

ARTICLE

Common coding variants in the HLA-DQB1 region confer susceptibility to age-related macular degeneration

Eric Jorgenson^{*1}, Ronald B Melles², Thomas J Hoffmann^{3,4}, Xiaoming Jia⁵, Lori C Sakoda¹, Mark N Kvale³, Yambazi Banda³, Catherine Schaefer¹, Neil Risch^{1,3,4} and Ling Shen¹

Age-related macular degeneration (AMD) risk variants in the complement system point to the important role of immune response and inflammation in the pathogenesis of AMD. Although the human leukocyte antigen (HLA) region has a central role in regulating immune response, previous studies of genetic variation in HLA genes and AMD have been limited by sample size or incomplete coverage of the HLA region by first-generation genotyping arrays and imputation panels. Here, we conducted a large-scale HLA fine-mapping study with 4841 AMD cases and 23 790 controls of non-Hispanic white ancestry from the Kaiser Permanente Genetic Epidemiology Research on Adult Health and Aging cohort. Genotyping was conducted using custom Affymetrix Axiom arrays, with dense coverage of the HLA region. Classic HLA polymorphisms were imputed using SNP2HLA, which utilizes a large reference panel to provide improved imputation accuracy of variants in this region. We examined a total of 6937 SNPs and 172 classical HLA alleles, conditioning on established AMD risk variants, which revealed novel associations with two non-synonymous SNPs in perfect linkage disequilibrium, rs9274390 and rs41563814 (odds ratio (OR) = 1.21; $P = 1.4 \times 10^{-11}$) corresponding to amino-acid changes at position 66 and 67 in *HLA-DQB1*, respectively, and the DQB1*02 classical HLA allele (OR = 1.22; $P = 3.9 \times 10^{-10}$) with the risk of AMD. We confirmed these association signals, again conditioning on established risk variants, in the MMAP data set of subjects with advanced AMD (rs9274390/rs41563814: OR = 1.28; $P = 1.30 \times 10^{-3}$, DQB1*02: OR = 1.32; $P = 9.00 \times 10^{-4}$). These findings support a role of HLA class II alleles in the risk of AMD.

European Journal of Human Genetics (2016) 24, 1049–1055; doi:10.1038/ejhg.2015.247; published online 6 January 2016

INTRODUCTION

Age-related macular degeneration (AMD) is a complex, late-onset vision disorder, which is the leading cause of blindness among the elderly in developed countries.^{1–4} Immune response and inflammation play a causal role in the pathogenesis and progression of AMD.^{5–11} The human leukocyte antigen (HLA) region on chromosome 6p21.31 encodes gene products that regulate immune response. Drusen, a pathologic hallmark of AMD, contain HLA class II antigens,¹² and increased HLA class II immunoreactivity has been associated with drusen formation,¹³ suggesting that HLA may influence the development of AMD. Although a few small candidate gene studies found that certain HLA class I and class II polymorphisms were associated with a predisposition to AMD,^{14,15} the largest genome-wide association study (GWAS) meta-analysis conducted by the AMD Gene Consortium, which included more than 77 000 subjects, did not identify any HLA polymorphisms that were independently associated with the risk of AMD.¹⁶

The AMD Consortium meta-analysis, although well powered to detect common variants of small effect, included a number of studies that utilized genotyping arrays, with array density ranging from 165 770 to 668 238 genotyped SNPs after QC, with incomplete coverage of genetic variants in the HLA region. In addition,

imputation was conducted using the HapMap reference panel, which does not include the majority of known HLA variants, limiting the ability of the individual studies to infer HLA genotypes. Because of this lack of coverage in the HLA region, and the small sample sizes of earlier candidate gene studies, and thus, limited statistical power, the effect of HLA polymorphisms on AMD susceptibility remains poorly understood.

In this study, we overcome the limitations of previous investigations of HLA polymorphisms and their association with AMD. We examined the association of genetic variants in HLA and the risk of AMD and its subtypes, that is, AMD unspecified, non-exudative AMD, and choroidal neovascularization (CNV), in the Kaiser Permanente Genetic Epidemiology Research on Adult Health and Aging (GERA) cohort. GERA subjects were genotyped using custom Affymetrix Axiom arrays, with dense coverage of the HLA region. We then imputed HLA amino acids and classical HLA alleles using SNP2HLA,¹⁷ an HLA-specific imputation tool that utilizes a much larger reference panel of subjects with European ancestry than the HapMap or 1000 genomes to achieve greater imputation accuracy, and consequently improve the power to investigate the association of genetic variants in this region. Finally, we confirmed significant

¹Kaiser Permanente Northern California Division of Research, Oakland, CA, USA; ²Department of Ophthalmology, Kaiser Permanente Northern California Redwood City Medical Center, Redwood City, CA, USA; ³Institute for Human Genetics, University of California–San Francisco, San Francisco, CA, USA; ⁴Department of Epidemiology and Biostatistics, University of California–San Francisco, San Francisco, CA, USA; ⁵Department of Neurology, University of California–San Francisco, San Francisco, CA, USA

*Correspondence: Dr E Jorgenson, Kaiser Permanente Northern California Division of Research, 2000 Broadway, Oakland, CA 94612, USA. Tel: +1 510 8 3473; Fax: +1 510 891 3836; E-mail: eric.jorgenson@kp.org

Received 25 March 2015; revised 21 September 2015; accepted 15 October 2015; published online 6 January 2016

associations in the publicly available MMAP data set of subjects with AMD diagnoses.

