HHS Public Access

Author manuscript

Arthritis Rheum. Author manuscript; available in PMC 2015 June 18.

Published in final edited form as:

Arthritis Rheum. 2013 September ; 65(9): 2457–2468. doi:10.1002/art.38036.

Association of Granulomatosis With Polyangiitis (Wegener's) With *HLA–DPB1*04* and *SEMA6A* Gene Variants Evidence From Genome-Wide Analysis

Gang Xie, MD, PhD[1], **Delnaz Roshandel, MD, MSc, PhD**[1], **Richard Sherva, PhD**[2], **Paul A. Monach, MD, PhD**[2], **Emily Yue Lu, MS**[3], **Tabitha Kung, MD, FRCPC**[1], **Keisha Carrington, HBSc**[1], **Steven S. Zhang, BSc**[1], **Sara L. Pulit, BA**[4], **Stephan Ripke, MD**[5], **Simon Carette, MD**[1], **Paul F. Dellaripa, MD**[6], **Jeffrey C. Edberg, PhD**[7], **Gary S. Hoffman, MD**[8], **Nader Khalidi, MD, FRCPC**[9], **Carol A. Langford, MD, MHS**[8], **Alfred D. Mahr, MD, PhD**[10], **E. William St.Clair, MD**[11], **Philip Seo, MD, MHS**[12], **Ulrich Specks, MD**[13], **Robert F. Spiera, MD**[14], **John H. Stone, MD, MPH**[5], **Steven R. Ytterberg, MD**[13], **Soumya Raychaudhuri, MD, PhD**[15], **Paul I. W. de Bakker, PhD**[4], **Lindsay A. Farrer, PhD**[2], **Christopher I. Amos, PhD**[16], **Peter A. Merkel, MD, MPH**[17], and **Katherine A. Siminovitch, MD, FRCPC, FACP**[18]

[1]Mount Sinai Hospital and University of Toronto, Toronto, Ontario, Canada [2]Boston University School of Medicine, Boston, Massachusetts [3]University of Texas M. D. Anderson Cancer Center, Houston [4]Brigham and Women's Hospital and Broad Institute, Boston, Massachusetts, and University Medical Center Utrecht, Utrecht, The Netherlands [5]Massachusetts General Hospital, Boston [6]Brigham and Women's Hospital, Boston, Massachusetts [7]University of Alabama at Birmingham [8]Cleveland Clinic, Cleveland, Ohio [9]McMaster University, Hamilton, Ontario, Canada [10]Hôpital Saint-Louis and Université Paris 7–Diderot, Paris, France [11]Duke University Medical Center, Durham, North Carolina [12]Johns Hopkins University, Baltimore, Maryland [13]Mayo Clinic, Rochester, Minnesota [14]Hospital for Special Surgery, New York, New York [15]Brigham and Women's Hospital, Broad Institute, and Partners Center for Personalized Genetic Medicine, Boston, Massachusetts, and University of Manchester, Manchester, UK [16]University of Texas M. D. Anderson Cancer Center, Houston, and Geisel School of Medicine at Dartmouth, Hanover, New Hampshire [17]Boston University School of Medicine, Boston, Massachusetts, and University of Pennsylvania School of Medicine, Philadelphia [18]Mount Sinai Hospital Samuel Lunenfeld Research Institute, Toronto General Research Institute, and University of Toronto, Toronto, Ontario, Canada

Address correspondence to Katherine A. Siminovitch, MD, FRCPC, FACP, Mount Sinai Hospital, 600 University Avenue, Room 778D, Toronto, Ontario M5G 1X5, Canada. ksimin@mshri.on.ca.
Drs. Xie and Roshandel contributed equally to this work.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Siminovitch had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Xie, Monach, Zhang, Ripke, Khalidi, Specks, Spiera, Stone, Raychaudhuri, Merkel, Siminovitch.

Acquisition of data. Xie, Monach, Carrington, Ripke, Carette, Dellaripa, Edberg, Hoffman, Khalidi, Langford, Mahr, St.Clair, Seo, Specks, Spiera, Stone, Ytterberg, Farrer, Merkel, Siminovitch.

Analysis and interpretation of data. Xie, Roshandel, Sherva, Lu, Kung, Pulit, Ripke, Hoffman, Langford, Seo, Spiera, Raychaudhuri, de Bakker, Farrer, Amos, Merkel, Siminovitch.

Abstract

Objective—To identify genetic determinants of granulomatosis with polyangiitis (Wegener's) (GPA).

Methods—We carried out a genome-wide association study (GWAS) of 492 GPA cases and 1,506 healthy controls (white subjects of European descent), followed by replication analysis of the most strongly associated signals in an independent cohort of 528 GPA cases and 1,228 controls.

Results—Genome-wide significant associations were identified in 32 single-nucleotide polymorphic (SNP) markers across the HLA region, the majority of which were located in the *HLA–DPB1* and *HLA–DPA1* genes encoding the class II major histocompatibility complex (MHC) DPβ chain 1 and DPα chain 1 proteins, respectively. Peak association signals in these 2 genes, emanating from SNPs rs9277554 (for DPβ chain 1) and rs9277341 (DPα chain 1) were strongly replicated in an independent cohort (in the combined analysis of the initial cohort and the replication cohort, $P = 1.92 \times 10^{-50}$ and 2.18×10^{-39}, respectively). Imputation of classic HLA alleles and conditional analyses revealed that the SNP association signal was fully accounted for by the classic *HLA–DPB1*04* allele. An independent single SNP, rs26595, near *SEMA6A* (the gene for semaphorin 6A) on chromosome 5, was also associated with GPA, reaching genome-wide significance in a combined analysis of the GWAS and replication cohorts ($P = 2.09 \times 10^{-8}$).

Conclusion—We identified the *SEMA6A* and *HLA–DP* loci as significant contributors to risk for GPA, with the *HLA–DPB1*04* allele almost completely accounting for the *MHC* association. These two associations confirm the critical role of immunogenetic factors in the development of GPA.

Granulomatosis with polyangiitis (Wegener's) (GPA) is a systemic vasculitis characterized by vessel inflammation and granuloma formation with involvement of the respiratory tract, kidneys, and other organs and tissues (1). It is a rare disease and is associated with high morbidity and mortality (2,3). GPA, microscopic polyangiitis, and eosinophilic granulomatosis with polyangiitis (Churg-Strauss) are collectively referred to as antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis (AAV) due to their association with the presence of cytoplasmic ANCA (cANCA) (4–6).

The primary etiology of GPA remains unknown, but epidemiologic data link its expression to both environmental and genetic factors (2,7–15). The HLA class II region is the best-established risk gene/locus in GPA, with an extended haplotype comprising alleles of *HLA–DPB1*0401* and *RXR*β (the gene for retinoid X receptor β) showing the strongest known association with risk for GPA ($P = 7.13 \times 10^{-17}$, odds ratio [OR] 6.41) (16). Other genes that have been implicated in the development of GPA based on candidate genetic association data include *SERPINA1* (17,18), *CTLA4* (19,20), *PTPN22* (21), *FCGR3B* (22), *PRTN3* (23), *CD226* (24), *LEPR* (25), *IL10* (20,26), *TNF* (20), *TNFAIP3* (27), *CDK6* (27), *IRF5* (27), *UCP2* (28), *INFG* (20), and *TGFβ1* (29). One of the non–major histocompatibility complex (non-MHC) loci, *SERPINA1*, emerged as a risk locus in a recent genome-wide association study (GWAS) of a European AAV cohort as well (30). However, the majority of non-MHC loci identified as candidates for GPA risk have not been independently replicated, and the genetic basis of GPA remains largely undefined (31). Accordingly, to further delineate the

genetic factors underlying GPA, we performed a GWAS of GPA cases ascertained from the US and Canada.

PATIENTS AND METHODS

Study subjects were all of European descent. All GPA cases fulfilled a modified version of the American College of Rheumatology criteria for the disease (6,32). The study was approved by the local institutional ethics committee at each center. Demographic data and blood and/or saliva samples for genomic DNA extraction were obtained from all study subjects, with written informed consent. A 2-stage approach was adopted in this study, as described below and summarized in Supplementary Figure 1 (on the *Arthritis & Rheumatism* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.38036/abstract).

Stage 1 GWAS

Cases initially entered into the stage 1 GWAS (n = 492) were recruited between 2001 and 2011, with 149 patients ascertained from Ontario and the remaining patients from regions across Canada and the US (see Supplementary Figure 1). Cases passing quality control measures (n = 459) were compared to controls from Canada and the US passing quality control measures (n = 1,503) (Supplementary Table 1 [http://onlinelibrary.wiley.com/doi/10.1002/art.38036/abstract]). The control subjects ("historical" controls) included 380 healthy volunteers with no prior history of autoimmune disease (ascertained in Canada between 2005 and 2008) and 1,123 healthy controls recruited to the M. D. Anderson Lung Cancer Study, a case–control study that has been ongoing since 1999 (33).

