Effect of *HLA-DRB1* alleles and genetic variants on the development of neutralizing antibodies to interferon beta in the BEYOND and BENEFIT trials

D. Buck[1,2,*], T.F.M. Andlauer[2,3,4,*], W. Igl[5], E.M. Wicklein[5], M. Mühlau[1,2], F. Weber[2,3,6], K. Köchert[5], C. Pohl[5,7,†], B. Arnason[8], G. Comi[9], S. Cook[10], M. Filippi[11], H.-P. Hartung[12], D. Jeffery[13], L. Kappos[14], F. Barkhof[15], G. Edan[16], M.S. Freedman[17], X Montalbán[18], B. Müller-Myhsok[2,3,4,19], and B. Hemmer[1,2,4], The BEYOND and BENEFIT Study Groups

[1] Department of Neurology, Klinikum rechts der Isar, Technische Universität München, Germany.

[2] German Competence Network Multiple Sclerosis (KKNMS), Germany.

[3] Max Planck Institute of Psychiatry, Munich, Germany.

[4] Munich Cluster for Systems Neurology (SyNergy), Germany.

[5] Bayer AG, Berlin, Germany.

[6] Neurological Clinic, Medical Park, Bad Camberg, Germany.

[7] Department of Neurology, University Hospital of Bonn, Bonn, Germany.

[8] Department of Neurology, University of Chicago Surgery Brain Research Institutes, Chicago, IL, United States.

[9] Department of Neurology and Institute of Experimental Neurology, Università Vita-Salute San Raffaele, Milan, Italy.

[10] Rutgers, The State University of New Jersey, Newark, NJ, United States.

[11] Neuroimaging Research Unit, Institute of Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute and Vita-Salute San Raffaele University, Milan, Italy.

[12] Department of Neurology, Heinrich-Heine University, Düsseldorf, Germany.

[13] Piedmont Health Care, Mooresville, NC, United States.

[14] University Hospital Basel, Basel, Switzerland.

[15] Radiology & Nuclear Medicine, VU University Medical Centre, Amsterdam, The Netherlands, and UCL Institutes of Neurology and Healthcare Engineering, London, United Kingdom.

[16] University of Rennes, Rennes, France.

[17] University of Ottawa and The Ottawa Hospital Research Institute, Canada.

[18] Department of Clinical Neuroimmunology, Hospital Vall d'Hebron, Barcelona, Spain.

[19] Institute of Translational Medicine, University of Liverpool, United Kingdom

* These authors contributed equally to the manuscript.

† Deceased

Corresponding Author:

Bernhard Hemmer

Postal address: Klinik und Poliklinik für Neurologie, Klinikum rechts der Isar, Ismaninger Str. 22, 81675 München

Phone: 0049 (0)89-4140-4601

Fax: 0049 (0)89-4140-7681

E-mail: hemmer@tum.de

Keywords:

multiple sclerosis, interferon beta, anti-drug antibodies, genetic variation, HLA-DRB1, genome-wide association study

Effect of *HLA-DRB1* alleles and genetic variants on the development of neutralizing antibodies to interferon beta in the BEYOND and BENEFIT trials

Abstract

Background: Treatment of multiple sclerosis with interferon β can lead to the development of antibodies directed against interferon β that interfere with treatment efficacy. Several observational studies have proposed different *HLA* alleles and genetic variants associated with the development of antibodies against interferon β.

Objective: To validate the proposed genetic markers and to identify new markers.

Methods: Associations of genetic candidate markers with antibody presence and development were examined in a post hoc analysis in 941 patients treated with interferon β-1b in the BEYOND and BENEFIT prospective phase III trials. All patients were treated with interferon β-1b for at least six months. In addition, a genome-wide association study was conducted to identify new genetic variants.

Results: We confirmed an increased risk for carriers of *HLA-DRB1*04:01* (OR=3.3, p=6.9×10^{-4}) and *HLA-DRB1*07:01* (OR=1.8, p=3.5×10^{-3}) for developing neutralizing antibodies (NAbs). Several additional, previously proposed *HLA* alleles and genetic variants showed nominally significant associations. In the exploratory analysis, variants in the *HLA* region were associated with NAb development at genome-wide significance (OR=2.6, p=2.30×10^{-15}).

Conclusion: The contribution of *HLA* alleles and *HLA*-associated SNPs to the development and titer of antibodies against interferon β was confirmed in the combined analysis of two multi-national, multi-center studies.

Introduction

Biopharmaceuticals provide new opportunities for treatment of severe diseases. In a subset of patients, however, these protein-based drugs induce an unintended immune response against the biopharmaceutical[1-3]. The development of antibodies against biopharmaceutical drugs (anti-drug antibodies, ADA) may thus be a major limitation of treatment efficacy[4,5]. Interferon β-preparations are widely used to treat multiple sclerosis (MS) or clinically isolated syndrome (CIS). Overall, up to 40% of patients will eventually develop antibodies against interferon β[6-9], of which a significant proportion neutralize the biological activity of interferon β (neutralizing antibodies, NAbs)[10,11]. Several factors related to either the interferon β compound and administration or the individual patient influence immunogenicity[6,12-14].

In two previous studies on German MS patients, *HLA* class II alleles were found to influence the development of ADA against interferon β. *HLA-DRB1*04:01* and *HLA-DRB1*04:08* showed significant, *HLA-DRB1*03:01*, *HLA-DRB1*04:04*, *HLA-DRB1*11:01*, *HLA-DRB1*11:04*, and *HLA-DRB1*16:01* nominal associations with ADA titer[15,16]. In addition, an association of *HLA-DRB1*07:01* with ADAs was described for 39 American patients[17]. A study conducted on 610 Spanish patients reported an increased risk of developing NAbs for the combined presence of the haplotype *HLA-*

*DRB1*07:01* with the *HLA* class I alleles *HLA-A*26* and *HLA-B*14* and replicated the association of *DRB1*04:01*[18]. In a Swedish analysis, the *HLA-DRB1*15* and *HLA-DQA1*05* alleles were associated with an increased risk for developing NAbs in patients treated with interferon β-1a[19], while *HLA-DRB1*04* was nominally associated in patients receiving interferon β-1b. In a genome-wide association study (GWAS) on German patients, two single nucleotide polymorphisms (SNPs), rs9272105 (linked to the *HLA* class II locus) and rs4961252 (linked to an intergenic region on chromosome 8), were significantly associated with ADA titers[20].

The aim of the current study was to validate these genetic markers in the setting of two multi-national phase III trials conducted on patients treated with interferon β-1b and carefully monitored for the occurrence of ADA.

Methods

Patient cohorts

Patients were treated with interferon β-1b in two phase III trials. These studies were the BEtaseron®/BEtaferon® in Newly Emerging multiple sclerosis For Initial Treatment trial (BENEFIT; clinicaltrials.gov ID: NCT00185211)[21] and the Betaferon® Efficacy Yielding Outcomes of a New Dose trial (BEYOND; clinicaltrials.gov ID: NCT00099502)[22]. In BENEFIT, 478 CIS patients were randomized to interferon β-1b 250 µg, administered subcutaneously every other day, or placebo (5:3 ratio). Patients received placebo for two years or until MS was diagnosed using the Poser criteria[23]. Patients were then offered interferon-β-1b for up to five years. In BEYOND, 2244 patients with relapsing-

remitting MS (RRMS) were randomized in a 2:2:1 ratio to receive either interferon β-1b 500 µg, interferon β-1b 250 µg, both administered subcutaneously every other day, or glatiramer acetate 20 mg, administered subcutaneously daily for a period of 2 or up to 3.5 years. Details of the study protocols and results of NAb testing were published elsewhere[22,24-26].

In the analyses presented in this manuscript, we included patients who had given informed consent for genetic testing and who were treated with interferon β-1b for at least six months with a dose of either 250 µg (n=567) or 500 µg (n=374). This dose was included as a covariate in the analyses. Because ADA development usually occurs within the first year of treatment[27], genetic association analyses were restricted to patients who had either received interferon β-1b for at least six months and showed a positive NAb result or who had continuously received interferon β-1b for at least twelve months, irrespective of NAb status. Patients with missing information in clinical covariates as well as genetic outliers were excluded, leading to a final number of 941 patients included in the analyses after quality control (QC).

Measurement of antibodies against interferon β

In both clinical trials, serum samples were collected at baseline and every six months after that to evaluate anti-interferon β-1b antibodies using the Myxovirus protein A (MxA) induction assay[28]. Measurement of anti-interferon β-1b antibodies was conducted in the same laboratory for both clinical trials (Rentschler Biotechnologie GmbH, Laupheim, Germany). A titer of 20 normalized units (NU)/mL was used as a cut-off between NAb negativity and positivity. Patients without a NAb-positive sample and with either a treatment

interruption of more than 90 days in the first treatment year or with a missing NAb measurement at year one and later were excluded from analyses.

Study Endpoints

The predefined primary endpoint (EP) was the presence of NAbs at any time during the studies (EP1). This endpoint was analyzed in all eligible patients. Secondary endpoints were mean NAb titer (EP2), the area under the NAb curve (NAb AUC) of four measurements performed during the first two treatment years (EP3), the maximum NAb titer during the first two years (EP4), and time to first NAb presence (EP5). EP5 was assessed in NAb-positive patients only. Because the duration of the BENEFIT trial was longer than the BEYOND trial, EP3 and EP4 potentially differed systematically between the studies. To prevent bias in these endpoints, EP3 and EP4 were only assessed during the first two years of either study. In addition to the examination of candidate variants, a hypothesis-free GWAS was conducted for each endpoint as an explorative analysis.

Sequencing, genotyping, and imputation

HLA-DRB1 alleles were sequenced by LGC-Genomics (formerly AGOWA) using the *HLA-DRB1* AlleleSEQR/HARP kits from Abbott Molecular. Genotyping was conducted using the Affymetrix Genome-Wide Human SNP Array 6.0. QC steps on samples included removal of individuals with a genotyping rate <98% and of genetic outliers using EIGENSTRAT[29] (for details see Supplementary Table e-1). 926 of the 941 remaining samples were of Caucasian ancestry (Supplementary Table e-2 and Supplementary Figure e-1). QC steps on variants included removal of variants with a call rate <98%, a MAF <1%, or a Hardy-Weinberg equilibrium test p-value $<10^{-6}$.

Imputation of genotype data was conducted to the 1000 Genomes Phase 3 reference panel in 5 Mbp chunks using SHAPEIT v2.r837 for phasing and IMPUTE2 v2.3.2 for imputation[30-32]. Imputed variants were filtered for MAF (≥1%) and INFO metric (≥0.8). The final dataset contained 8,671,751 variants and 941 patients.

Statistical Analyses

For EP1, the presence of NAbs, logistic regression was used. Endpoints EP2-EP4 were analyzed using linear regression. A Cox proportional hazards model was used to analyze EP5 using the R package *survival*[33]. Endpoints EP2-EP4 were inverse-normally transformed to normal distributions before analysis. The following covariates were used in all analyses: Study (BENEFIT / BEYOND), sex, age, interferon β-1b treatment dose (250 µg / 500 µg), interferon β-1b treatment duration, and the first six multidimensional scaling (MDS) components of the genetic similarity matrix to control for population substructure (Supplementary Figure e-1). For EP5, the study was included as a stratification variable to meet the proportional hazards assumption. Test statistics were not inflated in any model (genomic inflation $\lambda \leq 1.025$). For the analysis of candidate variants, the significance threshold α was Bonferroni-corrected for eleven tests ($\alpha=0.0045$). Because *HLA-DRB1*15* and *HLA-DRB1*15:01* are not independent of each other, they were not considered as separate alleles for calculation of this significance threshold. All analyses of candidate variants were conducted in *R* v3.2.1. The GWAS was conducted in PLINK2 v1.90b3.38 for EP1-EP4[34], and in *R* for EP5. The significance threshold for the GWAS was $\alpha=5\times10^{-8}$. In addition to the analyses in the complete set of patients, EP2-EP5 were analyzed separately using the same

methods and covariates within the subgroup of NAb-positive patients. The eQTL data was obtained from the Genotype-Tissue Expression (GTEx) project via the online GTEx Portal (http://www.gtexportal.org/) on December, 13th, 2017[35].

