Supplementary Material for

Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility

International Multiple Sclerosis Genetics Consortium

Corresponding author. Email: pld2115@cumc.columbia.edu

Published 27 September 2019, Science 365, eaav7188 (2019)
DOI: 10.1126/science.aav7188

This PDF file includes:

- Materials and Methods
- Supplementary Text
- Consortium Members
- Figs. S1 to S40
- References

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/content/365/6460/eaav7188/suppl/DC1)

- Tables S1 to S53 as separate Excel files
- Data Files S1 to S4 as zipped archives
Note: The order of the Supplementary Figures and Tables follow the order of appearance in the main text and then the supplementary material.

**Material and Methods**

**Discovery data sets**

**Cohorts**

For the discovery phase we analyzed all available to the IMSGC GWAS data as of end of 2011. These included 6 data sets included in Patsopoulos et al 2011 (5) and data from WTCCC2 and IMSGC 2011 (4) studies. The latter opted-in combining the data from 37 cohorts into 2 case-control sets: i) the UK, and ii) the non-UK. For the purpose of this project we developed an empirical approach using principal component analysis (PCA) and PCs’ projection to organize data from the 37 cohorts into 7 distinct case-control sets (Supplementary Table 1). In brief, we projected each of the 37 cohorts on the European samples of the 1000 Genomes Phase Ia populations, and we then clustered together cohorts that had overlapping cases and controls into 7 case-control datasets. Specifically, we compared different combinations of the component cohorts and used the estimated genomic inflation factor in the case-control analysis to identify the case-control data sets that had the lowest inflation factor (lambda). Besides these published data, we also analyzed 2 other data sets (Supplementary Table 2), one from US (Berkeley; Kaiser Permanente) and another one from Rotterdam, Netherlands (Rotterdam). The diagnosis of MS in all patients was evaluated according to the standard diagnostic criteria (43, 44). Supplementary Table 3 summarizes the number of subjects in each of 15 data sets of the discovery phase.

*Description of the Rotterdam study*

Illumina 610k genotyping was performed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. MS cases from the Netherlands were recruited and ascertained MS patients as part of a nationwide study on genetic susceptibility in MS, all examined and diagnosed at the outpatient clinic of the Rotterdam MS Centre ErasMS at Erasmus MC. Healthy controls were included in the Rotterdam Study cohort (RS-III) an ongoing prospective population-based cohort study in the Rotterdam region.
Description of the Berkeley study
MS cases and controls were recruited from the Kaiser Permanente Medical Care Plan, Northern California Region (KPNC). KPNC is an integrated health service delivery system with a membership of 4 million that comprises about 25–30% of the population of a 22-county service area and is the largest health care provider in northern California. Eligible KPNC cases were defined as current members with a diagnosis of MS by a neurologist (ICD-9 code 340.xx), age 18–69 years, and membership in KPNC at initial contact. Diagnoses were validated using electronic health record EHR review and individual interviews. Controls were current KPNC members without a diagnosis of MS or related condition (optic neuritis, transverse myelitis, or demyelinating disease; ICD-9 codes: 340, 341.0, 341.1, 341.2, 341.20, 341.21, 341.22, 341.8, 341.9, 377.3, 377.30, 377.39, and 328.82) confirmed through EHR data. All cases and controls were white/non-Hispanic based on self-report. Genotyping was performed with the Illumina Infinium 660K BeadChip Array (Illumina Inc., San Diego, California) or the Human Omni Express (Illumina Inc.) for cases and controls from the Kaiser Permanente Medical Care Plan of Northern California.

Quality check
In each of the 15 data sets of the discovery phase we applied the same quality check (QC) pipeline for the 22 autosomal chromosomes:

- We removed SNPs with missingness rate $\geq 0.01$
- We removed individuals with missingness rate $\geq 0.05$
- We removed SNPs that had differential missingness between cases and controls with p-value $< 0.001$.
- We removed SNPs that violated Hardy-Weinberg equilibrium (HWE) in controls only using a MAF-stratified p-value cut-offs:
  - MAF $> 0.3$: p-value $< 0.0001$
  - MAF 0.2 to 0.3: p-value $< 0.00001$
  - MAF 0.1 to 0.2: p-value $< 0.000001$
  - MAF $< 0.1$: p-value $< 0.000001$
- We removed SNPs with MAF $< 1\%$.
- We LD-pruned the autosomal chromosomes with the following settings in PLINK: --indep-pairwise 100 2 0.1 (100 SNPs window, 2 SNPs step, r$^2=0.1$).
• Using the LD-pruned set we removed samples with inbreeding coefficient F>=0.05 or <= -0.05.
• Using the LD-pruned set we calculated identity-by-descent (IBD) estimates, quantified as pi-hat (proportion of IBD) and we removed samples with values >= 0.1, corresponding to 3rd-4th degree relatives.
• Then we used EIGENSOFT (45) to calculate principal components (PCs) and remove individuals that were outside +/- 6 standard deviations in each of the 10 first PCs. We further projected the samples of each data set onto the HapMap 3 populations and we manually removed any individuals that fell outside the main cluster of the respective data set or the main European-ancestry cluster.

Supplementary Table 3 summarizes the cases and controls per data set and the number of SNPs that survived the QC.

**Imputation**

To increase the number of analyzed SNPs and facilitate the meta-analysis of the 15 data sets we used BEAGLE v3.2.1 (46) and the 1000 Genomes Phase Ia European panel to impute ~34 million genetic variants. We designed the imputation pipeline to impute genetic regions of 5Mbps with another 2Mbps around each side, to capture long range LD. We imputed the samples of each data set in random sets of ~300-400 individuals, preserving the case-control ratio. We reconstructed post-imputation the autosomal genome by merging the 5Mbps regions into the respective chromosomes.

**Analysis of autosomal genome**

We used the post-imputation probabilities to perform logistic regression within each of the 15 data sets, using the first 5 PCs to account for population stratification. Post-analysis we removed monomorphic SNPs, and SNPs with MAF less than 1% or imputation quality of 0.1 (INFO score; ratio of observed vs. expected post-imputation genotypes). We opted in for a liberal threshold of imputation quality (INFO>0.1) to allow for a larger overlap of imputed SNPs across the 15 data sets that used different genotyping arrays. Furthermore, the large-scale genotyped replication collection sets enabled the refinement of true associations, despite their level of
imputation quality in the discovery phase. The genomic inflation factor ranged from 0.96 to 1.09 (Supplementary Table 3).

**Meta-analysis of discovery set GWAS**

We combined the 15 discovery data sets using a fixed effects meta-analysis. To control for any residual inflation, we applied genomic control by penalizing the standard error of the natural logarithm of the odds ratio (OR).

**Strategy to identify independent effects**

In order to identify statistically independent effects, especially within loci that could foster more than one effects, we devised a divide-and-conquer stepwise procedure:

1. Removed from the whole autosomal genome an extended region of 12Mbps around the MHC region (from 24,000,000 to 35,000,000Mbps of chromosome 6; hg19).
2. Identify the most statistically significant SNP from the remaining autosomal genome and remove ±1Mbps around it, resulting in a 2Mbps region. We call this a lead SNP.
3. Repeat #2 until no SNP with p-value<0.05 is left.
4. Within each region and per data set, run a logistic regression model adding the lead SNP in the model, along with the previous covariates. Meta-analyze across the 15 data sets, applying genomic control, and identify the most statistically significant SNP in the region. We call these SNPs effect SNP, acknowledging that these represent a statistically independent effect that tags a possible true causal variant.
5. Repeat #4, adding in each step the previously identified effect SNP.
6. Stop when: i) there are no more SNPs with p-value < 0.05 and 10 or less effect SNPs have been identified in the region, or ii) there no more SNPs with p-value < 0.001 when more than 10 SNPs have been identified in the region, or iii) if the region contained one of the suggestive SNPs (p-value < 1x10^-5) reported in our ImmunoChip project (3), set the p-value threshold at 0.1 for the first 10 steps.

The above approach resulted in a list of seemingly statistically independent effects (see Results below). We decided to replicate these effects in two large-scale replication studies.

**Replication data sets: MS Chip**

In order to replicate the findings of the discovery set we conducted a large-scale replication effort, aiming to genotype more than 40,000 cases and controls with a custom-made array.
**Design of MS Chip**

We chose to design the MS Chip using the Illumina iSelect platform, adding ~90K custom selected SNPs to the Illumina Exome Core content (~200K SNPs). Given the extended list of prioritized effects (see Results below) it was not feasible to fine map all the potentially associated loci. Instead we aimed to include several tagging variants around prioritized effect SNPs, thus allowing replicating these effects by leveraging LD structure. Furthermore, we enriched the list of SNPs with genetic variants associated with other autoimmune diseases and reported in the GWAS Catalog. To allow correction for population stratification we added 2 panels of ancestry informative markers (AIMs) optimized for European ancestry samples and one panel for Latinos.

We included the following SNPs in the custom content of the MS Chip (the above SNP count numbers included overlapping sets):

- Prioritized SNPs from the discovery set (n=24,208).
- List of tagging SNPs for the genome-wide and suggestive effects of our ImmunoChip study (n=3,840).
- SNPs from a Swedish MS cohort (n=3,528)
- 5,000 SNPs near the centromeric regions of chromosome 1
- Two panels of AIMs optimized for European ancestry subpopulations (n=4,014).
- A panel of AIMs optimized for Latinos (n=1,989).
- SNPs prioritized from studies with Latino ancestry (n=1,372)
- SNPs that were included on the reference panels of HLA imputation tools: SNP2HLA (47) (n=2,548) and HLA*IMP (48) (1,967).
- 9,580 SNPs tagging the extended MHC.
- 1,884 SNPs in chromosome X.
- 1,734 SNPs in chromosome Y.
- SNPs prioritized from ImmunoChip studies of inflammatory bowel diseases (1,878), celiac disease (n=455), and rheumatoid arthritis (n=452).
- SNPs prioritized from studies of MS in sardinian populations (n=426) and mainland Italy (n=89).
• Genetic variants prioritized from targeted sequencing studies of IMGC members (n=6,670).
• SNPs under positive selection (n=1,309).
• SNPs tagging known cis-eQTLs in immune cells (n=5,987).
• GWAS Catalog list of variants (as of July 2012) and various other SNPs provided by IMGC members (n=7,091)
• Potentially common effects between MS and other common diseases (n=8,194)

Overall, 94,168 genetic variants were submitted to Illumina and 88,659 variants made it to the final MS Chip (including 242,901 variants from the exome core).