Stage 2 replication analysis

In stage 2, several independently collected populations were used in analyses to replicate significant genetic associations with disease that had been discovered in stage 1 (2 single-nucleotide polymorphisms [SNPs] in the *HLA* region and 16 SNPs in non-*HLA* regions). Cases used for replication analyses were recruited from 2 cohorts: the Wegener's Granulomatosis Genetic Repository (WGGER) and the Vasculitis Clinical Research Consortium (VCRC). The WGGER is an ancillary study emanating from the Wegener's Granulomatosis Etanercept Trial conducted at 8 academic centers in the US between 2001 and 2005 (34,35). The VCRC, a member of the National Institutes of Health Rare Disease Clinical Research Network (http://RareDiseasesNetwork.org/VCRC/), conducts research studies that include patients with systemic vasculitis followed up at 10 major referral centers in the US and Canada (31). Any cases included in both the VCRC and WGGER collections were retained in the WGGER and removed from the VCRC cohorts studied here. Controls used for replication analyses came from the WGGER cohort and the Multi-Institutional Research in Alzheimer's Genetic Epidemiology (MIRAGE) study, a genetic epidemiologic study of Alzheimer's disease (36). Replication set 1 (stage 2a, examining 2 *HLA* SNPs and 4 *non-HLA* SNPs) included 291 cases and 317 controls from the WGGER cohort. Replication set 2 (stage 2b, examining 12 non-*HLA* SNPs) included 528 cases (279 from the WGGER and 249 from the VCRC) and 1,228 controls (300 from the WGGER and 928 "historical" controls from the MIRAGE study) (Supplementary Table 1 and Supplementary Figure 1).

Data from all patients and controls in stages 1 and 2 were then included in a combined analysis.

Genotyping and quality control

DNA was prepared using standard methods. For stage 1, 492 samples from GPA cases and 382 from historical controls were genotyped using Illumina HumanHap370 BeadChips. An additional 1,124 historical control samples from the M. D. Anderson Cancer Center Lung Cancer Study were genotyped at the Johns Hopkins University Center for Inherited Disease Research, using Illumina HumanHap300 v1.1 BeadChips. Each SNP was determined using Illumina Genome Studio version 2010.3 with Genotyping Module version 1.8.4 and the combined intensity data from >90% of study samples. The cluster definition file was used to determine genotype calls and quality scores, for all study samples with genotype calls made when a genotype yielded a quality score (GenCall value) of ≥0.15. Stage 2 samples were genotyped at the Boston University Molecular Genetics Core Facility, on an ABI 7900 (real-time) platform, using SNP assay kits according to the protocol of the manufacturer (ABI). After setting both maximum per-person missingness and maximum per-SNP missingness to be no more than 0.05, minor allele frequency to be no less than 0.01, and Hardy-Weinberg equilibrium to be no less than 0.0001, we retained 459 samples and 332,290 SNPs from the GPA case group, 382 samples and 336,887 SNPs from the Canadian historical control group, and 1,124 samples and 312,412 SNPs from the US historical control group. Genotyping data from 459 patients, 382 Canadian historical controls, and 1,124 US historical controls were merged, and the same missingness settings were applied to all of the samples and SNPs. The final combined data set contained 1,962 subjects (459 patients with GPA, 380 controls from Canada, and 1,123 controls from the US), with 287,802 common SNPs. Quantile–quantile plots of the test statistics (*P* values) for the GPA GWAS are shown in Supplementary Figure 2 (on the *Arthritis & Rheumatism* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.38036/abstract).

Statistical power

Power calculations for the GWAS were performed using Quanto version 1.2.4 (http://hydra.usc.edu/gxe/), with the following parameters: disease prevalence for GPA 30/1,000,000, minor allele frequency 0.20, 459 cases and 1,503 controls, α 1.00×10^{-7} (0.05/287,802 SNPs tested) and (less conservative) 1.00×10^{-4} under a log-additive model. For $\alpha = 1.00 \times 10^{-7}$ or $\alpha = 1.00 \times 10^{-4}$, the power to detect associations with relative risks of ≥1.8 or ≥1.6, respectively, was estimated to be >80%. In the combined analysis, with 987 cases and 2,731 controls, $\alpha = 1.00 \times 10^{-7}$, and the same parameters, the power to detect SNP associations with relative risks of ≥1.5 was also estimated to be >80%.

Population structure

Genome-wide relatedness was explored using pairwise identity-by-state analysis implemented in Plink (version 1.07, http://pngu.mgh.harvard.edu/purcell/plink/) (37) to identify subjects with excess identity-by-descent sharing (PI_HAT >0.10). One subject from each of the 15 pairs exceeding this threshold was removed from the association analysis. Hierarchical cluster analysis was performed with Plink to identify subjects with similar

genotypes over the entire genome; samples that were >4 SD from a nearest neighbor were excluded from the analysis.

Association testing

We used Plink version 1.07 to perform the Cochran-Mantel-Haenszel test in order to adjust for stratification among groups of subjects identified by hierarchical clustering. To test for robustness of our analytical approach, we also performed association analyses with adjustment for the top 10 eigenvectors, on the basis of a principal components analysis implemented with the EigenStrat method (38). SNPs in regions of high linkage disequilibrium (e.g., the MHC region) were removed from the principal components analysis. The lambda values showed minimal inflation (1.029 and 1.022, respectively, before and after adjustment for eigenvectors (Supplementary Figure 2 [http://onlinelibrary.wiley.com/doi/10.1002/art.38036/abstract]), and were thus not used to adjust *P* values in this study.

Imputation

We imputed additional SNPs in the 1,962 samples that passed quality control standards in the GWAS, using Impute 2 version 2.1.2 (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html) (39) and the Centre d' Étude du Polymorphisme Humain study population of Utah residents with ancestry from northern and western Europe (CEU) 1000 Genomes reference set (version 37), containing whole-genome sequence data from 629 individuals including 19,097,376 SNP variants (40). To impute classic *HLA* alleles, we used a reference data set from the Type 1 Diabetes Genetics Consortium, which includes genotype data for 7,261 SNPs selected across the entire MHC, high-resolution 4-digit typing data for 424 classic alleles at *HLA–A, B, C, DRB1*, *DQA1*, *DBQ1*, *DPA1*, and *DPB1*, and 1,276 amino acids corresponding to serologically defined antigens in 5,225 individuals of European descent (41–43). From the GWAS data set, we removed SNPs with missingness of >0.01, MAF of <0.1, and Hardy-Weinberg equilibrium in controls at $P < 0.05$ before performing the HLA imputations. We imputed SNPs, classic *HLA* alleles (at both 4-and 2-digit resolution), and their amino acid polymorphisms in the 1,962 individuals from our GWAS, using Beagle software as previously described (44). After the HLA imputations, we removed samples with cumulative dosage of >2.5 across all 4-digit classic alleles at any *HLA* gene, resulting in omission of 24 subjects (5 cases and 19 controls) from further analysis.

Conditional analysis

Conditional analysis of imputed *HLA* alleles was performed to test for multiple independent effects within the region. We first used a logistic regression framework to test individual alleles for association, including the top 10 principal components as covariates to account for population stratification. After identifying the single most significant marker, we tested for additional independent effects by including the dosage of that marker as a covariate.

Population attributable fraction

Population attributable fraction (PAF) was estimated using ORs obtained with a multivariate logistic regression model incorporating SNPs from multiple loci. For an individual SNP the population attributable fraction is calculated from the OR, as

$$\mathrm{PAF}=\frac{[\mathrm{AF}_{\mathrm{ctrls}}\times(\mathrm{OR}-1)]}{[\mathrm{AF}_{\mathrm{ctrls}}\times(\mathrm{OR}-1)]+1}$$

where AF_{ctrls} is the risk allele frequency in controls. The combined population attributable fraction allowing for all SNPs jointly is 1 - $II_{allSNPs}$ (1 - PAF_i) where PAF_i is the population attributable fraction for SNP_i and II denotes multiplication (45).

RESULTS

Results of the stage 1 GWAS

A total of 287,802 SNPs and 1,962 subjects (459 cases and 1,503 controls) met all sample and SNP quality control standards with minimal inflation of the test statistic (λ_{GC} = 1.02) (Supplementary Figure 2 [http://onlinelibrary.wiley.com/doi/10.1002/art.38036/abstract]), suggesting that population stratification was limited. The study design and subject characteristics are summarized in Supplementary Figure 1 and Supplementary Table 1.

