Generalization of Associations of Kidney-Related Genetic Loci to American Indians

Nora Franceschini,* Karin Haack,† Laura Almasy,§ Sandra Laston,* Elisa T. Lee,* Lyle G. Best,‡ Richard R. Fabsitz,* Jean W. MacCluer,† Barbara V. Howard,¶ Jason G. Umans,** and Shelley A. Cole†

Summary

Background and objectives CKD disproportionately affects American Indians, who similar to other populations, show genetic susceptibility to kidney outcomes. Recent studies have identified several loci associated with kidney traits, but their relevance in American Indians is unknown.

Design, setting, participants, & measurements This study used data from a large, family-based genetic study of American Indians (the Strong Heart Family Study), which includes 94 multigenerational families enrolled from communities located in Oklahoma, the Dakotas, and Arizona. Individuals were recruited from the Strong Heart Study, a population-based study of cardiovascular disease in American Indians. This study selected 25 single nucleotide polymorphisms in 23 loci identified from recently published kidney-related genome-wide association studies in individuals of European ancestry to evaluate their associations with kidney function (estimated GFR; individuals 18 years or older, up to 3282 individuals) and albuminuria (urinary albumin to creatinine ratio; n=3552) in the Strong Heart Family Study. This study also examined the association of single nucleotide polymorphisms in the APOL1 region with estimated GFR in 1121 Strong Heart Family Study participants. GFR was estimated using the abbreviated Modification of Diet in Renal Disease Equation. Additive genetic models adjusted for age and sex were used.

Results This study identified significant associations of single nucleotide polymorphisms with estimated GFR in or nearby PRKAG2, SLC6A13, UBE2Q2, PIP5K1B, and WDR72 (P<2.1 × 10⁻⁵ to account for multiple testing). Single nucleotide polymorphisms in these loci explained 2.2% of the estimated GFR total variance and 2.9% of its heritability. An intronic variant of BCAS3 was significantly associated with urinary albumin to creatinine ratio. APOL1 single nucleotide polymorphisms were not associated with estimated GFR in a single variant test or haplotype analyses, and the at-risk variants identified in individuals with African ancestry were not detected in DNA sequencing of American Indians.

Conclusion This study extends the genetic associations of loci affecting kidney function to American Indians, a population at high risk of kidney disease, and provides additional support for a potential biologic relevance of these loci across ancestries.

Introduction

Familial aggregation of CKD (1,2) suggests a strong genetic susceptibility to CKD. Decreased estimated GFR (eGFR) and albuminuria (as assessed by urinary albumin to creatinine ratio [UACR]) are associated with increased risk of kidney disease and show high heritability, defined by the proportion of the phenotypic variance caused by genetic effects (36%–75% for eGFR and 16%–46% for UACR) (3–9). Recent genome-wide association (GWA) studies have highlighted the contribution of common variants (minor allele frequency≥0.05) to kidney traits. These studies have described several genomic regions (and genes) associated with CKD, eGFR, and UACR in the general population (10–15), and some of these loci are generalizable across ancestries. Additional loci, such as the region encoding MYH9 and APOL1 on chromosome 22, have been associated with nondiabetic kidney disease, FSGS, HIV-associated nephropathy, and hypertensive-attributed nephropathy (16–18). Two alleles of APOL1 confer risk to individuals of African ancestry (16), but it is unknown if this gene is associated with kidney traits in American Indians.

CKD disproportionally affects racial and ethnic minorities in the United States, especially American Indians, who also experience greater disparities in cardiometabolic disorders, including diabetes and obesity. Earlier stages of CKD are highly prevalent in American Indians, even among nondiabetic individuals. In the Strong Heart Study (SHS), which is a cohort of American Indians ages 45–74 years, 18% of the nondiabetic participants had CKD at baseline. Both the SHS and the Zuni Kidney Project, a population-based cross-sectional survey of the Pueblo of Zuni,
have shown microalbuminuria, a marker of kidney damage, at a rate of 10%–20%, but only one half of the individuals with albuminuria have diabetes (19,20). The Kidney Early Evaluation Program, a national kidney disease screening program, reported that 29% of self-identified American Indians and Alaskan Natives had either reduced kidney function or microalbuminuria (21). We have previously shown that MYH9 variants and haplotypes are not associated with kidney function or UACR in American Indian participants of the Strong Heart Family Study (SHFS) (22). The effects of APOL1 and other GWA-identified loci on kidney traits have not been previously reported in American Indians.