Cohorts
All subjects genotyped with the MS chip in this replication effort gave valid informed consent in line with approval from local ethics committee/IRB. Cases were diagnosed by neurologists familiar with MS in accordance with standard criteria based on clinical and para-clinical information. Subtle differences between the demographics of the cohorts, summarized in Supplementary Table 2, reflect local interests of the contributing groups. Details of local ascertainment are outlined below.

Australia
Australian MS cases were recruited as part of the Tasmanian MS genes and prevalence study and the Ausimmune/Auslong study based at the Menzies Institute for Medical Research University of Tasmania. The control subjects were recruited as part of the Ausimmune study.

Denmark
Patients were recruited from The Danish Multiple Sclerosis Biobank at Rigshospitalet in Copenhagen, which includes patients recruited by neurologists from MS centers across Denmark, although most patients were recruited from the Copenhagen area. The large proportion of RRMS cases in this cohort reflects the frequent visits to the MS clinics by this group of patients for blood sampling and treatment, and is not as such representative for the frequency of RRMS in the general population. The control group consisted of healthy white Danish blood donors residing in the Copenhagen area and were recruited as part of the Danish Blood Donor Study from five major donor locations.

Greece
Greek MS patients and controls were recruited from the MS center of the University hospital of Larissa, Greece. Patients were included in the study if they had clinically definite MS and were ≥18 years of age. The control group consisted of healthy volunteers from the same geographical region and were matched with cases for age, gender and ethnicity.

**Italy**

Italian MS cases were recruited through the PROGEMUS (PROgnostic GEnetic factors in MUltiple Sclerosis) consortium (a cross-sectional multicenter study including all patients affected by MS, both at first diagnosis and already diagnosed, consecutively examined in 10 MS Centers mainly from Northern Italy and the Hospital San Raffaele (HSR). Healthy controls are geographically matched with MS patients and of continental Italian ancestry. They include medical and university staff and healthy donors.

**Sweden**

Patients were recruited through three population-based case-control studies of incident (Epidemiological Investigation of Risk Factors for MS, EIMS), prevalent (Genes and Environment in MS, GEMS) MS patients and a post-marketing study of novel immunomodulatory drugs in Sweden (IMSE) as well as a local collection of MS cases from the neurology clinic at Karolinska University hospital. The EIMS study (2005-2012) inclusion criteria were: age 16-70 years, diagnosed MS according to the McDonald criteria within two years, and ability to understand the Swedish language. GEMS study participants were identified from the Swedish National MS registry, fulfilled the McDonald criteria, and were recruited during 2009-2011. The IMSE study participants were requited between 2006 and 2012. For the EIMS, GEMS and IMSE studies, controls were randomly chosen from the population register and matched to cases by sex, age at inclusion in the study, and region of residence. Two controls were matched to each case in the EIMS and IMSE study and one control per case in the GEMS study.

**UK**

The UK cases were recruited through a nationwide network of MS centers supported by the National Institute for Health Research. Controls were obtained from three sources, the Cambridge Bioresource (http://www.cambridgebioresource.group.cam.ac.uk), the Oxford Bioresource (http://oxfordbioresource.org) and index individuals from the Twins UK study (http://www.twinsuk.ac.uk).
In the US, lead investigators at several sites recruited participants. At UCSF cases and controls were recruited from multi-disciplinary clinics, remote sample collections, and field trips. MS patients are located by reference to MS registries, by physician referrals, and by advertisements in relevant websites, newsletters and functions describing the research projects and enrolment criteria. We also implemented an electronic intake survey to facilitate recruitment of research subjects. In Miami majority of study participants were enrolled through the MS Center for Excellence, established within the University of Miami Health System. Additional participants were recruited from the local multiple sclerosis community through outreach events and support group meetings. Diagnosis of multiple sclerosis was confirmed via medical chart review by a board-certified neurologist. Control subjects were recruited through the same ascertainment mechanisms and included unaffected individuals from the local community as well as spouses and friends of affected participants. In Boston participants were recruited over a period of 3 years from 10 different MS centers in the United States. All subjects met the 2010 McDonald criteria for MS and were of self-declared non-hispanic white ancestry. Control subjects came from three different studies of aging - the Chicago Health and Aging Project, The Religious Order Study, and the Memory and Aging Project - and the Harvard NeuroDiscovery Center biosample repository.

**Genotyping calling and quality check**

Overall, we genotyped 44,083 samples. We used Illumina’s Genome Studio for genotype calling, along with manual correction of the clusters for ~10,000 priority SNPs. We excluded any individual in the 10th percentile of GenCall scores and less than 97% call rate. Next, we performed a fast-pass QC and estimated the genetic relationship of any sample in the MS Chip data against the samples in the discovery phase. We removed any overlapping samples ($\pi^2 \geq 0.3$) from the MS Chip data, since these were already analyzed in the discovery phase.

We initially organized the MS Chip in data sets based on the country of origin: Australia, Denmark, Italy, Greece, Sweden, UK, US. Then we applied a similar, to the discovery phase, QC pipeline:

- Remove SNP with missingness ($\geq 1\%$)
- Remove individuals with missingness ($\geq 5\%$)
- Remove SNPs with p-value <0.001 for differential missingness
- Remove SNPs not in HWE in controls only; MAF > 0.3: p-value<0.0001; MAF>0.2 & MAF<=0.3 then p-value <0.00001; MAF>0.1 & MAF<=0.2: p-value <0.000001; MAF<0.1: p-value <0.000001
- Remove individuals with inbreeding coefficient F>=0.05 or F<= -0.05 (based on LD-prune SNPs of the custom content)
- Remove individuals with a pi^ (proportion of identity-by-descent) >=0.1 (between 3rd and 4th degree of relatedness)

Next, we calculated PCs of the respective data set and also projected to the 1000 Genomes Population (Phase I v3). The MS Chip is a custom designed array to replicate MS hits. During the design process (see above) we included 3 AIM panels to enable us to efficiently controls for stratification. Furthermore, the exome core set of SNPs provided another source of SNPs that we could potentially leverage for the PCA. We performed a series of tests in which we used different approaches to select the SNPs for PCA. We also created a series of artificial data sets that included a mixture of Danish, Swedish, and UK samples. We decided to apply a strategy that used 2 different sets of SNPs to calculate PCs: i) the LD-pruned AIMS, ii) LD-pruned exonic variants that passed QC and had a MAF>1%.

During the principal component analysis of the US and Swedish cohort we observed that each of them had 2 distinct clusters (Supplementary Figures 11 and 12). We separated manually the 2 respective clusters and verified that these represented 2 separate subpopulations (Askenazi Jews in the US and Middle Eastern ancestry individuals in the Swedish cohort). We decided to separate these subpopulations and create 2 US and Swedish data sets, and we re-ran QC for each of these independently. Supplementary Table 4 summarizes the MS Chip cases/controls counts for the final 9 data sets.

**Analysis**

To analyze the non-MHC autosomal genome we applied a logistic regression model using the first 5 PCs as covariates per data set, in a similar fashion as the discovery phase data sets. Although the MS Chip had a sufficient number of independent SNPs to calculate PCs that differentiate subpopulations, it was still designed to be an array with SNPs selected to have high prior of association with MS, thus having the expectation of inflated statistic. The exome content
of the MS Chip was biased towards low-frequency variants and resulted in a heavily deflated lambda (<0.5). Selection of the exome variants with MAF>1% (~20-50% of the overall exonic) resulted in inflation (lambda >1) in all data sets, in some cases similar to the lambda of the custom content. It was unclear whether there was a reliable set of SNPs to estimate the correct genomic inflation factor, hence we chose not to apply genomic control to the per-data set association results of the MS Chip data. Instead, we meta-analyzed the per data sets natural logarithm of the ORs with the respective uncorrected standard errors.

**Replication data sets: ImmunoChip**

We have previously genotyped and analyzed 14,498 cases and 24,091 controls using the ImmunoChip platform that were organized in 11 data sets (3). To add another large-scale replication cohort to our discovery set results we decided to re-use these data. The ImmunoChip samples have already been checked for overlap and relatedness with our discovery samples, however the MS Chip data could have some overlapping individuals. We repeated the same approach as the comparison of the MS Chip against the GWAS data to remove any ImmunoChip samples with pi^ >=0.3. To differentiate this version of ImmunoChip from the published one we refer it as ICv2 throughout the text. In cases where the published version of the ImmunoChip was used, i.e. when MS Chip had no tagging SNP for the prioritized effect, we denote it by using the term ICv1.

**Cohorts**

The ImmunoChip samples have been described in detail previously (3). Supplementary Table 5 lists the case/control counts per data set for the ICv2 and the originally published ICv1.

**Quality check and Analysis**

We re-QCed the ICv2 data sets using the same pipeline as above. We re-estimated PCs and we analyzed the data per data set using a logistic regression. We used the first 5 PCs to control for population stratification. We meta-analyzed across the 11 data with a fixed effects meta-analysis model, correcting the standard error of the natural logarithm of the OR for the respective genomic inflation factor (see analysis of discovery phase data for details). Although ImmunoChip was also a targeted array we have previously shown that the estimate lambda can be used for genomic control (3). This approach it is rather conservative, and it warrants that any replicated effect will have a high posterior probability to be a true effect.
Analyses on non-MHC autosomal genome

Joint-analysis of discovery and replication collections

In order to replicate the discovery phase prioritized effects, we jointly analyzed the discovery results with the two replication cohorts. We used a fixed effects inverse variance meta-analysis model with genomic control corrected estimates for the discovery phase and the ImmunoChip results. In order to replicate the prioritized effects of the discovery phase we followed the same strategy, i.e. we re-created the same regions in the replication sets and we applied the same stepwise procedure. Within each region we used the following strategy to choose the best SNP in the replication sets’ stepwise models:

i) meta-analyze marginal effects across the discovery, MS Chip, and ImmunoChip.

ii) Identify the most statistically significant SNP within the set of SNPs that have a r2>=0.1 with the effect SNP. We used 0.1 as the r2 threshold to allow for haplotypic effects that could be due to lower frequency variants that were not imputed well.

iii) Use the most statistically SNP from the previous step to use in a conditional analysis of the replication sets. Replace the effect SNP with this newly selected SNP for the discovery set as well. This resulted in less statistically SNPs to be used in the discovery set, reducing potential winner’s curse phenomenon in the discovery set.

iv) Repeat steps i) to iii) until the last step.

v) Manually inspect the selection of SNPs in the stepwise models for effects that reached genome-side significance or had a p-value less than 1x10^{-5}.