Association analysis revealed that the SNPs most significantly associated with GPA mapped within the *HLA* region on chromosome 6 (Table 1, Figure 1, and Supplementary Table 2). Thirty-two SNPs across this region met a genome-wide significance threshold for association ($P \leq 5.00 \times 10^{-8}$), with the majority and most significantly associated of these SNPs located in the *HLA–DPB1* and *HLA–DPA1* genes encoding class II MHC DPβ chain 1 and DPβα chain 1, respectively (Figure 2A). Peak association signals in these 2 genes arose from SNPs rs9277554 ($P = 4.88 \times 10^{-38}$) and rs9277341 ($P = 1.84 \times 10^{-30}$), located 3' of *HLA–DPB1* and in intron 1 of *HLA–DPA1*, respectively. In addition to multiple *HLA–DPB1* and *HLA–DPA1* variants, SNPs significantly associated with GPA were found in the *HLA–DOA* and *HLA–DPB2* genes, in 4 other MHC region genes (*COL11A2*, *RXRB*, *HSD17B8*, and *RING1*), and in 4 non-MHC genes (*DCTD*, *COBL*, *CCDC86*, and *WSCD1*) (Table 1).

Results of the imputation analysis

HLA region—As multiple SNPs across the *HLA* region met a genome-wide significance threshold for association, an imputation procedure specific for this locus was used to finemap these signals. To this end, 424 classic *HLA* alleles and 1,276 corresponding amino acid polymorphisms at the *HLA–A, B, C, DRB1, DQA1, DQB1, DPA1*, and *DPB1* loci along with 7,261 SNPs across the MHC were imputed in 459 cases and 1,503 controls from the GWAS data set, and conditional analysis was then used to identify independent variants, with adjustment for the top 10 principal components in the GWAS.

In this analysis, 215 variants with genome-wide significant disease association were identified (Supplementary Table 3, on the *Arthritis & Rheumatism* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.38036/ abstract). Among these, the strongest signal

($P = 5.56 \times 10^{-47}$) mapped to the classic *HLA–DPB1*04* allele (Figure 2B). We did not observe a statistically significant difference in results between a model for *HLA–DPB1*04* (2-digit) and a model for both *HLA–DPB1*04:01* and *HLA-DPB1*04:02* (4-digit) ($P = 0.796$), suggesting that the association signal is common to all 4-digit alleles of the *HLA–DPB1*04* 2-digit (serologic) type. Conditioning on the *HLA–DPB1*04* allele dramatically diminished the association with other variants (including SNPs in 2 genes, *RXRB* and *RING1*, previously implicated in GPA susceptibility) (46,47), revealing that the association signal emanated almost exclusively from *HLA–DPB1* (Supplementary Table 3 and Supplementary Figure 3). However, the data do not exclude the possibility that this signal derives from an unobserved allele in a neighboring gene carried by *HLA–DPB1*04* haplotypes.

Non-HLA region—All SNPs meeting quality control standards in the GWAS were used to impute more than 19 million SNPs present at >1% allele frequency in the CEU 1000 Genomes reference panel. This analysis, performed using logistic regression with adjustment for the top 10 principal components as covariates, did not identify any new *non-HLA* loci with genome-wide significant association with GPA (data not shown).

Results of the independent replication (stage 2a) and combined analysis

In contrast to associations within the *HLA* region, each of the significant non-*HLA* associations emerging from the GWAS arose from a single SNP. For this reason and because of the modest size of the initial cohort surveyed, a second-stage analysis (stage 2a) was performed, involving genotyping of the most strongly associated SNPs at each of the *HLA–DPB1* and *HLA–DPA1* loci, together with the 4 disease-associated non-MHC SNPs, in an independently collected, US-derived cohort (WGGER: 291 patients with GPA and 317 controls) (see Supplementary Figure 1 [http://onlinelibrary.wiley.com/doi/10.1002/art.38036/abstract]). Both the *HLA–DPB1* and *HLA–DPA1* associations with GPA were strongly replicated in this population ($P = 1.40 \times 10^{-13}$ and $P = 3.84 \times 10^{-9}$, respectively), yielding combined (GWAS and replication) *P* values of 1.92×10^{-50} and 2.18×10^{-39}, respectively (Table 1). In contrast, none of the *non-HLA* loci were replicated in this cohort (Table 1).

Results of the independent replication (stage 2b) and combined analysis

To ascertain whether use of a reduced stringency threshold for significance might reveal any non-HLA SNPs of potential relevance to GPA, the replication analysis was expanded (stage 2b) to include 12 additional SNPs that were selected based on suggestive evidence for disease association in the stage 1 GWAS (*P* values between 1.00×10^{-7} and 1.00×10^{-4}) and proximity to at least 1 other SNP with nominal evidence for association (*P* value between 1.00×10^{-4} and 1.00×10^{-2}). These SNPs were genotyped in an enlarged set of cases (n = 528) and controls (n = 1,228) derived by combining the WGGER cohort with the VCRC cohort and the MIRAGE control collection (Supplementary Figure 1). Among the 12 variants tested, genome-wide statistical significance was achieved only for SNP rs26595 in a combined analysis of the GWAS and replication cohorts ($P = 2.09 \times 10^{-8}$) (Table 2). Located on chromosome 5q23.1, rs26595 lies 20 kb proximal to the 3' end of the *SEMA6A* gene, which is the most compelling candidate gene from this region since the next closest

gene (*COMMD10*) lies another 120 kb upstream and several imputed SNPs across *SEMA6A* also show suggestive evidence for disease association (Figure 3). No significant interaction between rs26595 at *SEMA6A* and *HLA–DPB1*04* was detected ($P > 0.66$).

The relevance of the *HLA* and *SEMA6A* associations to specific sub-phenotypes of GPA was also explored. Comparisons of the *HLA–DPB1*, *HLA–DPA1*, and *SEMA6A* peak risk allele frequencies between the proteinase 3 (PR3) cANCA–positive subgroup (88% of cases) and the ANCA-negative subgroup revealed that all 3 associations were restricted to the PR3 cANCA–positive group (Table 3), a finding consistent with other data suggesting genetic divergence between ANCA-positive and -negative GPA (30). Stratification based on specific organ involvement did not reveal any significant differences in associations between subgroups (data not shown).

DISCUSSION

Results of this GWAS provide compelling evidence of *HLA* class II region involvement in GPA, with *HLA–DPB1* and *HLA–DPA1* variants showing the strongest association signals. Data from previous candidate gene analyses have suggested that 2 other genes in this region, *RXRB* and *RING*, confer GPA risk independent of the *HLA–DPB1* gene (47). Three SNPs at this locus (1 in *HLA–DPB1* and 2 in *COL11A2*) were also associated with AAV at the genome-wide significance level in a recent study of patients with AAV from Europe (30). However, the results of our conditional analysis suggest that associations with these genes, which were also observed in the present study, are entirely or nearly entirely attributable to the *DPB1*04* allele. Our data also suggest negligible effects of *HLA* class I or class III genes in GPA, but do not pinpoint any single functional variant in the *DPB1* gene as the driver of the *HLA* effect. However, while the causal variant(s) at this locus remains unknown, the population attributable fraction (63%) for the *HLA–DPB1* rs9277554 C allele (Supplementary Table 4 [http://onlinelibrary.wiley.com/doi/10.1002/art.38036/abstract]) and the OR of 3.41 for the risk (C) allele indicate a substantive contribution of this locus to risk for GPA.

We also identified an association of the *SEMA6A* locus on chromosome 5q23.1 with GPA, with the 19% population attributable fraction for the rs26595 T risk allele (Supplementary Table 4) implying that this locus has an important role in disease risk. The *SEMA6A* gene codes for semaphorin 6A, a member of the semaphorin protein family (48). Semaphorins were originally identified as axon guidance factors that are needed to direct neuronal axons to their appropriate targets (49). However, further studies have demonstrated that they have diverse roles in many physiologic processes unrelated to axon guidance, including vasculogenesis, cardiogenesis, osteoclastogenesis, tumor metastasis, and immune regulation (49,50). The functions of *SEMA6A* are not well characterized, but this gene is expressed in dendritic cells and expression of its product, semaphorin 6A, appears to be increased in Langerhans' cells in patients with Langerhans' cell histiocytosis (48). The mechanisms by which *SEMA6A* may influence GPA risk are unclear, but a possible link to the disease is consistent with recognized roles of other semaphorins, e.g., semaphorins 3A, 4D, 6D, and 7A, in the immune response as well as in autoimmune and allergic disorders (50,51). Further

studies are needed, however, to identify the disease-causal allele at this locus and to delineate other non-MHC GPA risk loci.

