APOL1 Variants Increase Risk for FSGS and HIVAN but Not IgA Nephropathy


Departments of *Medicine and **Pathology, Columbia University College of Physicians and Surgeons, New York, New York; †Laboratory of Molecular Genetics, Department of Otolaryngology, New York University School of Medicine, New York, New York; ‡Departments of Microbiology and Medicine, University of Alabama at Birmingham, Birmingham, Alabama; §Children’s Foundation Research Center at the Le Bonheur Children’s Hospital, and the Division of Pediatric Nephropathy, University of Tennessee Health Sciences Center, Memphis, Tennessee; ‖Department of Medicine, Mount Sinai School of Medicine, New York, New York; ¶Department of Medicine, Duke University School of Medicine, Durham, North Carolina; and ††Department of Medicine, Baylor College of Medicine, Houston, Texas

ABSTRACT

A chromosome 22q13 locus strongly associates with increased risk for idiopathic focal segmental glomerulosclerosis (FSGS), HIV-1-associated nephropathy (HIVAN), and hypertensive ESRD among individuals of African descent. Although initial studies implicated MYH9, more recent analyses localized the strongest association within the neighboring APOL1 gene. In this replication study, we examined the six top-most associated variants in APOL1 and MYH9 in an independent cohort of African Americans with various nephropathies (44 with FSGS, 21 with HIVAN, 32 with IgA nephropathy, and 74 healthy controls). All six variants associated with FSGS and HIVAN (additive ORs, 1.8 to 3.0; P values 3 × 10⁻² to 5 × 10⁻⁵) but not with IgA nephropathy. In conditional and haplotype analyses, two APOL1 haplotypes accounted for virtually all of the association with FSGS and HIVAN on chromosome 22q13 (haplotype P value = 5.6 × 10⁻⁶). To assess the role of MYH9 deficiency in nephropathy, we crossbred Myh9-haploinsufficient mice (Myh9⁻/-) with HIV-1 transgenic mice. Myh9⁻/- mice were healthy and did not demonstrate overt proteinuria or nephropathy, irrespective of the presence of the HIV-1 transgene. These data further support the strong association of genetic variants in APOL1 with susceptibility to FSGS and HIVAN among African Americans.


African Americans have a three- to four-fold increased risk of focal segmental glomerulosclerosis (FSGS), HIV-associated nephropathy (HIVAN, a secondary form of FSGS associated with HIV infection), and nondiabetic ESRD compared with European-Americans.1–4 Family members of African Americans with ESRD also have a higher incidence of kidney failure, suggesting that genetic susceptibility in part accounts for this skewed epidemiology.5,6 In 2008, two independent studies demonstrated that variants on chromosome 22q13 were highly associated with the risk of FSGS, HIVAN, and nondiabetic ESRD among African Americans and explained much of the increased risk of kidney failure in this population (per alleles odds ratios [ORs] of 3 to 4).7,8 Initially, MYH9 was considered as the likely causal gene in this interval because it encodes a podocyte cytoskeletal protein and because MYH9 coding mutations cause rare Mendelian disorders with occasional glomerulopathy. Follow-up studies confirmed and refined the signal within the MYH9 locus,9–11 but most recently, a comprehensive analysis of the chromosome 22q13 interval indicated that the association signal originated in the neighboring APOL1 gene, and protein-altering variants (named G1 and G2) on two independent APOL1 haplotypes explained all of the associations in this region.12–14 The G1 haplotype encodes two linked missense variants (P.S342G and P.I384M), and the G2 haplotype encodes a two-amino-acid deletion (P.N388_Y389del) within the APOL1 carboxy-terminal domain; these variants are protective against Trypanosoma brucei rhod...
esense infection and therefore provide an evolutionary advantage in Africa where this parasite is an endemic cause of sleeping sickness. Accordingly, these nephropathy risk alleles harbor signatures of positive selection and are very common on African American chromosomes but absent from European chromosomes.

To date, a direct comparison of the most significantly associated single-nucleotide polymorphisms (SNPs) in MYH9 and APOL1 has not been performed in independent cohorts, and the association of APOL1 variants in the subgroups of HIVAN and IgA nephropathy (IgAN) has not been examined. Such replication studies are important for validation of genome-wide association studies and determination of the magnitude of risk imparted on different traits. In addition, examination of candidate genes in relevant animal models can provide biologic validation for statistical association in humans. In this study, we examined the most significantly associated SNPs in MYH9 and APOL1 in independent African American cohorts with FSGS, HIVAN, and IgAN. Next, we evaluated the effects of Myh9 haplinsufficiency on the development of nephropathy using a validated mouse model of HIVAN.