The above algorithmic process aims to imitate a procedure of targeted replication via genotyping of a single effect, however in our case we ended up with few thousand prioritized effects to replicate. We opted out from using imputation on the MS Chip and the ImmunoChip data to avoid any potential imputation errors or bias that the discovery phase data had. For both the MS Chip and ImmunoChip the genotypic data were used, minimizing the potential errors to genotypic calling and QC. Given that the genotyped replication collections are expected to be of higher quality than the imputed discovery set we allowed the replacement of the effect SNP by tagging SNPs. The replacement of the effect SNP but these tagging SNPs resulted in less statistically significant summary statistics from the imputed discovery phase to be include in the joint analysis. A direct result of this was the penalization of the imputed discovery set, compared
to the genotyped replication collections. Finally, the purpose of these tagging SNPs was not to perfectly tag the effect SNP that was identified in the imputed discovery set, acknowledging the winner’s curse of the discovery phase and the imputation noise, but rather to tag the underlying haplotype that carries the true causal variant. Thus, we allowed a liberal threshold of $r^2 > 0.1$. This also enabled the analysis of tagging variants that were present in both or at least one of the replication cohorts, whereas the effect SNP was not present. We provided detailed information of this process for all effects in the respective supplementary tables and an illustrative example in Supplementary Table 7 and Supplementary Figure 3.

**Sex chromosomes QC and analysis**

The sex chromosomes, X and Y, were processed and analyzed separately. We kept the same samples that survived QC for the autosomal genome and then we created sex chromosome specific files. We removed from either X and Y specific files any genetic variant that had:

- missingness $\geq 0.01$
- $p$-value for differentially missingness $< 1 \times 10^{-4}$
- HWE $p$-value $< 1 \times 10^{-4}$ (separately for males and females)

We also removed any individuals that had missingness for chromosome X or Y variants more than 1%. We applied these QC steps to all data sets that had available chromosome X or Y data. Specifically, for the MS Chip chromosome X and Y data we performed a manual visual inspection and correction of the cluster plots during genotypic calling with GenomeStudio. The ImmunoChip samples did not have readily usable data for either chromosome X or Y.

In the discovery phase samples data, we leveraged imputation to maximize the numbers of common SNPs to analyze in chromosome X. Post-imputation we compared the MAF and $p$-values of association for all the genotyped SNPs. We observed that the correlation of the $– \log_{10}(p$-value) of the genotyped SNPs before and after imputation varied substantially in all the WTCCC2 derived data sets (Pearson correlation coefficient from 0.048 to 0.24 for WTCCC2 data sets vs. 0.994 to 0.999 for all other GWAS data sets; Supplementary Table 27). We decided to exclude all WTCCC2 data sets from any chromosome X and Y analyses, assuming that there was an error in the respective genotypic calling process of sex chromosomes.

We fitted two separate logistic regressions, one for each gender, to analyze the imputed chromosome X SNPs of the discovery phase data (n=8), using the 5 top PCs as covariates. We used a dosage compensation model to account for chromosome X inactivation (49). This model
assigns chromosome X SNPs in males as either 0 or 2, i.e. treats males as homozygotes for with
the minor or major allele. Then we meta-analyzed the males and females results per data set to
estimate the common effect of chromosome X variants in both sexes. The MS Chip data had
extensive coverage of chromosome X variants and we analyzed the genotyped SNPs in a similar
fashion per data set (n=9).
Chromosome X has a much different LD structure compared to the rest of the genome, thus we
could not apply the same strategy we employed for the autosomal non-MHC genome. Instead we
performed a joint meta-analysis of all the data sets of the discovery phase (n=8) and the MS Chip
(n=9). We applied genomic control on the discovery phase data as described above. We reported
as genome-wide significant any chromosome X variant that reached a p-value < 5x10^{-8} in the
joint meta-analysis.
We performed chromosome Y analysis per data set. No single SNP had a p-value < 0.05, thus we
did not report any associations.

**MHC analysis**

The major histocompatibility complex (MHC) region harbors the strongest association with MS,
HLA-DRB1*15:01, and we have already reported that this region contains multiple independent
associations with MS, including HLA and non-HLA effects (17). This region has several loci
with long range LD, thus we decided to study this region with a different approach, fitting one
model with all the data. First, we imputed HLA alleles, respective amino acids, and SNPs in the
region using SNP2HLA (47) per data set. We imputed the discovery data sets (n=15) and all data
sets in the two replication sets (ICv2, n=11; MS Chip, n=9). Supplementary figure 1 displays
the imputation quality across all data sets. Then we ran a logistic regression model using the
post-imputation allelic dosages jointly for all data sets (47,429 cases and 68,374 controls),
including 5 first PCs and dummy variables for the data sets. We applied a forward stepwise
approach to identify independent effects, selecting the genetic variant in each step whose model
had the lowest Bayesian Information Criterion (BIC). We allowed the algorithm to include as
many steps until no genetic variant had a p-value less than the genome-wide significance
threshold (p-value < 5x10^{-8}).
Recently, we replicated a statistical interaction between HLA-DRB1*15:01 and HLA-
DQA*01:01 (16) and we aimed to refine this finding and explore how it affected the discovery of
downstream effects in the forward stepwise algorithm. Thus, we ran two parallel sensitivity
models: i) in the marginal model and each next iterative step we included an interaction term of
the analyzed variant with HLA-DRB1*15:01; we selected only marginal effects for inclusion in
each step, ii) same as the previous model but we included in the model the first interaction with
HLA-DRB1*15:01, if the respective model resulted in lowest BIC. The first model simply traces
the interaction of HLA-DRB1*15:01 with any other analyzed variant in MHC as the forward
stepwise algorithm identifies statistically independent effects. The second model includes the
most statistically significant interaction of HLA*DRB1*15:01 and traces how the downstream
selection of statistically independent effects is affected.

**Heritability analysis**

We estimated the narrow sense heritability ($h^2_g$) in the discovery data sets only. The 15
discovery data sets were the only analyzed data that had true genome-wide coverage, thus all
heritability estimations were based on the discovery phase only. To allow for a larger overlap
across the 15 data sets of the discovery phase and ensure that the prioritized effect SNPs, or their
proxies, were included in all analyses, we leveraged the post-imputation data. We created
genotype data from the post-imputation probabilities, assigning as the true genotype any of the
post-imputation probabilities that was larger than 50%. Then, we applied a set of filters to the
newly created post-imputation genotyped data: i) SNP missingness (> 0.001), differential
missingness (p < 0.05), HWE (p < 1x10^{-5}), and MAF (< 0.01). Next, we used LDAK (50) to
estimate SNPs that had zero weight in heritability contribution. We removed these SNPs and we
ended up with 421,512 autosomal SNPs overall, for a combined data set of 14,802 cases and
26,703 controls.

We used three different tools to estimate the narrow sense heritability in the whole genome data:
i) GCTA (20), ii) LDAK (50), iii) BOLT-LMM (51). All three tools calculated similar values of
$h^2_g$, hence we used GCTA for further analyses driven only by convenience to run all models.
For all GCTA estimations we used the data set of 421,512 SNPs and 41,505 samples described
above. GCTA’s model does not account for local LD, hence in a sensitivity analyses we ld-
pruned the data (--indep-pairwise 100 2 0.9), resulted in 263,789 genome-wide autosomal SNPs.
The pruned $h^2_g$ (19.17%; 95%CI: 18.49% - 19.85%) was not different than then non-pruned
estimate (19.06%; 95%CI: 18.38% - 19.74%), hence we used the non-pruned estimates for all
downstream analyses.
For the LDAK estimates we run two rounds to estimate the SNP weights for the heritability analysis, as per suggested best practices (50). For the BOLT-LMM analysis we used the data set with the 421,512 SNPs as well. In all 3 tools we set prevalence to be 0.001 and we included dummy variable for the included 15 data sets of the discovery phase and the respective 5 first principal components.

We created variance components (genetic relationship matrices; grms) for:

1. Regions: SNPs within the regions created by the stepwise approach (see above for details).
2. No-regions: SNPs outside the regions; all these SNPs had a p-value of > 0.05 by design.
3. Super extended MHC (SE MHC): SNPs that fall within the super extended MHC (from 24Mbps to 35Mbps; hg19).
4. Prioritized: SNPs that were in LD ($r^2 >= 0.1$) with any of the prioritized effect SNPs (see above for details).
5. Non-prioritized: SNPs within the region that were not in LD ($r^2 < 0.1$) with any of the prioritized effect SNPs.
6. Genome-wide (GW): SNPs in LD ($r^2 >= 0.1$) with any of the identified genome-wide effect SNPs
7. Suggestive (S): SNPs in LD ($r^2 >= 0.1$) with any of the identified suggestive effect SNPs
8. Non-replicated (NR): SNPs in LD ($r^2 >= 0.1$) with any of the non-replicated SNPs
9. No data (ND): SNPs in LD ($r^2 >= 0.1$) with any of the effect SNPs that had no replication data in both the MS Chip and the ImmunoChip replication collections.

We fitted the following models in GCTA, besides the overall genome:

1. Regions + SE MHC + no-regions
2. Prioritized + non-prioritized + SE MHC + no-regions
3. GW + S + NR + ND + non-prioritized + SE MHC + no-regions

For all analyses with more than one grm we report the percent of $h^2_g$ that a given variance component explains, using as the denominator the $h^2_g$ of the respective model.

**Enrichment analyses**

We used SNPsea (23) to test for tissue and cell specific enrichment of the genome-wide autosomal SNPs (GW) and the suggestive (S), non-replicated (NR) and no-data (ND) set of
SNPs. We leveraged two reference expression maps: i) the Genomics Institute of the Novartis Research Foundation (GNF) tissue atlas (expression profiles of 79 human tissues and cells types) (52) and ii) the Immunological Genome Project (ImmGen) (expression profiles of 223 sorted cells from immunological tissues and blood obtained from mice) (53).