Results

A total of 941 patients from the BENEFIT and BEYOND trials were analyzed in this study. Of these, 361 patients developed NAbs during the trials (Table 1). All 941 patients had genotyping data available, while *HLA* sequencing was available for only 678 of them. Eleven genetic markers, at least nominally associated with ADA in previous studies, were selected as candidates: *HLA-DRB1*04:01* and *HLA-DRB1*04:08*, significantly associated with ADA titers in two previous publications[15,16], as well as *HLA-DRB1*03:01*, *HLA-DRB1*04:04*, *HLA-DRB1*11:01*, *HLA-DRB1*11:04*, and *HLA-DRB1*16:01*, nominally associated in either of these studies. Markers selected from additional studies were the haplotype *HLA-DRB1*15*, associated with NAb development[19], *HLA-DRB1*07:01*, associated with ADAs[17], and two SNPs, rs9272105 and rs4961252, associated with ADA titers[20].

Candidate genetic markers associated with the presence of NAbs

Patients with a NAb titer of ≥ 20 NU/mL in any measurement were defined as NAb-positive. Considering NAb-positive patients as cases and NAb-negative ones as controls, we conducted a logistic regression on NAb presence. Of the candidate markers, *HLA-DRB1*04:01* and *HLA-DRB1*07:01* were significantly associated with the presence of NAbs after correction for multiple testing

(eleven tests, α=0.0045, Table 2). In addition, *HLA-DRB1*11:01* and the *HLA*-associated SNP rs9272105 were nominally associated (one-sided *p*-values) and all candidate markers except for rs4961252, *HLA-DRB1*15*, and *HLA-DRB1*15:01* showed the direction of effect expected from the previous publications (Table 2).

Association of candidate genetic markers with secondary endpoints

We performed analyses for the three secondary endpoints mean NAb titer (EP2), NAb AUC (EP3), and maximum NAb titer (EP4) using linear regression on all patients. These three measurements were highly correlated (Spearman's correlation coefficients ≥ 0.98, Supplementary Table e-3). Accordingly, results for EP2 (Table 3) were highly similar to EP3 and EP4 (Supplementary Tables e-5 and e-6), with quantitative differences in *p*-values. Qualitatively, these results were also very similar to the analysis of EP1 (Table 2). In addition to the variants already nominally or significantly associated with EP1, *HLA-DRB1*04:08* was nominally associated with EP2 (Table 3).

Association of candidate genetic markers in NAb-positive patients

Similar to our primary endpoint and previous studies, the analysis of EP2-EP4 in all patients addressed the question whether candidate markers influence the risk of developing NAbs. However, genetic variants might also influence the titer of NAbs once a patient has already developed ADA. To examine this hypothesis, we repeated analyses in the subset of patients that developed NAbs within the first two years, excluding NAb-negative patients. The sample size for this analysis was only approximately one-third of the complete number of available patients (Supplementary Table e-4). In this subset of

patients, *HLA-DRB1*04:08*, *HLA-DRB1*07:01*, and variant rs9272105 were nominally associated with mean NAb titer, but no variant remained associated after correcting for multiple testing (Supplementary Tables e-7 to e-9).
EP5, time to NAb presence, was only analyzed in NAb-positive patients, using a Cox proportional hazard model. Under the hypothesis that risk alleles should decrease the time to the presence of NAbs, markers rs9272105 and *HLA-DRB1*15* were nominally associated, yet no association of candidate markers remained significant after correction for multiple testing (Supplementary Table e-10).

Explorative GWAS for all endpoints

The main objective of our study was the replication of candidate variants from previous publications. However, as we had genome-wide genotyping data available for all patients, we also conducted exploratory GWAS to identify novel variants associated with the development of ADA. For EP1, the SNP showing the overall highest association was rs522308 (OR 2.60 (CI 2.05-3.29), p=2.30×10^{-15}), located in between *HLA-DRB1* and *HLA-DQA1* (Figure 1A, Supplementary Figure e-2, Supplementary Table e-11). Variant rs522308 was in weak linkage disequilibrium (LD) with the candidate SNP rs9272105 (r^2=0.15) and with *HLA* allele *HLA-DRB1*07:01* (r^2=0.28). In the GWAS for EP2-4, the SNP rs2454138 showed the overall highest association (EP2: β=0.41±0.05, p=5.29×10^{-19}; Figure 1B, Supplementary Figures e-3 to e-5, Supplementary Table e-11). This variant is in strong LD with the upstream EP1-associated SNP rs522308 (r^2=0.88). Test statistics were not inflated: the median genomic inflation factor λ ranged from 1.014 to 1.025 for the different endpoints. Variant rs522308 is significantly associated with expression of

several *HLA* class II genes in whole blood in the public expression quantitative trait locus (eQTL) database GTEx (Supplementary Table e-12)[35]. Neither rs2454138 nor rs9272105 was part of any eQTL in this database. Finally, we also conducted GWAS on all endpoints in NAb-positive patients only. However, no variant reached genome-wide significance here (Supplementary Figures e-6 to e-9).

Discussion

Interferon β is a well-established first-line treatment in MS. Up to 40% of patients treated with interferon β develop ADA, most of them with neutralizing activity, that may antagonize the therapeutic activity of the drug[6]. Although the immunogenicity of biopharmaceuticals is a known critical phenomenon, reasons for the large inter-individual variation in the occurrence of ADA are still poorly understood. In previous studies, nine *HLA-DRB1* alleles and two SNPs have been proposed to either influence the risk to develop anti-interferon β antibodies or to be associated with ADA titers[15-20]. Six of these markers are assumed to increase the risk for ADA development, the remaining five are candidates for protection from ADA.

In the current study, these genetic markers were evaluated in an independent dataset of patients with CIS or RRMS from two large international phase III clinical trials, carefully monitored for ADA titer and persistence of ADA over time. We confirmed the associations of *HLA-DRB1*04:01* and *HLA-DRB1*07:01* with NAb presence (Table 2) and NAb titer (Table 3). Seven of the remaining candidate markers showed an association following the published direction of the effect but were not significant after correction for

multiple testing. For some of the candidate alleles, the power in our study was very low due to their low frequencies (e.g., previously proposed *HLA-DRB1*04:08* was present in only 2.1% of patients). Non-significant associations in the expected directions thus do not imply that these alleles are not associated with NAb presence. Analyses of the secondary EPs showed very similar results (Table 3, Supplementary Tables e-5 and e-6). Overall, quantitative mean NAb titer (EP2) was more strongly associated with candidate variants than binary NAb presence (EP1). This increased sensitivity indicates that the pre-defined threshold for NAb positivity (20 NU/mL) was not optimally chosen. Indeed, these results suggest that for future genetic association studies, mean antibody titers might be better suited than NAb presence defined by a fixed cut-off.

Discordant results with and among published associations likely arose from different allele frequencies in the investigated populations, different distributions of interferon β preparations in the study cohorts, and different assays for ADA measurement. In general, the heterogeneous geographical and thus genetic background of BENEFIT and BEYOND patients likely led to a decrease in power in our analysis (Supplementary Figure e-1). Another reason for observed discrepancies might lie in the assay used. In our current study, NAb titers were measured through an MxA induction *in vitro* assay. In a previous Swedish study, an MxA protein assay, as well as an MxA gene expression assay, were used[19]. In previous German publications, identification of genetic risk alleles was based on ADA titers determined by capture ELISA and NAbs assessed by *in vivo* MxA measurements[15,16,20].

In the study on Swedish patients, the *HLA-DRB1*15* allele was associated with NAb development in interferon β 1a-treated patients[19]. However, in two previous studies on German patients, not stratified for interferon β preparations[15,16], as well as in our current analysis, restricted to interferon β-1b, *HLA-DRB1*15* was not more common among patients developing ADA (Table 2). Of note, Link *et al.* had examined the haplotype *HLA-DRB1*15* only[19]. In addition, we also included the most frequent *HLA-DRB1*15* allele *HLA-DRB1*15:01* in our analyses, which was weakly correlated with the candidate SNP rs9272105 (r^2=0.30). Like *HLA-DRB1*15*, *HLA-DRB1*15:01* was neither more common among patients with NAbs nor associated with increased NAb titers (Tables 2 and 3). Interestingly, both *HLA-DRB1*15* and rs9272105 were associated at nominal significance with time to maximum NAb titer (EP5, Supplementary Table e-10).

In a previously published German cohort, the minor allele G of the imputed SNP rs4961252 was associated with higher ADA titers[20]. However, in our international cohort, the SNP was negatively associated with NAbs. The most likely explanation for this phenomenon is a flip-flop effect[36]. Such an effect can occur if a tested candidate marker is a non-causal common variant in weak LD with the unknown real causal variant. Under these circumstances, the effect direction of the candidate variant can reverse due to sampling variation, especially when examining multi-ethnic populations. This means that rs4961252 is likely not the causal variant, but that it is in weak LD with the actual causal SNP, which remains to be identified.

In addition to analyses on the complete set of patients, we also examined the subset of NAb-positive patients. The significant association of *HLA-*

*DRB1*04:01* in all patients was not supported in analyses of NAb-positive patients (Supplementary Tables e-7 to e-10). By contrast, the allele *HLA-DRB1*04:08* showed the strongest nominal association with EP2-4 in NAb-positive patients. It is, therefore, possible that *HLA-DRB1*04:01* is important for the risk to develop NAbs and that *HLA-DRB1*04:08* is relevant for an increased NAb titer.

Conclusion

In summary, our study confirms a genetic association of the *HLA* class II locus with the development of ADA against interferon β-1b. This association was confirmed in both analyses of candidate variants and, indirectly, in hypothesis-free GWAS. Our validation of the importance of the *HLA* class II locus fits well into the pathophysiologic concept because *HLA* class II proteins are crucial for antigen presentation to $CD4^+$ T cells that are necessary for inducing the B cell response and subsequent antibody production. In future studies, a larger dataset needs to be examined, including several subsets of in each case genetically homogeneous samples selected from different populations and different interferon β preparations. Importantly, all patients will need to be evaluated using the same assay, ideally both for ADA and for NAbs. These studies will help to elucidate the contribution of genetic variations to ADA development further. The ultimate aim of such analyses should be the establishment of prediction algorithms for ADA development.

References

1. van Schouwenburg PA, Rispens T, Wolbink GJ. Immunogenicity of anti-TNF biologic therapies for rheumatoid arthritis. *Nat Rev Rheumatol*; 9:

164–172.

2. Jawa V, Cousens LP, Awwad M, et al. T-cell dependent immunogenicity of protein therapeutics: Preclinical assessment and mitigation. *Clin Immunol*; 149: 534–555.

3. Wroblewska A, Reipert BM, Pratt KP, et al. Dangerous liaisons: how the immune system deals with factor VIII. *J Thromb Haemost*; 11: 47–55.

4. Sørensen PS, Ross C, Clemmesen KM, et al. Antibodies to IFN-beta: the Danish National IFN-beta Project. *Neurology*; 61: S27–8.

5. Polman CH, Bertolotto A, Deisenhammer F, et al. Recommendations for clinical use of data on neutralising antibodies to interferon-beta therapy in multiple sclerosis. *Lancet Neurol*; 9: 740–750.

6. Bertolotto A, Deisenhammer F, Gallo P, et al. Immunogenicity of interferon beta: differences among products. *Journal of Neurology*; 251 Suppl 2: II15–II24.