In a human association study, we analyzed 32 patients with IgAN, 44 patients with FSGS, and 21 patients with HIVAN against a group of 74 healthy control individuals. All of the participants were of African American ancestry. We genotyped the six most significantly associated SNPs in APOL1 (rs73885319 and rs60910145 for the G2 risk allele) and MYH9 (rs11912763, rs4821481, and rs5750250 tagging the F-1, E-1, and S-1 haplotypes, respectively).

Power analysis demonstrated that the IgAN cohort and the combined HIVAN and FSGS cohort provided >90% power to detect additive ORs >3.0 at two-sided alpha of 0.05 (Supplemental Table 1). All SNPs were in Hardy-Weinberg equilibrium in controls, and their allele frequencies did not differ from those reported in HapMap or other African American control populations. After standard genotype quality control (see the Concise Methods section), we conducted association testing using a simple allelic test for each patient-control cohort (Table 1). None of the SNPs in APOL1 or MYH9 were associated with the risk of IgAN under an additive or a recessive model. In contrast, and consistent with previously reported data, five SNPs were significantly associated with the risk of FSGS (allelic ORs, 1.6 to 2.9) and HIVAN (allelic ORs, 2.1 to 3.4). Moreover, when the FSGS and HIVAN groups were combined, all six SNPs were significantly associated with risk of disease (allelic ORs, 1.8 to 3.0; P values, $3 \times 10^{-2}$ to $5 \times 10^{-5}$; Table 1). Interestingly, the two most-significant SNPs (rs60910145 and rs11912763) reside in APOL1 and MYH9, respectively, and thus the single SNP analysis could not clearly resolve the origin of the signal.

We performed a careful examination of the linkage disequilibrium (LD) structure among the six variants (Supplemental Table 2). The two APOL1 G1-defining SNPs (rs73885319 and rs60910145) were in tight but imperfect LD ($r^2 = 0.924$, $D' = 0.98$) and were also in LD with the G2-defining SNP (rs71785313, $r^2 = 0.13$, $D' = 1$). The complete LD between the G1- and G2-defining SNPs ($D' = 1$) is consistent with the presence of two mutually exclusive haplotypes at this locus. The three MYH9 variants were in partial LD with APOL1 variants ($r^2 = 0.02$ to 0.4, $D' = 0.725$ to 0.9). These data are consistent with prior reports of the LD structure of this region.

To better resolve the signal at this interval, we performed haplotype analysis with two- and three-SNP moving windows (Figure 1A). This analysis clearly localized the most significant signal within the APOL1 locus: the association statistics within APOL1 are within the threshold of genome-wide significance for genome-wide association studies and are nearly 4 orders of magnitude more significant compared with signals within the MYH9 locus (peak two-SNP haplotype P values = $5.6 \times 10^{-8}$ in APOL1 versus $2.7 \times 10^{-4}$ in MYH9 [Figure 1A and Supplemental Table 3]). Analysis of three-SNP haplotypes within APOL1 confirmed that the G1 and G2 alleles reside on mutually exclusive haplotypes, and each confers independent risk of dis-