It is noteworthy that our GWAS data also provide nominal evidence of an association of GPA with 4 of 15 loci that have been previously shown in candidate gene analyses to be associated with risk for GPA (16–20,28). These include the *SERPINAI* (rs1956707; P = 0.009), *CTLA4* (rs231726 and rs3096851; P = 0.003 for both SNPs), *PTPN22* (rs2476601; P = 0.006), and *TGFβ1* (rs8110090; P = 0.03) loci. Among these loci, *SERPINA1* was the only non-MHC locus that achieved genome-wide significance in a combined discovery and replication data set from the recent GWAS of a European AAV population (30). However, the SNP that was most strongly associated in that study (rs7151526) is located 557 bp from, and is not in linkage disequilibrium with, rs1956707 and was not independently associated with risk in that population. Suggestive associations of the *PTPN22* and *CTLA4* loci with risk for GPA were also observed in the European GWAS as in our study. However, the peak association signals at each of these loci again emanated from different SNPs in the respective studies, and, in the European study, did not include the *PTPN22* rs2476601(T) allele that is considered to be the disease-causal variant in other autoimmune diseases (30). Similarly, the peak association signal detected within the *HLA* region also differed between the two studies.

These differences in GWAS results may reflect several factors, including differences in the patient cohorts (i.e., ANCA-positive GPA and microscopic polyangiitis in the European study versus ANCA-positive and -negative GPA in our study), in the array platforms (e.g., the peak *HLA* region SNP in each of the 2 GWAS was not tested in the other study), and in statistical power. Although our GWAS had statistical power of >80% to detect association with disease with an average minor allele frequency of 0.20 and an OR of 1.8 for cases as compared to controls, the power to detect associations with more modest effects, even in the combined analysis, was diminished (combined analysis >80% power to detect associations with an OR of ≥1.5).

Our North American study and the European study also differ with regard to the finding of a *SEMA6A* variant being significantly associated with risk for GPA (current study) and the findings suggesting that the *HLA–DP* and several non-MHC loci are more strongly associated with ANCA positivity than with disease status (European study). While effects of ANCA status on genetic associations were also explored in our study (Table 3), the inclusion of relatively low numbers of ANCA-negative patients in our GWAS, and of only PR3 ANCA–positive, rather than PR3 ANCA and myeloperoxidase ANCA–positive, cases makes it difficult to gauge the extent to which phenotypic disparities between case populations contribute to differences in the findings of the respective GWAS.

Genetic analyses of additional case cohorts are needed to more fully delineate the genetic basis of GPA. However, the current data identify the *SEMA6A* and *HLA* loci as significant contributors to GPA risk and demonstrate that the strongest association is *HLA* derived and is almost entirely attributable to a single allele, *HLA–DPB1*04*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by an Erna Baird Memorial Grant, the Vasculitis Foundation Canada, the Ontario Research Fund (RE-01-061), the Vasculitis Foundation, and the Vasculitis Clinical Research Consortium (NIH grants U54-RR-019497, U54-AR-47785, R01-AR-047799, and R01-AG025259). Dr. Monach's work was supported by the Arthritis Foundation. Dr. Mahr's work was supported in part by a grant from the Société Nationale Française de Médecine Interne. Dr. de Bakker is recipient of a Vidi Award from The Netherlands Organization for Scientific Research. Dr. Merkel's work was supported by a Mid-Career Development Award in Clinical Investigation (NIH–National Institute of Arthritis and Musculoskeletal and Skin Diseases grant K24-AR-02224). Dr. Siminovitch is a Tier 1 Canada Research Chair and the Sherman Family Chair in Genomic Medicine.

Dr. Hoffman has received consulting fees, speaking fees, and/or honoraria from Sanofi-Aventis, Roche, and Genentech (less than $10,000 each).

REFERENCES

1. Cotch MF, Hoffman GS, Yerg DE, Kaufman GI, Targonski P, Kaslow RA. The epidemiology of Wegener's granulomatosis: estimates of the five-year period prevalence, annual mortality, and geographic disease distribution from population-based data sources. Arthritis Rheum. 1996; 39:87–92. [PubMed: 8546743]
2. Mahr AD, Neogi T, Merkel PA. Epidemiology of Wegener's granulomatosis: lessons from descriptive studies and analyses of genetic and environmental risk determinants. Clin Exp Rheumatol. 2006; 24:S82–S91. [PubMed: 16859601]
3. Mukhtyar C, Flossmann O, Hellmich B, Bacon P, Cid M, Cohen-Tervaert JW, et al. on behalf of the European Vasculitis Study Group (EUVAS). Outcomes from studies of antineutrophil cytoplasm antibody associated vasculitis: a systematic review by the European League Against Rheumatism Systemic Vasculitis Task Force. Ann Rheum Dis. 2008; 67:1004–1010. [PubMed: 17911225]
4. Csernok E, Lamprecht P, Gross WL. Diagnostic significance of ANCA in vasculitis. Nat Clin Pract Rheumatol. 2006; 2:174–175. [PubMed: 16932679]
5. Chen M, Kallenberg CG. ANCA-associated vasculitides: advances in pathogenesis and treatment. Nat Rev Rheumatol. 2010; 6:653–664. [PubMed: 20924413]
6. Jennette JC, Falk RJ, Bacon PA, Basu N, Cid MC, Ferrario F, et al. 2012 revised International Chapel Hill Consensus Conference nomenclature of vasculitides. Arthritis Rheum. 2013; 65:1–11. [PubMed: 23045170]
7. Muniain MA, Moreno JC, Gonzalez Campora R. Wegener's granulomatosis in two sisters. Ann Rheum Dis. 1986; 45:417–421. [PubMed: 3718017]
8. Knudsen BB, Joergensen T, Munch-Jensen B. Wegener's granulomatosis in a family: a short report. Scand J Rheumatol. 1988; 17:225–227. [PubMed: 3175551]
9. Hay EM, Beaman M, Ralston AJ, Ackrill P, Bernstein RM, Holt PJ. Wegener's granulomatosis occurring in siblings. Br J Rheumatol. 1991; 30:144–145. [PubMed: 2012946]
10. Stoney PJ, Davies W, Ho SF, Paterson IC, Griffith IP. Wegener's granulomatosis in two siblings: a family study. J Laryngol Otol. 1991; 105:123–124. [PubMed: 2013723]
11. Rottem M, Cotch MF, Fauci AS, Hoffman GS. Familial vasculitis: report of 2 families. J Rheumatol. 1994; 21:561–563. [PubMed: 7911836]
12. Nowack R, Lehmann H, Flores-Suarez LF, Nanhou A, van der Woude FJ. Familial occurrence of systemic vasculitis and rapidly progressive glomerulonephritis. Am J Kidney Dis. 1999; 34:364–373. [PubMed: 10430990]
13. Knight A, Sandin S, Askling J. Risks and relative risks of Wegener's granulomatosis among close relatives of patients with the disease. Arthritis Rheum. 2008; 58:302–307. [PubMed: 18163522]
14. Weiner SR, Kwan LW, Paulus HE, Caro XJ, Weisbart RH. Twins discordant for Wegener's granulomatosis. Clin Exp Rheumatol. 1986; 4:389–390. [PubMed: 3791724]

