ORIGINAL ARTICLE

Genome-wide significant association with seven novel multiple sclerosis risk loci

Christina M Lill,[1,2,3] Felix Luessi,[2] Antonio Alcina,[4] Ekaterina A Sokolova,[5,6] Nerea Ugidos,[7,8] Belén de la Hera,[9] Léna Guillot-Noël,[10] Sunny Malhotra,[11] Eva Reinthaler,[12] Brit-Maren M Schjeide,[3] Julia Y Mescheriakova,[13] Andriy Mashychev,[1] Inken Wohlers,[1] Denis A Akkad,[14] Orhan Aktas,[15] Iraide Alloza,[7,8,16] Alfredo Antigüedad,[17] Rafa Arroyo,[18] Ianire Astobiza,[7,8] Paul Blaschke,[19] Alexei N Boyko,[20] Mathias Buttmann,[21] Andrew Chan,[22] Thomas Dörner,[23] Joerg T Epplen,[13,24] Olga O Favorova,[25] Maria Fedetz,[4] Oscar Fernández,[26] Angel García-Martínez,[9] Lisa-Ann Gerdes,[27] Christiane Graetz,[2] Hans-Peter Hartung,[15] Sabine Hoffjan,[14] Guillermo Izquierdo,[28] Denis S Korobko,[29] Antje Kroner,[21,30] Christian Kubisch,[31] Tania Kümpfel,[27] Laura Leyva,[26] Peter Lohse,[32,33] Nadezhda A Malkova,[29,34] Xavier Montalban,[11] Ekaterina V Popova,[20] Peter Rieckmann,[21,35] Alexei S Rozhdestvenskii,[36] Christiane Schmied,[12] Inna V Smagina,[37,38] Ekaterina Y Tsareva,[25] Alexander Winkelmann,[19] Uwe K Zettl,[19] Harald Binder,[39] Isabelle Cournu-Rebeix,[10] Rogier Hintzen,[13] Alexander Zimprich,[12] Manuel Comabella,[11] Bertrand Fontaine,[10,40] Elena Urcelay,[9] Koen Vandenbroeck,[7,8,16] Maxim Filipenko,[5,6] Fuencisla Matesanz,[4] Frauke Zipp,[2] Lars Bertram[1,3,41]

▶ Additional material is published online only. To view these files please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2015-103442)

For numbered affiliations see end of article.

Correspondence to
Dr Christina M Lill, Institute of Neurogenetics, University of Lübeck, Maria-Goeppert-Str. 1, Lübeck 23562, Germany; christina.lill@uni-luebeck.de

Received 6 August 2015
Revised 3 September 2015
Accepted 17 September 2015
Published Online First 16 October 2015

To cite: Lill CM, Luessi F, Alcina A, *et al*. *J Med Genet* 2015;**52**:848–855.

ABSTRACT

Objective A recent large-scale study in multiple sclerosis (MS) using the ImmunoChip platform reported on 11 loci that showed suggestive genetic association with MS. Additional data in sufficiently sized and independent data sets are needed to assess whether these loci represent genuine MS risk factors.

Methods The lead SNPs of all 11 loci were genotyped in 10 796 MS cases and 10 793 controls from Germany, Spain, France, the Netherlands, Austria and Russia, that were independent from the previously reported cohorts. Association analyses were performed using logistic regression based on an additive model. Summary effect size estimates were calculated using fixed-effect meta-analysis.

Results Seven of the 11 tested SNPs showed significant association with MS susceptibility in the 21 589 individuals analysed here. Meta-analysis across our and previously published MS case-control data (total sample size n=101 683) revealed novel genome-wide significant association with MS susceptibility ($p<5\times10^{-8}$) for all seven variants. This included SNPs in or near *LOC100506457* (rs1534422, $p=4.03\times10^{-12}$), *CD28* (rs6435203, $p=1.35\times10^{-9}$), *LPP* (rs4686953, $p=3.35\times10^{-8}$), *ETS1* (rs3809006, $p=7.74\times10^{-9}$), *DLEU1* (rs806349, $p=8.14\times10^{-12}$), *LPIN3* (rs6072343, $p=7.16\times10^{-12}$) and *IFNGR2* (rs9808753, $p=4.40\times10^{-10}$). *Cis* expression quantitative locus effects were observed in silico for rs6435203 on *CD28* and for rs9808753 on several immunologically relevant genes in the *IFNGR2* locus.

Conclusions This study adds seven loci to the list of genuine MS genetic risk factors and further extends the list of established loci shared across autoimmune diseases.

INTRODUCTION

Multiple sclerosis (MS) is the most common auto-inflammatory disease of the central nervous system. It is caused by the action and interaction of genetic and environmental factors. Genome-wide association studies (GWAS) and other large-scale genotyping projects have revealed that only few common genetic variants exist that exert relatively large effects (ie, OR ranging from ~1.3 to 3), all of which are located in the *HLA* (human leucocyte antigen) locus (reviewed in ref. 1). The remainder of the genetic risk spectrum is likely determined by a large number of susceptibility variants exerting much smaller effects. The hitherto completed GWAS and follow-up projects have identified 110 independent SNPs outside the *HLA* locus showing genome-wide significant association ($p<5\times10^{-8}$) with MS risk (eg, refs 2–4). The most recent of these studies using the ImmunoChip (a customised genotyping array with extensive coverage of loci involved in immune system disorders, including MS[5]) in the discovery phase and previous GWAS data for validation of top results reported on 48 novel genome-wide significant ($p<5\times10^{-8}$) MS risk loci in a total of 29 300 MS cases and 50 794 controls.[4] Eleven additional loci showed suggestive evidence for association in the combined data ($p<1\times10^{-6}$) but failed to surpass the genome-wide

significance threshold.[4] To address the question whether and which of these suggestive loci possibly represent genuine MS susceptibility factors, we genotyped the most significant SNP originally highlighted[4] in each of the 11 loci (table 1) using an independent, multicentric case-control data set of 10 796 MS cases and 10 793 controls. This included 18 335 individuals of European descent from Austria, France, Germany, the Netherlands and Spain, as well as 3254 subjects from Russia (see online supplement 1: table e-1). None of these data sets were previously analysed for the 11 SNPs under scrutiny here. Association results from all 21 589 individuals were subsequently combined with those from the ImmunoChip study,[4] amounting to a total of 101 683 subjects, the largest data set collectively analysed in MS genetics to date.

METHODS

Subjects

The effective sample size available for analysis after quality control (QC) comprised 21 589 individuals including 9079 MS cases and 9256 unrelated controls of European descent from Austria, France, Germany, the Netherlands and Spain, as well as 1717 cases and 1537 controls from three regions in Russia (see online supplement 1: table e-1) and have been described previously.[3 6 7] Diagnosis of MS was established according to standard diagnostic criteria.[8 9] All samples were collected with informed written consent. None of the samples tested here were included in the previous ImmunoChip study.

Power analyses

Power analyses were performed using the Genetic Power Calculator[10] assuming a disease prevalence of 0.1% and no between-study heterogeneity. The combined effective validation data sets of 10 796 cases and 10 793 unrelated controls had >80% power to detect an OR of 1.10 at a one-sided type-1 error rate of $\alpha=4.5\times10^{-3}$ and down to allele frequencies of 0.16. The combined data sets of the original[4] and of this study across 101 683 individuals had >80% power to detect an OR of 1.10 at a two-sided genome-wide type-1 error rate of $\alpha=5\times10^{-8}$ down to allele frequencies of 0.10.