### Table 1. Results of single SNP association analysis

<table>
<thead>
<tr>
<th>Chr</th>
<th>Position (kb)</th>
<th>SNP</th>
<th>G</th>
<th>A</th>
<th>MAF</th>
<th>Minor</th>
<th>MAF</th>
<th>Cases OR</th>
<th>Controls OR</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>rs73885319</td>
<td>G</td>
<td>36,661</td>
<td>G</td>
<td>0.64</td>
<td>0.244</td>
<td>1.01</td>
<td>0.249</td>
<td>0.10</td>
<td>0.76</td>
</tr>
<tr>
<td>22</td>
<td>rs60910145</td>
<td>A</td>
<td>36,662</td>
<td>A</td>
<td>0.250</td>
<td>0.95</td>
<td>1.01</td>
<td>0.95</td>
<td>0.05</td>
<td>1.01</td>
</tr>
<tr>
<td>22</td>
<td>rs71785313</td>
<td>Del</td>
<td>36,662</td>
<td>Del</td>
<td>0.300</td>
<td>0.60</td>
<td>1.01</td>
<td>0.60</td>
<td>0.05</td>
<td>1.01</td>
</tr>
<tr>
<td>22</td>
<td>rs11912763</td>
<td>A</td>
<td>36,662</td>
<td>A</td>
<td>0.320</td>
<td>0.68</td>
<td>1.01</td>
<td>0.68</td>
<td>0.05</td>
<td>1.01</td>
</tr>
<tr>
<td>22</td>
<td>rs4821481</td>
<td>T</td>
<td>36,662</td>
<td>T</td>
<td>0.350</td>
<td>0.65</td>
<td>1.01</td>
<td>0.65</td>
<td>0.05</td>
<td>1.01</td>
</tr>
<tr>
<td>22</td>
<td>rs5750250</td>
<td>A</td>
<td>36,670</td>
<td>A</td>
<td>0.330</td>
<td>0.67</td>
<td>1.01</td>
<td>0.67</td>
<td>0.05</td>
<td>1.01</td>
</tr>
</tbody>
</table>

*Chr:* chromosome; *SNP:* single-nucleotide polymorphism; *MAF:* minor allele frequency; *NS:* not significant; *P-value:* corresponds to the allelic test of association.
ease (per haplotype ORs = 3.8 to 4.0; Figure 1B). The results of stepwise conditional analysis further support these conclusions (Table 2). After conditioning on the most significant SNP (APOL1 G1/ rs60910145), one APOL1 SNP (G2/ rs71785313, OR = 3.56, P = 3.9 × 10^{-4}) and one MYH9 SNP (rs11912763, OR = 1.9, P = 0.048) remained significant. The residual association with rs11912763 is due to its LD with G2-defining SNP (Supplemental Table 2). After accounting for rs71785313, there were no other significant signals (Table 2). These data clearly indicate that the APOL1 G1 and G2 variants explain all of the association observed with the MYH9 SNPs.

Because the APOL1 G1 and G2 haplotypes confer a comparable risk of disease but reside on distinct haplotypes (Figure 1B), we recoded rs60910145 and rs71785313 into a single locus, where each individual was considered to harbor 0, 1, or 2 APOL1 risk alleles (i.e. G1/+ and G2/+ genotypes were coded as heterozygotes, and G1/G1, G1/G2, and G2/G2 genotypes were coded as homozygotes). In this analysis, additive and recessive models were similarly supported and reached thresholds for genome-wide significance (ORADD = 3.89, PADD = 5 × 10^{-8} and ORREC = 10.9, PREC = 1 × 10^{-9}). Because eight individuals with FSGS or HIVAN did not carry any G1 or G2 alleles, we sequenced all APOL1 exons in these individuals but did not identify any rare coding mutations that could account for disease. In other secondary analyses, we examined the effect of APOL1 variants on the risk of renal failure and detected a suggestive association with increased risk of ESRD among the IgAN patients, but not among the FSGS/HIVAN cohorts (OR 3.2, P = 0.049; Supplemental Table 4).

In addition to the statistical evidence provided by human association studies, analysis of biologically relevant mouse models can provide independent assessment of candidate genes within association intervals. In this regard, we had previously shown that renal expression of Myh9 is reduced by approximately 30% in HIV-1 transgenic mice on the susceptible FVB/NJ genetic background,17 a validated model of HIVAN that exhibits proteinuria and classical features of collapsing glomerulopathy with cystic tubular dilation by 6 to 8 weeks of age.18,19 However, we could not determine whether this reduced expression played a direct role in the pathogenesis of disease or represented a secondary consequence of glomerulosclerosis. Homozygous inactivation of Myh9 results in embryonic lethality at embryonic day 6.5, but Myh9 haploinsufficient mice (Myh9^{+/-}), despite significantly reduced MYH9 protein level, are born normally and do not develop gross organ dysfunction.20,21 Therefore, we examined Myh9 haploinsufficient mice