15. Abdou NI, Kullman GJ, Hoffman GS, Sharp GC, Specks U, McDonald T, et al. Wegener's granulomatosis: survey of 701 patients in North America. Changes in outcome in the 1990s. J Rheumatol. 2002; 29:309–316. [PubMed: 11838848]
16. Jagiello P, Gencik M, Arning L, Wieczorek S, Kunstmann E, Csernok E, et al. New genomic region for Wegener's granulomatosis as revealed by an extended association screen with 202 apoptosis-related genes. Hum Genet. 2004; 114:468–477. [PubMed: 14968360]
17. Borgmann S, Endisch G, Urban S, Sitter T, Fricke H. A linkage disequilibrium between genes at the serine protease inhibitor gene cluster on chromosome 14q32.1 is associated with Wegener's granulomatosis. Clin Immunol. 2001; 98:244–248. [PubMed: 11161981]
18. Mahr AD, Edberg JC, Stone JH, Hoffman GS, St. Clair EW, Specks U, et al. for the Wegener's Granulomatosis Genetic Repository Research Group. Alpha1-antitrypsin deficiency-related alleles Z and S and the risk of Wegener's granulomatosis. Arthritis Rheum. 2010; 62:3760–3767. [PubMed: 20827781]
19. Giscombe R, Wang X, Huang D, Lefvert AK. Coding sequence 1 and promoter single nucleotide polymorphisms in the CTLA-4 gene in Wegener's granulomatosis. J Rheumatol. 2002; 29:950–953. [PubMed: 12022356]
20. Spriewald BM, Witzke O, Wassmuth R, Wenzel RR, Arnold ML, Philipp T, et al. Distinct tumour necrosis factor α, interferon γ, interleukin 10, and cytotoxic T cell antigen 4 gene polymorphisms in disease occurrence and end stage renal disease in Wegener's granulomatosis. Ann Rheum Dis. 2005; 64:457–461. [PubMed: 15708894]
21. Jagiello P, Aries P, Arning L, Wagenleiter SE, Csernok E, Hellmich B. The PTPN22 620W allele is a risk factor for Wegener's granulomatosis. Arthritis Rheum. 2005; 52:4039–4043. [PubMed: 16320352]
22. Fanciulli M, Norsworthy PJ, Petretto E, Dong R, Harper L, Kamesh L, et al. FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. Nat Genet. 2007; 39:721–723. [PubMed: 17529978]
23. Gencik M, Meller S, Borgmann S, Fricke H. Proteinase 3 gene polymorphisms and Wegener's granulomatosis. Kidney Int. 2000; 58:2473–2477. [PubMed: 11115080]
24. Wieczorek S, Hoffjan S, Chan A, Rey L, Harper L, Fricke H, et al. Novel association of the CD226 (DNAM-1) Gly307Ser polymorphism in Wegener's granulomatosis and confirmation for multiple sclerosis in German patients. Genes Immun. 2009; 10:591–595. [PubMed: 19536154]
25. Wieczorek S, Holle JU, Bremer JP, Wibisono D, Moosig F, Fricke H, et al. Contrasting association of a non-synonymous leptin receptor gene polymorphism with Wegener's granulomatosis and Churg-Strauss syndrome. Rheumatology (Oxford). 2010; 49:907–914. [PubMed: 20185531]
26. Bartfai Z, Gaede KI, Russell KA, Murakozy G, Muller-Quernheim J, Specks U. Different gender-associated genotype risks of Wegener's granulomatosis and microscopic polyangiitis. Clin Immunol. 2003; 109:330–337. [PubMed: 14697748]
27. Wieczorek S, Holle JU, Muller S, Fricke H, Gross WL, Epplen JT. A functionally relevant IRF5 haplotype is associated with reduced risk to Wegener's granulomatosis. J Mol Med. 2010; 88:413–421. [PubMed: 20049410]
28. Yu X, Wieczorek S, Franke A, Yin H, Pierer M, Sina C, et al. Association of UCP2 —866 G/A polymorphism with chronic inflammatory diseases. Genes Immun. 2009; 10:601–605. [PubMed: 19387457]
29. Murakozy G, Gaede KI, Ruprecht B, Gutzeit O, Schurmann M, Schnabel A, et al. Gene polymorphisms of immunoregulatory cytokines and angiotensin-converting enzyme in Wegener's granulomatosis. J Mol Med. 2001; 79:665–670. [PubMed: 11715070]
30. Lyons PA, Rayner TF, Trivedi S, Holle JU, Watts RA, Jayne DR, et al. Genetically distinct subsets within ANCA-associated vasculitis. N Engl J Med. 2012; 367:214–223. [PubMed: 22808956]
31. Chung SA, Xie G, Roshandel D, Sherva R, Edberg JC, Kravitz M, et al. Meta-analysis of genetic polymorphisms in granulomatosis with polyangiitis (Wegener's) reveals shared susceptibility loci with rheumatoid arthritis. Arthritis Rheum. 2012; 64:3463–3471. [PubMed: 22508400]
32. Leavitt RY, Fauci AS, Bloch DA, Michel BA, Hunder GG, Arend WP, et al. The American College of Rheumatology 1990 criteria for the classification of Wegener's granulomatosis. Arthritis Rheum. 1990; 33:1101–1107. [PubMed: 2202308]

33. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet. 2008; 40:616–622. [PubMed: 18385676]
34. The WGET Research Group. Design of the Wegener's Granulomatosis Etanercept Trial (WGET). Control Clin Trials. 2002; 23:450–468. [PubMed: 12161090]
35. The WGET Research Group. The Wegener's Granulomatosis Etanercept Trial (WGET): randomized, double-masked, placebo-controlled trial of etanercept for the induction and maintenance of remission. N Engl J Med. 2005; 352:351–361. [PubMed: 15673801]
36. Green RC, McNagny SE, Jayakumar P, Cupples LA, Benke K, Farrer LA, et al. Statin use and the risk of Alzheimer's disease: the MIRAGE study. Alzheimers Dement. 2006; 2:96–103. [PubMed: 19595865]
37. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81:559–575. [PubMed: 17701901]
38. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006; 38:904–909. [PubMed: 16862161]
39. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009; 5:e1000529. [PubMed: 19543373]
40. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, et al. Mapping copy number variation by population-scale genome sequencing. Nature. 2011; 470:59–65. [PubMed: 21293372]
41. Brown WM, Pierce J, Hilner JE, Perdue LH, Lohman K, Li L, et al. Overview of the MHC fine mapping data. Diabetes Obes Metab. 2009; 11(Suppl 1):2–7. [PubMed: 19143809]
42. Holdsworth R, Hurley CK, Marsh SG, Lau M, Noreen HJ, Kempenich JH, et al. The HLA dictionary 2008: a summary of HLA-A, -B, -C, -DRB1/3/4/5, and -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR, and -DQ antigens. Tissue Antigens. 2009; 73:95–170. [PubMed: 19140825]
43. Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD, et al. International HIV Controllers Study. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. Science. 2010; 330:1551–1557. [PubMed: 21051598]
44. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet. 2012; 44:291–296. [PubMed: 22286218]
45. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating the population attributable risk for multiple risk factors using case-control data. Am J Epidemiol. 1985; 122:904–914. [PubMed: 4050778]
46. Szyld P, Jagiello P, Csernok E, Gross WL, Epplen JT. On the Wegener granulomatosis associated region on chromosome 6p21.3. BMC Med Genet. 2006; 7:21. [PubMed: 16526951]
47. Heckmann M, Holle JU, Arning L, Knaup S, Hellmich B, Nothnagel M, et al. The Wegener's granulomatosis quantitative trait locus on chromosome 6p21.3 as characterised by tagSNP genotyping. Ann Rheum Dis. 2008; 67:972–979. [PubMed: 17967832]
48. Gautier G, de Saint-Vis B, Senechal B, Pin JJ, Bates EE, Caux C, et al. The class 6 semaphorin SEMA6A is induced by interferon-7 and defines an activation status of Langerhans cells observed in pathological situations. Am J Pathol. 2006; 168:453–465. [PubMed: 16436660]
49. Kumanogoh A, Kikutani H. Semaphorins and their receptors: novel features of neural guidance molecules. Proc Jpn Acad Ser B Phys Biol Sci. 2010; 86:611–620.
50. Ji JD, Ivashkiv LB. Roles of semaphorins in the immune and hematopoietic system. Rheumatol Int. 2009; 29:727–734. [PubMed: 19139899]
51. Takamatsu H, Okuno T, Kumanogoh A. Regulation of immune cell responses by semaphorins and their receptors. Cell Mol Immunol. 2010; 7:83–88. [PubMed: 20118971]