Genotyping and quality control

Genotyping for all samples except the Dutch data set (see below) was performed at the individual sites using single-assay allelic discrimination assays based on TaqMan chemistry following the manufacturer's instructions (Life Technologies). Taqman assays and reagents were ordered centrally by the Berlin site and either used there or distributed to all participating centres. The three Russian data sets were all genotyped at the site in Novosibirsk, Russia, and all German samples were genotyped in Berlin. Up to 5% of samples were genotyped in duplicate across plates to assess genotyping accuracy. Genotypes in the Dutch sample were generated on the Human610-Quad Bead GWAS array (Illumina) and subjected to standard QC using the GenABEL package in R (http://www.genabel.org/packages/GenABEL). This entailed excluding samples showing ≥2% missingness, cryptic relatedness, and excess heterozygosity as well as SNPs showing ≥2% missingness, or deviation from Hardy-Weinberg equilibrium (HWE) at $p>1\times10^{-6}$. Principal component analyses did not reveal ethnic outliers when plotting Dutch and CEU HapMap samples together. For all data sets included in this study, individuals with missing genotypes for more than three SNPs were excluded prior to analysis (applicable to a total of 193 samples (~0.9%) across all data sets). The threshold for genotyping efficiency per SNP and data set was set to >95%. SNPs falling under this threshold were rs6072343 (in the Russia Novosibirsk and Yakutsk data sets), rs727098 and rs9808753 (Russia-Yakutsk). HWE was assessed in controls and deviations from HWE were defined as $p<4.5\times10^{-3}$ in each data set (ie, applying a Bonferroni correction for 11 tests) based on Pearson's χ^2 as implemented in PLINK V.1.07,[11] which led to the exclusion of rs806349 from the French data set. All other SNPs passed these QC thresholds and were included in the actual association analyses. Genotyping accuracy based on comparison with duplicated samples was >99.8% and total genotyping efficiency was >98% for each SNP after QC.

Association analyses

Association analyses were performed using PLINK and in the R environment and were based on logistic regression using an additive genetic model. The Dutch data set was adjusted for the first four principal components. Based on recent recommendations for rare diseases including MS,[12] the logistic regression analyses were not adjusted for known covariates such as age and sex. Meta-analyses across all data sets included here were based on fixed-effect models. The experiment-wide significance threshold was set to $p<4.5\times10^{-3}$ (ie, applying a Bonferroni correction for 11 tests). Between-study heterogeneity was quantified using the I^2 metric, and statistical significance was assessed using the Q test statistic. Significant evidence for heterogeneity was defined as $p<0.1$, a threshold commonly used in this context.[13 14] Forest plots were generated using a customised version of the 'rmeta' package in R.[15] The data sets of the original study[4] and this study were combined using fixed-effect meta-analysis. For this purpose, ancestry-adjusted summary ORs and 95% CIs of the discovery and replication data sets were extracted from the online supplementary table S2 of the original study.[4] In addition, differences between the summary effect size estimates of the eight western European data sets and the three data sets from Russia were assessed by interaction analysis as previously described.[16] For these analyses, significance was defined as $p<4.5\times10^{-3}$ (ie, Bonferroni-corrected for 11 tests). All meta-analysis p values of the validation data sets only are reported as one-sided with regards to the direction of effect reported in the original study.[4] p Values for all other tests (ie, heterogeneity, the meta-analysis of the original study[4] and this study, interaction analyses) are two-sided.

In silico fine-mapping of putative causal variants

We applied a recently developed in silico fine-mapping algorithm (the 'probabilistic identification of causal SNPs' approach (PICS), http://www.broadinstitute.org/pubs/finemapping/?q=pics)[17] to all seven genome-wide significantly associated MS risk SNPs. PICS was developed based on ImmunoChip data to estimate a SNP's probability of being 'causal' in densely mapped genotyping data based on the association result of the most significant SNP.[17]

RESULTS

Among the 11 candidate SNPs assessed in 21 589 individuals across 11 data sets, six showed significant association with MS susceptibility after Bonferroni correction (ie, $p_{corr}<4.5\times10^{-3}$, table 1, figure 1). These six SNPs were: rs1534422 in *LOC100506457* (OR=1.10, $p=1.39\times10^{-6}$), rs6435203 downstream of *CD28* (OR=1.08, $p=7.20\times10^{-4}$), rs3809006 in *ETS1* (OR=1.07, $p=1.26\times10^{-3}$), rs806349 in *DLEU1* (OR=1.12, $p=9.80\times10^{-8}$), rs6072343 upstream of *LPIN3* (OR=1.14, $p=1.14\times10^{-5}$) and rs9808753 in *IFNGR2*

Table 1 Association results of the 11 loci and MS risk assessed in 21 589 subjects of European descent

			This study						Original study [4]		Combined	
SNP	Location (hg38)	Nearest gene	N	MAF	OR (95% CI)	p Value*	I^2 (95% CI)	P_Q	OR	p Value†	OR	p Value†
rs1534422 (G/A)	chr2:12,500,615	*LOC100506457*	11	44.3	1.10 (1.06 to 1.15)	1.39×10^{-6}	0 (0 to 56)	0.530	1.06	2.49×10^{-7}	1.07	4.03×10^{-12}
rs6435203 (G/A)	chr2:203,746,472	*CD28*	11	29.5	1.08 (1.03 to 1.13)	7.20×10^{-4}	32 (0 to 67)	0.142	1.07	1.23×10^{-7}	1.07	1.35×10^{-9}
rs9846396 (T/C)	chr3:141,422,126	*ZBTB38*	11	40.4	1.03 (0.99 to 1.08)	0.0560	0 (0 to 54)	0.575	1.07	7.77×10^{-8}	1.10	5.94×10^{-8}
rs4686953 (G/A)‡	chr3:188,365,131	*LPP*	11	45.5	1.05 (1.01 to 1.10)	6.00×10^{-3}	0 (0 to 52)	0.603	1.06	2.26×10^{-7}	1.06	3.35×10^{-8}
rs727098 (C/T)	chr6:139,590,185	*DQ571824*	10	24.8	1.02 (0.97 to 1.07)	0.254	0 (0 to 53)	0.609	1.07	6.49×10^{-7}	1.06	2.20×10^{-6}
rs13260060 (A/G)	chr8:70,306,125	*NCOA2*	11	10.8	1.06 (0.99 to 1.13)	0.0444	9 (0 to 48)	0.357	1.10	7.40×10^{-7}	1.09	2.37×10^{-7}
rs3004212 (T/C)	chr10:43,147,362	*CSGALNACT2*	11	26.1	1.04 (0.99 to 1.09)	0.0640	0 (0 to 0)	0.976	1.07	2.99×10^{-7}	1.06	3.01×10^{-7}
rs3809006 (G/A)	chr11:128,540,941	*ETS1*	11	48.2	1.07 (1.02 to 1.11)	1.26×10^{-3}	38 (0 to 69)	0.0978	1.06	3.51×10^{-7}	1.06	7.74×10^{-9}
rs806349 (T/C)	chr13:50,285,854	*DLEU1*	10	46.7	1.12 (1.07 to 1.16)	9.80×10^{-8}	54 (6 to 77)	0.0213	1.06	7.85×10^{-7}	1.07	8.14×10^{-12}
rs6072343 (A/G)	chr20:41,339,548	*LPIN3*	9	13.7	1.14 (1.07 to 1.22)	1.14×10^{-5}	38 (0 to 72)	0.112	1.09	7.13×10^{-8}	1.11	7.16×10^{-12}
rs9808753 (G/A)	chr21:33,415,005	*IFNGR2*	10	11.7	1.12 (1.05 to 1.19)	2.62×10^{-4}	0 (0 to 61)	0.423	1.09	1.26×10^{-7}	1.10	4.40×10^{-10}