7. Deisenhammer F. Interferon-Beta: Neutralizing Antibodies, Binding Antibodies, Pharmacokinetics and Pharmacodynamics, and Clinical Outcomes. *J Interferon Cytokine Res*; 34: 938–945.

8. Sørensen PS, Ross C, Clemmesen KM, et al. Clinical importance of neutralising antibodies against interferon beta in patients with relapsing-remitting multiple sclerosis. *Lancet*; 362: 1184–1191.

9. Bertolotto A, Gilli F, Sala A, et al. Persistent neutralizing antibodies abolish the interferon beta bioavailability in MS patients. *Neurology*; 60: 634–639.

10. Kappos L, Clanet M, Sandberg-Wollheim M, et al. Neutralizing antibodies and efficacy of interferon beta-1a: a 4-year controlled study. *Neurology*; 65: 40–47.

11. Pachner AR, Cadavid D, Wolansky L, et al. Effect of anti-IFN{beta} antibodies on MRI lesions of MS patients in the BECOME study. *Neurology*; 73: 1485–1492.

12. Torosantucci R, Sharov VS, van Beers M, et al. Identification of oxidation sites and covalent cross-links in metal catalyzed oxidized interferon Beta-1a: potential implications for protein aggregation and immunogenicity. *Mol Pharm*; 10: 2311–2322.

13. Barnard JG, Babcock K, Carpenter JF. Characterization and quantitation of aggregates and particles in interferon-β products: potential links between product quality attributes and immunogenicity. *J Pharm Sci*; 102: 915–928.

14. Bozhinov A, Handzhiyski Y, Genov K, et al. Advanced glycation end products contribute to the immunogenicity of IFN-β pharmaceuticals. *J*

Allergy Clin Immunol; 129: 855–858.e6.

15. Buck D, Cepok S, Hoffmann S, et al. Influence of the HLA-DRB1 genotype on antibody development to interferon beta in multiple sclerosis. *Arch Neurol*; 68: 480–487.

16. Hoffmann S, Cepok S, Grummel V, et al. HLA-DRB1*0401 and HLA-DRB1*0408 are strongly associated with the development of antibodies against interferon-beta therapy in multiple sclerosis. *Am J Hum Genet*; 83: 219–227.

17. Barbosa MDFS, Vielmetter J, Chu S, et al. Clinical link between MHC class II haplotype and interferon-beta (IFN-beta) immunogenicity. *Clin Immunol*; 118: 42–50.

18. Núñez C, Cénit MC, Alvarez-Lafuente R, et al. HLA alleles as biomarkers of high-titre neutralising antibodies to interferon-β therapy in multiple sclerosis. *J Med Genet*; 51: 395–400.

19. Link J, Lundkvist Ryner M, Fink K, et al. Human leukocyte antigen genes and interferon beta preparations influence risk of developing neutralizing anti-drug antibodies in multiple sclerosis. *PLoS ONE*; 9: e90479.

20. Weber F, Cepok S, Wolf C, et al. Single-nucleotide polymorphisms in HLA- and non-HLA genes associated with the development of antibodies to interferon-β therapy in multiple sclerosis patients. *The Pharmacogenomics Journal*; 12: 238–245.

21. Kappos L, Polman CH, Freedman MS, et al. Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes. *Neurology*; 67: 1242–1249.

22. O'Connor P, Filippi M, Arnason B, et al. 250 µg or 500 µg interferon beta-1b versus 20 mg glatiramer acetate in relapsing-remitting multiple sclerosis: a prospective, randomised, multicentre study. *Lancet Neurol*; 8: 889–897.

23. Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol*; 13: 227–231.

24. Goodin DS, Hartung H-P, O'Connor P, et al. Neutralizing antibodies to interferon beta-1b multiple sclerosis: a clinico-radiographic paradox in the BEYOND trial. *Mult Scler*; 18: 181–195.

25. Hartung HP, Freedman MS, Polman CH, et al. Interferon β-1b-neutralizing antibodies 5 years after clinically isolated syndrome. *Neurology*; 77: 835–843.

26. Kappos L, Freedman MS, Polman CH, et al. Long-term effect of early treatment with interferon beta-1b after a first clinical event suggestive of

multiple sclerosis: 5-year active treatment extension of the phase 3 BENEFIT trial. *Lancet Neurol*; 8: 987–997.

27. Sorensen PS, Deisenhammer F, Duda P, et al. Guidelines on use of anti-IFN-beta antibody measurements in multiple sclerosis: report of an EFNS Task Force on IFN-beta antibodies in multiple sclerosis. *European journal of neurology*; 12: 817–827.

28. Pungor E, Files JG, Gabe JD, et al. A novel bioassay for the determination of neutralizing antibodies to IFN-beta1b. *J Interferon Cytokine Res*; 18: 1025–1030.

29. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*; 38: 904–909.

30. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*; 5: e1000529.

31. Howie B, Fuchsberger C, Stephens M, et al. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics*; 44: 955–959.

32. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods*; 10: 5–6.

33. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. Springer, 2010.

34. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*; 4: 7.

35. Melé M, Ferreira PG, Reverter F, et al. Human genomics. The human transcriptome across tissues and individuals. *Science*; 348: 660–665.

36. Lin P-I, Vance JM, Pericak-Vance MA, et al. No gene is an island: the flip-flop phenomenon. *Am J Hum Genet*; 80: 531–538.

Acknowledgements:

The study was supported by the German Competence Network Multiple Sclerosis (KKNMS). BH and DB were supported by the IMI project ABIRISK and the EU project MultipleMS.

Author Contributions:

DB: Conception and design, drafting and revising the manuscript. TFMA: Conception and design, data analysis, drafting and revising the manuscript. WI: Data analysis, revision of the manuscript. MM: Conception and design, revision of the manuscript. FW: Initiation, revision of the manuscript. EMW, KK: Analysis, revision of the manuscript. CP: Initiation, conception and design. BMM: Conception and design, revision of the manuscript. BH: Initiation, conception and design, drafting and revision of the manuscript. BA, GC, SC, MF, H-PH, DJ, LK, FB, GE, MSF, XM: Members of the study steering committee and study investigators.

Conflict of Interest Disclosures:

Dorothea Buck has received compensation for activities with Bayer HealthCare, Biogen Idec, Merck Serono, and Novartis. She is supported by the ABIRISK Consortium. **Wilmar Igl** was working as a contractor for Bayer Healthcare Pharmaceuticals, Berlin, Germany, between 2010/07 and 2012/06 and has been employed by AstraZeneca, Cambridge, UK, since 2015/05. **Mark Mühlau** has received research support from Merck Serono and Novartis. **Frank Weber** received honoraria from Genzyme, Novartis, Teva, and Biogen for speaking or for serving on a scientific advisory board; he received travel grants for the attention of a scientific meeting from Merck-Serono and Novartis and grant support from Merck Serono, Novartis, and from the Federal Ministry of Education and Research (BMBF, Projects Biobanking and Omics in ControlMS as part of the KKNMS). **Karl Köchert** and **Eva-Maria Wicklein** are salaried employees of Bayer AG, **Christoph**

Pohl was a salaried employee. **Barry Arnason** received grant support from Mallinckrodt Pharmaceuticals. In the last years, **Giancarlo Comi** received personal compensation for activities such as consulting, scientific advisory boards, or speaking from: Bayer HealthCare, Biogen-Dompè, Biogen Idec, Merck-Serono, Novartis, Roche, Sanofi-Aventis, Almirall, Teva, Actelion, Receptors. **Stuart Cook** has received grants from and held lectures with Merck Serono and Bayer HealthCare. **Massimo Filippi** is Editor-in-Chief of the Journal of Neurology; he serves on a scientific advisory board for Teva Pharmaceutical Industries; he has received compensation for consulting services and/or speaking activities from Biogen Idec, Merck Serono, Novartis, and Teva Pharmaceutical Industries; and he receives research support from Biogen Idec, Teva Pharmaceutical Industries, Novartis, the Italian Ministry of Health, Fondazione Italiana Sclerosi Multipla, Cure PSP, Alzheimer's Drug Discovery Foundation (ADDF), the Jacques and Gloria Gossweiler Foundation (Switzerland), and ARiSLA (Fondazione Italiana di Ricerca per la SLA). **Hans-Peter Hartung** received fees for consulting, serving on steering and data monitoring committees, and speaking from Bayer, Biogen, Geneuro, Genzyme, Medimmune, Merck, Novartis, Opexa, Receptos Celgene, Roche, and Teva, with approval by the rector of Heinrich-Heine University. **Douglas Jeffery** has received honoraria for speaking and consulting for Bayer HealthCare, Biogen, Teva, Serono, Pfizer, Novartis, Genzyme, Genentech, Acorda, and Mallinckrodt Pharmaceuticals. He has received research support from Bayer HealthCare, Biogen, Teva, Serono, Pfizer, Genentech, and Novartis. **Ludwig Kappos** does not have conflicts of interest relevant to the contents of this manuscript. **Frederik Barkhof** is a member of the editorial

board of Brain, European Radiology, Neurology, Multiple Sclerosis Journal, and Radiology. He received financial compensation as a consultant for Bayer HealthCare, Biogen Idec, Teva, Merck Serono, Novartis, Roche, Synthon BV, Jansen Research, and Genzyme. He received a grant from the Dutch MS Society (EU-FP7/H2020). He or his institution received payment for the development of educational presentations by Biogen Idec and IXICO. **Mark Freedman** received research or educational grants from Genzyme, honoraria or consultation fees from Actelion, Bayer HealthCare, Biogen Idec, Chugai, EMD Canada, Genzyme, Merck Serono, Novartis, Hoffman La-Roche, Sanofi-Aventis, and Teva Canada Innovation. He is a member of the company advisory board, the board of directors, or another similar group for Actelion, Bayer HealthCare, Biogen Idec, Hoffman La-Roche, Merck Serono, MedDay, Novartis, and Sanofi-Aventis, and he participates in the company-sponsored speaker's bureau of Genzyme. **Xavier Montalban** has received speaking honoraria and travel expenses for participation in scientific meetings, has been a steering committee member of clinical trials or participated in advisory boards of clinical trials in the past years with Actelion, Almirall, Bayer AG, Biogen, Celgene, Hoffmann-La Roche, Merck, Novartis, Oryzon Genomics, Sanofi-Genzyme and Teva Pharmaceutical. **Bernhard Hemmer** has served on scientific advisory boards for F. Hoffmann-La Roche Ltd, Novartis, Bayer AG, and Genentech; he has served as DMSC member for AllergyCare; he or his institution have received speaker honoraria from Biogen Idec, Teva Neuroscience, Merck Serono, Medimmune, Novartis, Desitin, and F. Hoffmann-La Roche Ltd; his institution has received research support from Chugai Pharmaceuticals and Hoffmann-La-Roche; he holds part of two

patents, one for the detection of antibodies and T cells against KIR4.1 in a subpopulation of MS patients and one for genetic determinants of neutralizing antibodies to interferon β. **Till Andlauer**, **Gilles Edan**, and **Bertram Müller-Myhsok:** Nothing to disclose.

Table 1: Characteristics of BEYOND and BENEFIT patients included in the analyses. Only a subset of patients had *HLA* data available. For all median values, the median absolute deviation is indicated in brackets. Additional characteristics of these patients as well as characteristics for the subset of patients that developed NAbs within the first two years are shown in Supplementary Tables e-3 and e-4, respectively.