Figure 1. Haplotype analysis localized the most significant signal within the APOL1 locus. (A) Two-SNP and three-SNP haplotype sliding-window analysis localized the strongest signal within the APOL1 locus. The y axis shows the −log (P value) for the association statistic. The x axis shows the tested SNP and their location within the APOL1 and MYH9 genes. (B) Haplotype analysis of the APOL1 SNPs demonstrated that G1 and G2 reside on distinct haplotypes and each confers an independent risk of disease compared with the reference haplotype. G1 is defined by the rs73885319-G and rs60910145-G alleles, and G2 is defined by the rs71785313 deletion allele (D). The odds ratios (ORs) were calculated in reference to the most common haplotype, which does not contain any risk variants (ATI). *P value < 0.05 versus the reference haplotype. The P value for the global haplotype test is also indicated.
and crossbred them with HIV-1 transgenic mice on the C57BL/6J (B6) genetic background. We chose the B6 background because it is protective against HIV-1 nephropathy in mice; consequently, this breeding scheme could clearly determine whether a primary reduction in MYH9 expression, in conjunction with a biologically relevant stressor, would overcome genetic resistance and produce kidney disease.

We assessed renal histopathology at 5 to 6 months of age (19 to 29 mice per group; Figure 2) and quantified urine protein levels using an albumin-to-creatinine ratio in a random subset (four to five mice per group). As previously reported, MYH9 haploinsufficiency does not induce nephropathy in HIV-1 transgenic mice on the B6 genetic background but was still within the range of normal for inbred strains (reported range 3 to 142 μg of albumin/mg creatinine).

We confirmed that MYH9 expression was significantly decreased (by 30%) in MYH9 haploinsufficient mice compared with wild-type littermates (Supplemental Figure 1), which was comparable to the reduction observed in the TgFVB strain. This suggests that reduced MYH9 expression in the TgFVB strain occurs secondary to glomerulosclerosis and is likely not causal in the development of nephropathy.

In conclusion, analysis of the most significantly associated SNPs in APOL1 and MYH9 in independent human cohorts confirmed that APOL1 haplotypes explain most of the association signal at the Chr. 22q13 locus, suggesting little or no association with HIVAN risk. The results also suggest that MYH9 may not be a major contributor to HIVAN risk, at least in the populations studied.
no remaining effect of the MYH9 locus on the risk of FSGS and HIVAN. These conclusions are further supported by the finding that Myh9 haploinsufficiency alone does not cause kidney disease, and crossbreeding with a validated mouse model of HIVAN produces only a negligible increase in proteinuria. We also confirm that the APOL1 G1 and G2 alleles are independent risk factors for FSGS and HIVAN, with a magnitude and direction of effects that was nearly identical to previously reported findings.12,13 The APOL1 variants did not impart the same large risk on IgAN, although we found a nominal association with increased risk of ESRD in this group, suggesting that APOL1 risk alleles may play a role in progression of some glomerulopathies; these findings will require confirmation in larger cohorts. Contrary to prior studies, however, the current analysis did not clearly favor the recessive over the additive model and showed similar odds ratios for the risk of HIVAN and FSGS. This suggests that APOL1-associated FSGS and HIVAN do not conform to a simple Mendelian recessive model. Given the significant risk conferred by the G1 and G2 haplotypes, better delineation of the genetic model can facilitate the development and interpretation of APOL1 genotyping as a predictive tool for nephropathy. In addition, about 12 to 15% of individuals with FSGS or HIVAN carry neither G1 nor G2 APOL1 risk alleles,12 suggesting that additional risk factors account for disease in this subgroup. Further investigation of APOL1 and downstream pathways will likely clarify novel pathogenetic mechanisms common to various forms of nephropathy.

CONCISE METHODS

Human Studies
We recruited African American patients with biopsy-documented idiopathic FSGS, HIVAN, or IgAN. FSGS was diagnosed by the presence of focal and segmental glomerular lesions in a proteinuric patient without known secondary causes such as drug toxicity or sickle cell disease; HIVAN was defined by the finding of collapsing FSGS in the setting of HIV infection; IgAN was defined by mesangial proliferation combined with predominant glomerular IgA deposition on immunofluorescence. In total, we studied 32 patients with IgAN, 44 patients with idiopathic FSGS (including eight collapsing cases), and 21 patients with HIVAN, as well as a control group of 74 healthy African American individuals with no proteinuria on urine dipstick. Participants were recruited from three medical centers: Columbia University (New York, NY), Mount Sinai Medical Center (New York, NY), and University of Alabama (Birmingham, AL). The study protocol was approved by the Columbia University, University of Alabama at Birmingham, and Mount Sinai Medical Center Institutional Review Board committees.