Fixed effect meta-analysis results for the SNPs tested across all validation data sets and after combining the results of the original[4] and of this study were performed using R. Allele names are displayed as minor/major allele based on control frequencies across all validation data sets. The underlined allele name corresponds to the risk allele. The Bonferroni-corrected threshold for significant association in the validation data sets was set to $p=4.5\times10^{-3}$ to account for 11 tests; upon combining all data sets from the original[4] and validation study, the threshold to define significance was set to $p=5\times10^{-8}$, that is, corresponding to genome-wide significance. Underlined p values indicate experiment-wide significance for the validation data sets or genome-wide significance for the meta-analysis of all available data. The locations are annotated based on the human genome build 38 (hg38) and the nearest gene has been determined according to RefGene as annotated on the UCSC (University of California, Southern California) Genome Browser. Note that the nearest gene does not necessarily represent the functional element underlying the genetic association.

*One-sided with respect to the effect direction of the original study,[4]

†Two-sided.

‡This SNP was listed with archive ID rs66756607 in the original publication.[4]

I^2, estimate of percentage of between-study heterogeneity that is beyond chance; MAF, minor allele frequency in controls across all validation datasets (in %); MS, multiple sclerosis; N, number of validation data sets after quality control; P_Q, p value derived from Q statistic to assess between-study heterogeneity.

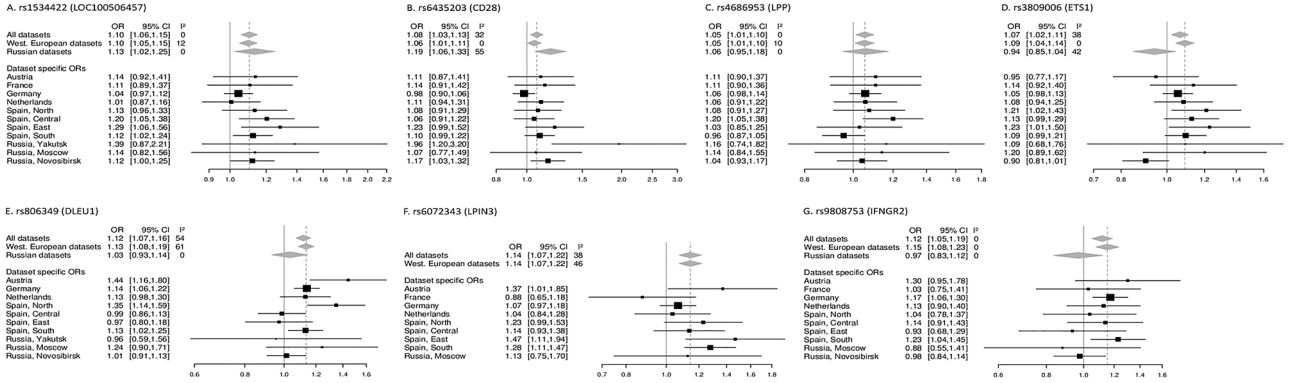

Figure 1 Meta-analyses of validation data sets showing genome-wide significant association between the putative loci and multiple sclerosis (MS) risk upon combination of all data. The x axis depicts the OR. Study-specific ORs (black squares) and 95% CIs (lines) were calculated using an additive model. The summary ORs and 95% CIs (grey diamonds) were calculated based on fixed-effect meta-analysis.

(OR=1.12, $p=2.62\times10^{-4}$). In addition, rs4686953 in *LPP* (OR=1.05, $p=6.00\times10^{-3}$) missed the threshold for experiment-wide multiple testing correction ($p_{corr}<4.5\times10^{-3}$) by a small margin. All of these seven loci showed genome-wide significant ($p<5\times10^{-8}$) association with MS risk upon meta-analysis of all available data, that is, after combining the results from the original study[4] and from the 11 independent data sets tested here, amounting to a total of 101 683 individuals (table 1).

Furthermore, while the effect size estimates of the four remaining loci showed non-significant effects in the combined validation data sets after multiple testing correction (with p values ranging from 0.254 to 0.0444, table 1, see online supplement 1: figure e-1), all pointed into the same direction of effect as in the original study (4) ($p=0.0625$ based on the underlying binomial distribution) with ORs ranging from 1.02 to 1.06. The corresponding 95% CIs of the ORs computed here included the effect size estimates of the original study[4] in all instances. However, none of the four variants reached genome-wide significance upon meta-analysis across all 101 683 available individuals (table 1).

For two of the 11 tested SNPs, there was weak evidence for between-study heterogeneity in the validation data, that is, rs806349 in *DLEU1* ($p_Q=0.0210$) and rs3809006 in *ETS1* ($p_Q=0.0978$; table 1). Accordingly, the 95% CIs of the I^2 estimates were large in both instances (ie, 6–77 for rs806349 and 0–69 for rs3809006). The heterogeneity for rs806349 was primarily due to heterogeneity of effect size estimates pointing into the same direction of effect rather than heterogeneity of ORs on either side of the null (figure 1E) and for rs3809006, this was due to the single outlying estimate of the Russian-Novosibirsk study pointing into the opposite direction of effect than the majority of data sets ($I^2=0$ (95% CI 0 to 46), $p_Q=0.709$ upon exclusion of this data set). Finally, interaction analyses of the effect size estimates of the stratified western European and Russian data sets did not yield any significant differences after correction for 11 tests ($p_{corr}<4.5\times10^{-3}$, see online supplement 1: table e-2). The only SNP approaching this threshold was rs3809006 in *ETS1* ($p_{interaction}=9.7\times10^{-3}$) showing ORs of 1.09 (95% CI 1.04 to 1.14) in the western European and 0.94 (95% CI 0.85 to 1.04) in the Russian data sets (which was due to the single outlying effect estimate from the Russia-Novosibirsk data set as described above).

To pinpoint the putative causal variant(s) underlying the newly identified association signals we applied the PICS algorithm[17] to all seven SNPs showing genome-wide significant association with MS. This revealed a number of variants (with $r^2\geq0.8$ to the index SNPs) that had 80% or more cumulative probability of including the causal variant per locus (median number of SNPs per locus 16, range 5–31; see online supplement 2). Interestingly, in all seven instances the index SNPs showed the highest probability of representing the causal variant (median probability=23.3%, range 7.2–29.1%, see online supplement 2). Gene ontology (using category 'biological process') terms and/or Biocarta pathways were available for four (ie, *CD28, LPP, IFNGR2* and *ETS1*) of the nearest genes for all seven loci. Three of these (ie, all except *LPP* to which generally only few gene ontology terms could be attributed) suggested a functional involvement in processes related to the immune system (see online supplement 1: table e-3).

DISCUSSION

Our study represents the first to show genome-wide significant association between risk for MS and common genetic variants in or near *LOC100506457*, *CD28*, *ETS1*, *DLEU1*, *LPIN3*, *IFNGR2* and *LPP*. All of these SNPs showed experiment-wide and/or genome-wide significant association in the large collection of newly analysed samples (21 589 subjects) and upon combining our data with those from previous efforts[4] (resulting in a total of 101 683 subjects), respectively. Overall, these results provide compelling evidence that the highlighted loci represent genuine genetic risk factors for MS and should be considered in future genetic, functional and clinical studies. Moreover, our study represents one of the few MS genetic association studies in data sets from Russia. Our analyses show that effect size estimates in this population are not substantially different from those from western European data sets, at least not for the 11 SNPs that were analysed here.