Measurement	**With SNP data**	**With *HLA* data**
Number of samples	941	678
Study: BEYOND (n (%))	766 (81.4)	669 (98.7)
Sex: Female (n (%))	661 (70.2)	472 (69.6)
Median age (years)	35 (10.4)	36 (10.4)
Median disease duration (years)	2.3 (3.5)	3.3 (3.9)
Median EDSS	2 (1.5)	2 (0.7)
Median cumulative newly active lesions during the studies	1 (1.5)	1 (1.5)
Interferon β-1b dose: 250 µg (n (%))	567 (60.3)	352 (51.9)
Median duration of treatment (years)	2.6 (0.6)	2.6 (0.5)
Minimum duration of treatment with interferon β-1b (years)	0.5	1.5
Median total dose of interferon β-1b (mg)	182.5 (81)	151.8 (71.9)
Presence of NAbs at least during one measurement (%), EP1	361 (38.4)	252 (37.2)
Median mean NAb titer (NU/ml), EP2 (including NAb-negative patients)	0 (0)	0 (0)

Table 2: Summary statistics for the association analysis of the presence of NAbs (NAb titer ≥ 20 NU/mL) with candidate markers (logistic regression). The Bonferroni-corrected significance threshold was $\alpha=4.5\times10^{-3}$ (eleven tests, because the tests for *HLA-DRB1*15* and for *HLA-DRB1*15:01* were not independent of each other). Chr = chromosome, Freq = frequency, n = sample size, OR = odds ratio, CI = 95 % confidence interval, *p*-value (1-sided) = one-sided *p*-value for replication based on the direction of effect expected from the literature (the published effect). Alleles significant after multiple testing are labeled in bold font, nominally significant results in italics.

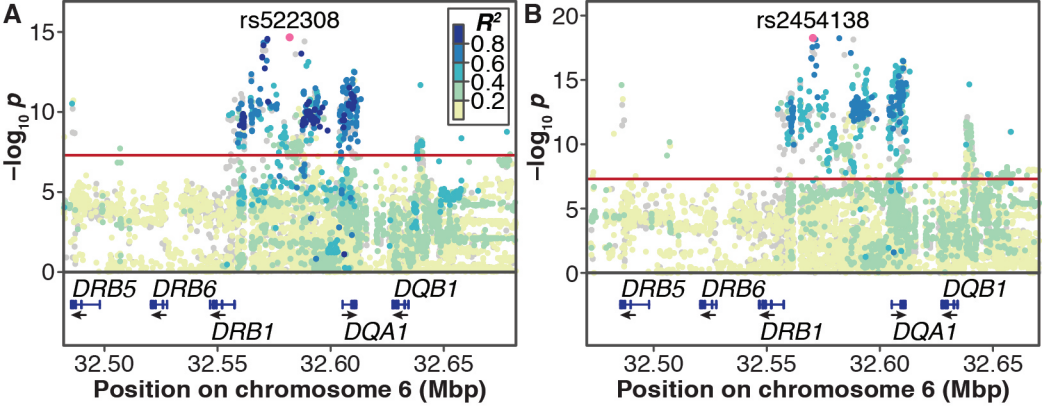
Chr.	Gene	Allele / Variant	Freq. (%)	n	OR (CI)	Published effect	*p*-value	*p*-value (1-sided)
6	*HLA-DRB1*	***04:01**	11.9	678	3.33 (1.66-6.67)	>1	**6.9×10^{-04}**	**3.5×10^{-04}**
6	*HLA-DRB1*	*04:08	2.1	678	2.15 (0.42-10.95)	>1	3.6×10^{-01}	1.8×10^{-01}
6	*HLA-DRB1*	***07:01**	9.5	678	1.78 (1.21-2.61)	>1	**3.5×10^{-03}**	**1.7×10^{-03}**
6	*HLA-DRB1*	*15	25.5	678	0.73 (0.56-0.96)	>1	2.6×10^{-02}	9.9×10^{-01}
6	*HLA-DRB1*	*15:01	24.0	678	0.75 (0.57-0.99)	>1	4.2×10^{-02}	9.8×10^{-01}
6	*HLA-DRB1*	*16:01	2.6	678	1.62 (0.6-4.37)	>1	3.4×10^{-01}	1.7×10^{-01}
6	*HLA-DRB1*	*03:01	2.9	678	0.87 (0.61-1.23)	<1	4.2×10^{-01}	2.1×10^{-01}
6	*HLA-DRB1*	*04:04	0.4	678	0.58 (0.25-1.34)	<1	2.0×10^{-01}	1.0×10^{-01}
6	*HLA-DRB1*	*11:01*	1.2	678	0.64 (0.4-1.02)	<1	6.3×10^{-02}	*3.1×10^{-02}*
6	*HLA-DRB1*	*11:04	6.9	678	0.66 (0.32-1.38)	<1	2.7×10^{-01}	1.3×10^{-01}
8	intergenic	rs4961252_G	40.6	941	0.92 (0.74-1.14)	>1	4.5×10^{-01}	7.7×10^{-01}
6	intergenic	*rs9272105_A*	49.0	941	0.82 (0.68-0.99)	<1	*4.2×10^{-02}*	*2.1×10^{-02}*

Table 3: Summary statistics for the association analysis of mean NAb titer with candidate markers (linear regression). Significance thresholds and abbreviations are as described in the legend of Table 2. SE = standard error (shown instead of CI because of the quantitative phenotype).

Chr.	**Gene**	**Allele / Variant**	**Freq. (%)**	**n**	**Effect**	**SE**	**Published effect**	***p*-value**	***p*-value (1-sided)**
6	*HLA-DRB1*	***04:01**	11.9	678	0.39	0.13	>0	**1.8×10^{-03}**	**8.9×10^{-04}**
6	*HLA-DRB1*	**04:08*	2.1	678	0.75	0.33	>0	*2.1×10^{-02}*	*1.1×10^{-02}*
6	*HLA-DRB1*	***07:01**	9.5	678	0.25	0.07	>0	**9.1×10^{-04}**	**4.6×10^{-04}**
6	*HLA-DRB1*	*15	25.5	678	-0.12	0.05	>0	1.8×10^{-02}	9.9×10^{-01}
6	*HLA-DRB1*	*15:01	24.0	678	-0.12	0.05	>0	1.6×10^{-02}	9.9×10^{-01}
6	*HLA-DRB1*	*16:01	2.6	678	0.22	0.19	>0	2.5×10^{-01}	1.2×10^{-01}
6	*HLA-DRB1*	*03:01	2.9	678	-0.08	0.06	<0	2.1×10^{-01}	1.0×10^{-01}
6	*HLA-DRB1*	*04:04	0.4	678	-0.11	0.15	<0	4.7×10^{-01}	2.3×10^{-01}
6	*HLA-DRB1*	**11:01*	1.2	678	-0.18	0.08	<0	*3.5×10^{-02}*	*1.7×10^{-02}*
6	*HLA-DRB1*	*11:04	6.9	678	-0.20	0.13	<0	1.1×10^{-01}	5.7×10^{-02}
8	intergenic	rs4961252_G	40.6	941	-0.07	0.04	>0	1.0×10^{-01}	9.5×10^{-01}
6	intergenic	*rs9272105_A*	49.0	941	-0.09	0.04	<0	*1.5×10^{-02}*	*7.3×10^{-03}*

Figure legends

Figure 1: Regional association plots for lead variants identified in the GWAS (see Supplementary Table e-11). The red line marks the genome-wide significance level (5×10^{-8}). Color of dots indicates LD with the lead variant (pink); note that all variants showing genome-wide significance were in LD with the respective lead variants. Gray dots represent signals with missing R^2 values. The bottom box indicates the positions of *HLA* genes. Mbp = Mega base pairs. **A:** GWAS of EP1 in all patients, showing the strength of evidence for association with NAb presence (rs522308 $p=2.30\times10^{-15}$, genomic inflation $\lambda = 1.025$). **B:** GWAS of EP2 in all patients, showing the strength of evidence for association with mean NAb titer (rs2454138 $p=5.29\times10^{-19}$, $\lambda = 1.015$).


A
rs522308
R^2
0.8
0.6
0.4
0.2
$-\log_{10} p$
15
10
5
0
DRB5
DRB6
DRB1
DQA1
DQB1
32.50
32.55
32.60
32.65
Position on chromosome 6 (Mbp)
B
rs2454138
$-\log_{10} p$
20
15
10
5
0
DRB5
DRB6
DRB1
DQA1
DQB1
32.50
32.55
32.60
32.65
Position on chromosome 6 (Mbp)

Effect of *HLA-DRB1* alleles and genetic variants on the development of neutralizing antibodies to interferon beta in the BEYOND and BENEFIT trials: Supplement

Table of Contents:

Supplementary Table e-1: Quality control (QC) of genotyping data. QC of genotyped data was conducted in two separate iterations. The first QC took place following genotyping using PLINK, a second, refined round of QC before imputation using PLINK2 v1.90b3s. QC of imputed probabilities was conducted in QCTOOL v1.4 (http://www.well.ox.ac.uk/~gav/qctool/).

Samples were removed according to the following criteria:		**Variants were removed according to the following criteria:**	
Individual genotyping rate	< 98 %	Variant call rate	< 98 %
Gender mismatches		Minor allele frequency	< 1 %
Cryptic relatedness (PI-HAT)	> 0.05	HWE test *p*-value (first round)	$< 10^{-6}$
Removal of genetic outlier via EIGENSTRAT/SMARTPCA	> 6 SD	HWE test *p*-value (second round, after removal of individuals)	$< 10^{-7}$
Significant deviation of autosomal heterozygosity from the mean		Variants on non-autosomal chromosomes	
		Ambivalent SNPs (A/T and G/C)	
		Variants not present in the 1000 Genomes Phase 3 reference panel	
		IMPUTE2 info metric	< 0.8

Supplementary Table e-2: Ethnicities of samples. Based on genotype data, samples of non-Caucasian ancestry clustered together with samples of Caucasian ancestry (Supplementary Figure e-1).

Ethnicity	**Number of samples**
Caucasian	926
Hispanic	7
Asian	1
Black	1
Other	4
Unknown	2

Supplementary Table e-3: Additional characteristics of BEYOND and BENEFIT patients included in the analyses. For all median values, the median absolute deviation is indicated in brackets.

Measurement	**With SNP data**	**With *HLA* data**
Median NAb AUC during the first two years (Years*NU/mL), EP3 (including NAb-negative patients)	0 (0)	0 (0)
Median maximum NAb titer during the first two years (NU/ml), EP4 (including NAb-negative patients)	0 (0)	0 (0)
Minimal maximum NAb titer in any individual (NU/ml), including NAb-negative patients	0	0
Median time to the presence of a positive NAb titer, if NAbs were developed (months), EP5	6 (0)	6 (0)
Spearman correlation of mean NAb titer to NAb positivity	0.949	0.952
Spearman correlation of mean NAb titer to NAb AUC	0.997	0.999
Spearman correlation of mean NAb titer to max. NAb titer	0.998	0.999
Spearman correlation of mean NAb titer to the time of a positive NAb titer	-0.347	-0.368

Supplementary Table e-4: Characteristics of NAb-positive patients having developed NAb within two years (secondary analysis set). NAb-positive patients were defined as having a titer >20 NU/mL during the first two years of either study. Only a subset of patients had *HLA* data available. For all median values, the median absolute deviation is indicated in brackets.

Measurement	With SNP data	With *HLA* data
Number of samples	353	247
Study: BEYOND (n (%))	289 (81.9)	246 (99.6)
Sex: Female (n (%))	246 (69.7)	168 (68.0)
Median age (years)	38 (10.4)	39 (10.4)
Median disease duration (years)	2.5 (3.7)	3.4 (4)
Median EDSS	2 (1.5)	2 (1.5)
Median cumulative newly active lesions during the studies	2 (3)	2 (3)
Interferon β-1b dose: 250 µg (n (%))	202 (57.2)	121 (49.0)
Median duration of treatment (years)	2.7 (0.6)	2.6 (0.5)
Minimum duration of treatment with interferon β-1b (years)	0.5	1.5
Median total dose of interferon β-1b (mg)	187.2 (84.5)	184.2 (90.1)
Presence of NAbs (n (%)), EP1	353 (100)	247 (100)
Median mean NAb titer (NU/ml), EP2	55 (69.7)	58.5 (76.0)
Median NAb AUC during the first two years (Years*NU/mL), EP3	1013.5 (1272.1)	1016.5 (1285.4)
Median maximum NAb titer during the first two years (NU/ml), EP4	102 (112.7)	118 (136.4)
Minimal maximum NAb titer in any individual (NU/ml)	20	20
Median time to the presence of a positive NAb titer, if NAbs were developed (months), EP5	6 (0)	6 (0)
Spearman correlation of mean NAb titer to NAb AUC	0.978	0.985
Spearman correlation of mean NAb titer to max. NAb titer	0.981	0.984
Spearman correlation of mean NAb titer to the time of a positive NAb titer	-0.3	-0.323

Supplementary Table e-5: Summary statistics for the association analysis of NAb AUC with candidate markers (linear regression). Significance thresholds and abbreviations are as described in the legend of Table 2; SE = standard error.