MATERIALS AND METHODS

Setting

Kaiser Permanente Medical Care Plan, Northern California Region (KPNC), is a non-profit integrated health-care delivery organization with an active membership of 3.5 million people covering about 30% of the population. The membership is representative of the population of northern California, with the exception of extremes of the socioeconomic spectrum.¹⁸ KPNC provides comprehensive care, including coverage of optometry and ophthalmologic care. In 1995, KPNC instituted a comprehensive electronic health record (EHR) system, which records diagnoses, prescriptions, and lab results from all inpatient and outpatient encounters. KPNC has high membership retention, with over 90% of those over age 65 years, and 66% of all active members as of June 2012, having five or more years of retrospective membership.

Kaiser permanente GERA cohort

The GERA cohort is comprised of 110 266 KPNC adult members. A detailed description of the cohort and study design and genetic variant and AMD phenotype data can be found in dbGaP (Study Accession: phs000674.v1.p1). Briefly, subjects enrolled through participation in a health survey of all adult members of KPNC in 2007 and sent in (Oragene) saliva samples with a written consent authorizing use of their biospecimens, survey data, and data from EHR for studies of genetic and environmental influences on health and disease. Survey data included information on demographic factors, behaviors, and self-reported health. A total of 102 998 samples were successfully assayed and passed genotyping quality control (QC).^{19,20} The average age of the GERA cohort at the saliva sample collection was 62.9 (SD=13.8) and had high KPNC membership retention (Supplementary Table 1). In this study, we focused on unrelated non-Hispanic white subjects in the GERA cohort who were at least 65 years of age as of 30 June 2013 ($N = 53\,449$). Relatives were identified from kinship analysis with KING²¹ and removed such that no first-degree relationships remained. The AMD cases ($n = 4841$) were preferentially included if the relative pair was case and control and otherwise related individuals in pairs were removed at random. This resulted in a total of 28 631 unrelated individuals.

AMD phenotypes and control definitions

AMD was diagnosed by KPNC ophthalmologists after a dilated eye examination of the retina. Optical coherence tomography (OCT) and fluorescein angiography were typically performed when there was evidence of subretinal fluid or hemorrhage in the retina. Ophthalmologists determined AMD subtype based on patient age, drusen number, drusen size, evidence of outer retinal atrophy and the results from OCT or fluorescein angiography.

We established four case categories: overall AMD, AMD unspecified, non-exudative AMD, and CNV. Subjects were classified as overall AMD cases if they had two or more AMD-related diagnoses during 1995–2013 – AMD unspecified (ICD9 362.50), non-exudative macular degeneration (ICD9 362.51), or exudative macular degeneration (ICD9 362.52) that characterizes the presence of CNV – and at least one of those diagnoses was made by a KPNC ophthalmologist.²² Overall, AMD cases with one or more diagnosis of 362.52 by KPNC ophthalmologists were then sub-categorized as CNV. Remaining cases with a diagnosis of 362.51 by ophthalmologists were classified as non-exudative AMD. To assess the genetic risks associated with these diagnoses and validate the AMD phenotypes defined through EHR data, cases identified through diagnosis of AMD unspecified (ICD9 362.50) alone, who were likely to have mild non-exudative AMD, were categorized in the AMD unspecified category, but not in the non-exudative AMD category. Controls were free of any AMD-related diagnoses, including drusen (ICD9 362.57), had one or more visits to ophthalmology clinic during 2008–2013 and were at least 65 years of age as of 30 June 2013. On average, these AMD cases had 14.7 (SD=15.7) diagnoses and 75% had 5 or more AMD diagnoses. The consistent diagnoses from repeated examinations support the reliability of the diagnostic categories used by the study. Furthermore, independent, blinded chart review of 336

subjects with AMD conducted by an ophthalmologist (RBM) confirmed all of the 336 overall AMD cases, including 145 of the 154 CNV (94.2%, 95% CI: 89.2 to 97.3%) and 80 of the 82 non-exudative AMD cases without CNV (97.6% CI: 91.5–99.7%). Eight of the non-exudative AMD cases were found to have progressed to geographic atrophy. The chart audit to confirm the accuracy of the AMD phenotype was based on clinical notes, a history of intraocular injection treatment with ranibizumab or bevacizumab, and a review of imaging studies including fundus photographs, OCT or fluorescein angiography. 98% of CNV, 66% of the non-exudative AMD without CNV, and 78% of overall AMD had at least one of the aforementioned studies available for review. Overall, the validation study demonstrated a high positive-predictive value for the phenotype definitions in this study.

Genotyping, QC, imputation, and genetic ancestry of the GERA cohort

Genotyping was based on the custom Affymetrix Axiom EUR array with 674 518 SNPs as previously described (dbGaP Study Accession: phs000674.v1.p1).^{19,20} Samples with dish QC <0.82 or initial genotype call rate <0.97 were excluded, resulting in a total of 83 285 individuals of European ancestry.²³ To improve genotype calls, SNPs were re-called within packages of plates assayed under similar conditions (array type, reagent, hybridization time, and DNA concentration). SNPs were removed if either package call rate or overall call rate (across packages) was below 90%. Additional SNP exclusion criteria were (1) large allele frequency (FRQ) variance across packages – defined as the ratio of overall variance of the SNP allele FRQ across packages to the sample SNP heterozygosity (total sample variance) (<31); (2) large allele FRQ differences between males and females (>0.15) for autosomal SNPs; and (3) poor concordance among duplicates (<0.24). These QC measures removed a total of 4431 SNPs.