In addition to considerably extending the list of established genetic risk factors for MS, our study also represents an important step forward in extending the list of genetic factors shared among autoimmune diseases. This is based on the observation that six of the seven SNPs showing genome-wide significant association with MS here also show considerable evidence for association with several other autoimmune diseases, for example, coeliac disease, autoimmune thyroiditis and type 1 diabetes (table 2). This observation is in line with previous reports also suggesting putative involvements of the then established MS risk loci in other autoimmune diseases (eg, refs 1 4 18). This entirely independent evidence provides further indirect support for a genuine involvement of these loci in MS pathogenesis. Interestingly, the direction of effect was not always the

Table 2 Genetic association results of the 11 putative MS loci in other autoimmune diseases

Locus	Index SNP	Index SNP associations in other autoimmune diseases with p≤0.05	Best genome-wide significant SNP per locus across other autoimmune diseases
LOC100506457	rs1534422 (G)	ATD: OR=1.17 (G), p=1.76E-07 T1D: OR=1.09 (G), p=9.82E-06 CEL: OR=1.05 (G), p=0.0122 JIA: OR=1.13 (G), p=0.0215	–
CD28	rs6435203 (A)	CEL: OR=1.14 (G), p=2.63E-10 RA :OR=1.06 (G), p=2.55E-03 T1D: OR=1.06 (G), p=4.27E-03 NAR: OR=1.10 (G), p=0.0134	ATD: rs11571297, OR=1.37 (A), p=2.09E-23 T1D: rs3087243, OR=1.19 (G), p=7.36E-21 CEL: rs1980422, OR=1.19 (C), p=1.43E-15
ZBTB38	rs9846396 (C)	PBC: OR=1.11 (C), p=1.07E-03	–
LPP	rs4686953* (G)	CEL: OR=1.30 (A), p=2.979E-43 ATD: p=5.77E-06	CEL: rs2030519, OR=1.32 (A), p=3.00E-49 ATD: rs13093110, p=3.69E-08
DQ571824	rs727098 (C)	T1D: OR=1.04 (C), p=0.0361	–
NCOA2	rs13260060 (A)	PSO: OR=1.08 (A), p=0.0275 RA: OR=1.07 (A), p=0.0253	–
CSGALNACT2	rs3004212† (C)	–	–
ETS1	rs3809006 (G)	CEL: OR=1.10 (G), p=5.36E-07 PSO: OR=1.14 (G), p=5.32E-04 ATD: p=0.0118 NAR: OR=1.09 (A), p=0.0187	CEL: rs61907765, OR=1.18 (T), p=3.43E-13
DLEU1	rs806349 (T)	CEL: OR=1.05 (C), p=0.0119 ATD: p=0.0228 PBC: OR=1.07 (T), p=0.0255	–
LPIN3	rs6072343 (A)	PBC: OR=1.14 (A), p=1.90E-03 T1D: OR=1.07 (A), p=0.0124 ATD: p=0.0314	–
IFNGR2	rs9808753 (G)	PBC: OR=1.14 (G), p=2.46E-03 RA: OR=1.08 (A), p=3.93E-03 PSO: OR=1.12 (A), p=0.0141	–

This table displays genetic association results of the 11 loci from ImmunoChip studies of eight other autoimmune diseases as available on the ImmunoBase database (http://www.immunobase.org). The 'index SNP' indicates the most significant SNP in the MS ImmunoChip study.[4] The risk allele for each specific disease is listed in brackets if available at ImmunoBase (https://www.immunobase.org/downloads/protected_data/iChip_Data/). Underlined allele names of the index SNP indicate directions of effect inverse to the one observed for MS. Effect size estimates and effect directions were not available for autoimmune thyroid disease (ATD) on Immunobase; instead, for SNPs rs1534422 and rs11571297 they were retrieved from the original study.[26] The last column lists the most significant SNP per disease for loci (defined as ±1 Mb around the index SNP) that show genome-wide significant association with the specific disease. Dark grey background indicates loci that did not replicate in this study.
*For the diseases ATD and T1D, results for a proxy SNP were available in ImmunoBase (rs2030516, r^2>0.9 based on 1000 Genomes CEU data).
†For the diseases PSO and T1D only results for a proxy SNP of the index SNP (rs2460554, r^2>0.6) were available in ImmunoBase; no data were available for CEL and NAR for the index SNP or proxies down to r^2=0.3; for the remaining five diseases, association results for the index SNP were available.
CEL, coeliac disease; JIA, juvenile idiopathic arthritis; MS, multiple sclerosis; NAR, narcolepsy; PBC, primary biliary cirrhosis; PSO, psoriasis; RA, rheumatoid arthritis; T1D, type 1 diabetes.

same as for MS. This was most noticeable for *CD28* where the effect direction was inverse in MS compared with all other four autoimmune diseases (table 2). Such inverse effect directions across autoimmune diseases have been described previously.[18 19] Furthermore, for *CD28*, *LPP* and *ETS1* several SNPs beyond the index SNP show genome-wide significant association with coeliac disease, autoimmune thyroiditis and/or type 1 diabetes. In all cases, the association of the non-index SNPs was statistically stronger than the association of the index SNP possibly suggesting allelic heterogeneity across the disease entities assessed here (table 2).

Despite the compelling accumulated genetic evidence for 7 out of the 11 loci tested and highlighted here, there are several potential limitations to our study. First, despite having examined a sample size of over 21 500 individuals in the data sets newly genotyped here and over 101 500 subjects in the analyses combining our data with those from the ImmunoChip study,[4] power was still limited to detect very small effect sizes (ie, ORs<1.10), especially for minor allele frequencies below 0.20. In this context it is noteworthy that the effect size estimates of the four loci that did not reach genome-wide significance in the combined meta-analyses all pointed into the same direction in the validation data sets when compared with the original study.[4] Thus, based on the currently available data, we cannot exclude the possibility that some or all of the loci currently not displaying genome-wide significant evidence for association with MS risk may eventually also prove to be genuine disease loci once tested in even larger data sets. Second, this is the first study to examine the 11 putative MS risk loci of interest here in data sets of Russian origin. While to our knowledge we examined all MS case-control data sets currently available from Russia, their combined size of 3254 individuals is still comparatively small. Thus, these results need to be interpreted carefully. This also includes the analyses comparing effect size estimates between these and western European samples: currently there is—with the potential exception of rs3809006 in *ETS1*—no strong evidence suggesting that the genetic effects exerted by these 11 SNPs are different in the Russian population when compared with the western European population, but larger data sets are needed to assess this question more thoroughly. In this context, it can also not be excluded that population substructure may have affected some of the association results. However, genome-wide SNP data to assess and adjust for population substructure in the Russian data sets tested here are currently not available. This also applies to most of the western European data sets as only the Dutch data set was adjusted for ethnic substructure. However, the Dutch effect size estimates were similar to the remaining western European data sets suggesting that

undetected population stratification may not have substantially biased the results. In addition, the western European data sets have been used previously in multiple similar studies nearly always resulting in association findings in line with those from independent GWAS (eg, refs 3 20 21). Furthermore, none of the meta-analysis results was driven by a single outlying study effect estimate. Thus, the presence of type 1 error in our results due to uncorrected population stratification would imply the same direction of bias for the majority of data sets analysed. This would include the validation data sets tested here as well as those included in the original study,[4] the latter of which were all adjusted for population stratification. It appears rather unlikely that such a widespread bias would be present in the majority of all independently recruited data sets from regionally distinct populations. However, we cannot exclude that some results in our validation study were affected by bias towards the null due to subtle population stratification.