Genomic DNA was isolated from blood using DNeasy kit (Qiagen). We selected SNPs with the strongest reported associations in APOL112,13 and MYH910 and genotyped them by direct Sanger sequencing. The PCR primer sequences used are provided in Supplemental Table 5. The genotypes were called by using the SEQUENCHER software (GeneCodes). Genotyping accuracy was confirmed by bidirectional sequencing in approximately 10% of individuals. All of the genotypes underwent standard quality control procedures before association analysis: all SNPs had >99% genotyping rates, and all passed the Hardy-Weinberg equilibrium test in controls (P value >0.05).

We performed a standard allelic test of association using PLINK v.1.07.26 We estimated the per-allele odds ratios and 95% confidence intervals for all tested SNPs. We calculated the pairwise LD statistics (r2 and D’) for all analyzed SNPs after estimating haplotype frequencies using the EM algorithm. In the conditional analyses, we controlled for the genotypes of the conditioning SNPs using logistic regression under additive and recessive models. Because this is a replication study, we report nominal two-sided P values and considered P values <0.05 as significant. A Bonferroni adjustment for multiple testing can be obtained by multiplying P values by 6 (the number of SNPs tested in this study), but this would be an overly stringent correction because these loci are not independent (Supplemental Table 2).

The haplotypes were phased using the EM algorithm as implemented in PLINK v.1.0728; the haplotype frequencies were estimated in the patients and controls separately, as well as jointly in the entire cohort, to detect the likelihood of a true difference in frequency between patients and controls. Only the haplotypes with the overall frequency greater than 1% were included in association analyses. The odds ratios and the corresponding 95% confidence intervals were estimated in reference to the most common haplotype that carried no putative risk alleles.

Mouse Studies
Myh9 haploinsufficient mice (Myh9+/−) have been previously characterized in detail.21 The mice were maintained by backcrossing to B6 at Columbia University. For the HIV-1 transgenic mouse line, we backcrossed the well-characterized TgN(pNL434d1)26Lom 26 mice27 to B6 mice for at least seven generations to generate heterozygous HIV-1 transgenic strain on the B6 genetic background (TgB6). We confirmed that both Myh9+/− and TgB6 mice had a homogeneous B6 background by genotyping 82 informative loci across the genome (Khioscience, Hoddesdon, UK). Myh9 haploinsufficient mice (Myh9+/−, n = 29) were compared with wild-type littermates (B6, n = 23), TgB6 mice (n = 19), and Myh9+/− − TgB6 (n = 19) mice at 5 to 6 months of age. The animal protocol was approved by the institutional animal care and use committee at Columbia University.

Total RNA was extracted from mouse kidneys using TRIzol reagent (Invitrogen) followed by treatment with DNasel and cleanup using the RNeasy kit (Qiagen) according to the manufacturers’ protocols. cDNA was generated with the Omni-Script kit (Qiagen). Gene expression was quantitated by quantitative PCR using SYBR-Green mix and IQ5 thermal Cycler (Bio-Rad). β-actin was used as a housekeeping control.

We used albumin-to-creatinine ratio (Albuwell M and creatinine ELISA kits; Exocell) for quantification of urine albumin in a random subset (four to five mice in each genotype group). The degree of renal injury was evaluated by examination of kidney histopathology (Periodic acid–Schiff staining). Pathologic evaluation included the severity of tubulointerstitial disease (tubular casts, tubular dilation, and epithelial regeneration/regeneration), as well as the type and extent of glomerular injury (segmental and global glomerulosclerosis and collapse, podocyte hyperplasia, and collapse).
ACKNOWLEDGMENTS

This study was supported by National Institutes of Health Grants DK087626 (to A.G.G.), DK082753 (to A.G.G. and J.N.), and DK056492 (to P.E.K.), the Columbia Center for Glomerular Diseases, and the Daniel Midler Research Fund in Nephrology. K.K. is supported by National Institutes of Health Grant 1K23DK090207, and R.S. is supported by a fellowship from the Belgian-American Education Foundation.

DISCLOSURES

None.

REFERENCES

26. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham, P.C. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575, 2007


Supplemental information for this article is available online at http://www.jasn.org/.