Following QC, we conducted imputation of classical HLA alleles and amino-acid polymorphisms at class I (HLA-A, -B, -C) and class II (-DPA1, -DPB1, -DQA1, -DQB1, and -DRB1) loci using SNP2HLA, which utilizes a reference panel of 5225 European descent individuals from the T1DGC reference panel that has been described in detail by Jia *et al.*¹⁷ In total, we examined 6937 SNPs, 1000 amino-acid changes, and 172 classical HLA alleles in the HLA region with minor allele FRQ >0.01 and imputation info metrics >0.80. Because a number of AMD risk loci have large effects on the risk of disease, we also imputed previously established AMD risk variants of large effect in *CFH*, *ARMS2/HTRA1*, *CFB/SKIV2L*, and *C3* gene regions based on the 1000 Genomes Project²⁴ reference panel (March 2012 release) using IMPUTE2 v2.3.0 and standard procedures.²⁵ The *CFB/SKIV2L* locus is located within the HLA gene region, which may potentially obscure other association signals within this region because of linkage disequilibrium.

EIGENSTRAT²⁶ was used to compute eigenvectors with 41 228 high-quality SNPs that were common among all arrays and the Human Genome Diversity Project, as has been described (dbGaP phs000674.v1.p1).²⁷ These principal components (PCs) were used in the analysis to adjust for genetic ancestry (Supplementary Figure 1).

Replication data set

To confirm findings from the GERA cohort, we utilized data from the publicly available MMAP study (dbGaP phs000182.v3.p1), with 2185 AMD cases and 1155 controls. As described above, we conducted imputation of the HLA region using both the 1000 genomes and SNP2HLA T1DGC reference panels. We then used the imputed variants association analyses to replicate findings.

Statistical analysis

Analyses were conducted using PLINK²⁸ v1.07 and R.²⁹ For each AMD category, we tested single-marker associations in a logistic regression model adjusted for age, sex, and the first 10 ancestry PCs using allele counts for typed SNPs and imputed dosages for the imputed SNPs and a log-additive genetic model. In a previous paper,²² we examined the effect of previously reported AMD risk SNPs in the GERA cohort, and we provide the genome-wide genomic control lambda value (1.095) and QQ plot (Supplementary Figure 2) here.

Because the purpose of this study was to determine whether there were variants in the HLA region that were associated independently from the previously established AMD loci, we conditioned on the previously reported SNPs in the *CFH*, *ARMS2/HTRA1*, *CFB/SKIV2L*, and *C3* gene regions, specifically, by including rs1061170, rs10490924, rs429608, and rs2230199 in the regression model. We used the same set of covariates to test multi-allelic variants, including classical HLA alleles, using an omnibus test as described by Jia *et al*. We then included the genome-wide significant HLA risk SNPs identified in this analysis (rs9274390 and rs41563814 are in perfect LD, so one SNP was included in this step), in a second conditional model to determine whether there was evidence for additional signals in the region. We examined the top associations of genotyped SNPs by inspecting the cluster plots, call rates, and Hardy–Weinberg Equilibrium *P*-values of the genotyped SNPs. We conducted a sensitivity analysis of our top SNP association findings in a more homogeneous subset of subjects (> -0.02 on PC1 and < 0.02 on PC2, Supplementary Figure 1).

RESULTS

Among non-Hispanic white GERA cohort participants, we identified 4841 overall AMD cases and 23 790 controls. In all, 56% of the overall AMD cases were classified as AMD unspecified, 20% were non-exudative, and 24% were CNV. Cases were older and more likely to be female compared with controls (Table 1). The most strongly associated SNP in the HLA region was rs429608 (Figure 1a and Supplementary Figure 3A), an established AMD risk locus in the *CFB/SKIV2L* gene region, which showed an odds ratio of 1.55 with overall AMD, with the strongest effect on CNV (odds ratio (OR) = 2.03), followed by non-exudative AMD (OR = 1.56), and AMD unspecified (OR = 1.40).

To determine whether additional SNPs were associated with overall AMD, independent of rs429608, we conducted a conditional analysis, adjusting for previously reported risk variants in the gene regions of *CFH*, *ARMS2/HTRA1*, *CFB/SKIV2L*, and *C3*. In this analysis, we identified statistically significant associations with missense SNPs and corresponding amino-acid changes at position 66 and 67 in *HLA-DQB1* and at position 73 in *HLA-DRB1* (Table 2, Figure 1b, and Supplementary Figure 3B). Although all of these SNPs were in high LD with each other, rs9274390 and rs41563814 in *HLA-DQB1*, which are just one base pair apart and in perfect LD in our sample, showed the strongest association (OR = 1.21, $P = 1.4 \times 10^{-11}$). Both the *HLA-DQB1* and *HLA-DRB1* genes are located ~700 kb away from the most strongly associated genetic risk variant in the neighboring *CFB/SKIV2L* gene region. The association between rs9274390 and overall AMD was stronger after conditioning on previously established AMD risk SNPs, especially rs429608, (unconditional: OR = 1.16; $P = 5.17 \times 10^{-8}$; conditioning on rs429608, but not other previously established AMD risk loci: OR = 1.19; $P = 5.32 \times 10^{-10}$; conditioning on previously