Finally, it is well known that the lead SNPs emerging from GWAS are not necessarily the variants exerting the pathogenic functional effects. To this end, we used PICS to generate a list of potentially causal variants (see online supplement 2). These variants were further assessed regarding their potential functional impact by calculating the scaled 'Combined Annotation Dependent Depletion' (CADD; http://cadd.gs.washington.edu/) score,[22] by annotating variants based on the Encyclopedia of DNA Elements (ENCODE) project using HaploReg V.3 (http://www.broadinstitute.org/mammals/haploreg/haploreg_v3.php),[23] and by interrogating a recently described large-scale expression quantitative trait locus (eQTL) database (http://genenetwork.nl/bloodeqtlbrowser/)[24] with information on transcriptome-wide microarray-based expression data from 5311 blood samples.[24] Possibly the most interesting result was obtained with SNP rs3809006, located intronically in *ETS1*, which carried the highest PICS probability of all index SNPs (29.1%) and at the same time showed one of the highest CADD scores (10.3). CADD scores >10 are among the top 10% of potentially pathogenic variants of all theoretically possible 8.6 billion single nucleotide exchanges in the human genome. A number of SNPs in linkage disequilibrium with the seven index SNPs also showed CADD scores >10, some of which were located in proximity to promoter or enhancer histone marks in the thymus, blood or brain (see online supplement 2); for instance, this included several SNPs located 3′ downstream of *CD28*. Notably, genetic association—albeit at subgenome-wide significance—between MS and *CD28* had already been reported in the candidate-gene era owing to this gene's involvement in T cell activation.[25] Interestingly, *CD28* index SNP rs6435203 highly significantly correlated with the expression of *CD28* ($p=1.15\times10^{-15}$, see online supplement 1: table e-4). Another of the novel genome-wide MS SNPs identified here is rs9808753, which is located in a region on chromosome 21q22.11 characterised by the presence of several genes involved in the immune system response. In the eQTL database rs9808753 was found to correlate with the expression of several of these genes, including *IL10RB* ($p=3.22\times10^{-14}$), *IFNAR1* ($p=2.57\times10^{-7}$), *TMEM50B* ($p=3.55\times10^{-7}$) and *IFNAR2* ($p=9.77\times10^{-4}$; see online supplement 1: table e-4). Online table e-5 (supplement 1) highlights previously described functional implications of these eQTL genes potentially underlying autoimmune disease pathophysiology and progression. Clearly, further experimental work is needed to characterise the potential functional impact of these and the other MS loci newly nominated here.

In conclusion, our study is the first to establish genome-wide significant association between risk for MS and seven new loci, of which several show strong association with other autoimmune diseases. Using previously generated transcriptome data, we observed strong eQTL effects for the MS-associated SNPs in *CD28* and *IFNGR2*. Further fine-mapping and functional studies are required to elucidate mechanisms underlying the newly highlighted disease associations.

Author affiliations
[1]Platform for Genome Analytics, Institutes of Neurogenetics & Integrative and Experimental Genomics, University of Lübeck, Lübeck, Germany
[2]Department of Neurology, Focus Program Translational Neuroscience, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany
[3]Department of Vertebrate Genomics, Max Planck Institute for Molecular Genetics, Berlin, Germany
[4]Department of Cell Biology and Immunology, Instituto de Parasitología y Biomedicina López Neyra (IPBLN), CSIC, Granada, Spain
[5]Laboratory of Pharmacogenomics, Institute of Chemical Biology and Fundamental Medicine, Siberian Division, Russian Academy of Sciences, Novosibirsk, Russia
[6]Novosibirsk State University, Novosibirsk, Russia
[7]Neurogenomiks Laboratory, University of the Basque Country (UPV/EHU), Leioa, Spain
[8]Achucarro Basque Center for Neuroscience, Zamudio, Spain
[9]Immunology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[10]Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Inserm U 1127, CNRS UMR 7225, ICM F-75013, Paris, France
[11]Department of Neurology-Neuroimmunology, Centre d'Esclerosi Múltiple de Catalunya, Cemcat, Hospital Universitari Vall d'Hebron (HUVH), Barcelona, Spain
[12]Department of Neurology, Medical University of Vienna, Vienna, Austria
[13]Department of Neurology, MS Centre ErasMS, Erasmus Medical Centre, Rotterdam, The Netherlands
[14]Department of Human Genetics, Ruhr University, Bochum, Germany
[15]Department of Neurology, Medical Faculty, Heinrich Heine University, Düsseldorf, Germany
[16]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[17]Servicio de Neurología, Hospital Universitario de Basurto, Bilbao, Spain
[18]Multiple Sclerosis Unit, Neurology Department, Hospital Clínico San Carlos, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), Madrid, Spain
[19]Department of Neurology, University of Rostock, Rostock, Germany
[20]Department of Neurology and Neurosurgery, Pirogov Russian National Research Medical University (RNRMU), Moscow, Russia
[21]Department of Neurology, University of Würzburg, Würzburg, Germany
[22]Department of Neurology, St. Josef-Hospital, Ruhr-University, Bochum, Germany
[23]Department of Medicine, Rheumatology, and Clinical Immunology & DRFZ, Charité University Medicine, Berlin, Germany
[24]Faculty of Health, University Witten/Herdecke, Witten, Germany
[25]Department of Molecular Biology and Medical Biotechnology, Pirogov Russian National Research Medical University (RNRMU), Moscow, Russia
[26]Unidad de Gestión Clínica de Neurociencias, Instituto de Biomedicina de Málaga (IBIMA), Hospital Regional Universitario de Málaga, Málaga, Spain
[27]Institute of Clinical Neuroimmunology, Campus Grosshadern, Ludwig Maximilian University, Munich, Germany
[28]Unidad de Esclerosis Múltiple, Hospital Universitario Virgen Macarena, Sevilla, Spain
[29]Multiple Sclerosis Center, Novosibirsk Regional State Clinical Hospital, Novosibirsk, Russia
[30]Medical College of Wisconsin, Milwaukee, Wisconsin, USA
[31]Institute of Human Genetics, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany
[32]Department of Clinical Chemistry, Ludwig Maximilian University, Munich, Germany
[33]CeGaT GmbH, Tuebingen, Germany
[34]Department of Clinical Neurology and Algology, Novosibirsk State Medical University, Novosibirsk, Russia
[35]Department of Neurology, Sozialstiftung Bamberg Hospital, Bamberg, Germany
[36]Department of Neurology of Postgraduate Education, Omsk State Medical Academy, Omsk, Russia
[37]Territorial Clinical Hospital, Barnaul, Russia
[38]Department of Nervous Diseases, Altai State Medical University, Barnaul, Russia
[39]Institute for Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center Mainz, Mainz, Germany
[40]AP-HP, Département of des Maladies du Système Nerveux, Pitié-Salpêtrère Hospital, Paris, France
[41]Department of Medicine, School of Public Health, Imperial College London, London, UK

Acknowledgements The authors thank ICM, Généthon, for their help and support. The authors also thank Drs Elena G Arefyeva, Svetlana A Elchaninova, Fedor A Platonov, Ilja Demuth, Elisabeth Steinhagen-Thiessen and Ulman Lindenberger for recruiting individuals included in this study.

