).
2. Ye, S.K. *et al.* Induction of germline transcription in the TCRgamma locus by Stat5: implications for accessibility control by the IL-7 receptor. *Immunity* **11**, 213–223 (1999).
3. Ebers, G.C. *et al.* A full genome search in multiple sclerosis. *Nat. Genet.* **13**, 472–476 (1996).
4. Zhang, Z. *et al.* Two genes encoding immune-regulatory molecules (LAG3 and IL7R) confer susceptibility to multiple sclerosis. *Genes Immun.* **6**, 145–152 (2005).

5. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
6. Gregory, S.G. *et al.* Interleukin 7 receptor α chain (*IL7R*) shows allelic and functional association with multiple sclerosis. *Nat. Genet.* advance online publication 29 July 2007 (doi:10.1038/ng2103).
7. de Bakker, P.I. *et al.* Efficiency and power in genetic association studies. *Nat. Genet.* **37**, 1217–1223 (2005).
8. Gabriel, S.B. *et al.* The structure of haplotype blocks in the human genome. *Science* **296**, 2225–2229 (2002).
9. Svejgaard, A. & Ryder, L.P. HLA and disease association: detecting the strongest association. *Tissue Antigens* **43**, 18–27 (1994).
10. Sawcer, S. *et al.* A high-density screen for linkage in multiple sclerosis. *Am. J. Hum. Genet.* **77**, 454–467 (2005).
11. Burwick, R.M. *et al.* APOE epsilon variation in multiple sclerosis susceptibility and disease severity: some answers. *Neurology* **66**, 1373–1383 (2006).
12. Roxburgh, R.H. *et al.* No evidence of a significant role for CTLA-4 in multiple sclerosis. *J. Neuroimmunol.* **171**, 193–197 (2006).
13. Teutsch, S.M., Booth, D.R., Bennetts, B.H., Heard, R.N. & Stewart, G.J. Identification of 11 novel and common single nucleotide polymorphisms in the interleukin-7 receptor-alpha gene and their associations with multiple sclerosis. *Eur. J. Hum. Genet.* **11**, 509–515 (2003).
14. Booth, D.R. *et al.* Gene expression and genotyping studies implicate the interleukin 7 receptor in the pathogenesis of primary progressive multiple sclerosis. *J. Mol. Med.* **83**, 822–830 (2005).
15. Kurz, T. *et al.* Fine mapping and positional candidate studies on chromosome 5p13 identify multiple asthma susceptibility loci. *J. Allergy Clin. Immunol.* **118**, 396–402 (2006).
16. Smyth, D.J. *et al.* A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (*IFIH1*) region. *Nat. Genet.* **38**, 617–619 (2006).
17. Liu, W. *et al.* CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells. *J. Exp. Med.* **203**, 1701–1711 (2006).
18. Hugot, J.P. *et al.* Association of NOD2 leucine-rich repeat variant with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).
19. Ogura, Y. *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).
20. Hampe, J. *et al.* Association between insertion mutation in *NOD2* gene and Crohn's disease in German and British populations. *Lancet* **357**, 1925–1928 (2001).
21. Sigurdsson, S. *et al.* Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am. J. Hum. Genet.* **76**, 528–537 (2005).
22. Bottini, N. *et al.* A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nat. Genet.* **36**, 337–338 (2004).
23. Begovich, A.B. *et al.* A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *Am. J. Hum. Genet.* **75**, 330–337 (2004).