To investigate the role of known genetic variants in accounting for interindividual variation of kidney traits in a population with a high prevalence of CKD, we genotyped several GWA-identified single nucleotide polymorphisms (SNPs) in American Indians of the SHFS. Additionally, because of the results from studies of African Americans, we hypothesized that APOL1 variants are associated with kidney traits in American Indians. We examined the association of 15 APOL1 SNPs available for 1121 SHFS individuals. Here, we report findings for the generalization of the association of these loci with kidney traits in American Indians.

Materials and Methods

Study Population: The SHFS

The National Heart, Lung, and Blood Institute-funded SHFS is a large, family-based genetic study of cardiovascular disease in American Indians enrolled from 13 tribal communities located in Oklahoma, the Dakotas, and Arizona. Individuals were recruited from the parent study, the SHS, a population-based cohort study of American Indians enrolled from 13 tribal communities located in Oklahoma, the Dakotas, and Arizona. Individuals were recruited from the parent study, the SHS, a population-based cohort study of American Indians 45 years or older. Briefly, the SHFS began as a pilot study in 1998, when ~900 members of extended families of the SHS parent study were examined. Additional family members were recruited in 2001–2003 for a total sample of 3798 individuals in 94 multigenerational families (mean family size of 40 individuals, range=5–110 people). Extensive and detailed lifestyle, clinical, and laboratory measures and cardiovascular outcomes are available in the SHFS. The SHFS protocols were approved by the Institutional Review Board of all institutions, and the SHFS Institutional Review Board, Institutional Review Boards of all institutions, and the participating Indian tribes (23,24). All participants gave informed consent for genetic testing. This study includes SHFS American Indians seen during exam phase 4. We excluded individuals with missing serum creatinine or urine albumin/creatinine and individuals with unknown dialysis/transport status. Individuals younger than 18 years of age were also excluded for eGFR analysis. The final sample for eGFR was 3282 and 3218 individuals for eGFR analyses with and without including individuals with ESRD, respectively (before genotyping exclusions).

Clinical and Outcome Measures

Standardized anthropometric, clinical, and 12-hour fasting laboratory measures (serum and urine) were collected at each visit (including serum creatinine) as well as spot urine samples for albumin and creatinine measurement. Demographic data (age, sex, and education), lifestyle and behaviors (smoking and alcohol intake), medical history, and medications were obtained through questionnaires (24,25). BP was measured in resting individuals, and the last two of three measures were averaged. Body weight (kilograms) and height (meters) were used to estimate body mass index (kilograms per meter²). DNA samples were obtained. All laboratory tests were performed in a centralized laboratory. Serum and urine creatinine were assayed by a kinetic alkaline picrate method, and urine albumin was assayed by a sensitive nephelometric method; both were performed on the Hitachi 717 Platform (Roche Diagnostics, Indianapolis, IN). Type 2 diabetes is defined as a fasting glucose=126 mg/dl (7.0 mmol/L), history of diabetes, or use of diabetic medications (26). Hypertension is defined as BP>140/90 mmHg or use of antihypertensive medications (27).

Urinary albumin excretion was estimated as UACR (milligrams per gram), and eGFR was calculated using the abbreviated Modification of Diet in Renal Disease Equation (28).

SNP Selection, Genotyping, and Quality Control

SNPs were selected from GWA publications before January of 2011 for studies reporting findings for eGFR, serum creatinine, CKD, and UACR traits. Loci were retrieved from the National Human Genome Research Institute Genome Catalog (www.genome.gov/gwastudies/) for SNPs showing a genome-wide association of $P<5.0 \times 10^{-8}$. SHROOM3 variants were not available for analysis. Several SNPs in the APOL1 region were also available in a subset of SHFS participants, some of whom were related (16 tag SNPs; $n=1121$, with 753 half-sibling relative pairs [kinship coefficient=0.25] and 2891 first cousin relationships or less [kinship coefficient=0.125]). These SNPs were genotyped for a previous study of genetic variation affecting cardiovascular disease in this chromosomal region, and they were selected from HapMap CEU and CHB/JPT. SNPs were typed using the multiplex VeraCode technology from Illumina according to the manufacturer’s protocol (Illumina, San Diego, CA). Cluster calls were checked for accuracy, and genotypes were exported as text files for additional use in association analysis. Replica samples were included as controls for genotyping and allele calling consistencies.