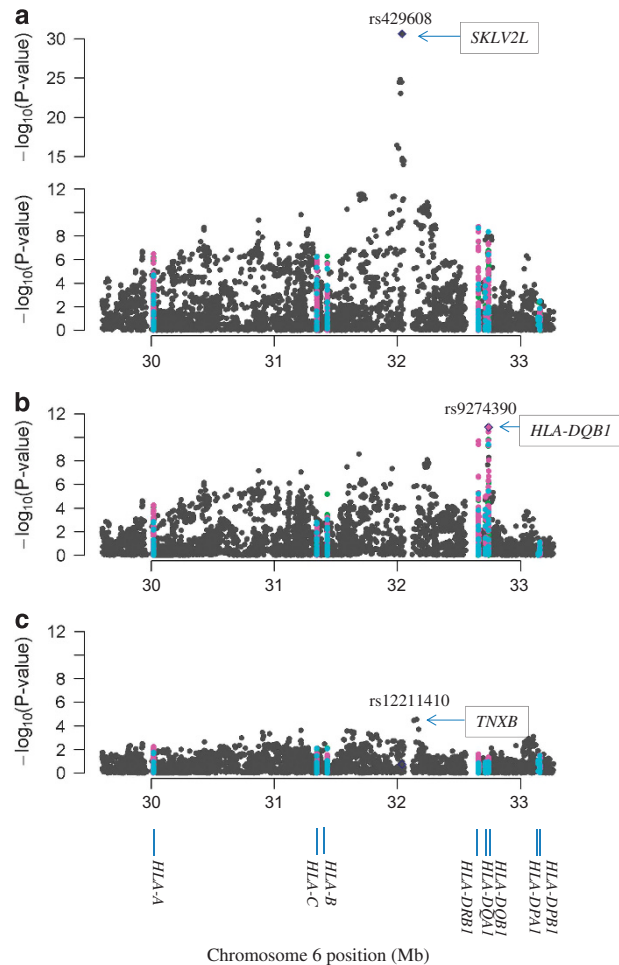


Figure 1 Manhattan plots for the HLA region: The lead SNP at each locus is shown in the shape of diamond. Each circle represents a SNP, amino-acid change, or classical HLA allele. We plotted SNPs in HLA genes (*HLA-A*, *-B*, *-C*, *-DPA1*, *-DPB1*, *-DQA1*, *-DQB1*, and *-DRB1*) in green, amino-acid changes in magenta, classical HLA alleles in light blue, and rest of the SNPs in gray. The red dotted line represents the genome-wide significance level ($P = 5 \times 10^{-8}$) and purple line represents $P = 1 \times 10^{-4}$. (a) The most strongly associated SNP was rs429608, an established association in the *CFB/SKIV2L* gene region with overall AMD adjusted for age, sex, and the first 10 ancestry PCs; (b) The top association is at rs9274390 in *HLA-DQB1* (dark blue triangle) with overall AMD adjusting for age, sex, first 10 ancestry PCs, and established *CFH*, *ARMS2/HTRA1*, *CFB/SKIV2L*, *C3*, and *LIPC* gene regions, including rs1061170, rs10490924, rs429608, rs2230199, and rs173539; (c) Results after adding rs9274390 as a covariate into the model in (b).

Table 1 Characteristics of non-Hispanic white AMD cases and controls, age ≥ 65 years as of 30 June 2013

	Overall AMD	AMD unspecified	Non-exudative AMD	CNV	Controls
Number of GERA subjects ^a	4841	2706	987	1148	23 790
Male (%)	2015 (41.62)	1166 (43.09)	392 (39.72)	457 (39.81)	10 384 (43.65)
Female (%)	2 826 (58.38)	1540 (56.91)	595 (60.28)	691 (60.19)	13 406 (56.35)
Age ^b (mean \pm SD), years	78.60 \pm 7.31	78.16 \pm 7.47	78.76 (7.23)	79.51 \pm 6.91	71.58 \pm 7.22

Abbreviations: AMD, age-related macular degeneration; CNV, choroidal neovascularization; GERA, Genetic Epidemiology Research on Adult Health and Aging.

Overall AMD includes all three subtypes of AMD.

^aThe subjects included in this study were unrelated non-Hispanic whites subjects from the GERA cohort who were at least 65 years of age as of 30 June 2013.

^bAge at biospecimen collection in 2008–2009.

Table 2 Conditional analysis of overall AMD, AMD unspecified, non-exudative AMD, and CNV