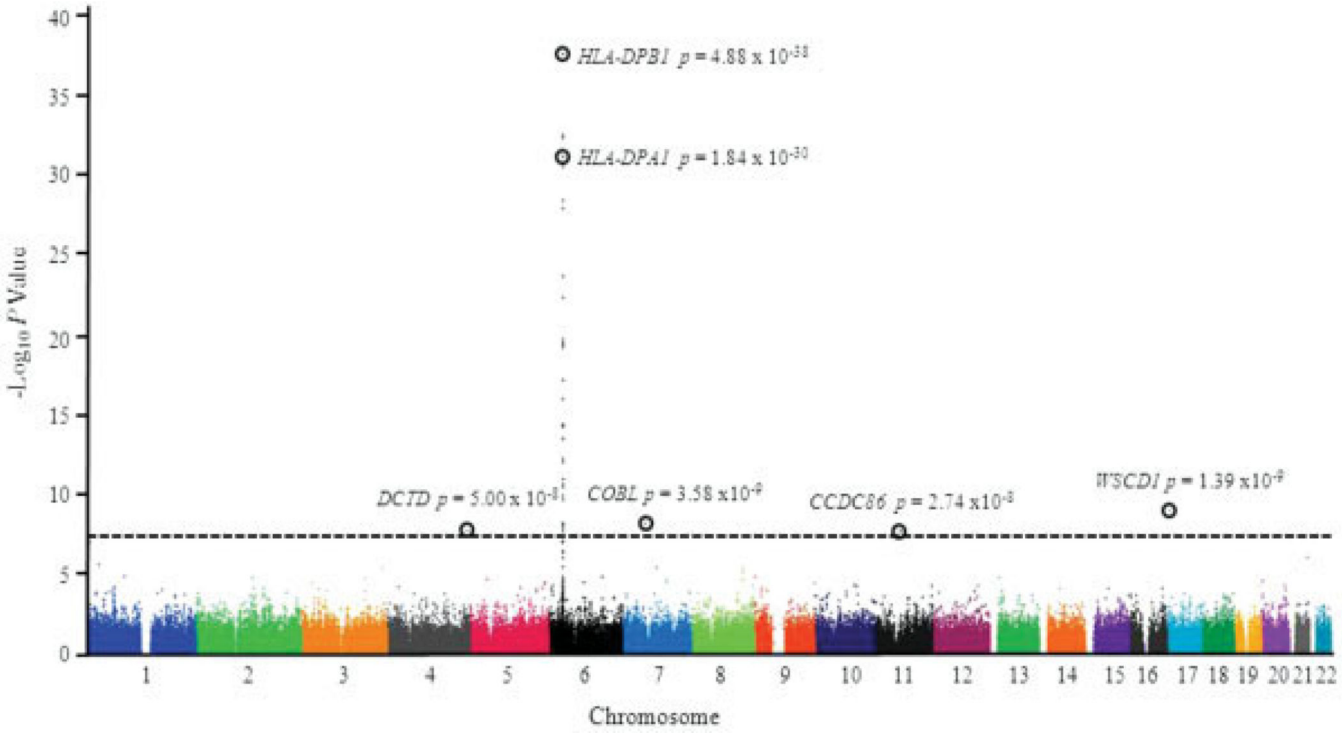

Figure 1.

Results of the granulomatosis with polyangiitis (Wegener's) genome-wide association screen. The $-\log_{10} P$ values for association of 287,802 single-nucleotide polymorphisms (SNPs) in 1,962 subjects are shown based on chromosomal locations. The y-axis represents the level of significance (determined using Plink) for each SNP on each chromosome shown along the x-axis. Dashed line shows the threshold for significance of the genome-wide association ($P < 5 \times 10^{-8}$). Circles with dots indicate the SNPs most significantly associated with disease, in the major histocompatibility complex (MHC) region genes *HLA–DPB1* and *HLA–DPA1* (encoding class II MHC DP β chain 1 and α chain 1, respectively) and in the non-MHC region genes *DCTD* (encoding deoxycytidylate deaminase), *COBL* (encoding protein cordon bleu), *CCDC86* (encoding coiled-coil domain-containing protein 86), and *WSCD1* (encoding water-soluble carbodiimide domain-containing protein 1). Color figure can be viewed in the online issue, which is available at http:onlinelibrary.wiley.com/doi/10.1002/acr.38036/abstract.

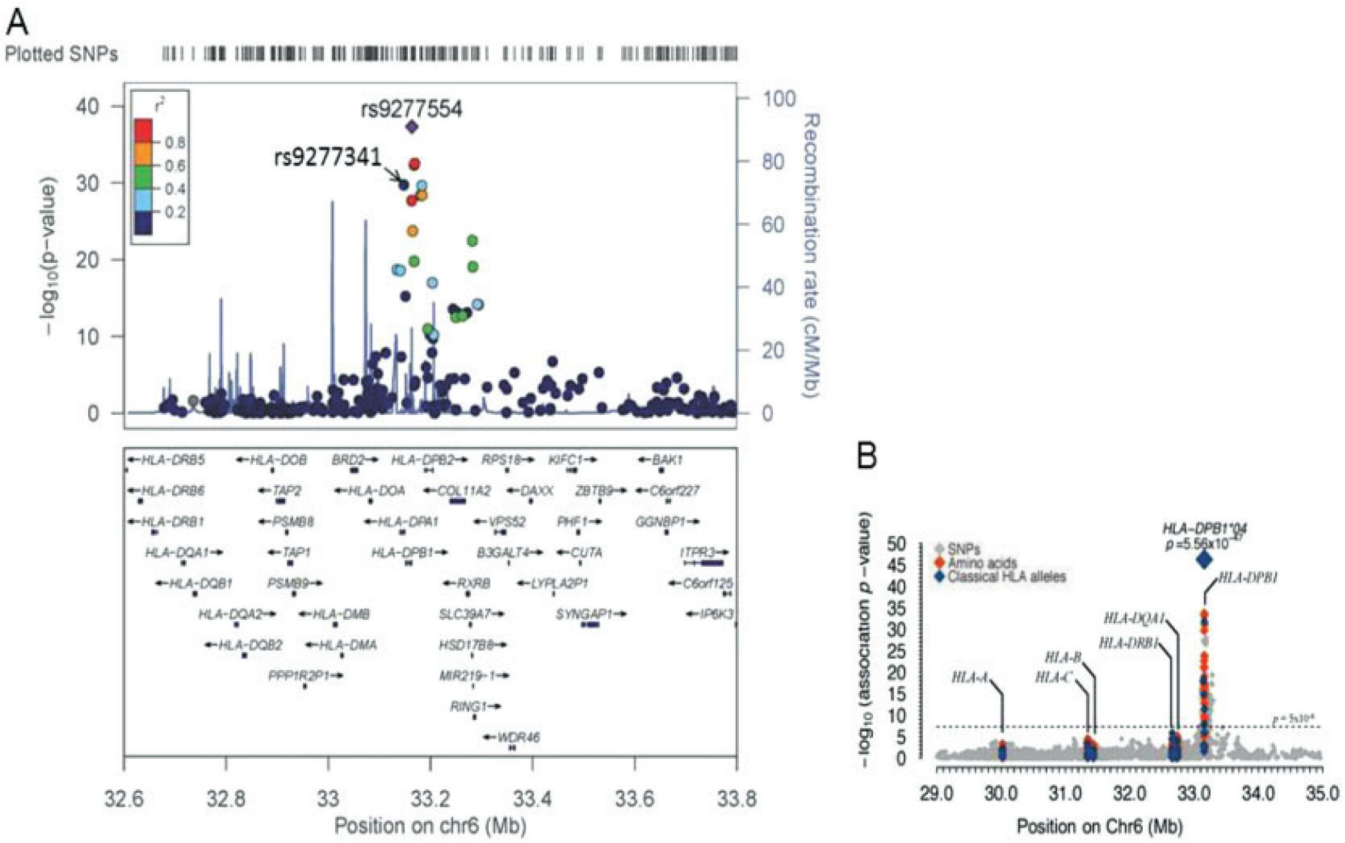

Figure 2.

Association plots from genome-wide analyses across the major histocompatibility complex (MHC) region. **A**, The -$\log_{10}$ *P* values of 283 genotyped single-nucleotide polymorphisms (SNPs) across the class II MHC region are plotted against their physical position on chromosome 6 (chr6). LocusZoom software was used for analysis. Estimated recombination rates from the HapMap phase II Centre d' Étude du Polymorphisme Humain study population of Utah residents with ancestry from northern and western Europe (CEU) population show the local linkage disequilibrium (LD) structure. Colors in the upper panel indicate the extent of LD with the top *HLA–DPB1* SNP (rs9277554) (♦), according to a scale of r^2= 0 to r^2= 1 based on pairwise r^2 values from HapMap CEU. The *HLA–DPA1* rs9277341 is indicated by an **arrow**. The lower panel shows gene annotation data from the University of California at Santa Cruz genome browser. **B**, The-$\log_{10}$ *P* values of 1,160 genotyped SNPs and 7,261 imputed SNPs, 424 classic HLA alleles, and 1,276 corresponding amino acid polymorphisms across the MHC region are plotted against their physical position on chromosome 6. Dotted line shows the threshold for significance of the genome-wide association ($P < 5 \times 10^{-8}$).