Furthermore, we performed enrichment analysis using DNase I hypersensitivity data. We obtained processed DNase I hypersensitivity (BED format) sequencing reads for 350 Roadmap Epigenomics Project (REP) samples (54, 55) corresponding to 73 cell types from http://www.genboree.org/EdaccData/Current-Release/experiment-sample/Chromatin\_Accessibility. For each sample, we called 150bp DNase I hypersensitive sites (DHS) passing a 1% FDR threshold (56). We found 56 tissues with at least two replicates and limited our analysis to these. Where more than two replicates were available, we chose the two replicates with the smallest Jaccard distance between their DHS peaks positions on the genome. For each tissue, we obtained the union of DHS peaks for the two replicates to form the consensus of the regions of DNase I hypersensitivity active in each tissue. We used Fisher’s exact test to determine if any set of SNPs are enriched in DHS regions active in each tissue. For each tissue, we compared the proportion of SNPs in the list that were located on DHS regions to the proportion of all SNPs located on DHS regions active in that tissue.

**Additional CNS Cell type enrichment analysis**

For these analyses, we used RNA sequence data generated from induced pluripotent stem cells (iPSC) differentiated into neurons and purified human astrocytes. (57) In short, iPSCs were maintained in media containing 400mL DMEM/F12, 100mL Knockout Serum Replacement, 5mL pencillin/streptomycin/glutamine, 5mL MEM-NEAA, and 500uL 2-mercaptoethanol (all from Invitrogen) with fresh addition of 10ug/mL bFGF (Millipore). The induced neuron (iN) differentiation was accomplished as previously reported with minor modifications (58). iNs were plated at DIV4 on matrigel coated 96-well plates at a density of 25,000 cells/well and maintained in media consisting of 485mL neurobasal medium (Gibco), 5mL Glutamax, 7.5mL 20% Dextrose, 2.5mL MEM NEAA with 1:50 B27, BDNF, CNTF, GDNF and doxycline. Human cortical astrocytes were purchased from Sciencell and cultured in media containing 2% FBS, 1% penicillin/streptomycin, and 1% astrocyte growth supplement (Sciencell). These data are available at: https://www.synapse.org/#!Synapse:syn2580853/wiki/409844. The human microglia data are generated from 10 subjects of the Memory and Aging Project from a sample
of dorsolateral prefrontal cortex collected at the time of autopsy; these data are described in
detail elsewhere (59). In short, the fresh autopsy sample is mechanically dissociated, and live
CD11b+, CD45+ cells are selected for RNA sequencing by flow cytometry-based cell sorting.
The Smartseq2 protocol was used to generate libraries that were then sequenced. Full details are
available in the primary report of these data. For each cell type, we calculated the mean level of
expression for each gene from the available samples, and this average measure of gene
expression was used in downstream analyses for each cell type.

We compared the expression level of each ensembl gene that could be mapped to a protein in
uniprot in at least one of the three cell types to the level of expression observed in 541 subjects
with dorsolateral prefrontal cortex data, which contains the data used in the brain eQTL analyses
described above (60). The microglial data Genes with a four-fold increase in expression in the
cell type of interest vs. the cortex profile were deemed to be enriched for that cell type. We then
took the 551 genes that emerge from our gene prioritization efforts for the 201 MS susceptibility
SNPs. We assessed whether they were over-represented in these cell-type enriched gene sets
using a hypergeometric test. If we consider all genes that display a four-fold increase in
expression in microglia, the enrichment is significant, with a p-value of 0.00012. However, there
is a large number of MS genes that are expressed at very low level in microglia, and we therefore
impose a second threshold of FPKM>5 to evaluate the subset of genes that are robustly
expressed in human microglia, which yields the result reported in Figure 6D. The astrocyte and
neuronal analyses are non-significant both before and after imposing an expression cut-off in
addition to the fold-change threshold.

**Cis-expression quantitative trait locus (cis-eQTL) analyses**

For the T cell and Monocyte eQTLs used in this manuscript, we repurposed results from the
ImmVar project that have been described in detail in our earlier publication (26). In brief, we
limited our analyses to data generated from the 211 subjects of European ancestry; each subject
underwent a blood draw from which CD4+ naive T cells and CD14+ monocytes were purified by
flow cytometry-based cell sorting. The selected cells were collected in RNA lysis buffer, and
data were generated using an Affymetrix GeneChip Human Exon ST 1.0 array (the data is
available on the Gene Expression Omnibus website: series GSE56035). Associations between
SNP genotypes and adjusted expression values were conducted using Spearman Rank
Correlation. For the cis analyses that we use here, we considered only SNPs within a 1MB
window from the transcript start site (TSS) of genes. Significance of the nominal $p$-values was determined by comparing the distribution of the most significant $p$-values generated by permuting expression phenotypes 10,000 times independently for each gene. We call a cis-eQTL significant if the FDR<0.05.

The eQTL analysis of dorsolateral prefrontal cortex RNA sequencing profiles from two longitudinal cohort studies of aging (n=494) is described in detail elsewhere.(61) In brief, 494 European ancestry individuals from the ROS and MAP cohorts (Religious Orders Study and Memory and Aging Project) were used in the eQTL analysis. Post-QC genetic data (62) were imputed with the 1000 Genomes reference were imputed using version 3.3.2 version of the BEAGLE software18 (1000 Genomes Project Consortium interim phase I haplotypes, 2011 phase 1b data freeze). Imputed SNPs were filtered based on minor allele frequency (MAF) > 0.01 and imputation INFO score > 0.3, resulting in 7,321,515 SNPs available for analysis. Processed RNA-seq data (62) were used to estimate gene expression FPKM values using “rsem-calculate-expression” from RSEM.(63)

The COMBAT algorithm (64) was used with log-transformed FPKM (Fragments Per Kilobase of transcript per Million mapped reads) data to account for the effect of batch and remove the effects of RIN, postmortem interval (PMI), sequencing depth, study index (ROS sample or MAP sample), genotyping PCs, age at death and sex. The eQTL analyses were performed for 13,484 expressed genes and SNPs that were up to 1 Mb upstream or downstream of the TSS of each gene. Significance of the nominal $p$-values was determined by comparing the distribution of the most significant $p$-values generated by permuting expression phenotypes 10,000 times independently for each gene (65). We call a cis-eQTL significant if the FDR<0.05. Data and samples from the ROSMAP study are available on the RADC Research Resource Sharing Hub at http://www.radc.rush.edu/ and https://www.synapse.org/#!Synapse:syn3219045. RNA-seq data are available at http://dx.doi.org/10.7303/syn3388564

Finally, gene expression levels were quantified using mRNA derived from peripheral blood mononuclear cells (PBMCs) of 225 subjects of European ancestry with Relapsing Remitting (RR) Multiple Sclerosis (MS) using an Affymetrix Human Genome U133 Plus 2.0 Array. These data were collected between July 2002 and October 2007, as part of the Comprehensive Longitudinal Investigation of MS at the Brigham and Women’s Hospital. The expression levels were adjusted for confounding factors, such as subject’s use of immunomodulatory drugs, age,
gender, and batch effects, as well as 10 principal components (PCs), from the expression data, to account for unmeasured confounders. The data is available on the Gene Expression Omnibus website (GSE16214) (2). DNA from each individual was genotyped on the Affymetrix 550K GeneChip 6.0 platform as a part of a previously published study (2). The genotyped datasets were imputed as described above for the discovery set samples. Associations between SNP genotypes and adjusted expression values were conducted using Spearman Rank Correlation (SRC). For the cis analysis, we considered only SNPs within a 1MB window from the transcript start site (TSS) of genes. Significance of the nominal p-values was determined by comparing the distribution of the most significant p-values generated by permuting expression phenotypes 10,000 times independently for each gene (65). We call a cis-eQTL significant if the FDR<0.05.

**Analyses to identify cell-type specific eQTLs in brain RNAseq data**

We estimated the cell-specificity of an eQTL using a statistical model that tests for an interaction effect between the SNP dosage and the proportion of a cell type of interest (66). We applied this approach by estimating the proportions of neurons, astrocytes, microglia, oligodendrocytes, and endothelial cells based on known cell type markers for these cells. Specifically, ENO2 was used as the marker for neurons, CD68 for microglia, OLIG2 for oligodendrocytes, GFAP for astrocytes, and CD34 for endothelial cells. For each MS genome-wide significant lead SNP/brain eQTL pair available in the ROS/MAP brain data, we performed this interaction analysis in order to test the cell-specificity of lead SNPs that impacted the expression levels of a nearby gene. For this analysis, we transformed FPKM expression values of each gene to more normally distributed log2(gene+1) values, adjusting for technical factors such as RIN (RNA integrity number) score, total reads, post mortem interval, ribosomal bases, genotype chip, as well as age, sex, study, genotype principle components. A Bonferroni correction was implemented to assess the significance of the results.

**Gene prioritization**

Acknowledging that a single universal method to identify putatively causal genes that mediate the effect of identified genome-wide genetic variants does not exist, we implemented an ensemble of methods to prioritize potentially casual genes:

1) **Genes with exonic variants.** We used the 1000 Genomes European panel to identify SNPs in LD (r2>=0.1) with any of the 200 non-autosomal genome-wide lead SNPs and then we used
annovar (67) to identify genes that contain any of these SNPs. This step makes the assumption that the identified genome-wide effects could mediate their effect via changes in protein structure. Although it is well documented that the vast majority of the common variants associated with common diseases affect and alter gene regulation, the identification of low frequency variants in this study motivated us to include genes with exonic variants. The expectation is that a true causal exonic variant will have a lower allele frequency than the respective identified lead SNP, thus the r2 will be smaller even if these 2 variants are on the same haplotype. In sensitivity analyses we applied two different r2 thresholds: i) r2≥0.8, ii) r2≥ 0.5. The former assumes a strong posterior that the 200 genome-wide SNPs tag well the true casual variants, whereas the latter allows for larger uncertainty but is more conservative than the r2=0.1 threshold.

2) *Cis-eQTL genes*: We included any genes that has at least one *cis-eQTL* association (FDR<5%) with any of the 200 effect SNPs or SNPs in r2 ≥ 0.5 in the 4 eQTL studies described previously, i.e. PBMC of MS patients, CD4+ naïve T cells and monocytes from healthy controls, and brain tissue from an aging cohort. The assumption here is that the true causal SNP is mediating its effect via change of expression of nearby genes. The selected eQTL studies represent large-scale interrogations of immune system cell-specific populations, individuals with MS, and the target organ of inflammation. In sensitivity analyses we report the *cis-eQTL* genes using 2 different r2 thresholds: i) r2≥0.8, r2>0.1.