Variant ^a	Gene function or HLA amino-acid change				Associations with AMD in GERA					Replication in MMAP (advanced AMD)					
	EA ^b	Info ^b	Control FRQ	Case FRQ	OR	95% CI	P	Info ^b	Control FRQ	Case FRQ	OR	95% CI	P		
<i>HLA-DQB1</i>															
rs9274390	Glu(E) ≥ Asp(D) at position 66	G	1	0.252	Overall AMD	0.274	1.2	(1.14, 1.28)	1.4 × 10⁻¹¹	0.98	0.247	0.281	1.28	(1.10, 1.49)	1.3 × 10 ⁻³
rs41563814	Val(V) ≥ Ile(I) at position 67	T			AMD unspecified	0.276	1.19	(1.11, 1.27)	5.1 × 10 ⁻⁷						
					Non-exudative AMD	0.275	1.24	(1.11, 1.38)	8.3 × 10 ⁻⁵						
					CNV	0.27	1.19	(1.08, 1.32)	8.6 × 10 ⁻⁴						
<i>HLA-DRB1</i>															
rs17878857	Ala(A) ≥ Gly(G) at position 73	C	1	0.244	Overall AMD	0.267	1.2	(1.13, 1.27)	2.1 × 10⁻¹⁰	0.99	0.245	0.297	1.39	(1.18, 1.64)	8.4 × 10 ⁻⁴
					AMD unspecified	0.269	1.19	(1.11, 1.27)	4.3 × 10 ⁻⁷						
					Non-exudative AMD	0.264	1.21	(1.08, 1.34)	6.5 × 10 ⁻⁴						
					CNV	0.265	1.18	(1.06, 1.31)	1.8 × 10 ⁻³						
<i>Sensitivity analysis of a homogeneous subset of subjects based on the first two principal components of ancestry</i>															
<i>HLA-DQB1</i>															
rs9274390	Glu(E) ≥ Asp(D) at position 66	G	1	0.256	Overall AMD	0.278	1.2	(1.14, 1.28)	1.1 × 10⁻¹⁰						
rs41563814	Val(V) ≥ Ile(I) at position 67	T			AMD unspecified	0.277	1.17	(1.09, 1.26)	1.7 × 10 ⁻⁵						
					Non-exudative AMD	0.284	1.27	(1.14, 1.43)	2.3 × 10 ⁻⁵						
					CNV	0.276	1.21	(1.08, 1.35)	6.5 × 10 ⁻⁴						
<i>HLA-DRB1</i>															
rs17878857	Ala(A) ≥ Gly(G) at position 73	C	1	0.251	Overall AMD	0.275	1.2	(1.14, 1.28)	5.0 × 10⁻¹⁰						
					AMD unspecified	0.274	1.17	(1.09, 1.26)	1.3 × 10 ⁻⁵						
					Non-exudative AMD	0.278	1.26	(1.12, 1.41)	7.2 × 10 ⁻⁵						
					CNV	0.275	1.2	(1.08, 1.35)	9.7 × 10 ⁻⁴						

Abbreviations: AMD, age-related macular degeneration; CNV, choroidal neovascularization; GERA, Genetic Epidemiology Research on Adult Health and Aging. Bold indicates genome-wide significant association result ($P < 10^{-9}$).

^aTop associated SNPs in the region from a logistic regression analysis, adjusting for age, sex, the first 10 ancestry PCs, and established *CFH*, *ARMS2/HTRA1*, *CFB/SKIV2L*, and *C3* gene region SNPs, specifically rs1061170, rs10490924, rs429608, and rs230199, rs9274390 and rs41563814 are in perfect linkage disequilibrium in our sample. Sensitivity analysis: restricting to a homogeneous subset of subjects (>0.02 on PC1 and <0.02 on PC2, Supplementary Figure 1). Overall AMD includes all three subtypes of AMD.

^bEffect Allele. Alternate alleles were rs9274390: A, rs41563814: A, and rs17878857: G.

^cThe info metric from SNP2HLA is an estimated correlation of the imputed genotype to the true genotype.

established AMD risk loci: $OR=1.21$; $P=1.4 \times 10^{-11}$), and no interaction between rs9274390 and rs429608 was detected ($P=0.46$). A sensitivity analysis of a homogeneous subset of subjects based on the first two PCs of ancestry revealed similar results (Table 2).

We then examined the association of the imputed classical HLA alleles in the same conditional model, and we found that the DQB1*02 allele was most strongly associated with overall AMD ($OR=1.22$, $P=3.9 \times 10^{-10}$; Table 3), driven by the G-allele of rs9274390, which resides mainly on the haplotypes of DQB1*02:01 and DQB1*02:02. Among the 3.7% of rs9274390 G-allele carriers that did not also carry a DQB1*02 haplotype, the variant had a similar effect ($OR=1.17$, $P=0.03$) on the risk of overall AMD. DRB1*03 was the next most significant association ($OR=1.19$, $P=6.1 \times 10^{-6}$), due in part to the C-allele of rs17878857, which resides partly on the DRB1*0301 haplotype.

In our second round of conditional analyses, in which we added rs9274390 as a covariate, the strongest association that we observed was with rs12211410 ($OR_{\text{conditional}}=1.19$; $P_{\text{conditional}}=2.91 \times 10^{-5}$; Figure 1c), a missense SNP in *TNXB*. We did not observe any significant associations of classical HLA alleles ($P > 0.05$).

We attempted to confirm the SNP associations in or near *HLA-DQB1* using HapMap imputed discovery sample results available in the large ($n > 77\,000$) AMD Gene Consortium GWAS meta-analysis (<http://www.sph.umich.edu/csg/abecasis/public/amdgene2012/>),¹⁶ but neither rs9274390 nor SNPs in moderate-to-high LD with our top association were included in that analysis, likely due to insufficient coverage in HapMap in the region. Although our analysis included 5640 SNPs (2116 SNPs were directly genotyped) within $\pm 250\,000$ base pairs around rs9274390, 166 SNPs were found in the Consortium meta-analysis results.

As an alternative, we conducted imputation in the MMAP data set ($n=3340$) using both the 1000 genomes and SNP2HLA T1DGC references panels and conducted association analyses of variants in this region, conditioning on the same previously reported AMD risk SNPs. We observed evidence of replication for both rs9274390 ($OR=1.28$; $P=1.30 \times 10^{-3}$, Table 2) and the HLA-DQB1*02 allele ($OR=1.32$; $P=9.00 \times 10^{-4}$, Table 3).