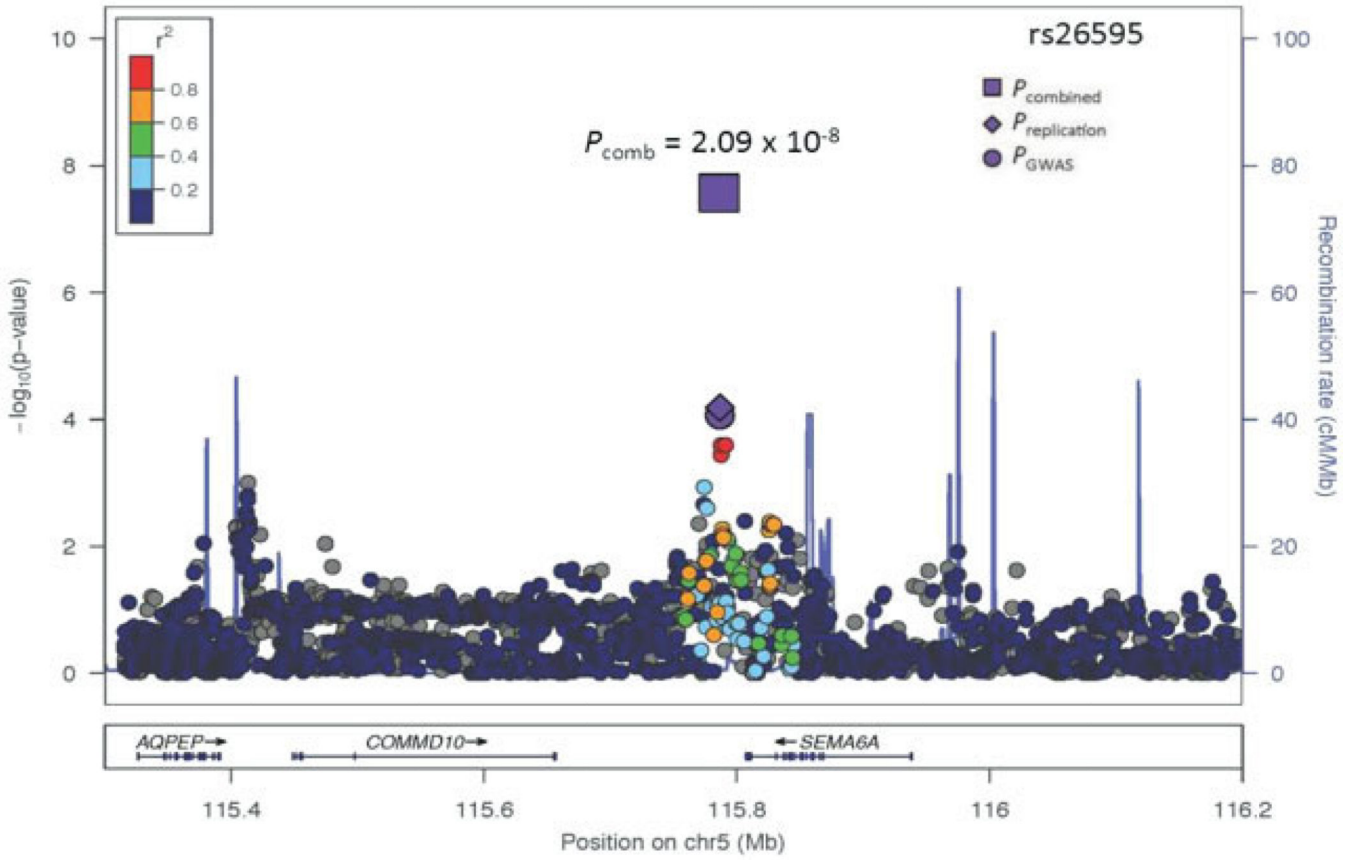

Figure 3. cctf

Association plot and LD structure of the disease-associated region surrounding *SEMA6A*. The -$\log_{10}P$ values of 135 genotyped SNPs and 3,880 imputed SNPs surrounding *SEMA6A* are plotted against their physical position on chromosome 5. LocusZoom software was used for analysis. Estimated recombination rates from the HapMap phase II CEU population show the local LD structure. Colors in the upper panel indicate the extent of LD with the top SNP (rs26595), according to a scale of r^2= 0 to r^2= 1 based on pairwise r^2 values from HapMap CEU. The lower panel shows gene annotation data from the University of California at Santa Cruz genome browser. GWAS = genome-wide association study (see Figure 2 for other definitions). Color figure can be viewed in the online issue, which is available at http:onlinelibrary.wiley.com/doi/10.1002/acr.38036/abstract.

Table 1

Results of genome-wide association (stage 1) and replication (stage 2a) analyses of GPA associations*

						Genome-wide association analysis (stage 1)				Replication analysis, US cohort (stage 2a)				Combined analysis (n = 750 patients, 1,820 controls)	
						Minor allele frequency				Minor allele frequency					
SNP	Chr.	Location, bp	Gene	Minor allele	Major allele	Patients (n = 459)	Controls (n = 1,503)	Plink *P* value	OR (95% CI)	Patients (n = 291)	Controls (n = 317)	Plink *P* value	OR (95% CI)	*P*	OR (95% CI)
HLA region															
rs9277554	6p21.3	33163516	*HLA–DPB1*	T	C	0.083	0.292	4.88×10^{-38}	0.22 (0.17–0.28)	0.106	0.274	1.40×10^{-13}	0.32 (0.23–0.43)	1.92×10^{-50}	0.24 (0.20–0.30)
rs3117222	6p21.3	33168927	*HLA–DPB1*	A	G	0.070	0.254	3.05×10^{-33}	0.22 (0.17–0.29)						
rs3128917	6p21.3	33167974	*HLA–DPB1*	G	T	0.070	0.253	4.92×10^{-33}	0.22 (0.17–0.29)						
rs9277341	6p21.3	33147603	*HLA–DPA1*	C	T	0.124	0.320	1.84×10^{-30}	0.30 (0.25–0.38)	0.149	0.29	3.84×10^{-9}	0.43 (0.32–0.57)	2.18×10^{-39}	0.33 (0.28–0.39)
rs3130215	6p21.3	33182941	*HLA–DPB2*	A	G	0.625	0.406	2.37×10^{-30}	2.42 (2.08–2.82)						
rs2064478	6p21.3	33180244	*HLA–DPB2*	A	G	0.061	0.225	4.29×10^{-29}	0.22 (0.17–0.30)						
rs3117230	6p21.3	33183613	*HLA–DPB2*	C	T	0.061	0.225	4.29×10^{-29}	0.22 (0.17–0.30)						
rs9277535	6p21.3	33162839	*HLA–DPB1*	G	A	0.071	0.237	2.12×10^{-28}	0.24 (0.19–0.32)						
rs9277565	6p21.3	33164875	*HLA–DPB1*	T	C	0.057	0.199	1.91×10^{-24}	0.24 (0.18–0.32)						
rs439205	6p21.3	33281820	*HSD17B8*	T	C	0.092	0.242	3.51×10^{-23}	0.31 (0.24–0.39)						
rs2281389	6p21.3	33167774	*HLA–DPB1*	C	T	0.048	0.168	1.69×10^{-20}	0.24 (0.18–0.34)						
rs421446	6p21.3	33282761	*HSD17B8*	C	T	0.132	0.279	8.90×10^{-20}	0.39 (0.31–0.48)						
rs2395309	6p21.3	33134224	*HLA–DPA1*	G	A	0.053	0.176	2.15×10^{-19}	0.27 (0.20–0.36)						
rs3077	6p21.3	33141000	*HLA–DPA1*	C	T	0.053	0.175	2.68×10^{-19}	0.27 (0.20–0.36)						
rs3117016	6p21.3	33203494	*HLA–DPB2*	T	C	0.246	0.401	1.09×10^{-17}	0.48 (0.41–0.57)						
rs987870	6p21.3	33150858	*HLA–DPB1*	C	T	0.041	0.141	6.09×10^{-16}	0.26 (0.19–0.37)						
rs213213	6p21.3	33291708	*RING1*	A	G	0.445	0.305	6.98×10^{-15}	1.83 (1.57–2.14)						
rs213212	6p21.3	33293896	*RING1*	G	T	0.406	0.269	7.63×10^{-15}	1.85 (1.58–2.17)						
rs986521	6p21.3	33244123	*COL11A2*	C	T	0.376	0.245	2.91×10^{-14}	1.85 (1.57–2.16)						
rs2855425	6p21.3	33252351	*COL11A2*	C	T	0.410	0.276	7.77×10^{-14}	1.80 (1.54–2.11)						
rs6531	6p21.3	33271429	*RXRB*	C	T	0.410	0.276	8.48×10^{-14}	1.80 (1.54–2.11)						
rs2855459	6p21.3	33262634	*COL11A2*	T	C	0.050	0.136	2.14×10^{-13}	0.32 (0.23–0.44)						