3) *Network approach*: We integrated thousands of statistically associated variants in associated genomic regions with regulatory information from ENCODE (68) and Roadmap Epigenomics Project (54) data to identify the most likely genes causative of the association signals in 10 different cell populations (T and B lymphocytes, monocytes, brain cells, etc). For each associated locus, an aggregate score including all available regulatory information was computed and a genes-by-cell matrix was generated. Overall, more than 5 million data points were used to conduct this analysis, which enabled prioritization of likely regulated genes in each region. Further integration of these with a highly curated human protein interaction network, allowed for the identification of specific disease-associated interactomes.

4) *DEPICT*: We used DEPICT (33), that uses predefined gene sets reconstituted using coexpression data, to prioritize genes that have similar predicted function.
In a secondary analysis we also included genes that had an intronic variant in LD (r²>0.5) with any of the 200 autosomal non-MHC SNPs. This makes the assumption that any of these intronic variants could affect the respective gene, e.g. via splicing. Finally, we applied a leave-one-out strategy to remove genes prioritized by only one method to explore the robustness of the downstream analyses and the relative influence of each method (see Supplementary Results for the results the secondary and sensitivity analyses).

**Pathway analyses**

We used the above gene lists to identify enriched known pathways. We used two different sources to this objective:

1) the Molecular Signatures Database (MSigDB), as it is available from the Gene Set Enrichment Analysis website (http://software.broadinstitute.org/gsea/msigdb). We run the Canonical Pathways, Biocarta, KEGG, and Reactome gene sets categories independently. We estimated statistical significance using the hypergenometric distribution and applied false discovery correction.

2) The Ingenuity Pathway Analysis (IPA) knowledge database. We used the following settings for the IPA analysis:
   a. Reference set: Ingenuity Knowledge Base (genes only)
   b. Species: Human

All results report p-value and false discovery rate.

**Protein-protein interaction analysis (GeNets)**

We used GeNets (https://apps.broadinstitute.org/genets) to leverage known protein-protein interactions (PPI) of our prioritized genome-wide genes. GeNets uses a random forest classified, trained in PPI data with 18 parameters that capture information about centrality and clustering. It creates *communities* of genes, sets of genes (nodes) that are connected to each other more than genes outside this community. Furthermore, it uses the random forest classifier and the connectivity to the tested gene set to propose candidate genes. For each described network the p-value is estimated by testing whether the number of observed edges divided by the numbers of possible edges using permutations.
Detailed eQTL results

Supplementary Table 53 provides information for the text files that contain detailed eQTL results.
Supplementary Results
Supplementary Figure 1 provides an overview of the study and the different type of analyses.

Autosomal non-MHC discovery results

The non-MHC autosomal genome was split into regions of 2Mbp in an iterative process starting from the most statistically significant effect and then moving to the second best and so forth. This process resulted in some regions to have smaller length, since the nearby genome was already included in already created regions. In the 1,961 regions that were created the median length was 1,447,260 bps (interquartile range: 912,286 to 1,998,939 (Supplementary Figure 14). The median number of SNPs within these regions was 4,274 (IQR: 2,287 to 5,916, Supplementary Figures 15 and 16). The part of the non-MHC autosomal genome that was not included in any of the above regions included 363,146 SNPs.

There were 1,961 lead SNPs prioritized for replication, one from each region, and another 2,881 SNPs prioritized via the within regions’ stepwise models. Most of the effects were identified either at step 0 (marginal) or step 1 (median: 1, IQR: 0 to 1), however the maximum number of steps within a region was up to 48 (49 including the lead SNP), in a region that included an effect identified in our previous study (3). P-values estimated from conditional models can fluctuate compared to the respective ones from the marginal model due to statistical artifacts besides true differences; hence we quantified the degree of that difference. The median difference in –log10 scale of the marginal minus the conditional p-values was -0.04859 (IQR: -0.26714 to 0.06011), which corresponds to a conditional p-value that is 0.89x more statistical significant than the marginal one, e.g. a marginal p-value of $1\times 10^{-4}$ will become $8.94\times 10^{-5}$ in the conditional model. The median value of the difference was similar when considering only the effects identified at step 1, i.e. when only the most statistically significant SNP in the region was also included in the model (-0.04387, IQ: -0.20684 to 0.04570). The majority of the conditional effects (88.6%) had a conditional p-value within a fold difference in either direction from the respective marginal p-value of the same SNP. Supplementary Figure 17 plots the histogram of the difference between the marginal and the conditional p-values at the –log10 scale, whereas Supplementary Figure 18 plots the scatterplot between the –log10 of these p-values.

The typical effect size, i.e. odds ratio (OR), of a genome-wide association variant in complex diseases is usually between 1.1-1.2. The median OR of included SNPs was 1.158 with an IQR of 1.080 to 1.414. Next, we used annovar (67) to annotate the known genetic function of the above-
prioritized effects, acknowledging that these identified SNPs are not expected to be the true causal variants and that most of the will be false positive. The vast majority (53.3%, 2,583 out of 4,842) was mapped to intergenic regions, with the next most frequent category being intronic (37%, 1,793 out of 4,842; Supplementary Figure 19). Exonic SNPs represent less than 1% (43 out of 4,842) of the prioritized effects.

We examined the overall imputation quality of the prioritized effect SNPs (Supplementary Figure 20). We observed different levels of imputation quality across data sets and prioritized SNPs. The mean imputation quality of the prioritized SNPs followed a bimodal distribution (Hartigans' dip test p-value < 2.2x10^-16) with the 2 modes near the bad and perfect imputation quality scores (Supplementary Figure 21). About a third (36.96%) of the prioritized SNPs had a mean imputation quality across the 15 data sets >= 0.9, and almost half (47.84%) had a mean imputation quality less than 0.5. Given that we had available 2 large-scale replication cohorts with genotyped data, one of which (MS Chip) was designed to replicate these prioritized effects, we decided to: i) not impute any of the replication cohorts, as to avoid any noise from imputation uncertainty, ii) manually inspect regions and prioritized effects that were replicated and reached genome-wide significance in the joint analysis of discovery and replication cohorts.

**Autosomal non-MHC joint analysis results**

The discovery phase prioritized 4,842 effects for further replication in the 2 replications sets, the MS Chip and the ImmunoChip (IC). In order to replicate these effects we followed the same strategy as in the discovery phase, i.e. we re-created the same regions in the replication sets and we performed the same stepwise modeling. Within each region we used the following strategy to choose the best SNP in the replications sets’ stepwise models:

i) meta-analyze marginal effects across the discovery, MS Chip, and ImmunoChip.

ii) Identify the most statistically significant SNP within the set of SNPs that have a r2>=0.1 with the effect SNP. We used 0.1 as the r2 threshold to allow for haplotypic effects that can be due to lower frequency variants that were not that well imputed. Furthermore, this allowed for a manual investigation of a wider set of SNPs around the initially identified effect SNP in the discovery set.

iii) Use the most statistically significant SNP from the previous step to use in a conditional analysis of the replication sets. Replace the effect SNP with this newly selected SNP for the discovery set as well. This resulted in less statistically significant
SNPs to be used in the discovery set, reducing potential winner’s curse phenomenon in the discovery set.

iv) Repeat steps i) to iii) until the last step.

v) Manually inspect the selection of SNPs in the stepwise models for effects that reached genome-side significance or had a p-value less than 1x10^{-5}.

The above algorithmic process resulted in the “replacement” of the discovery-prioritized SNP in 2,844 out of the 4,842 effects (58.7%). The median r² and D’ for the SNPs that replaced a discovery-prioritized SNP in any of the 4,842 effects was 0.62 and 1 respectively (IQR: 0.296 to 0.896, and 0.923 to 1 respectively), whereas 288 of these (10.1%) were perfect proxies (r²=1 and D’=1). The median drop of statistical significance in terms of –log10(p-value) in the discovery phase between the effect SNP and the newly selected SNP for the joint analysis was 0.6550 (IQR: 0.2305 to 1.3530), representing a 0.22 times less statistically significant result, e.g. a p-value of 1x10^{-4} for the effect SNP would be 4.52x10^{-3} for the replacement SNP. The OR was also affected, with a median change of -0.000751 in the natural log scale (IQR: -0.035998 to 0.018826).

**Genome-wide effects (GW)**

The joint analysis of the discovery and 2 replication sets resulted in 198 effects with a p-value less than 5x10^{-8}. We manually inspected all these effects to identify any false positive replication effect and to examine whether these were indeed statistically independent.

We used the following criteria to declare an effect genome-wide significant:

1. Joint p-value <= 5x10^{-8}
2. Joint p-value < discovery p-value
3. At least one replication set’s, i.e. MS Chip and ImmunoChip, p-value should be < 0.05, and the direction of effect was consistent in both replication sets, if data were available for both of them.

We set criteria #2-3 due to the large-scale of both our replication sets.