Contributors Study concept: CML and LB. Study supervision or coordination: CML, FL, IC-R, RH, AZ, MC, BF, EU, KV, MF, FM, FZ and LB. Acquisition of data: CML, FL, AA, EAS, NU, BdlH, LG-N, SM, ER, B-MMS, JYM, AM, IW, DAA, OA, IAl, AA, RA, IAs, PB, ANB, MB, AC, TD, JTE, OOF, MF, OF, AG-M, L-AG, CG, H-PH, SH, GI, DSK, AK, CK, TK, LL, PL, NAM, XM, EVP, PR, ASR, CS, IVS, EYT, AW, UKZ, IC-R, RH, AZ, MC, BF, EU, KV, MF, FM and FZ. Statistical analysis: CML. Analysis and interpretation of data: CML, FL, HB, FM and LB. Drafting the manuscript: CML and LB. Critical revision of the manuscript for content: FL, AA, EAS, NU, BdlH, LG-N, SM, ER, B-MMS, JYM, AM, IW, DAA, OA, IAl, AA, RA, IAs, PB, ANB, MB, AC, TD, JTE, OOF, MF, OF, AG-M, L-AG, CG, H-PH, SH, GI, DSK, AK, CK, TK, LL, PL, NAM, XM, EVP, PR, ASR, CS, IVS, EYT, AW, UKZ, HB, IC-R, RH, AZ, MC, BF, EU, KV, MF, FM and FZ.

Funding The German Ministry for Education and Research (grant 16SV5538 to LB, KKNMS to FZ, NBL3 to UKZ, 01GM1203A to H-PH and OA), the Johannes Gutenberg University Mainz (grants MAIFOR and "Inneruniversitäre Forschungsförderung Stufe I" to FL), the Walter- and Ilse-Rose-Stiftung (to H-PH and OA), the Instituto de Salud Carlos III-Fondo Europeo de Desarrollo Regional (Feder), the Fondo de Investigaciones Sanitarias FIS (grant numbers P12/00555 to FM, PI13/01527 to AA, PI13/01466 to GI] and Junta de Andalucía (JA)- Fondos Europeos de Desarrollo Regional (FEDER) [grant number CTS2704 to FM), (PI13/0879) and Fundación Alicia Koplowitz (to EU), INSERM, ARSEP, AFM and ANR-10-IAIHU-06 (to BF), "REEM: Red Española de Esclerosis Múltiple" (RETICS-REEM RD12/0032/009, http://www.reem.es, to AG-M), "Fondo de Investigaciones Sanitarias" (FI11/00560 to BdlH), the Dutch MS Research Foundation (to RH), the grant Ayudas para Grupos de Investigación del Sistema Universitario Vasco—Gobierno Vasco (Ref. IT512-10 to KV), the Russian Foundation for Basic Research (13-04-40281-H and 15-04-04866 to ANB and EVP) and the Russian Science Foundation (14-14-00605 to EYT). The funding sources did not have an influence on the design and conduct of the study; the collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication. EU works for the Fundación para la Investigación Biomédica-Hospital Clínico San Carlos- IdISSC. Samples and data from Northern Spanish patients with MS were processed by the Basque BioBank for Research-OEHUN (http://www.biobanco.vasco). The French DNA samples were provided by the BRC-REFGENSEP.

Competing interests None declared.

Ethics approval Local Ethics Committees.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All analysis results have been made available in this manuscript.

REFERENCES

1 Sawcer S, Franklin RJM, Ban M. Multiple sclerosis genetics. *Lancet Neurol* 2014;13:700–9.

2 International Multiple Sclerosis Genetics Consortium, Wellcome Trust Case Control Consortium 2, Sawcer S, Hellenthal G, Pirinen M, Spencer CCA, Patsopoulos NA, Moutsianas L, Dilthey A, Su Z, Freeman C, Hunt SE, Edkins S, Gray E, Booth DR, Potter SC, Goris A, Band G, Oturai AB, Strange A, Saarela J, Bellenguez C, Fontaine B, Gillman M, Hemmer B, Gwilliam R, Zipp F, Jayakumar A, Martin R, Leslie S, Hawkins S, Giannoulatou E, D'alfonso S, Blackburn H, Martinelli Boneschi F, Liddle J, Harbo HF, Perez ML, Spurkland A, Waller MJ, Mycko MP, Ricketts M, Comabella M, Hammond N, Kockum I, McCann OT, Ban M, Whittaker P, Kemppinen A, Weston P, Hawkins C, Widaa S, Zajicek J, Dronov S, Robertson N, Bumpstead SJ, Barcellos LF, Ravindrarajah R, Abraham R, Alfredsson L, Ardlie K, Aubin C, Baker A, Baker K, Baranzini SE, Bergamaschi L, Bergamaschi R, Bernstein A, Berthele A, Boggild M, Bradfield JP, Brassat D, Broadley SA, Buck D, Butzkueven H, Capra R, Carroll WM, Cavalla P, Celius EG, Cepok S, Chiavacci R, Clerget-Darpoux F, Clysters K, Comi G, Cossburn M, Cournu-Rebeix I, Cox MB, Cozen W, Cree BAC, Cross AH, Cusi D, Daly MJ, Davis E, de Bakker PIW, Debouverie M, D'hooghe MB, Dixon K, Dobosi R, Dubois B, Ellinghaus D, Elovaara I, Esposito F, Fontenille C, Foote S, Franke A, Galimberti D, Ghezzi A, Glessner J, Gomez R, Gout O, Graham C, Grant SFA, Guerini FR, Hakonarson H, Hall P, Hamsten A, Hartung H-P, Heard RN, Heath S, Hobart J, Hoshi M, Infante-Duarte C, Ingram G, Ingram W, Islam T, Jagodic M, Kabesch M, Kermode AG, Kilpatrick TJ, Kim C, Klopp N, Koivisto K, Larsson M, Lathrop M, Lechner-Scott JS, Leone MA, Leppä V, Liljedahl U, Bomfim IL, Lincoln RR, Link J, Liu J, Lorentzen AR, Lupoli S, Macciardi F, Mack T, Marriott M, Martinelli V, Mason D, McCauley JL, Mentch F, Mero I-L, Mihalova T, Montalban X, Mottershead J, Myhr K-M, Naldi P, Ollier W, Page A, Palotie A, Pelletier J, Piccio L, Pickersgill T, Piehl F, Pobywajlo S, Quach HL, Ramsay PP, Reunanen M, Reynolds R, Rioux JD, Rodegher M, Roesner S, Rubio JP, Rückert I-M, Salvetti M, Salvi E, Santaniello A, Schaefer CA, Schreiber S, Schulze C, Scott RJ, Sellebjerg F, Selmaj KW, Sexton D, Shen L, Simms-Acuna B, Skidmore S, Sleiman PMA, Smestad C, Sørensen PS, Søndergaard HB, Stankovich J, Strange RC, Sulonen A-M, Sundqvist E, Syvänen A-C, Taddeo F, Taylor B, Blackwell JM, Tienari P, Bramon E, Tourbah A, Brown MA, Tronczynska E, Casas JP, Tubridy N, Corvin A, Vickery J, Jankowski J, Villoslada P, Markus HS, Wang K, Mathew CG, Wason J, Palmer CNA, Wichmann H-E, Plomin R, Willoughby E, Rautanen A, Winkelmann J, Wittig M, Trembath RC, Yaouanq J, Viswanathan AC, Zhang H, Wood NW, Zuvich R, Deloukas P, Langford C, Duncanson A, Oksenberg JR, Pericak-Vance MA, Haines JL, Olsson T, Hillert J, Ivinson AJ, De Jager PL, Peltonen L, Stewart GJ, Hafler DA, Hauser SL, McVean G, Donnelly P, Compston A. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011;476:214–19.