Chr.	Gene	Allele / Variant	Freq. (%)	n	Effect	SE	Published effect	*p*-value	*p*-value (1-sided)
6	*HLA-DRB1*	***04:01**	11.9	678	0.41	0.13	>0	**1.2×10^{-03}**	**6.0×10^{-04}**
6	*HLA-DRB1*	**04:08*	2.1	678	0.78	0.33	>0	*1.7×10^{-02}*	*8.6×10^{-03}*
6	*HLA-DRB1*	***07:01**	9.5	678	0.24	0.07	>0	**1.5×10^{-03}**	**7.7×10^{-04}**
6	*HLA-DRB1*	*15	25.5	678	-0.12	0.05	>0	1.6×10^{-02}	9.9×10^{-01}
6	*HLA-DRB1*	*15:01	24.0	678	-0.12	0.05	>0	1.5×10^{-02}	9.9×10^{-01}
6	*HLA-DRB1*	*16:01	2.6	678	0.22	0.19	>0	2.7×10^{-01}	1.3×10^{-01}
6	*HLA-DRB1*	*03:01	2.9	678	-0.08	0.06	<0	2.3×10^{-01}	1.1×10^{-01}
6	*HLA-DRB1*	*04:04	0.4	678	-0.12	0.15	<0	4.1×10^{-01}	2.0×10^{-01}
6	*HLA-DRB1*	**11:01*	1.2	678	-0.18	0.08	<0	*3.6×10^{-02}*	*1.8×10^{-02}*
6	*HLA-DRB1*	*11:04	6.9	678	-0.2.0	0.13	<0	1.2×10^{-01}	5.8×10^{-02}
8	intergenic	rs4961252_G	40.6	941	-0.07	0.04	>0	9.6×10^{-02}	9.5×10^{-01}
6	intergenic	*rs9272105_A*	49.0	941	-0.08	0.04	<0	*2.0×10^{-02}*	*9.8×10^{-03}*

Supplementary Table e-6: Summary statistics for the association analysis of maximum NAb titer with candidate markers (linear regression). Significance thresholds and abbreviations are as described in the legend of Table 2; SE = standard error.

Chr.	Gene	Allele / Variant	Freq. (%)	n	Effect	SE	Published effect	*p*-value	*p*-value (1-sided)
6	*HLA-DRB1*	***04:01**	11.9	678	0.38	0.13	>0	**2.8×10^{-03}**	**1.4×10^{-03}**
6	*HLA-DRB1*	**04:08*	2.1	678	0.74	0.33	>0	*2.3×10^{-02}*	*1.2×10^{-02}*
6	*HLA-DRB1*	***07:01**	9.5	678	0.25	0.07	>0	**7.3×10^{-04}**	**3.7×10^{-04}**
6	*HLA-DRB1*	*15	25.5	678	-0.11	0.05	>0	2.4×10^{-02}	9.9×10^{-01}
6	*HLA-DRB1*	*15:01	24.0	678	-0.12	0.05	>0	1.9×10^{-02}	9.9×10^{-01}
6	*HLA-DRB1*	*16:01	2.6	678	0.20	0.19	>0	3.0×10^{-01}	1.5×10^{-01}
6	*HLA-DRB1*	*03:01	2.9	678	-0.08	0.06	<0	2.0×10^{-01}	1.0×10^{-01}
6	*HLA-DRB1*	*04:04	0.4	678	-0.09	0.15	<0	5.5×10^{-01}	2.8×10^{-01}
6	*HLA-DRB1*	**11:01*	1.2	678	-0.18	0.08	<0	*3.1×10^{-02}*	*1.6×10^{-02}*
6	*HLA-DRB1*	*11:04	6.9	678	-0.20	0.13	<0	1.1×10^{-01}	5.7×10^{-02}
8	intergenic	rs4961252_G	40.6	941	-0.07	0.04	>0	7.8×10^{-02}	9.6×10^{-01}
6	intergenic	*rs9272105_A*	49.0	941	-0.09	0.04	<0	*1.4×10^{-02}*	*6.9×10^{-03}*

Supplementary Table e-7: Summary statistics for the association analysis of mean NAb titer with candidate markers **in NAb-positive patients** (linear regression). Significance thresholds and abbreviations are as described in the legend of Table 2; SE = standard error.

Chr.	Gene	Allele / Variant	Freq. (%)	n	Effect	SE	Published effect	*p*-value	*p*-value (1-sided)
6	*HLA-DRB1*	*04:01	11.3	247	0	0.21	>0	9.9×10^{-01}	5.0×10^{-01}
6	*HLA-DRB1*	*04:08*	1.8	247	1.58	0.60	>0	9.4×10^{-03}	4.7×10^{-03}
6	*HLA-DRB1*	*07:01*	11.9	247	0.32	0.14	>0	2.7×10^{-02}	1.4×10^{-02}
6	*HLA-DRB1*	*15	22.1	247	-0.06	0.11	>0	5.7×10^{-01}	7.2×10^{-01}
6	*HLA-DRB1*	*15:01	20.6	247	-0.11	0.11	>0	3.3×10^{-01}	8.4×10^{-01}
6	*HLA-DRB1*	*16:01	1.8	247	0.25	0.43	>0	5.7×10^{-01}	2.8×10^{-01}
6	*HLA-DRB1*	*03:01	5.1	247	-0.15	0.14	<0	2.8×10^{-01}	1.4×10^{-01}
6	*HLA-DRB1*	*04:04	0.6	247	0.35	0.36	<0	3.4×10^{-01}	8.3×10^{-01}
6	*HLA-DRB1*	*11:01	1.4	247	-0.31	0.21	<0	1.5×10^{-01}	7.3×10^{-02}
6	*HLA-DRB1*	*11:04	5.3	247	-0.21	0.30	<0	4.7×10^{-01}	2.4×10^{-01}
8	intergenic	rs4961252_G	38.8	353	-0.16	0.09	>0	6.4×10^{-02}	9.7×10^{-01}
6	intergenic	*rs9272105_A*	46.6	353	-0.16	0.07	<0	2.9×10^{-02}	1.4×10^{-02}

Supplementary Table e-8: Summary statistics for the association analysis of NAb AUC with candidate markers **in NAb-positive patients** (linear regression). Significance thresholds and abbreviations are as described in the legend of Table 2; SE = standard error.

Chr.	Gene	Allele / Variant	Freq. (%)	n	Effect	SE	Published effect	*p*-value	*p*-value (1-sided)
6	*HLA-DRB1*	*04:01	11.3	247	0.07	0.21	>0	7.4×10^{-01}	3.7×10^{-01}
6	*HLA-DRB1*	*04:08*	1.8	247	1.58	0.60	>0	9.3×10^{-03}	4.6×10^{-03}
6	*HLA-DRB1*	*07:01	11.9	247	0.22	0.15	>0	1.2×10^{-01}	6.2×10^{-02}
6	*HLA-DRB1*	*15	22.1	247	-0.09	0.11	>0	4.4×10^{-01}	7.8×10^{-01}
6	*HLA-DRB1*	*15:01	20.6	247	-0.13	0.11	>0	2.4×10^{-01}	8.8×10^{-01}
6	*HLA-DRB1*	*16:01	1.8	247	0.08	0.44	>0	8.6×10^{-01}	4.3×10^{-01}
6	*HLA-DRB1*	*03:01	5.1	247	-0.14	0.14	<0	3.1×10^{-01}	1.5×10^{-01}
6	*HLA-DRB1*	*04:04	0.6	247	0.26	0.36	<0	4.8×10^{-01}	7.6×10^{-01}
6	*HLA-DRB1*	*11:01	1.4	247	-0.33	0.21	<0	1.2×10^{-01}	6.0×10^{-02}
6	*HLA-DRB1*	*11:04	5.3	247	-0.23	0.30	<0	4.4×10^{-01}	2.2×10^{-01}
8	intergenic	rs4961252_G	38.8	353	-0.17	0.09	>0	4.9×10^{-02}	9.8×10^{-01}
6	intergenic	*rs9272105_A*	46.6	353	-0.13	0.07	<0	7.9×10^{-02}	4.0×10^{-02}

Supplementary Table e-9: Summary statistics for the association analysis of maximum NAb titer with candidate markers **in NAb-positive patients** (linear regression). Significance thresholds and abbreviations are as described in the legend of Table 2; SE = standard error.

Chr.	Gene	Allele / Variant	Freq. (%)	n	Effect	SE	Published effect	*p*-value	*p*-value (1-sided)
6	*HLA-DRB1*	*04:01	11.3	247	-0.03	0.21	>0	9.0×10^{-01}	5.5×10^{-01}
6	*HLA-DRB1*	*04:08*	1.8	247	1.57	0.60	>0	9.9×10^{-03}	4.9×10^{-03}
6	*HLA-DRB1*	*07:01*	11.9	247	0.33	0.14	>0	2.2×10^{-02}	1.1×10^{-02}
6	*HLA-DRB1*	*15	22.1	247	-0.03	0.11	>0	7.6×10^{-01}	6.2×10^{-01}
6	*HLA-DRB1*	*15:01	20.6	247	-0.09	0.11	>0	4.1×10^{-01}	7.9×10^{-01}
6	*HLA-DRB1*	*16:01	1.8	247	0.05	0.44	>0	9.0×10^{-01}	4.5×10^{-01}
6	*HLA-DRB1*	*03:01	5.1	247	-0.14	0.14	<0	3.0×10^{-01}	1.5×10^{-01}
6	*HLA-DRB1*	*04:04	0.6	247	0.43	0.36	<0	2.4×10^{-01}	8.8×10^{-01}
6	*HLA-DRB1*	*11:01	1.4	247	-0.33	0.21	<0	1.2×10^{-01}	5.8×10^{-02}
6	*HLA-DRB1*	*11:04	5.3	247	-0.19	0.30	<0	5.3×10^{-01}	2.6×10^{-01}
8	intergenic	rs4961252_G	38.8	353	-0.18	0.09	>0	4.3×10^{-02}	9.8×10^{-01}
6	intergenic	*rs9272105_A*	46.6	353	-0.18	0.07	<0	1.6×10^{-02}	7.8×10^{-03}

Supplementary Table e-10: Summary statistics for the association analysis of time to first NAb presence with candidate markers **in NAb-positive patients** (Cox proportional hazard model). Because of genomic inflation (λ = 1.095), *p*-values were adjusted using genomic control. Significance thresholds and abbreviations are as described in the legend of Table 2; SE = standard error.