We identified 2 effects, rs198398 (step 1 in Region 192) rs1801133 (marginal in Region 556) that were capturing the same effect. Rs1801133 has frequency of 0.349 for the A allele and rs198398 frequency of 0.086 for the T allele, and they have an r² of 0.1768 and D’ of 1 in the European panel of 1000 Genomes (Supplementary Table 28). The AT haplotype is capturing
both effects and has frequency of 0.086. Using the MS Chip data, to avoid imputation error from the discovery sets, we tested for independence of these 2 variants; the effect of rs198398 could be explained by rs1801133, thus we decided to merge these 2 under the same effect (Supplementary Table 29). For downstream analyses, e.g. gene prioritization, we treated them separately and then merged the results to a single effect. One effect, rs4361438, was genome-wide in the discovery set (OR= 0.4436, p-value= 3.62x10^{-11}) however the joint p-value, after replication, was still genome-wide significant but less statistically significant (p-value = 1.89x10^{-9}). In further inspection we observed that only the MS Chip had data to replicate this effect, however the p-value was not even nominal (OR = 0.75, p-value=0.1353, for tagging SNP chr4:157588252; r2=1, D'=1). Given the lack of replication for this effect we downgraded it to the non-replicated set of effects (see below). Another effect, rs802730, was genome-wide in the discovery set (p-value =3.24x10^{-10}), however the joint p-value, after replication, was still genome-wide significant but less statistically significant (p-value=1.08x10^{-9}). Part of this drop in p-value was due to the use of a tagging SNP for replication (rs802725, p-value= 9.23x10^{-10}), however the main reason was lack of replication in the ImmunoChip set (p-value= 0.1917). The MS Chip p-value was 0.002592 and both replication cohorts had the same direction of effect (OR>1). The imputation quality of both rs802730 and rs802725 was relatively good (Supplementary Table 30), thus we decided to keep this effect in the list of the genome-wide loci. Next, we compared the known list of genome-wide variants from the ImmunoChip project (3) to identify any effects that we might have missed in our replication. 107 out of the 110 genome-wide effects of the ImmunoChip and previously published studies were identified as genome-wide significant in our analysis. Of the remaining 3, rs2150702 (5) was not replicated even in the discovery data (p-value = 0.06279; Supplementary Table 31), although sub-setting to the common studies between our discovery set and the original publication there was evidence of a potentially significant effect (OR = 1.14, p-value = 2.50x10^{-5}; Supplementary Table 31). The MS Chip data for this variant corroborated the lack of replication (OR = 0.99, p-value = 0.4668). The second variant was rs1886700. This variant was captured by the effect chr16:68694818 (r2=0.732, D’=0.880) in the discovery data. This variant, chr16:68694818, was not present in the ImmunoChip data and in our joint analysis it was replaced by rs3114397 across the 3 data sets (Supplementary Table 32). Rs3114397 is in LD with both these variants (with chr16:68694818:
r2 = 0.709, D’ = 1, with rs1886700: r2 = 0.75, D’ = 1). None of these variants had a p-value < 0.05 in the MS Chip results, which resulted in loss of genome-wide significance (Supplementary Table 32). Despite the previously reported genome-wide status of this effect, we do not claim it as genome-wide significant but we list it in the list of suggestive effects (see below). The third one was rs17785991 (p-value= 4.20x10^{-8} in the ImmunoChip analysis) and could not be tagged to any of the 4,842 effects. This SNP was included in Region 731 that had 11 prioritized effects, however it was not in LD with any of these effects (Supplementary Table 33). The marginal p-value of rs17785991 in the discovery set was 5.917x10^{-3}, conditioning with the lead SNP of the region (rs61109982) the p-value was 6.881x10^{-3}, and conditioning on the next SNP (rs6020055) the p-value became 0.05631. The p-value remained >0.05 in all subsequent steps (Supplementary Table 34). Although rs17785991 is not in LD with any of the 11 prioritized effects, its significance could be explained by conditioning on rs6020055 (Step 1). Rs6020055 has a joint p-value of 5.118x10^{-8} and given the genome-wide status of rs17785991 in previous publications we declared rs6020055 to be genome-wide significant (Supplementary Table 7).

The comparison with the previously published effects motivated us to compare our findings with two newer studies that used independent cohorts to either replicate targeted ImmunoChip variants (3, 69) or performed new GWAS (6). The study from Lill et al (69) followed up our ImmunoChip paper (3) by genotyping 11 SNPs that had p-value < 1x10^{-6} but were not genome-wide significant in 10,796 cases and 10,793 controls unrelated with ImmunoChip samples. Seven of these reached genome-wide significance in the joint ImmunoChip and replication results (Supplementary Table 35). Four of these were already identified as genome-wide significant in our study. One, rs4686953 was not present in the discovery set or the MS Chip data, and it appears to not be in LD with any other known variant in 1000 Genomes data. This variant is intronic in the LPP gene and could potentially capture the genome-wide identified effect in the same gene, rs13066789 (rs13066789 and rs4686953 are 95,295 base pairs apart). The other two, rs6072343 and rs9808753, were explained in part by identified non-genome-wide effects. Rs6072343 was included in Region 174 that had 11 prioritized effects, with highest of r2=0.131 with rs4812492 that was prioritized in step 7 (Supplementary Table 36). The marginal p-value of rs6072343 in the discovery set was 8.91x10^{-4}, conditioning with the lead SNP of the region (rs3795131 the p-value was 0.02706, and conditioning on the next SNP (rs62208470) the p-value became 0.1098. The p-value
remained >0.05 in all subsequent steps (Supplementary Table 37). The imputation of some of the
effect SNPs in this region was relatively low (Supplementary Table 38). Rs4812492, which is in
some LD with rs6072343, had high INFO score in most of the studies, with the lowest value
being 0.7829 (Supplementary Table 37). Rs3795131 did not replicate in the MS Chip results (p-
value = 0.2864; Supplementary Table 39), whereas the association with rs62208470 (via
rs11698662 as tagging SNP; r² = 0.648, D’ = 1) was nominally significant (p-value = 0.00799).
None of the Region 174 effects were replicated in MS Chip and reached genome-wide
significance overall (Supplementary Table 39). Examination of rs6072343 in our data did
replicate at genome-wide level the Lill et al results (Supplementary Table 39). Given that we did
not have compelling evidence that the effect of rs6072343 is explained by any of the prioritized
effects in the region, we decided to claim rs6072343 as the most probable candidate and genome-
wide significant. To clearly mark this effect from the ones identified from the prioritized ones we
used the term “manual step” in the respective supporting files (Supplementary Table 7).
Rs9808753 was included in Region 274 that had 11 prioritized effects, with highest LD of
r²=0.088 and D’=0.842 with chr21:34584055 that was prioritized in step 2 (Supplementary
Table 40). The marginal p-value of rs9808753 in the discovery set was 4.81x10⁻⁴, conditioning
with the lead SNP of the region (rs71314176) the p-value was 7.55x10⁻³, and conditioning on the
next SNP (rs743309) the p-value became 7.14x10⁻³. The p-value remained <0.05 in all
subsequent steps, after conditioning with chr21:34584055 (Supplementary Table 41). The above
suggest that chr21:34584055 and rs9808753 could potentially capture the same true effect. The
imputation quality of chr21:34584055 was low, with the best data set having an INFO score of
0.45, whereas the imputation quality of rs9808753 was high (Supplementary Table 42). The
coverage of both the MS Chip and the ImmunoChip of the prioritized effects of region 274 was
extremely low, with only 4 effects have data in MS Chip and none in ImmunoChip
(Supplementary Table 43). Of the effects with replication data, only rs71314176 had nominally
significant replication p-value (0.0232) but including only 3 studies of the overall 9 in the MS
Chip. In contrast rs9808753 had data present in both replications sets and the joint p-value
reached genome-wide significance (p-value = 1.60 x 10⁻⁹; Supplementary Table 43). We decided
to claim rs9808753 as the most probable candidate and genome-wide significant, indicated as
“manual step” as well.
The study from Andlauer et al (6) performed a new GWAS on 3,934 cases and 8,455 controls of German ancestry (DE1) and a joint analysis with another 954 cases and 1,940 controls of German origin (DE2) that were also used in our study (Central European discovery data set). In their study they identified 4 novel genome-wide loci, of which 2 were also identified as genome-wide in our study, and a 3rd one, rs4925166, did not replicate in our data (OR=0.98, p-value=0.16; Supplementary Table 44). Rs4364506, an intronic variant in L3MBTL3, had a p-value of 0.01241 in our discovery set. It was included in Region 129 that had 4 prioritized effects. Rs4364506 could potentially tag the variant identified at step 1 of the stepwise analysis, chr6:130348257 (r2=0.025, D’ = 0.656; Supplementary Table 45). Conditioning on chr6:130348257 resulted in a p-value of 0.2362 for rs4364506 (Supplementary Table 46).

Rs4364506 is in weak LD (r2=0.34, D’=0.92) with the variant we used to replicate to replicate the chr6:130348257 effect, rs6928313. All 3 SNPs (chr6:130348257, rs6928313, and rs4364506) are intronic in L3MBTL3. Although the LD is extremely weak, we believe that rs4364506 and chr6:130348257 tag the same true causal variant. The effect captured by chr6:130348257 had a joint p-value of 2.59x10^{-6}, with the discovery (p-value= 4.17x10^{-3}) and MS Chip (1.87x10^{-4}) only contributing data. Given the genome-wide status of the rs4364506 variant in Andlauer et al and the consistent signal in both discovery and MS Chip data, we declared the effect in chr6:130348257 to be genome-wide significant, noting that the true casual effect is rather in weak LD with this variant. For consistency we used the joint p-value of 2.59x10^{-6} for all downstream analyses.

Overall, we declare 200 autosomal non-MHC genome-wide variants (Supplementary Table 7). The minor allele frequency (MAF) of these 200 variants ranged from 1.58% to 50.00% in the European panel of the 1000 Genome Project, with a median of 30.48%. About one out of 10 (n=23; 11.5%) had a MAF lower than 10% and 10 effects had MAF less than 5% (Supplementary Table 8 and Supplementary Figure 4). Of these 200, 43.5% were annotated as intergenic (compared to 53.3% of the 4,842 prioritized effects; p-value for equality of proportions = 7.78x10^{-3}), with intronic (41.5%) and non-coding RNA intronic (5%) been the two next most popular categories (Supplementary Figure 22).
**Suggestive effects (S)**

The vast majority of the prioritized effects neither reached the genome-wide significance threshold nor had sufficient external validation evidence. However, some of these effects were statistically replicated in the replication cohorts. To further characterize the non-genome-wide effects we identified the ones that could be considered as statistically replicated with the following criteria:

- The joint p-value was less than the discovery but not genome-wide. For this comparison we used the joint p-value for the replicated SNP and the discovery p-value of the respective prioritized effect, regardless if it was replaced in the replication process by another SNP, thus opting-in for a more conservative approach.
- The p-value in at least one of the replication collections is <0.05, but not necessary in both; thus, allowing for the possibility of failure to replicate a true effect.
- The ORs are in the same direction.

We identified 417 such effects and we further classified them into 2 categories:

i) **strongly suggestive**
   a. joint p-value less than $1 \times 10^{-5}$, but not genome-wide significant
   b. both replication sets p-value < 0.05

ii) **weakly suggestive; all other suggestive**

Strongly suggestive were 117 out of the 416 suggestive effects (28.3%; Supplementary Table 14). The MAF of these effects ranged from 0.66% to 49.87% in the European subpopulation of the 1000 Genomes project, with a median of 28.23% (Supplementary Table 47 and Supplementary Figure 23). Intergenic were 53 of the effects (44.9%), with intronic (41.5%) been the next most popular category (Supplementary Figure 24).