Quality control included requiring sample call rates of >95%, concordance of blinded replicates, and removal of those individuals that deviated from Hardy–Weinberg equilibrium among the subsample of 1121 individuals ($P<0.001$). Individuals with more than 10% of missing genotypes were excluded. For SNPs in the same loci, we estimated pairwise linkage disequilibrium (LD) and report only independent signals using $r^2<0.9$. LD was plotted using Haplovieview version 4.2 (29).

Statistical Analyses

Markedly skewed quantitative traits were log-transformed to achieve normality. We first obtained center-specific residuals of eGFR and log UACR using linear regression models adjusted for age, age², sex, and age-by-sex interactions. For eGFR analyses, we tested two models: (1) one
model excluding individuals on renal replacement therapy (dialysis) and individuals with a kidney transplant (model 1, n=66 excluded) and (2) one model imputing an eGFR value of 10 ml/min per 1.73 m² for individuals on dialysis and individuals with a kidney transplant (model 2). Trait residuals were regressed into genotypes of each SNP, which were coded as the number of copies of the minor allele that each individual carries (additive genetic model). For single test association analyses, we used the measured genotype approach, which is a mixed model that accounts for family relatedness as random effects and covariates as fixed effects (30). Center-specific summary results were combined using fixed effects meta-analyses (31). To assess the evidence for between-center heterogeneity, we also estimated the between-center variance (tested at P<0.05) and the I² metric (percent of the total variation between studies because of heterogeneity rather than chance; >50% indicates large heterogeneity) (32). The total variance of eGFR and the percent heritability explained by genetic variants significantly associated with the trait were obtained using maximum likelihood variance decomposition methods (33) by comparing adjusted models with and without the SNPs.

Family data are robust to population stratification, and relatedness is adjusted using an already estimated kinship matrix. However, we also assessed the evidence for population stratification using methods described in the work by Havill et al. (34), which decomposes genotype scores into between-pairs (b) and within-pairs (w) components. Population stratification should affect only the parameter $\beta_b$. Stratification is inferred when estimates of $\beta_b$ and $\beta_w$ differ significantly. SNPs showing significant evidence for population stratification ($P<0.05$) were considered false positives.

$APOL1$ risk alleles from African ancestry were not available, and therefore, we also performed haplotype analyses to better capture potentially associated alleles in the region. Haplotypes were inferred using Phase (version 2.1), which implements a Bayesian statistical method for reconstructing haplotypes from population genotype data (35). We then evaluated the association of haplotype copies with kidney function (eGFR) using linear regression models.

We used the approach of Cheverud (36) and Nyholt (37) to estimate the effective number of tests to correct based on the intra-SNP correlations: $\alpha=2.1 \times 10^{-3}$ for 24 independent tests for GWA SNP associations (rs7805747 and rs10224210 were highly correlated) (Supplemental Table 1) and $\alpha=7.1 \times 10^{-3}$ for $APOL1$ SNPs (seven independent SNPs) (Figure 1, LD).

Figure 1. | Pairwise linkage disequilibrium at the $APOL1$ locus in American Indians by geographic region. (A) Arizona, (B) Oklahoma, and (C) the Dakotas. Graphs were generated from the American Indian data using Haploview 4.2. Colors represent $D^*$ values: blue, statistically ambiguous $D^*$; dark red, high inter-single nucleotide polymorphism (inter-SNP) $D^*$; white, low inter-SNP $D^*$. Linkage disequilibrium blocks are also shown.
Results

Characteristics of study participants are shown in Table 1. Exclusions are described in Materials and Methods. Hypertension and diabetes were more common in American Indians recruited from Arizona followed by Oklahoma and the Dakotas. Individuals from Arizona also had higher mean eGFR and median UACR than individuals from Oklahoma and the Dakotas. Of 25 investigated SNPs in 23 loci, 7 SNPs (6 loci) were significantly associated with eGFR in meta-analysis of the three centers ($P<2.1 \times 10^{-3}$ to account for multiple testing) (Table 2, model 1). These SNPs were in or nearby PRKAG2 (two SNPs), SLC6A13, UBE2Q2, SYPL2/PSMA5, PIP5K1B, and WDR72. Study-specific allele frequencies are shown in Supplemental Table 2. Only rs1933182 (SYPL2/PSMA5) showed significant evidence for population stratification ($P<0.001$) and was considered a false positive. The remaining five SNPs explained 2.2% of the total variance and 2.9% of the eGFR heritability (we included only the SNP with lowest $P$ at PRKAG2). In analysis including dialysis and transplant individuals and imputed values for eGFR, only rs10774021 and rs10224210 associations remained significant (Table 2, model 2). An intronic SNP of BCAS3 was significantly associated with UACR in meta-analysis (Table 2). Additional SNPs showing nominal associations in center-specific or meta-analysis for eGFR and UACR are shown in Supplemental Tables 2 and 3, respectively.