3 International Multiple Sclerosis Genetics Consortium, Lill CM, Schjeide B-MM, Graetz C, Ban M, Alcina A, Ortiz MA, Pérez J, Damotte V, Booth D, Lopez de Lapuente A, Broer L, Schilling M, Akkad DA, Aktas O, Alloza I, Antigüedad A, Arroyo R, Blaschke P, Buttmann M, Chan A, Compston A, Cournu-Rebeix I, Dörner T, Epplen JT, Fernández Ó, Gerdes L-A, Guillot-Noël L, Hartung H-P, Hoffjan S, Izquierdo G, Kemppinen A, Kroner A, Kubisch C, Kümpfel T, Li S-C, Lindenberger U, Lohse P, Lubetzki C, Luessi F, Malhotra S, Mescheriakova J, Montalban X, Papeix C, Paredes LF, Rieckmann P, Steinhagen-Thiessen E, Winkelmann A, Zettl UK, Hintzen R, Vandenbroeck K, Stewart G, Fontaine B, Comabella M, Urcelay E, Matesanz F, Sawcer S, Bertram L, Zipp F. MANBA, CXCR5, SOX8, RPS6KB1 and ZBTB46 are genetic risk loci for multiple sclerosis. *Brain* 2013;136:1778–82.

4 International Multiple Sclerosis Genetics Consortium (IMSGC), Beecham AH, Patsopoulos NA, Xifara DK, Davis MF, Kemppinen A, Cotsapas C, Shah TS, Spencer C, Booth D, Goris A, Oturai A, Saarela J, Fontaine B, Hemmer B, Martin C, Zipp F, D'Alfonso S, Martinelli-Boneschi F, Taylor B, Harbo HF, Kockum I, Hillert J, Olsson T, Ban M, Oksenberg JR, Hintzen R, Barcellos LF, Wellcome Trust Case Control Consortium 2 (WTCCC2), International IBD Genetics Consortium (IIBDGC), Agliardi C, Alfredsson L, Alizadeh M, Anderson C, Andrews R, Søndergaard HB, Baker A, Band G, Baranzini SE, Barizzone N, Barrett J, Bellenguez C, Bergamaschi L, Bernardinelli L, Berthele A, Biberacher V, Binder TMC, Blackburn H, Bomfim IL, Brambilla P, Broadley S, Brochet B, Brundin L, Buck D, Butzkueven H, Caillier SJ, Camu W, Carpentier W, Cavalla P, Celius EG, Coman I, Comi G, Corrado L, Cosemans L, Cournu-Rebeix I, Cree BAC, Cusi D, Damotte V, Defer G, Delgado SR, Deloukas P, di Sapio A, Dilthey AT, Donnelly P, Dubois B, Duddy M, Edkins S, Elovaara I, Esposito F, Evangelou N, Fiddes B, Field J, Franke A, Freeman C, Frohlich IY, Galimberti D, Gieger C, Gourraud P-A, Graetz C, Graham A, Grummel V, Guaschino C, Hadjixenofontos A, Hakonarson H, Halfpenny C, Hall G, Hall P, Hamsten A, Harley J, Harrower T, Hawkins C, Hellenthal G, Hillier C, Hobart J, Hoshi M, Hunt SE, Jagodic M, Jelčić I, Jochim A, Kendall B, Kermode A, Kilpatrick T, Koivisto K, Konidari I, Korn T, Kronsbein H, Langford C, Larsson M, Lathrop M, Lebrun-Frenay C, Lechner-Scott J, Lee MH, Leone MA, Leppä V, Liberatore G, Lie BA, Lill CM, Lindén M, Link J, Luessi F, Lycke J, Macciardi F, Männistö S, Manrique CP, Martin R, Martinelli V, Mason D, Mazibrada G, McCabe C, Mero I-L, Mescheriakova J, Moutsianas L, Myhr K-M, Nagels G, Nicholas R, Nilsson P, Piehl F, Pirinen M, Price SE, Quach H, Reunanen M, Robberecht W, Robertson NP, Rodegher M, Rog D, Salvetti M, Schnetz-Boutaud NC, Sellebjerg F, Selter RC, Schaefer C, Shaunak S, Shen L, Shields S, Siffrin V, Slee M, Sorensen PS, Sorosina M, Sospedra M, Spurkland A, Strange A, Sundqvist E, Thijs V, Thorpe J, Ticca A, Tienari P, van Duijn C, Visser EM, Vucic S, Westerlind H, Wiley JS, Wilkins A, Wilson JF, Winkelmann J, Zajicek J, Zindler E, Haines JL, Pericak-Vance MA, Ivinson AJ, Stewart G, Hafler D, Hauser SL, Compston A, McVean G, De Jager P, Sawcer SJ, McCauley JL. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. *Nat Genet* 2013;45:1353–60.

5 Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* 2011;13:101.

6 Schmied MC, Zehetmayer S, Reindl M, Ehling R, Bajer-Kornek B, Leutmezer F, Zebenholzer K, Hotzy C, Lichtner P, Meitinger T, Wichmann H-E, Illig T, Gieger C, Huber K, Khalil M, Fuchs S, Schmidt H, Auff E, Kristoferitsch W, Fazekas F, Berger T, Vass K, Zimprich A. Replication study of multiple sclerosis (MS) susceptibility alleles and correlation of DNA-variants with disease features in a cohort of Austrian MS patients. *Neurogenetics* 2012;13:181–7.

7 Sokolova EA, Malkova NA, Korobko DS, Rozhdestvenskii AS, Kakulya AV, Khanokh EV, Delov RA, Platonov FA, Popova TY, Aref' eva EG, Zagorskaya NN, Alifirova VM, Titova MA, Smagina IV, El' chaninova SA, Popovtseva AV, Puzyrev VP, Kulakova OG, Tsareva EY, Favorova OO, Shchur SG, Lashch NY, Popova NF, Popova EV, Gusev EI, Boyko AN, Aulchenko YS, Filipenko ML. Association of SNPs of CD40 gene with multiple sclerosis in Russians. *PLoS ONE* 2013;8:e61032.

8 Poser CM, Paty DW, Scheinberg L, McDonald WI, Davis FA, Ebers GC, Johnson KP, Sibley WA, Silberberg DH, Tourtellotte WW. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol* 1983;13:227–31.

9 McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, McFarland HF, Paty DW, Polman CH, Reingold SC, Sandberg-Wollheim M, Sibley W, Thompson A, van den Noort S, Weinshenker BY, Wolinsky JS. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001;50:121–7.