Weakly suggestive were the other 299 effects (Supplementary Table 14). Supplementary Table 48 and Supplementary Figure 25 provide information of the MAF and Supplementary Figure 26 the functional annotation of these effects.

**Non-replicated effects (NR)**

All the other effects that had available data in at least one of the 2 replications sets and did not belong in any of the previous categories were categorized as non-replicated (NR; Supplementary Table 49). Generally, these included effects that their joint p-value was less statistically...
significant than the discovery one. Overall, 3,695 effects were grouped into this category (Supplementary Table 49). It is important to note that few of these effects could be truly associated with MS and were falsely non-replicated in our study. Given our strict replication criteria (see above in Supplementary Methods) one of the reasons why a true effect could erroneously not replicate is winner’s curse in the discovery set. In cases that the effect size in the discovery set is overestimated, i.e. we observe an OR that much larger than the true one, and the replication sets capture the true, much smaller effect, then the resulting joint p-value will be less statistically significant in the majority of these cases. However, the vast majority of these are due to imputation and statistical artifacts.

Supplementary Table 50 provides the allele frequencies of these effects and Supplementary Figure 27 plots the MAFs. Supplementary Figure 28A illustrates the imputation quality. Supplementary Figure 29 displays the annotated function of these effects.

**Effects with no data (ND)**

532 prioritized effects had no data in either of the two replication sets. For these effects we cannot make any inference on replications status. Supplementary Table 51 lists these. The imputation quality of most of these variants was extremely low (Supplementary Figure 30). Allele frequencies of these effects are listed in Supplementary Table 52 and Supplementary Figure 31 displays the MAF distribution. The function of these effects is illustrated in Supplementary Figure 32.

**Major Histocompatibility Complex region joint analysis results**

Supplementary Figures 33-40 illustrate the imputation quality per data set and variant. Supplementary Table 11 lists the selected variant in each step of the classical approach. Supplementary Table 12 lists the best model that includes an interaction with DRB1*15:01 following the selection order of the classical approach. Supplementary Table 13 lists the best model of the alternative approach. Supplementary Figure 5 plots the rank of the model that includes the previously reported interaction between DRB1*15:01 and DQA1*01:01.
Supplementary Figures

Supplementary Figure 1. Brief study overview. On the left we present the study design and main findings of the genetic analyses. On the right we enumerate the different types of analyses conducted, at order of introduction in the Online version. The main results of the protein-protein interaction (PPI) analyses are presented in ST22-25. The list of putative MS genes that was utilized in the PPI analyses is available as a public gene set to allow the replication and further extension of our analyses.

**Study overview**

### Autosomal non-MHC genome

**Discovery phase**
- 14,802 cases and 26,703 controls grouped in 15 data sets (Supplementary Table 3)
- Genome partitioning of SNPs with p-value < 0.05 to 1,961 regions (Supplementary Figure 2)
  - Step-wise model within regions to identify statistically significant results: > 4,842 effects (Supplementary Table 6)

**Replication phase**

A. **MS Chip (newly genotyped)**
- 20,360 cases and 19,047 controls grouped in 9 data sets (Supplementary Table 4)
- Designed to replicate the 4,842 effects

B. **ImmunoChip** (re-analysis of available data)
- 12,267 cases and 22,625 controls grouped in 11 data sets (Supplementary Table 5)

**Joint analysis** (Supplementary Tables 7, 14, 49, 51)
- 280 statistically independent effects with p-value < 5e-8 (GW)
- 416 suggestive independent effects (SR)
- 3,654 non-replicated effects (NR)
- 332 effects with no replication data (ND)

### Heritability analysis

- We can explain ~48% of the narrow sense heritability (Figure 4 in Online version)

### Cell type and tissue enrichment

- Enrichment of several immune cells but not brain tissue (Figure 5; Supplementary Tables 15; Supplementary Figure 7)

### Gene prioritization

- 551 putative MS genes
- Ensemble approach (Supplementary Table 18)
  - Cis-eQTL CD4+ T cells (n=415 healthy controls)
  - Cis-eQTL CD14+ monocytes cells (n=415 healthy controls)
  - Cis-eQTL PBMCs (n=225 multiple sclerosis subjects)
  - Cis-eQTL Brain tissue (n=494, aging cohorts)
  - Network approach of regulatory data (RegulomeDB)
  - Depict
  - Exonic variants

### Pathway analyses

(Supplementary Tables 20-21)

### Protein-protein interaction networks

(Supplementary Tables 22-25; Supplementary Figures 9-10)
Available online: [https://apps.broadinstitute.org/genets](https://apps.broadinstitute.org/genets), public gene set “The Multiple Sclerosis Genomic Map Genome-wide variants”

---

Chromosome X analysis: 1 genome-wide effect (Supplementary Table 9)

MHC imputation and analysis: 32 genome-wide independent effects (Supplementary Table 11-12)
Supplementary Figure 2. Genome partitioning approach to identify statistically independent autosomal non-MHC effects.
**Discovery phase**

**Marginal analyses**
Most statistically significant SNP is identified and a 2Mbp region around it is removed. The process is repeated until no SNP with p-value < 0.05 is left in the non-MHC autosomal genome. [meta-analysis of 15 GWAS data sets]

**Within region conditional analyses**
For each region the most statistically significant SNP is included as a covariate in the model and all SNPs in the regions are analyzed. The most statistically significant SNP is identified and the process is repeated until no SNP within the region has p-value < 0.01. [meta-analysis of 15 GWAS data sets]

End of discovery phase: 1,961 Regions created

End of stepwise approach: 4,842 effect SNPs within the 1,961 regions prioritized
Supplementary Figure 3. Illustration of the process to select the tagging SNP (r²>0.1) per prioritized effect SNP during the joint analysis of the discovery and replication phases.

Supplementary Figure 4. Minor allele frequency of non-replicated autosomal non-MHC effects.
Minor allele frequency distribution of autosomal non-MCH effects with no replication data.
Supplementary Figure 5. Rank of model with DRB1*15:01 and DQA1*01:01 interaction term. Rank is based on BIC compared to models with interactions terms of all other variants with DRB1*15:01.

![Graph showing rank of model with HLA-DRB1*15:01 and HLA-DQA1*01:01 interaction in different steps.](image-url)
Supplementary Figure 6. Rank of models including interaction term of plotted variants with DRB1*15:01. The steps in the X axis represent the steps of the classical model (Supplementary Table 11). Each of the plotted variants ranks at least once as the best model, in terms of Bayesian Information Criterion (BIC).
Supplementary Figure 7. Tissue and cell type enrichment analyses of the suggestive (S) and non-replicated (NR) lists of variants. Gene Atlas tissues and cell types gene expression enrichment is displayed. Rows are sorted from immune cells/tissues to central nervous system related ones. X axis displays $-\log_{10}$ of Benjamini & Hochberg p-values (false discovery rate).
Supplementary Figure 8. Dissection of cortical RNAseq data in astrocytes and neurons. (A) Here we plot the level of expression, transcriptome-wide, for each measured gene in our cortical RNAseq dataset (n=455)(y-axis) and purified human astrocytes (x-axis). In blue, we highlight those genes with > 4 fold increased expression in astrocytes relative to bulk cortical tissue and are expressed at a reasonable level in astrocytes. Each dot is one gene. Gray dots denote the 551 putative MS genes from our integrated analysis. SLC12A5 and CLECL1 are highlighted in red. There is no enrichment of MS genes among the astrocyte-enriched genes. (B) We repeat these analyses using RNA profiles derived from human induced pluripotent stem cell-derived neurons. There is no enrichment of MS genes among the neuron-enriched genes.
Supplementary Figure 9. Prediction of candidate genes via protein-protein interaction maps for the previously known list of MS-associated loci. The protein-protein interaction network depicts prior MS susceptibility genes with node degree ≥ 1 (purple color nodes, n=109). Interactions (graph edges), communities (i.e. groups of nodes sharing functional characteristics, n=9) and candidate genes (i.e. genes identified based on connectivity to MS susceptibility genes, coral and yellow color nodes, n=70) were identified using GeNets. Candidate genes prioritized in the newly identified list of MS susceptibility loci are marked with yellow (n=5). Connection widths between modules are analogous to the number of edges with nodes belonging to distinct communities. Note: the communities that include the predicted candidate genes are different than the communities drawn only by the protein-protein interactions between the MS genes.
Supplementary Figure 10. Prediction of candidate genes via protein-protein interaction maps for the newly identified list of MS-associated loci. Protein-protein interaction map depicting MS susceptibility genes and predicted candidate genes with node degree $\geq 1$, prioritized from the list of 200 autosomal non-MHC genome-wide loci (purple color nodes, n=222). Interactions (graph edges), communities (modules, n=10), and candidate genes (coral and yellow color nodes, n=102), were identified using GeNets. Candidate genes from Supplementary Figure 5 that were part of the prioritized MS genes are marked with yellow (n=5). The candidate genes that belong to the suggestive list of loci are marked with olive green. Connection widths between modules are analogous to the number of edges with nodes belonging to distinct communities. Note: the communities that include the predicted candidate genes are different than the communities drawn only by the protein-protein interactions between the MS genes.
Supplementary Figure 11. PCA plots of US data of MS Chip. There is a second cluster on the PC scatterplots, illustrated with a circle.
Supplementary Figure 12. PCA plot of Swedish data of MS Chip. There is a second cluster on the PC1 vs. PC2 scatterplot, illustrated with a circle.
Supplementary Figure 13. Imputation quality of the analyzed MHC variants. The scale displays INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 14. Histogram of the length of the 1,961 regions of the non-MHC autosomal genome. The units in the X axis are base pairs (bps).
Supplementary Figure 15. Histogram of the number of the SNPs in the regions of the non-MHC autosomal genome. The units in the X axis are number of SNPs.
Supplementary Figure 16. Scatterplot of the number of the SNPs vs. the length of the regions of the non-MHC autosomal genome. The units in the Y axis are number of SNPs and in X axis base pairs.
Supplementary Figure 17. Histogram of difference between \(-\log_{10}(P)\) of marginal model vs. \(-\log_{10}(P)\) of the conditional model.
Supplementary Figure 18. Scatterplot of $-\log_{10}(P)$ of marginal model vs. $-\log_{10}(P)$ of the conditional model.
Supplementary Figure 19. Frequency of the annotated function of the 4,842 prioritized effects.
Supplementary Figure 20. Imputation quality (INFO score) across the 15 discovery data sets for the 4,842 prioritized effect SNPs.
Supplementary Figure 21. Mean imputation quality (INFO score) for the 4,842 prioritized SNPs.
Supplementary Figure 22. Frequency of the annotated function of the 200 genome-wide autosomal non-MHC effects.