Chr.	Gene	Allele / Variant	Freq. (%)	n	Effect	SE	Published effect	*p*-value	*p*-value (1-sided)
6	*HLA-DRB1*	*04:01	11.3	247/247	0.17	0.21	<0	4.2×10^{-01}	7.9×10^{-01}
6	*HLA-DRB1*	*04:08	1.8	247/247	1.03	0.59	<0	8.3×10^{-02}	9.6×10^{-01}
6	*HLA-DRB1*	*07:01	11.9	247/247	-0.11	0.15	<0	4.5×10^{-01}	2.2×10^{-01}
6	*HLA-DRB1*	*15	22.1	247/247	-0.18	0.10	<0	7.5×10^{-02}	3.7×10^{-02}
6	*HLA-DRB1*	*15:01	20.6	247/247	-0.15	0.10	<0	1.4×10^{-01}	6.8×10^{-02}
6	*HLA-DRB1*	*16:01	1.8	247/247	-0.29	0.43	<0	5.0×10^{-01}	2.5×10^{-01}
6	*HLA-DRB1*	*03:01	5.1	247/247	0.13	0.13	>0	3.3×10^{-01}	1.6×10^{-01}
6	*HLA-DRB1*	*04:04	0.6	247/247	-0.40	0.36	>0	2.6×10^{-01}	8.7×10^{-01}
6	*HLA-DRB1*	*11:01	1.4	247/247	-0.11	0.22	>0	6.0×10^{-01}	7.0×10^{-01}
6	*HLA-DRB1*	*11:04	5.3	247/247	0.12	0.28	>0	6.6×10^{-01}	3.3×10^{-01}
8	intergenic	rs4961252_G	38.8	353/353	-0.10	0.09	<0	2.9×10^{-01}	1.4×10^{-01}
6	intergenic	rs9272105_A	46.6	353/353	0.15	0.07	>0	4.4×10^{-02}	2.2×10^{-02}

Supplementary Table e-11: Summary statistics of genome-wide significant variants for the association analysis of EP1-EP4 in all patients. Significance threshold $\alpha=5\times10^{-8}$; abbreviations are as described in the legend of Table 2; SE = standard error (EP2-4 are quantitative).

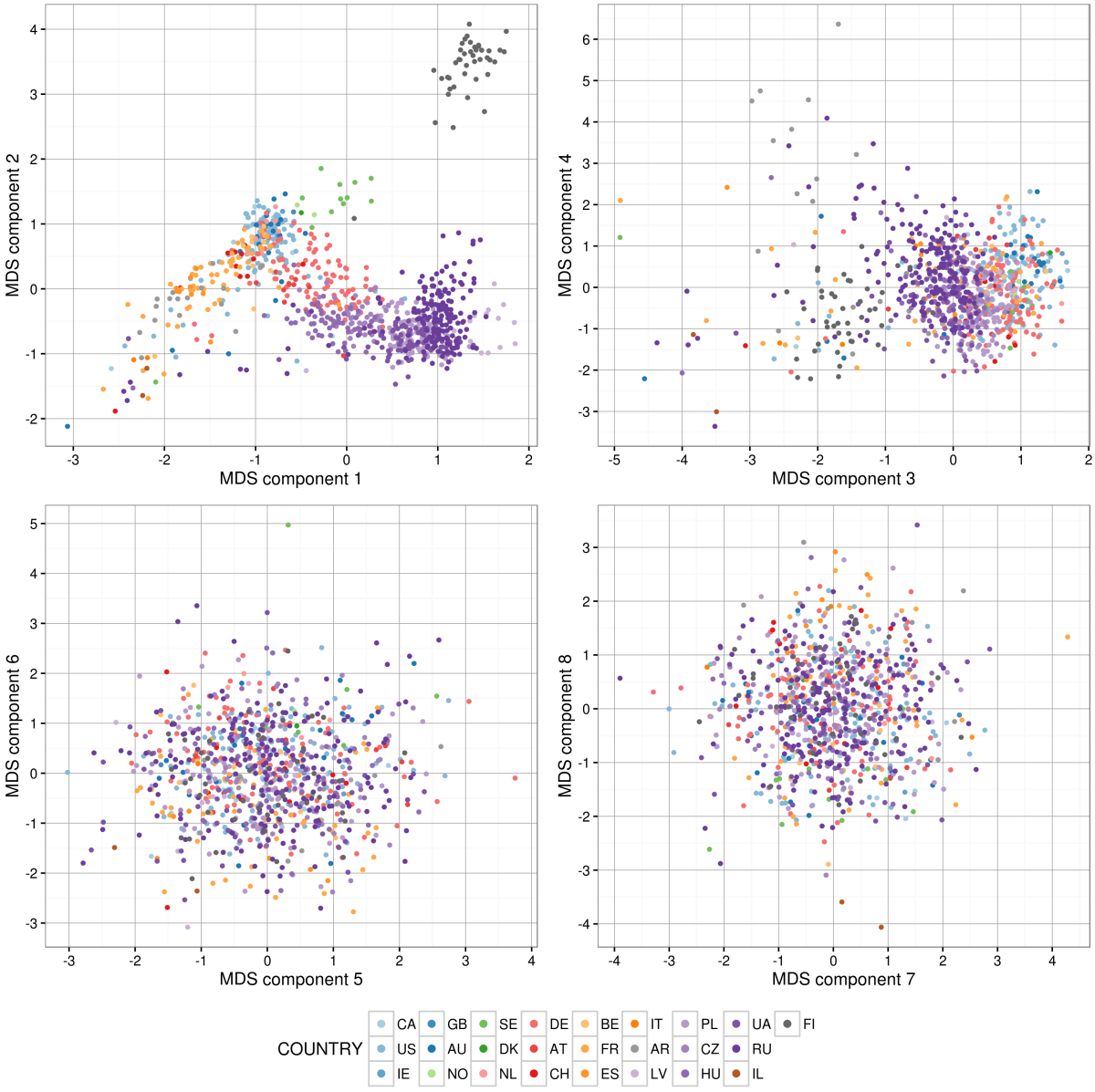
EP	Chr.	Gene	Variant	Freq. (%)	n	OR / Effect	CI / SE	*p*-value	λ
EP1	6	intergenic	**rs522308_T**	25.0	941	2.60	2.05-3.29	**2.30×10^{-15}**	1.025
EP2	6	intergenic	**rs2454138_A**	22.0	941	0.41	0.05	**5.29×10^{-19}**	1.015
EP3	6	intergenic	**rs2454138_A**	22.0	941	0.41	0.05	**8.44×10^{-19}**	1.015
EP4	6	intergenic	**rs2454138_A**	22.0	941	0.41	0.05	**1.00×10^{-18}**	1.014

Supplementary Table e-12: Significant single-tissue eQTLs of variant rs522308 in whole blood, GTEx Analysis Release V7.

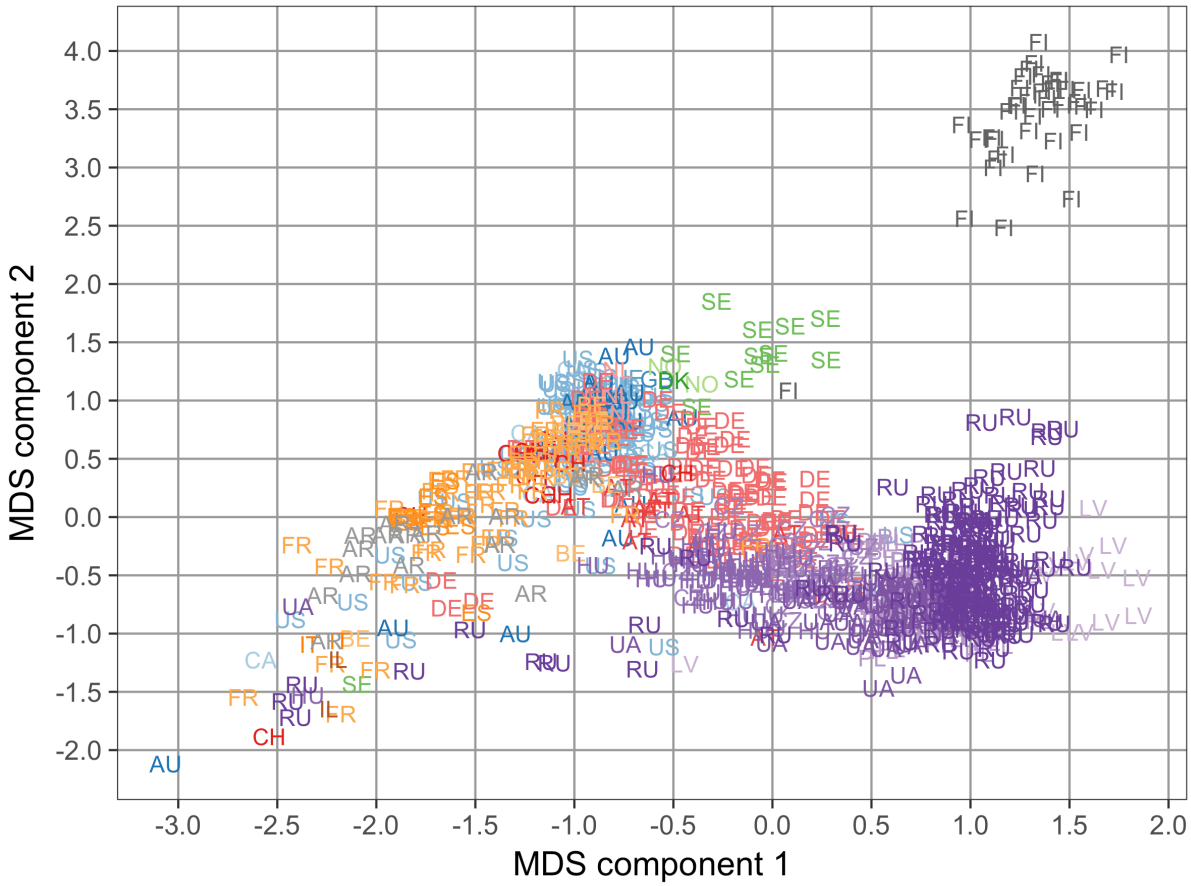
SNP	Gene	Effect	*p*-value
rs522308	***HLA-DQA2***	1	**2.00×10^{-54}**
rs522308	***HLA-DRB6***	0.72	**1.10×10^{-32}**
rs522308	***HLA-DQB2***	0.62	**4.50×10^{-21}**
rs522308	***HLA-DRB1***	-0.21	**3.30×10^{-16}**
rs522308	***HLA-DQB1***	-0.44	**5.70×10^{-16}**
rs522308	***HLA-DQA1***	-0.23	**1.90×10^{-11}**
rs522308	***HLA-DQB1-AS1***	-0.24	**4.80×10^{-07}**
rs522308	***HLA-DRB9***	0.31	**6.40×10^{-07}**

Supplementary Figure e-1: Population substructure analysis. The plot shows scaled MDS components of the genetic similarity matrix. MDS components were calculated in PLINK2 v1.90b3.27 using the eigendecomposition-based algorithm. IBS/IBD computation was conducted using the command *--genome* after filtering of genotyped variants (MAF ≥ 0.05, HWE *p*-value ≥ 10^{-3}, removal of the extended MHC region (chromosome 6, 25-35 Mbp) and a typical inversion site on chromosome 8 (7-13 Mbp)), and pruning (command *--indep-pairwise 200 100 0.2*). Country codes: CA = Canada, GB = Great Britain, SE = Sweden, DE = Germany, BE = Belgium, IT = Italy, PL = Poland, UA = Ukraine, FI = Finland, US = United States, AU = Australia, DK = Denmark, AT = Austria, FR = France, AR = Argentina, CZ = Czech Republic, RU = Russia, IE = Ireland, NO = Norway, NL = Netherlands, CH = Switzerland, ES = Spain, LV = Latvia, HU = Hungary, IL = Israel.
A: MDS components 1-8. **B:** MDS components 1 and 2, labeled by country codes.