DISCUSSION

This is the largest study to investigate in depth the association of HLA risk alleles and the risk of AMD. We identified a significant association between variants in *HLA-DQB1* and the risk of overall AMD. The association of rs9274390 and rs41563814, and the corresponding amino-acid changes at positions 66 and 67, was independent of the most strongly associated SNPs from previously established risk loci. These two SNPs reside mainly on the DQB1*02 classical HLA allele, which was also significantly associated with the risk of overall AMD.

The variants rs9274390 and rs41563814 are non-synonymous SNPs, encoding amino-acid changes at positions 66 and 67 in *HLA-DQB1*. *HLA-DQB1* is part of the HLA class II beta chain. Proteins encoded by *HLA-DQB1* and *HLA-DQA1* attach together to form antigen-binding DQ $\alpha\beta$ heterodimers that present foreign peptides to CD4(+) T cells and trigger immune response. Variations in the amino-acid sequences of the HLA class II genes enable the immune system to recognize and react to a wide range of foreign entities. These two amino-acid changes occur most often on the DQB1*02 classical HLA allele, which has previously been associated with celiac disease.³⁰ Drusen, the earliest clinical detectable presentation of AMD, contain HLA class II antigens¹² and many gene products such as apolipoprotein B and apolipoprotein E, amyloid, vitronectin, complement receptor 1, and immunoglobulin that are known to modulate immune responses.^{6–11}

Increased HLA class II immunoreactivity related to drusen formation has been found in the retina in AMD patients.¹³ For these reasons, HLA polymorphisms have been hypothesized to modulate susceptibility to AMD.^{14,15} Drusen formation and AMD progression could result from an inflammatory response to RPE injury that involves HLA and the complement systems.^{11,31,32} Our finding is the first to implicate the association of an HLA class II polymorphism with AMD pathogenesis.

Although we observed the most significant associations at rs9274390 and rs41563814 in *HLA-DQB1* with overall AMD, we also found similar effects on non-exudative AMD and CNV (Table 2), suggesting these MHC class II polymorphisms may be involved in both drusen formation and accumulation. In contrast to the similar effects of DQB1*02:01 and DQB1*02:02 on overall AMD, we observed larger effect of DQB1*02:02 on non-exudative AMD and DQB1*02:01 or DRB1*03 on CNV, indicating multiple causal variants may influence different forms of AMD. We also cannot rule out the possibility that the observed associations were a result of LD with causal polymorphisms in other genes. Although it is possible that high LD may extend further in the HLA region than other parts of the genome, our top reported SNPs were not in high LD with previously reported SNPs in *CFB/SKIV2L* (Supplementary Figure 3). Future investigation in an independent population, particularly one with a different pattern of LD in this region, and functional studies will be important to refine the signals observed here.

These results should be interpreted in the light of the following limitations. Non-exudative AMD presents with a wide spectrum of clinical manifestations. The ICD9 codes do not distinguish GA or do they differentiate stages or severities of non-exudative AMD. Some early AMD cases may have not been captured in the EHR because of the asymptomatic nature of the initial changes in the macula. Such misclassification would lead to reduced strength of the association of genetic factors with AMD. In addition, although the SNP2HLA method has been shown to perform well for variants with a minor allele FRQ greater than 2.5% in previous studies,³³ concerns have also been raised about the potential for bias in the estimation of variant frequencies in the HLA region.³⁴ We attempted to address these concerns in two ways. First, to determine whether imputation artifacts might be responsible for the observed signals in this region, we identified two genotyped SNPs (rs9274251 and rs927453), which were associated with overall AMD ($P < 5 \times 10^{-8}$) and were strongly correlated with our top imputed SNP rs9274390 ($r^2=0.895$). We examined visually the probe intensity plots for these genotyped SNPs and determined that the genotyping calls were of high quality. Second, to determine whether the genotyping calls were potentially biased due to the genotyping method used, we conducted replication analyses in the MMAP data set, whose participants were genotyped on a different platform, which confirmed our findings. Nevertheless, any association in this region should be greeted with a degree of caution, and confirmatory studies using other methods to assay genetic variants should be conducted.

This study has several important strengths: The membership of KPNC is representative of the population of northern California,¹⁸ enabling the generalization of risk estimates to the general population of northern California non-Hispanic whites. Second, leveraging HLA imputation using SNP2HLA,¹⁷ this is the first large study to implicate variants in *HLA-DQB1* and the DQB1*02 classical HLA allele in AMD susceptibility. Our findings here demonstrate how improved coverage of genotyping arrays and imputation to more comprehensive reference panels can lead to novel findings, even for traits that have been investigated in depth in previous genetic association studies.³⁵

Table 3 The top classical HLA allele associations^a with overall AMD and its subtypes