Of 15 SNPs available in the APOL1 region, 5 SNPs were located in an exon, and 3 SNPs were missense variants. Allele frequencies by center are shown in Table 3, and the LD in the region for each center is shown in Figure 1. In single SNP analysis, none of the variants were significantly associated with eGFR. Using seven highly correlated SNPs (rs136159, rs713929, rs2239785, rs136174, rs136175, rs136176, and rs136177), we identified two common haplotypes (GGAAAGA: frequency=0.93; AAGCGGG: frequency=0.06). We then used the most likely pairs of haplotypes for each individual (posterior probability=0.06). Among the loci associated with eGFR, SNPs in the SLC6A13 locus have been recently associated with eGFR in individuals of European ancestry (10) and East Asians (39). Of interest is the high proportion of the interindividual variation in eGFR explained by the SNPs in the five associated loci (2.2%), which corresponds to 2.9% of heritability of eGFR in our pedigrees. This finding is in contrast to the discovery population studies, in which 16 validated SNPs explain only UACR. Although significant, the association at the SYPL2/PSMA5 locus is likely attributable to population stratification. Several SNPs in other loci showed either significant or nominal associations with these traits within a geographic region. All alleles with eGFR were the same alleles identified in the meta-analysis of individuals of European ancestry, and an additional SNP in the PRKAG2 gene was in LD with the published SNP in HapMap CEU samples. Although there were regional allele frequency differences for SNPs, most of the associations showed no evidence for heterogeneity across regions. Comparison of the effect estimates for these SNPs with those SNPs from European populations showed the same direction of the at-risk allele estimates, supporting the relevance of these loci to American Indians (the magnitude of the effect cannot be compared, because eGFR residuals were estimated in our analysis).

The SNPs examined in this study were identified in analysis of eGFR in population studies, which usually have a small number of individuals with CKD. The SNP at UBE2Q2 and the SNP at PRKAG2 have been previously associated with a prospective reduction of eGFR (<60 ml/min per 1.73 m²) in individuals of European ancestry but not with ESRD (38). Of interest, the inclusion of individuals with ESRD ($n=66$) in our analysis attenuated the SNP associations with eGFR, and only the associations of the SNP at SLC6A13 and a SNP at PRKAG2 remained significant. This finding suggests a role of these loci on normal variation of kidney function rather than disease, although our results could be driven by the methods used for imputation of eGFR missing data in individuals with ESRD.

Among the loci associated with eGFR, SNPs in the SLC6A13 locus have been recently associated with eGFR in African Americans (12), and WDR72 was associated with eGFR in East Asians (39), suggesting that genetic variation at these loci may have relevance for kidney function across multiple ethnic groups/ancestries. The PRKAG2 locus was also associated with hematoctrit values (40) and CKD in individuals of European ancestry (10). BCAS3, associated with UACR in our data, was associated with eGFR, serum creatinine, and BUN in studies of individuals of European ancestry (10) and East Asians (39). Of interest is the high proportion of the interindividual variation in eGFR explained by the SNPs in the five associated loci (2.2%), which corresponds to 2.9% of heritability of eGFR in our pedigrees. This finding is in contrast to the discovery population studies, in which 16 validated SNPs explain only

Table 1. Descriptive characteristics of American Indians at baseline: Strong Heart Family Study 2001–2003