						Genome-wide association analysis (stage 1)				Replication analysis, US cohort (stage 2a)				Combined analysis (n = 750 patients 18,820 controls)	
						Minor allele frequency				Minor allele frequency					
SNP	Chr.	Location, bp	Gene	Minor allele	Major allele	Patients (n = 459)	Controls (n = 1,503)	Plink *P* value	OR (95% CI)	Patients (n = 291)	Controls (n = 317)	Plink *P* value	OR (95% CI)	*P*	OR (95%
rs2855430	6p21.3	33249258	*COL11A2*	T	C	0.051	0.137	3.28×10^{-13}	0.33 (0.24–0.45)						
rsl883414	6p21.3	33194426	*HLA–DPB2*	T	C	0.193	0.309	1.13×10^{-11}	0.53 (0.44–0.64)						
rs3117008	6p21.3	33204252	*HLA–DPB2*	T	C	0.373	0.498	4.90×10^{-11}	0.60 (0.51–0.70)						
rs6901221	6p21.3	33206254	*HLA–DPB2*	C	A	0.073	0.157	6.08×10^{-11}	0.42 (0.32–0.55)						
rs4713607	6p21.3	33198814	*HLA–DPB2*	A	G	0.373	0.497	6.70×10^{-11}	0.60 (0.52–0.70)						
rs3117004	6p21.3	33204744	*HLA–DPB2*	C	T	0.218	0.327	1.90×10^{-10}	0.57 (0.48–0.68)						
rs3129274	6p21.3	33202847	*HLA–DPB2*	G	A	0.422	0.317	1.35×10^{-8}	1.56 (1.34–1.82)						
rs763469	6p21.3	33112365	*HLA–DOA*	A	G	0.235	0.151	1.46×10^{-8}	1.70 (1.41–2.04)						
rs3130604	6p21.3	33093030	*HLA–DOA*	G	A	0.231	0.150	4.39×10^{-8}	1.67 (1.39–2.02)						
rs2301226	6p21.3	33142574	*HLA–DPA1*	T	C	0.071	0.140	4.85×10^{-8}	0.48 (0.36–0.62)						
Non-HLA region															
rs7503953	17pl3.2	6082401	*WSCD1*	A	C	0.264	0.171	1.39×10^{-9}	1.72 (1.44–2.06)	0.142	0.145	0.881	0.98 (0.71–1.35)	1.93×10^{-7}	1.50 (1.2
rsl949829	7pl2.1	51505381	*COBL*	T	C	0.112	0.055	3.58×10^{-9}	2.19 (1.68–2.86)	0.064	0.060	0.777	1.07 (0.67–1.71)	4.19×10^{-7}	1.78 (1.4
rs595018	11ql2.2	60348852	*CCDC86*	A	G	0.304	0.216	2.74×10^{-8}	1.61 (1.36–1.90)	0.221	0.199	0.354	1.14 (0.86–1.51)	1.60×10^{-7}	1.46 (1.2
rs4862110	4q35.1	183988023	*DCTD*	C	T	0.261	0.174	5.00×10^{-8}	1.63 (1.36–1.94)	0.174	0.175	0.966	0.99 (0.74–1.34)	2.14×10^{-6}	1.44 (1.2

*The single-nucleotide polymorphisms (SNPs) are listed in order of decreasing significance as indicated by Plink *P* values; only those with *P* values of less than 5.00×10^{-8} are shown. *P* values and odds ratios (ORs) are for the comparison of allele frequency between patients and controls, as calculated with the Plink method. For the combined analysis (stages 1 and 2a), *P* values and ORs were calculated using the Cochran-Mantel-Haenszel method of combining allele frequency counts. GPA = granulomatosis with polyangiitis (Wegener's); Chr. = chromosome; 95% CI = 95% confidence interval.

Table 2

Results of the genome-wide association (stage 1) and replication (stage 2b) analyses of GPA associations*

						Genome-wide association analysis (stage 1)				Replication analysis, US cohort (stage 2b)				Combined analysis (n = 987 patients, 2,731 controls)	
						Minor allele frequency				Minor allele frequency					
SNP	Chr.	Location, bp	Gene	Minor allele	Major allele	Patients (n = 459)	Controls (n = 1,503)	Plink *P* value	OR (95% CI)	Patients (n = 528)	Controls (n = 1,228)	Plink *P* value	OR (95% CI)	Plink *P* value	OR (95% CI)
rs7585252	2	134256798	*FLJ34870*	G	A	0.448	0.375	9.69×10^{-5}	1.35 (1.16–1.57)	0.420	0.382	0.034	1.17 (1.01–1.36)	1.74×10^{-5}	1.26 (1.13–1.40)
rs1579908	3	26966364	*NEK10*	T	G	0.197	0.141	3.97×10^{-5}	1.51 (1.24–1.85)	0.120	0.161	1.96×10^{-3}	0.71 (0.57–0.88)	0.498	1.05 (0.91–1.21)
rs9842552	3	41372045	*CTNNB1*	T	C	0.261	0.200	1.63×10^{-5}	1.46 (1.23–1.74)	0.224	0.227	0.859	0.98 (0.83–1.17)	9.70×10^{-3}	1.17 (1.04–1.33)
rs1512770	3	145949619	*C3orf58*	C	A	0.338	0.413	2.86×10^{-5}	0.72 (0.61–0.84)	0.416	0.396	0.284	1.04 (0.94–1.25)	0.026	0.89 (0.79–0.99)
rs26595	**5**	**115787389**	***SEMA6A***	**C**	**I**	**0.363**	**0.440**	$\mathbf{9.58 \times 10^{-5}}$	**0.74 (0.63–0.86)**	**0.438**	**0.511**	$\mathbf{7.89 \times 10^{-5}}$	**0.75 (0.65–0.86)**	$\mathbf{2.09 \times 10^{-8}}$	**0.74 (0.67–0.82)**
rs10515667	5	152616986	*GRIA1*	T	C	0.145	0.098	8.11×10^{-5}	1.56 (1.25–1.94)	0.102	0.090	0.265	1.15 (0.90–1.47)	3.01×10^{-4}	1.35 (1.15–1.59)
rs3853242	5	166349953	*WWC1*	G	A	0.396	0.472	8.55×10^{-5}	0.74 (0.63–0.86)	0.480	0.427	3.64×10^{-3}	1.24 (1.08–1.43)	0.480	1.04 (0.94–1.15)
rs2447436	8	140580708	*KCNK9*	T	C	0.150	0.102	3.25×10^{-5}	1.58 (1.27–1.97)	0.113	0.128	0.204	0.86 (0.69–1.08)	0.066	1.16 (0.99–1.35)
rs705669	9	137608215	*PAEP*	G	A	0.184	0.249	8.53×10^{-5}	0.69 (0.57–0.83)	0.258	0.290	0.055	0.86 (0.73–1.00)	2.52×10^{-5}	0.77 (0.68–0.87)
rs4937769	11	132880858	*OPCML*	C	A	0.086	0.136	5.75×10^{-5}	0.60 (0.46–0.77)	0.122	0.115	0.510	1.08 (0.86–1.35)	0.016	0.82 (0.69–0.96)
rs6140836	20	903412	*RSP04*	C	T	0.071	0.121	7.26×10^{-5}	0.58 (0.44–0.76)	0.124	0.134	0.444	0.94 (0.74–1.14)	7.09×10^{-4}	0.75 (0.63–0.89)
rs6023640	20	52892212	*DOK4*	T	G	0.281	0.217	8.98×10^{-5}	1.40 (1.18–1.66)	0.266	0.236	0.054	1.18 (1.00–1.39)	2.73×10^{-5}	1.29 (1.14–1.45)

*The replication SNPs are listed in order of genomic location on the chromosome. *P* values and ORs are for the comparison of allele frequency between patients and controls, as calculated with the Plink method. For the combined analysis (stages 1 and 2b), *P* values and ORs were calculated using the Cochran-Mantel-Haenszel method of combining allele frequency counts. The most significant association SNP is shown in boldface. See Table 1 for definitions.

Table 3

Effects of cANCA status on *HLA–DPB1*, *HLA–DPA1*, and *SEMA6A* associations with GPA*

Variant	Chr.	Position	Gene	Minor allele	Major allele	Minor allele frequency			OR(P)	
						cANCA-positive patients (n = 578/768)	cANCA-negative patients (n = 73/125)	Controls (n = 1,820/2,731)	cANCA-positive patients vs. controls	cANCA-negative patients vs. controls
Replication stage 2a										
rs9277554	6p21.3	33163516	*HLA–DPB1*	T	C	0.061	0.292	0.289	0.16 (4.77×10^{-57})	1.02 (0.945)
rs9277341	6p21.3	33147603	*HLA–DPA1*	C	T	0.112	0.295	0.315	0.27 (2.30×10^{-42})	0.91 (0.599)
Replication stage 2b										
rs26595	5q23.1	115787389	*SEMA6A*	C	T	0.405	0.452	0.472	0.76 (3.40×10^{-6})	0.92 (0.532)

* n values are the number of subjects in replication analysis stage 2a/stage 2b. *P* values were calculated from allele distribution. cANCA = cytoplasmic antineutrophil cytoplasmic antibody (see Table 1 for other definitions).