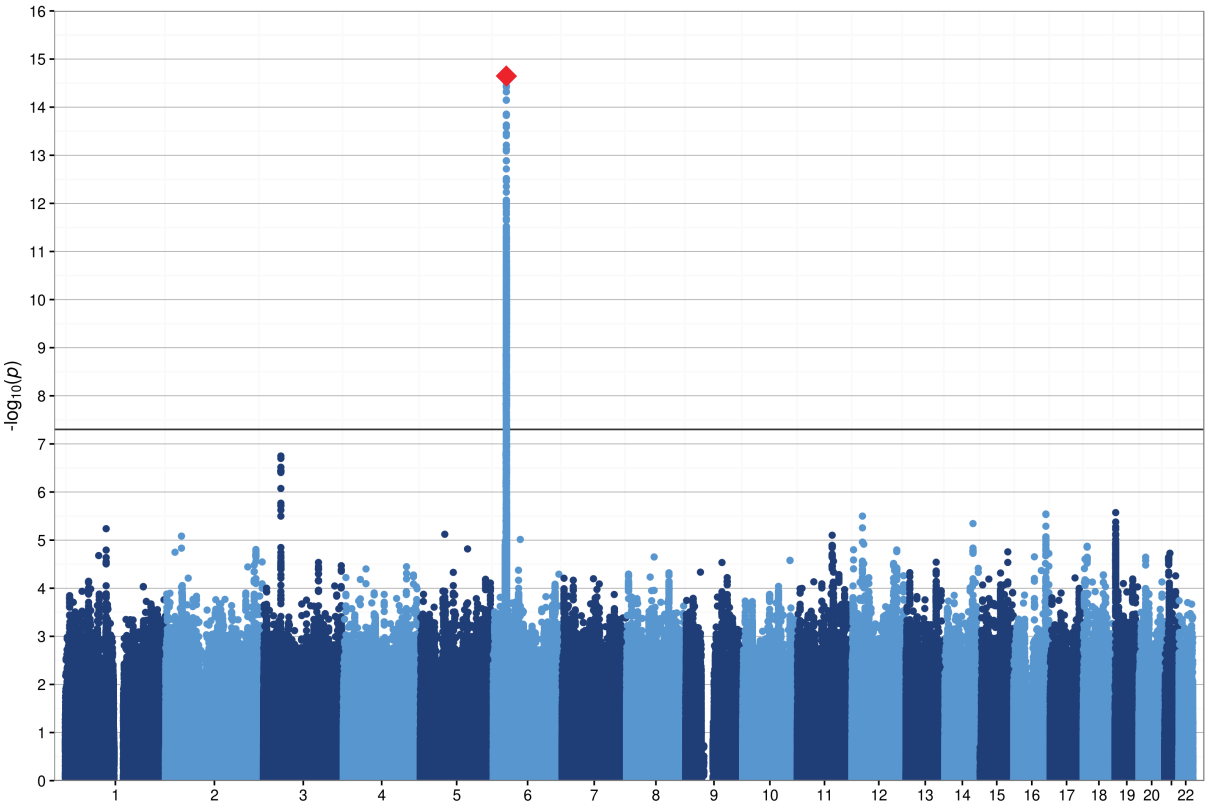
A



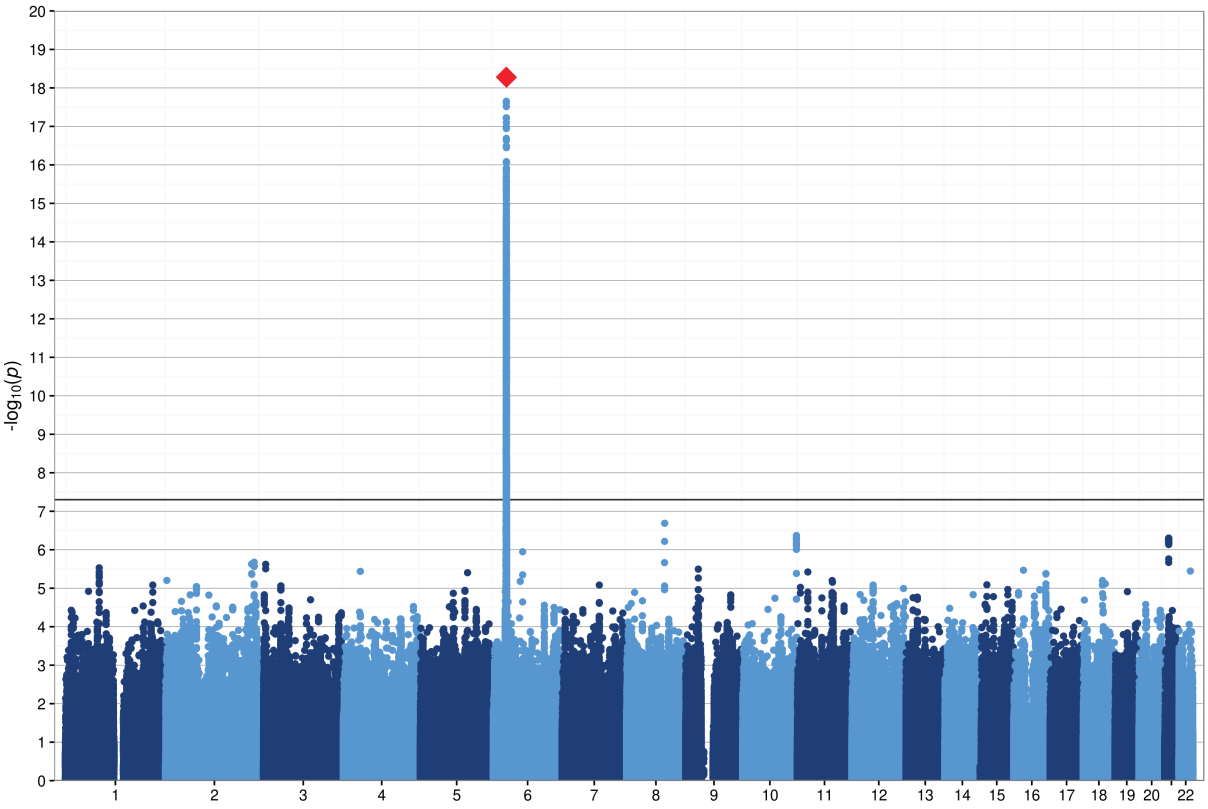
B


MDS component 2
MDS component 1
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-3.0
-2.5
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0

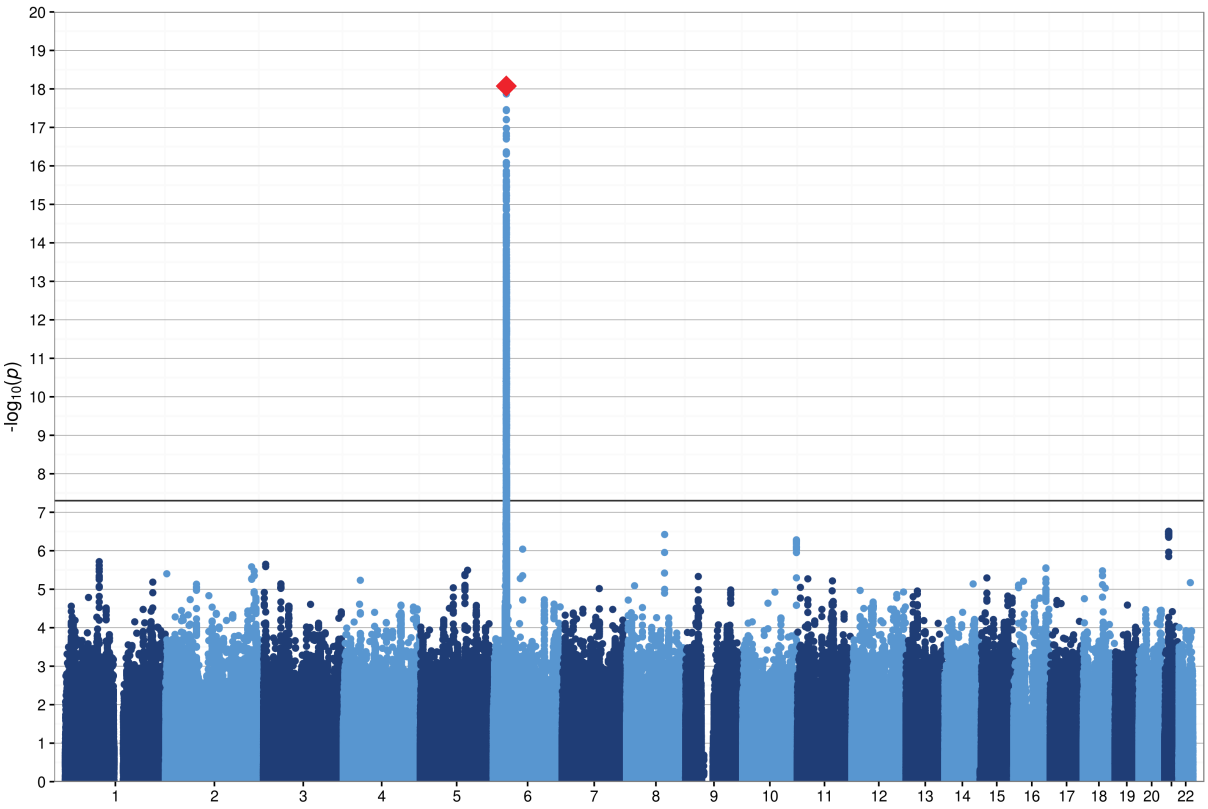
Supplementary Figure e-2: Manhattan plot of GWAS of EP1 in all patients, showing the strength of evidence for association with NAb presence. The gray line marks the genome-wide significance level and the red diamond the top SNP rs522308. Genomic inflation λ = 1.025.



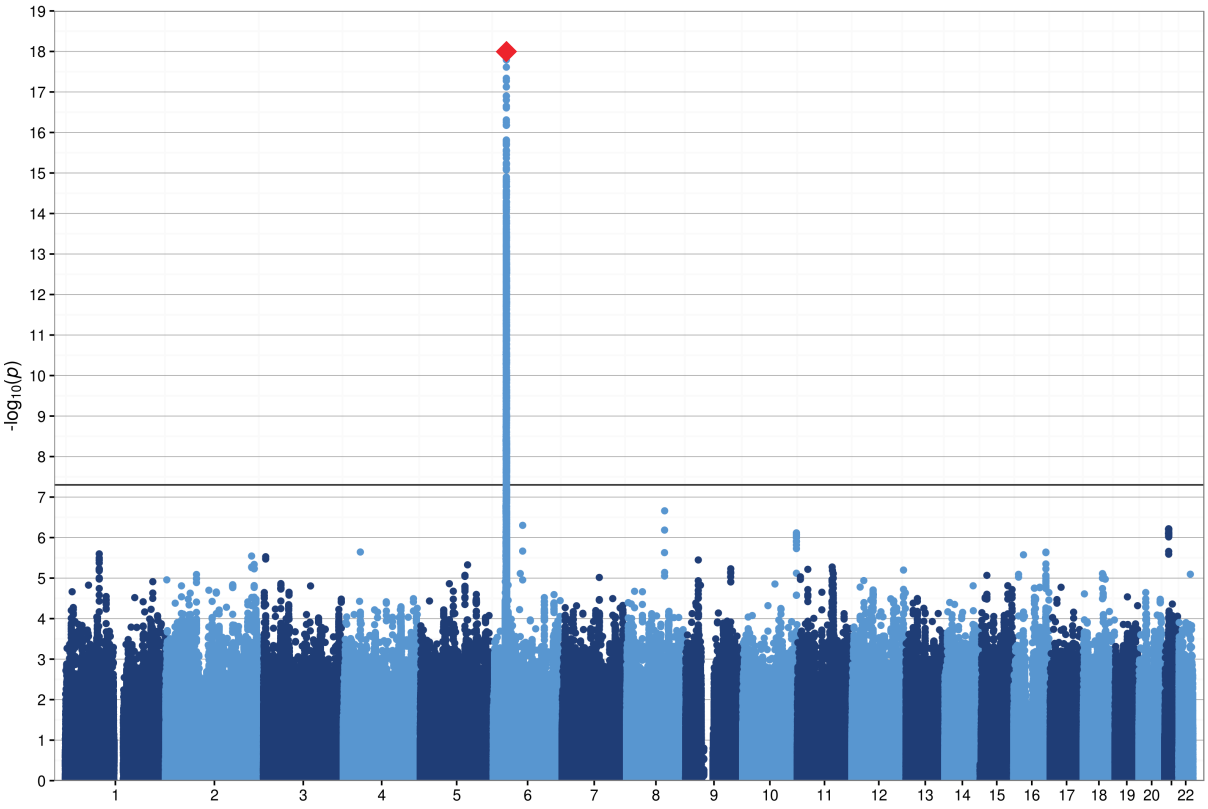
Supplementary Figure e-3: Manhattan plot of GWAS of EP2 in all patients, showing the strength of evidence for association with mean NAb titer. The gray line marks the genome-wide significance level and the red diamond the top SNP rs2454138. Genomic inflation λ = 1.015.