Frequency of function for the 200 autosomal non-MCH genome-wide effects.

Number of SNPs

<table>
<thead>
<tr>
<th>Function</th>
<th>Number of SNPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>UTR3</td>
<td>3</td>
</tr>
<tr>
<td>UTR5</td>
<td>4</td>
</tr>
<tr>
<td>downstream</td>
<td>3</td>
</tr>
<tr>
<td>exonic</td>
<td>5</td>
</tr>
<tr>
<td>intergenic</td>
<td>87</td>
</tr>
<tr>
<td>intronic</td>
<td>83</td>
</tr>
<tr>
<td>ncRNA_exonic</td>
<td>1</td>
</tr>
<tr>
<td>ncRNA_intronic</td>
<td>10</td>
</tr>
<tr>
<td>upstream</td>
<td>4</td>
</tr>
</tbody>
</table>
Supplementary Figure 23. Minor allele frequency of the strongly suggestive autosomal non-MHC effects.
Supplementary Figure 24. Frequency of the annotated function of the strongly suggestive autosomal non-MHC effects.
Supplementary Figure 25. Minor allele frequency of the weakly suggestive autosomal non-MHC effects.
Supplementary Figure 26. Frequency of the annotated function of the weakly suggestive autosomal non-MHC effects.
Supplementary Figure 27. Minor allele frequency of non-replicated autosomal non-MHC effects.
Supplementary Figure 28. Imputation quality of non-replicated (A), genome-wide (B), and suggestive (C) autosomal non-MHC effects.
Supplementary Figure 29. Frequency of the annotated function of the non-replicated autosomal non-MHC effects.
Supplementary Figure 30. Imputation quality of the prioritized effects with no replication data.
Supplementary Figure 31. Minor allele frequency of non-replicated autosomal non-MHC effects.
**Supplementary Figure 32.** Frequency of function of the autosomal non-MHC effects with no replication data.
Supplementary Figure 33. Imputation quality of HLA-A variants. Values represent INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 34. Imputation quality of HLA-B variants. Values represent INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 35. Imputation quality of HLA-C variants. Values represent INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 36. Imputation quality of HLA-DPA1 variants. Values represent INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 37. Imputation quality of HLA-DPB1 variants. Values represent INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 38. Imputation quality of HLA-DQA1 variants. Values represent INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 39. Imputation quality of HLA-DQB1 variants. Values represent INFO score. Metav3.0 refers to the discovery set.
Supplementary Figure 40. Imputation quality of HLA-DRB1 variants. Values represent INFO score. Metav3.0 refers to the discovery set.
ANZgene Membership


1Hunter Medical Research Institute, University of Newcastle NSW Australia. 2Westmead Institute for Medical Research, University of Sydney NSW Australia. 3Brain and Mind Centre, University of Sydney NSW Australia. 4Canterbury District Health Board, Christchurch, New Zealand. 5Griffith Institute of Health and Medical Research, Griffith University, Gold Coast QLD Australia. 6School of Medicine, Griffith University, Gold Coast QLD Australia. 7Bond University, Gold Coast QLD Australia. 8James Cook University, Townsville, QLD Australia. 9Flinders University, SA Australia. 10Menzies Research Institute Tasmania, University of Tasmania Australia. 11University of Melbourne, VIC Australia. 12Monash University, VIC Australia. 13Sir Charles Gairdner Hospital, Perth, WA Australia.

International IBD Genetics Consortium (IIBDGC) contributing members

Murray Barclay4, Laurent Peyrin-Biroulet2, Mathias Chamaillard4, Jean-Frederick Colombe4, Mario Cottone4, Anthony Croft4, Renata D'Incà7, Jonas Halfvarson8,9, Katherine Hanigan6, Paul Henderson10,11, Jean-Pierre Hugot12,13, Amir Karban14, Nicholas A Kennedy15, Mohammed Azam Khan16, Marc Lémann17, Arie Levine18, Duncan Massey19, Monica Milla20, Grant W Montgomery21, Sok Meng Evelyn Ng22, Ioannis Oikonomou22, Harald Peeters23, Deborah D. Proctor22, Jean-Francois Rahier24, Rebecca Roberts7, Paul Rutgeerts25, Frank Seibold26, Laura Stronati27, Kirstin M Taylor28, Leif Törkvist29, Kullak Ublick30, Johan Van Limbergen31, Andre Van Gossum32, Morten H. Vatn33, Hu Zhang20, Wei Zhang22, Australia and New Zealand IBDGC*, Belgium Genetic Consortium†, Initiative on Crohn and Colitis, NIDDK IBDGC‡, United Kingdom IBDGC, Wellcome Trust Case Control Consortium§

1Department of Medicine, University of Otago, Christchurch, New Zealand. 2Gastroenterology Unit, INSERM U954, Nancy University and Hospital, France. 3INSERM, U1019, Lille, France. 4Univ Lille Nord de France, CHU Lille and Lille-2 University, Gastroenterology Unit, France. 5Division of Internal Medicine, Villa Sofia-V. Cervello Hospital, University of Palermo, Palermo, Italy. 6Inflammatory Bowel Diseases, Genetic Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia. 7Department of Surgical and Gastroenterological Sciences, University of Padua, Padua, Italy. 8Department of Medicine, Örebro University Hospital, Örebro, Sweden. 9School of Health and Medical Sciences, Örebro University, Örebro, Sweden. 10Royal Hospital for Sick Children, Pediatric Gastroenterology and Nutrition, Edinburgh, UK. 11Child Life and Health, University of Edinburgh, Edinburgh, UK. 12INSERM U843, Paris, France. 13Univ-Paris Diderot Sorbonne Paris-Cité, Paris France. 14Department of Gastroenterology, Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel. 15Gastrointestinal Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. 16Genetic Medicine, MAHSC, University of Manchester, Manchester, UK. 17Université Paris Diderot, GETAID group, Paris, France. 18Pediatric Gastroenterology Unit, Wolfson Medical Center and Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel. 19Department of Medicine, National University Hospital, Singapore. 20National University of Singapore, Singapore. 21Institut Pasteur, Paris, France. 22Genetic Medicine, Monash University, Melbourne, Australia. 23Menzies Research Institute Tasmania, University of Tasmania, Hobart, Tasmania, Australia. 24Istituto di Genetica Medica, Università di Pisa, Pisa, Italy. 25School of Biomedical Sciences, Murdoch University, Perth, WA Australia. 26Department of Pediatrics, Saint Louis University, Saint Louis, Missouri, United States. 27Department of Gastroenterology and Nutrition, University of Western Australia, Perth, WA Australia. 28University of Paris 8, Paris, France. 29University Hospital of Uppsala, Uppsala, Sweden. 30Meningitis and Autoimmune Disease Research Group, Royal Adelaide Hospital, Adelaide, Australia. 31University of New South Wales, Sydney, Australia. 32Division of Gastroenterology, North York General Hospital, Toronto, Ontario, Canada. 33University of Edinburgh, Edinburgh, Scotland.
Membership of Wellcome Trust Case Control Consortium 2 (WTCCC2)

Management Committee

Peter Donnelly (Chair)1,2, Ines Barroso (Deputy Chair)3, Jenefer M Blackwell4,5, Elvira Bramon6, Matthew A. Brown7, Juan P Casas8, Aiden Corvin9, Panos Deloukas3, Audrey Duncanson10, Janusz Jankowski11, Hugh S. Markus12, Christopher G. Mathew13, Colin N. A. Palmer14, Robert Plomin15, Anna Rautanen1, Stephen J. Sawcer16, Richard C. Trembath13, Ananth C. Viswanathan17, Nicholas W. Wood18

Data and Analysis Group

Chris C. A. Spencer1, Gavin Band1, Céline Bellenguez1, Colin Freeman1, Garrett Hellenthal1, Eleni Giannoulatou1, Matti Pirinen1, Richard Pearson1, Amy Strange1, Zhan Su1, Damjan Vukcevic1, Peter Donnelly1,2

DNA, Genotyping, Data QC and Informatics Group


Publications Committee

Christopher G. Mathew (Chair)13, Jenefer M Blackwell4,5, Matthew A Brown7, Aiden Corvin9, Chris C. A. Spencer1

1Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. 2Department of Statistics, University of Oxford, Oxford OX1 3TG, UK. 3Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK. 4Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, 100 Roberts Road, Subiaco, Western Australia 6008. 5Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge CB2 0XY, UK. 6Department of Psychosis Studies, NIHR Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King’s College London and The South London and Maudsley NHS Foundation Trust, Denmark Hill, London SE5 8AF, UK. 7University of Queensland Diamantina Institute, Brisbane, Queensland, Australia. 8Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine,
London WC1E 7HT and Department of Epidemiology and Public Health, University College London WC1E 6BT, UK. 9Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Eire. 10Molecular and Physiological Sciences, The Wellcome Trust, London NW1 2BE. 11Department of Oncology, Old Road Campus, University of Oxford, Oxford OX3 7DQ, UK, Digestive Diseases Centre, Leicester Royal Infirmary, Leicester LE7 7HH, UK and Centre for Digestive Diseases, Queen Mary University of London, London E1 2AD, UK. 12Clinical Neurosciences, St George's University of London, London SW17 0RE. 13King’s College London Department of Medical and Molecular Genetics, King’s Health Partners, Guy’s Hospital, London SE1 9RT, UK. 14Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK. 15King’s College London Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK. 16University of Cambridge Department of Clinical Neurosciences, Addenbrooke’s Hospital, Cambridge CB2 0QQ, UK. 17NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London EC1V 2PD, UK. 18Department of Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK.
References and Notes


9. Materials and methods are available as supplementary materials.