10 Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003;19:149–50.
11 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75.
12 Pirinen M, Donnelly P, Spencer CCA. Including known covariates can reduce power to detect genetic effects in case-control studies. *Nat Genet* 2012;44:848–51.
13 Evangelou E, Ioannidis JPA. Meta-analysis methods for genome-wide association studies and beyond. *Nat Rev Genet* 2013;14:379–89.
14 Chatzinasiou F, Lill CM, Kypreou K, Stefanaki I, Nicolaou V, Spyrou G, Evangelou E, Roehr JT, Kodela E, Katsambas A, Tsao H, Ioannidis JPA, Bertram L, Stratigos AJ. Comprehensive field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma. *J Natl Cancer Inst* 2011;103:1227–35.
15 Lill CM, Roehr JT, McQueen MB, Kavvoura FK, Bagade S, Schjeide B-MM, Schjeide LM, Meissner E, Zauft U, Allen NC, Liu T, Schilling M, Anderson KJ, Beecham G, Berg D, Biernacka JM, Brice A, DeStefano AL, Do CB, Eriksson N, Factor SA, Farrer MJ, Foroud T, Gasser T, Hamza T, Hardy JA, Heutink P, Hill-Burns EM, Klein C, Latourelle JC, Maraganore DM, Martin ER, Martinez M, Myers RH, Nalls MA, Pankratz N, Payami H, Satake W, Scott WK, Sharma M, Singleton AB, Stefansson K, Toda T, Tung JY, Vance J, Wood NW, Zabetian CP, 23andMe Genetic Epidemiology of Parkinson's Disease Consortium, International Parkinson's Disease Genomics Consortium, Parkinson's Disease GWAS Consortium, Wellcome Trust Case Control Consortium 2), Young P, Tanzi RE, Khoury MJ, Zipp F, Lehrach H, Ioannidis JPA, Bertram L. Comprehensive research synopsis and systematic meta-analyses in Parkinson's disease genetics: The PDGene database. *PLoS Genet* 2012;8:e1002548.
16 Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;326:219.
17 Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, Shoresh N, Whitton H, Ryan RJH, Shishkin AA, Hatan M, Carrasco-Alfonso MJ, Mayer D, Luckey CJ, Patsopoulos NA, De Jager PL, Kuchroo VK, Epstein CB, Daly MJ, Hafler DA, Bernstein BE. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015;518:337–43.
18 Cotsapas C, Voight BF, Rossin E, Lage K, Neale BM, Wallace C, Abecasis GR, Barrett JC, Behrens T, Cho J, De Jager PL, Elder JT, Graham RR, Gregersen P, Klareskog L, Siminovitch KA, van Heel DA, Wijmenga C, Worthington J, Todd JA, Hafler DA, Rich SS, Daly MJ, FOCiS Network of Consortia. Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet* 2011;7:e1002254.
19 Sirota M, Schaub MA, Batzoglou S, Robinson WH, Butte AJ. Autoimmune disease classification by inverse association with SNP alleles. *PLoS Genet* 2009;5:e1000792.
20 Lill CM, Schjeide B-MM, Graetz C, Liu T, Damotte V, Akkad DA, Blaschke P, Gerdes L-A, Kroner A, Luessi F, Cournu-Rebeix I, Hoffjan S, Winkelmann A, Touze E, Pico F, Corcia P, Otaegui D, Antigüedad A, Alcina A, Comabella M, Montalban X, Olascoaga J, Matesanz F, Dörner T, Li S-C, Steinhagen-Thiessen E, Lindenberger U, Chan A, Rieckmann P, Hartung H-P, Aktas O, Lohse P, Buttmann M, Kümpfel T, Kubisch C, Zettl UK, Epplen JT, Fontaine B, Zipp F, Vandenbroeck K, Bertram L. Genome-wide significant association of ANKRD55 rs6859219 and multiple sclerosis risk. *J Med Genet* 2013;50:140–3.
21 Lill CM, Schilling M, Ansaloni S, Schröder J, Jaedicke M, Luessi F, Schjeide B-MM, Mashychev A, Graetz C, Akkad DA, Gerdes L-A, Kroner A, Blaschke P, Hoffjan S, Winkelmann A, Dörner T, Rieckmann P, Steinhagen-Thiessen E, Lindenberger U, Chan A, Hartung H-P, Aktas O, Lohse P, Buttmann M, Kümpfel T, Kubisch C, Zettl UK, Epplen JT, Zipp F, Bertram L. Assessment of microRNA-related SNP effects in the 3′ untranslated region of the IL22RA2 risk locus in multiple sclerosis. *Neurogenetics* 2014;15:129–34.
22 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 2014;46:310–15.
23 Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012;40:D930–4.
24 Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, Christiansen MW, Fairfax BP, Schramm K, Powell JE, Zhernakova A, Zhernakova DV, Veldink JH, Van den Berg LH, Karjalainen J, Withoff S, Uitterlinden AG, Hofman A, Rivadeneira F, 't Hoen PAC, Reinmaa E, Fischer K, Nelis M, Milani L, Melzer D, Ferrucci L, Singleton AB, Hernandez DG, Nalls MA, Homuth G, Nauck M, Radke D, Völker U, Perola M, Salomaa V, Brody J, Suchy-Dicey A, Gharib SA, Enquobahrie DA, Lumley T, Montgomery GW, Makino S, Prokisch H, Herder C, Roden M, Grallert H, Meitinger T, Strauch K, Li Y, Jansen RC, Visscher PM, Knight JC, Psaty BM, Ripatti S, Teumer A, Frayling TM, Metspalu A, van Meurs JBJ, Franke L. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013;45:1238–43.
25 Teutsch SM, Booth DR, Bennetts BH, Heard RNS, Stewart GJ. Association of common T cell activation gene polymorphisms with multiple sclerosis in Australian patients. *J Neuroimmunol* 2004;148:218–30.
26 Cooper JD, Simmonds MJ, Walker NM, Burren O, Brand OJ, Guo H, Wallace C, Stevens H, Coleman G, Wellcome Trust Case Control Consortium, Franklyn JA, Todd JA, Gough SCL. Seven newly identified loci for autoimmune thyroid disease. *Hum Mol Genet* 2012;21:5202–8.

Genome-wide significant association with seven novel multiple sclerosis risk loci

Christina M Lill, Felix Luessi, Antonio Alcina, Ekaterina A Sokolova, Nerea Ugidos, Belén de la Hera, Léna Guillot-Noël, Sunny Malhotra, Eva Reinthaler, Brit-Maren M Schjeide, Julia Y Mescheriakova, Andriy Mashychev, Inken Wohlers, Denis A Akkad, Orhan Aktas, Iraide Alloza, Alfredo Antigüedad, Rafa Arroyo, Ianire Astobiza, Paul Blaschke, Alexei N Boyko, Mathias Buttmann, Andrew Chan, Thomas Dörner, Joerg T Epplen, Olga O Favorova, Maria Fedetz, Oscar Fernández, Angel García-Martínez, Lisa-Ann Gerdes, Christiane Graetz, Hans-Peter Hartung, Sabine Hoffjan, Guillermo Izquierdo, Denis S Korobko, Antje Kroner, Christian Kubisch, Tania Kümpfel, Laura Leyva, Peter Lohse, Nadezhda A Malkova, Xavier Montalban, Ekaterina V Popova, Peter Rieckmann, Alexei S Rozhdestvenskii, Christiane Schmied, Inna V Smagina, Ekaterina Y Tsareva, Alexander Winkelmann, Uwe K Zettl, Harald Binder, Isabelle Cournu-Rebeix, Rogier Hintzen, Alexander Zimprich, Manuel Comabella, Bertrand Fontaine, Elena Urcelay, Koen Vandenbroeck, Maxim Filipenko, Fuencisla Matesanz, Frauke Zipp and Lars Bertram

J Med Genet 2015 52: 848-855 originally published online October 16, 2015
doi: 10.1136/jmedgenet-2015-103442

Notes