HLA alleles	Associations in GERA					Replication in MMAP (advanced AMD)							
	Inf ^b	Control FRQ	AMD category	Case FRQ	OR	95% CI	P	Inf ^b	Control FRQ	Case FRQ	OR	95% CI	P
HLA-DQB1													
DQB1*02	1	0.215	Overall AMD	0.237	1.2	(1.13, 1.27)	3.9 × 10⁻¹⁰	0.99	0.211	0.243	1.32	(1.12, 1.56)	9.0 × 10 ⁻⁴
			AMD unspecified	0.237	1.18	(1.10, 1.26)	5.1 × 10 ⁻⁶						
			Non-exudative AMD	0.237	1.24	(1.11, 1.39)	1.2 × 10 ⁻⁴						
			CNV	0.235	1.2	(1.08, 1.33)	1.1 × 10 ⁻³						
DQB1*02:01	1	0.118	Overall AMD	0.137	1.18	(1.10, 1.27)	3.7 × 10 ⁻⁶	0.99	0.112	0.143	1.34	(1.05, 1.72)	0.02
			AMD unspecified	0.135	1.15	(1.06, 1.26)	1.6 × 10 ⁻³						
			Non-exudative AMD	0.132	1.17	(1.02, 1.35)	0.024						
			CNV	0.146	1.32	(1.16, 1.51)	2.5 × 10 ⁻⁵						
DQB1*02:02	0.99	0.097	Overall AMD	0.01	1.17	(1.07, 1.27)	2.3 × 10 ⁻⁴	0.99	0.099	0.1	1.3	(1.04, 1.61)	0.02
			AMD Unspecified	0.103	1.16	(1.05, 1.28)	3.5 × 10 ⁻³						
			Non-exudative AMD	0.105	1.26	(1.08, 1.47)	3.8 × 10 ⁻³						
			CNV	0.089	0.99	(0.84, 1.16)	0.88						
HLA-DRB1													
DRB1*03	0.98	0.11	Overall AMD	0.128	1.19	(1.10, 1.28)	6.1 × 10 ⁻⁶	0.99	0.112	0.145	1.4	(1.08, 1.65)	9.6 × 10 ⁻³
			AMD Unspecified	0.127	1.17	(1.07, 1.28)	7.9 × 10 ⁻⁴						
			Non-exudative AMD	0.124	1.18	(1.02, 1.36)	0.023						
			CNV	0.136	1.31	(1.14, 1.49)	3.1 × 10 ⁻⁴						

Abbreviations: AMD, age-related macular degeneration; CNV, choroidal neovascularization; GERA, Genetic Epidemiology Research on Adult Health and Aging; HLA, human leukocyte antigen.

Overall AMD includes all three subtypes of AMD. Bold indicates genome-wide significant association result ($P < 10^{-8}$).

^aResults from an omnibus test described by Jia et al, comparing the presence of each allele on an additive scale to all other alleles, adjusting for age, sex, first 10 ancestry PCs and established CFH, ARMS2/HTRA1, CFB/SKIV2L, and C3 gene regions, including rs1061170, rs10490924, rs429608, and rs2230199.

^bThe info metric from SNP2HLA is the estimated correlation of the imputed genotype to the true genotype.

In summary, we identified significant associations between HLA class II alleles and AMD. Variants in *HLA-DQB1* influence the risk of developing AMD independent of previously identified AMD risk variants in the region.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by a research grant from KPNC Community Benefit, and grants RC2 AG036607 and R21 AG046616 from the National Institutes of Health.

SNP GENOME BUILD REFERENCE

rs9274390:G>A, hg18 chr6:g.32740388C>T; rs41563814:A>T, hg18 chr6:g.32740386A>T; rs17878857:C>G, hg18 chr6:g.32659679C>G; rs1061170:C>T, hg18 chr1:g.194925610C>T; rs10490924:G>T, hg18 chr10:g.124204188G>T; rs429608:A>G, hg18 chr6:g.32032191A>G; rs2230199:C>G, hg18 chr19:g.6669137C>G; rs173539:C>T, hg18 chr16:g.55545295C>T; rs12211410:C>T, hg18 chr6:g.32157151C>T; rs9274251:C>T, hg18 chr6:g.32739071C>T; rs927453:A>G, hg18 chr6:g.32739076A>G.

ACCESSION NUMBERS

The dbGap accession number for the genotype information of the GERA cohort presented in this paper is phs000674.v1.p1.