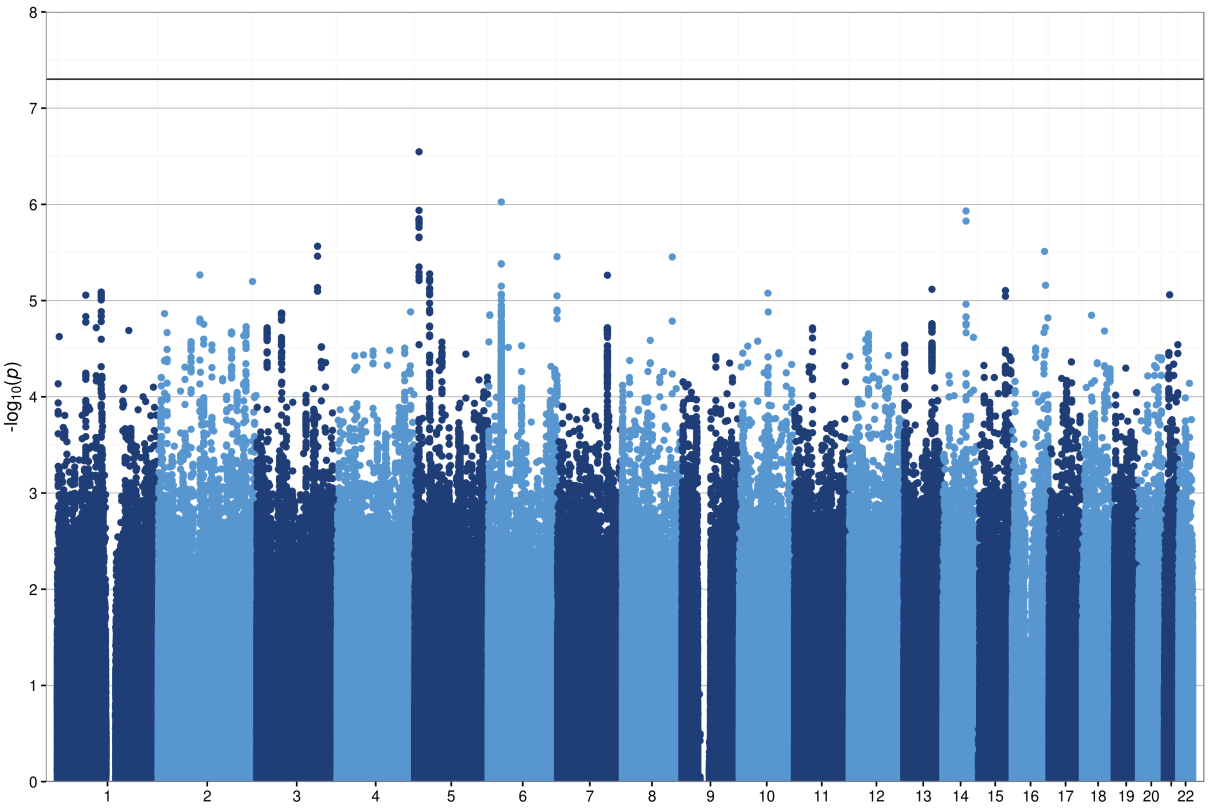
Supplementary Figure e-4: Manhattan plot of GWAS of EP3 in all patients, showing the strength of evidence for association with NAb AUC. The gray line marks the genome-wide significance level and the red diamond the top SNP rs2454138. Genomic inflation λ = 1.015.



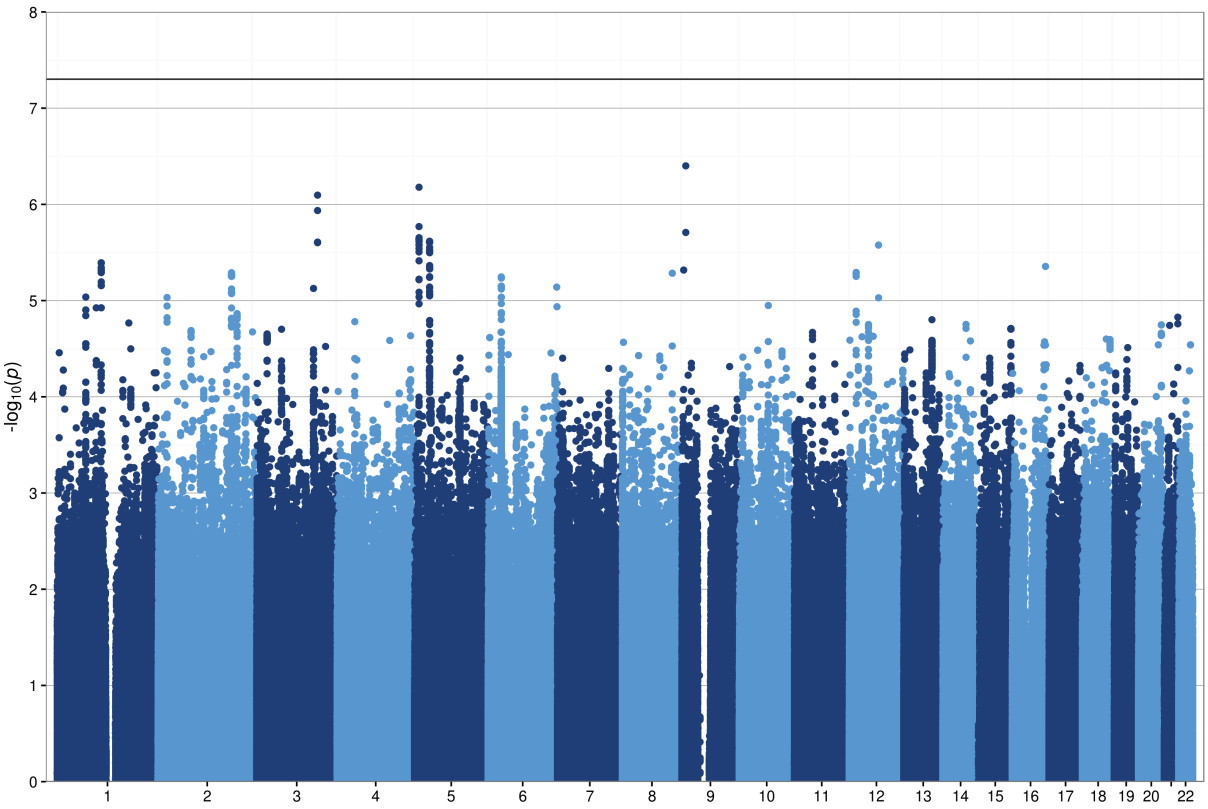
Supplementary Figure e-5: Manhattan plot of GWAS of EP4 in all patients, showing the strength of evidence for association with the maximum NAb titer. The gray line marks the genome-wide significance level and the red diamond the top SNP rs2454138. Genomic inflation λ = 1.014.



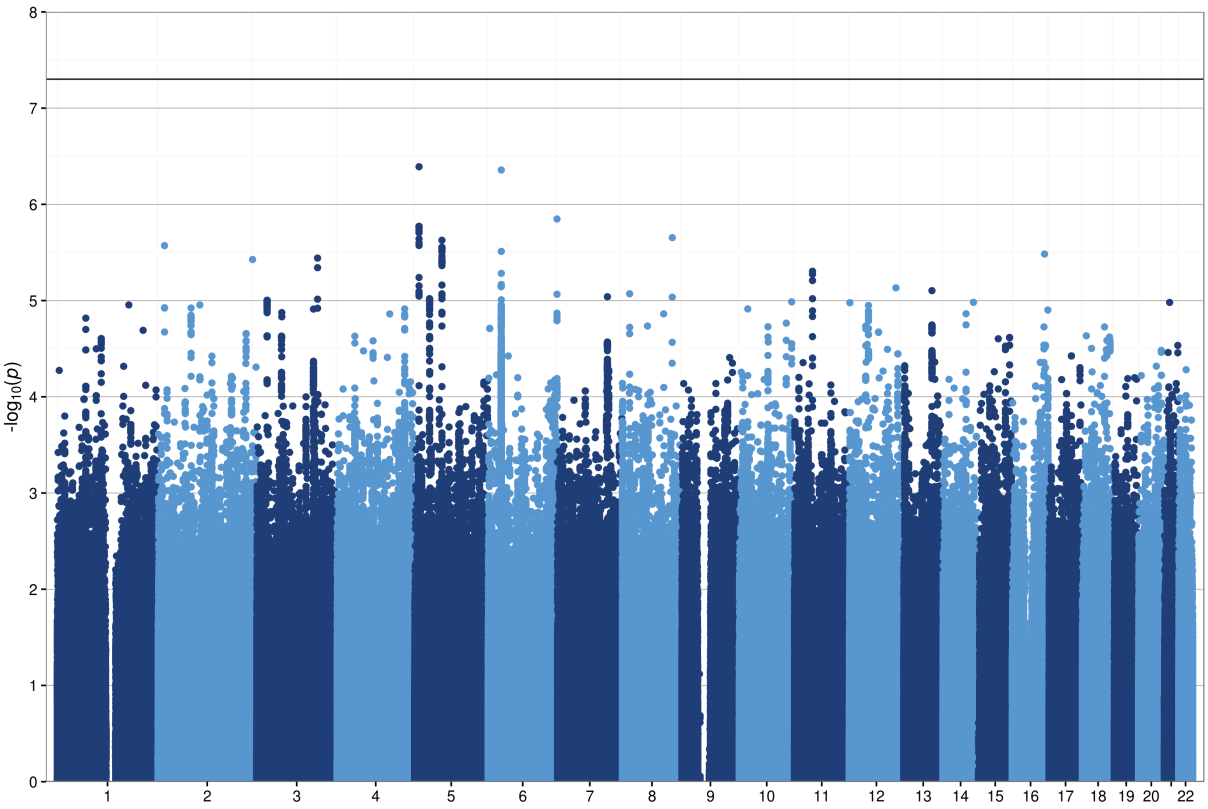
Supplementary Figure e-6: Manhattan plot of GWAS of EP2 in **NAb-positive patients**, showing the strength of evidence for association with mean NAb titer. The gray line marks the genome-wide significance level. Genomic inflation λ = 1.020.



Supplementary Figure e-7: Manhattan plot of GWAS of EP3 in **NAb-positive patients**, showing the strength of evidence for association with NAb AUC. The gray line marks the genome-wide significance level. Genomic inflation λ = 1.019.



Supplementary Figure e-8: Manhattan plot of GWAS of EP4 in **NAb-positive patients**, showing the strength of evidence for association with maximum NAb titer. The gray line marks the genome-wide significance level. Genomic inflation λ = 1.017.



Supplementary Figure e-9: Manhattan plot of GWAS of EP5 in **NAb-positive patients**, showing the strength of evidence for association with time to first NAb presence. The gray line marks the genome-wide significance level. Genomic inflation λ = 0.974.