- 1 Finger RP, Fimmers R, Holz FG, Scholl HP: Prevalence and causes of registered blindness in the largest federal state of Germany. *Br J Ophthalmol* 2011; **95**: 1061–1067.
- 2 Friedman DS, O'Colmain BJ, Munoz B et al: Prevalence of age-related macular degeneration in the United States. *Arch Ophthalmol* 2004; **122**: 564–572.
- 3 Klein R, Chou CF, Klein BE, Zhang X, Meuer SM, Saaddine JB: Prevalence of age-related macular degeneration in the US population. *Arch Ophthalmol* 2011; **129**: 75–80.
- 4 Wong WL, Su X, Li X et al: Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health* 2014; **2**: e106–116.
- 5 Wang Y, Wang VM, Chan CC: The role of anti-inflammatory agents in age-related macular degeneration (AMD) treatment. *Eye (Lond)* 2011; **25**: 127–139.
- 6 Johnson LV, Forest DL, Banna CD et al: Cell culture model that mimics drusen formation and triggers complement activation associated with age-related macular degeneration. *Proc Natl Acad Sci USA* 2011; **108**: 18277–18282.
- 7 Johnson LV, Ozaki S, Staples MK, Erickson PA, Anderson DH: A potential role for immune complex pathogenesis in drusen formation. *Exp Eye Res* 2000; **70**: 441–449.
- 8 Hageman GS, Mullins RF, Russell SR, Johnson LV, Anderson DH: Vitronectin is a constituent of ocular drusen and the vitronectin gene is expressed in human retinal pigmented epithelial cells. *FASEB J* 1999; **13**: 477–484.
- 9 Luibl V, Isas JM, Kaye R, Glabe CG, Langen R, Chen J: Drusen deposits associated with aging and age-related macular degeneration contain nonfibrillar amyloid oligomers. *J Clin Invest* 2006; **116**: 378–385.
- 10 Anderson DH, Ozaki S, Nealon M et al: Local cellular sources of apolipoprotein E in the human retina and retinal pigmented epithelium: implications for the process of drusen formation. *Am J Ophthalmol* 2001; **131**: 767–781.
- 11 Johnson LV, Leitner WP, Staples MK, Anderson DH: Complement activation and inflammatory processes in Drusen formation and age related macular degeneration. *Exp Eye Res* 2001; **73**: 887–896.
- 12 Mullins RF, Russell SR, Anderson DH, Hageman GS: Drusen associated with aging and age-related macular degeneration contain proteins common to extracellular deposits associated with atherosclerosis, elastosis, amyloidosis, and dense deposit disease. *FASEB J* 2000; **14**: 835–846.
- 13 Penfold PL, Liew SC, Madigan MC, Provis JM: Modulation of major histocompatibility complex class II expression in retinas with age-related macular degeneration. *Invest Ophthalmol Vis Sci* 1997; **38**: 2125–2133.
- 14 Goverdhan SV, Howell MW, Mullins RF et al: Association of HLA class I and class II polymorphisms with age-related macular degeneration. *Invest Ophthalmol Vis Sci* 2005; **46**: 1726–1734.
- 15 Goverdhan SV, Khakoo SI, Gaston H, Chen X, Lotery AJ: Age-related macular degeneration is associated with the HLA-Cw*0701 Genotype and the natural killer cell receptor AA haplotype. *Invest Ophthalmol Vis Sci* 2008; **49**: 5077–5082.
- 16 Fritsche LG, Chen W, Schu M et al: Seven new loci associated with age-related macular degeneration. *Nat Genet* 2013; **45**: 433–439, 439e431–432.
- 17 Jia X, Han B, Onengut-Gumuscu S et al: Imputing amino acid polymorphisms in human leukocyte antigens. *PLoS One* 2013; **8**: e64683.
- 18 Krieger N: Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health* 1992; **82**: 703–710.
- 19 Hoffmann TJ, Kvale MN, Hesselson SE et al: Next generation genome-wide association tool: design and coverage of a high-throughput European-optimized SNP array. *Genomics* 2011; **98**: 79–89.
- 20 Hoffmann TJ, Zhan Y, Kvale MN et al: Design and coverage of high throughput genotyping arrays optimized for individuals of East Asian, African American, and Latino race/ethnicity using imputation and a novel hybrid SNP selection algorithm. *Genomics* 2011; **98**: 422–430.
- 21 Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM: Robust relationship inference in genome-wide association studies. *Bioinformatics* 2010; **26**: 2867–2873.
- 22 Shen L, Hoffmann TJ, Melles RB et al: Differences in the genetic susceptibility to age-related macular degeneration clinical subtypes. *Invest Ophthalmol Vis Sci* 2015; **56**: 4290–4299.
- 23 Kvale MN, Hesselson S, Hoffmann TJ et al: Genotyping Informatics and Quality Control for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics* 2015; **200**: 1051–60.
- 24 Howie B, Marchini J, Stephens M: Genotype imputation with thousands of genomes. *G3 (Bethesda)* 2011; **1**: 457–470.
- 25 Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR: Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* 2012; **44**: 955–959.
- 26 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.
- 27 Banda Y, Kvale MN, Hoffmann TJ et al: Characterizing Race/Ethnicity and Genetic Ancestry for 100,000 Subjects in the Genetic Epidemiology Research on Adult Health and Aging (GERA) Cohort. *Genetics* 2015; **200**: 1285–95.
- 28 Purcell S, Neale B, Todd-Brown K et al: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
- 29 Team RDC: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria 2008.
- 30 Murray JA, Moore SB, Van Dyke CT et al: HLA DQ gene dosage and risk and severity of celiac disease. *Clin Gastroenterol Hepatol* 2007; **5**: 1406–1412.
- 31 Anderson DH, Mullins RF, Hageman GS, Johnson LV: A role for local inflammation in the formation of drusen in the aging eye. *Am J Ophthalmol* 2002; **134**: 411–431.
- 32 Hageman GS, Luthert PJ, Victor Chong NH, Johnson LV, Anderson DH, Mullins RF: An integrated hypothesis that considers drusen as biomarkers of immune-mediated processes at the RPE-Bruch's membrane interface in aging and age-related macular degeneration. *Prog Retin Eye Res* 2001; **20**: 705–732.
- 33 Raychaudhuri S, Sandor C, Stahl EA 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.
- 34 Brandt DY, Aguiar VR, Bitarello BD, Nunes K, Goudet J, Meyer D: Mapping Bias Overestimates Reference Allele Frequencies at the HLA Genes in the 1000 Genomes Project Phase I Data. *G3 (Bethesda)* 2015; **5**: 931–941.
- 35 Lindquist KJ, Jorgenson E, Hoffmann TJ, Witte JS: The impact of improved microarray coverage and larger sample sizes on future genome-wide association studies. *Genet Epidemiol* 2013; **37**: 383–392.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)