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Overall ($n=3599$)</th>
<th>Arizona ($n=1199$)</th>
<th>Dakotas ($n=1204$)</th>
<th>Oklahoma ($n=1196$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (SD), yr</td>
<td>39.6 (16.9)</td>
<td>36.5 (15.7)</td>
<td>38.8 (17.1)</td>
<td>43.5 (17.2)</td>
</tr>
<tr>
<td>Women, %</td>
<td>60.0</td>
<td>62.1</td>
<td>58.7</td>
<td>59.2</td>
</tr>
<tr>
<td>Mean eGFR (SD)a</td>
<td>98.5 (25.7)</td>
<td>111.6 (28.4)</td>
<td>92.8 (21.2)</td>
<td>91.6 (21.8)</td>
</tr>
<tr>
<td>Median UACR (interval)</td>
<td>7.8 (4.7–18.4)</td>
<td>10.3 (5.5–27.6)</td>
<td>7.1 (4.5–15.0)</td>
<td>7.1 (4.4–15.0)</td>
</tr>
<tr>
<td>Hypertension, %</td>
<td>30.6</td>
<td>32.4</td>
<td>22.8</td>
<td>36.7</td>
</tr>
<tr>
<td>Type 2 diabetes, %</td>
<td>21.6</td>
<td>31.4</td>
<td>13.5</td>
<td>20.1</td>
</tr>
</tbody>
</table>

eGFR, estimated GFR, ml/min per 1.73 m²; UACR, urine albumin to creatinine ratio.
aIndividuals ages 18 years or older.
Table 2. Significant results for association of estimated GFR and log-urine albumin to creatinine ratio with candidate single nucleotide polymorphisms

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chromosome Region</th>
<th>Gene (Function)</th>
<th>Effect Allele Frequency</th>
<th>Effect/Other Allele</th>
<th>Model 1</th>
<th>Model 2: P Value ESRD&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β (SE)</td>
<td>P Value</td>
<td>Total Number</td>
</tr>
<tr>
<td>eGFR</td>
<td>rs10774021</td>
<td>12p13.33</td>
<td>SLC6A13 (intron)</td>
<td>0.05–0.27</td>
<td>A/G</td>
<td>−0.16 (0.04)</td>
</tr>
<tr>
<td></td>
<td>rs1394125</td>
<td>15q24.2</td>
<td>UBE2Q2 (intron)</td>
<td>0.05–0.16</td>
<td>A/G</td>
<td>−0.17 (0.04)</td>
</tr>
<tr>
<td></td>
<td>rs10224210&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7q36.1</td>
<td>PRKAG2 (intron)</td>
<td>0.02–0.15</td>
<td>G/A</td>
<td>−0.19 (0.05)</td>
</tr>
<tr>
<td></td>
<td>rs491567</td>
<td>15q21.3</td>
<td>WDR72 (intron)</td>
<td>0.37–0.45</td>
<td>A/C</td>
<td>−0.09 (0.03)</td>
</tr>
<tr>
<td></td>
<td>rs1933182&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1p13.3</td>
<td>SYPL2/PSMA5</td>
<td>0.02–0.14</td>
<td>A/C</td>
<td>−0.16 (0.05)</td>
</tr>
<tr>
<td></td>
<td>rs7805747&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7q36.1</td>
<td>PRKAG2 (intron)</td>
<td>0.02–0.15</td>
<td>A/G</td>
<td>−0.17 (0.05)</td>
</tr>
<tr>
<td></td>
<td>rs4744712</td>
<td>9q21.11</td>
<td>PIP5K1B (intron)</td>
<td>0.04–0.28</td>
<td>A/C</td>
<td>−0.13 (0.04)</td>
</tr>
<tr>
<td>LUACR</td>
<td>rs9895661</td>
<td>17q23.2</td>
<td>BCAS3 (intron)</td>
<td>0.16–0.54</td>
<td>A/G</td>
<td>−0.11 (0.03)</td>
</tr>
</tbody>
</table>

SNPs are listed by the order of P values (lowest to highest). Effect estimates (β) are the change in eGFR or LUACR residuals per each copy of the effect allele from meta-analyses of three centers. eGFR and LUACR residuals were obtained from linear regression models adjusted for age, age<sup>2</sup>, sex, and age-by-sex interactions (Materials and Methods). A P<2.1×10<sup>-3</sup> was considered significant to account for multiple testing. SNP, single nucleotide polymorphism; eGFR, estimated GFR; LUACR, log urine albumin to creatinine ratio.

<sup>a</sup>Up to 3282 individuals.

<sup>b</sup>Note two significant SNPs at the PRKAG2 locus. These SNPs are 5.3 kb apart and in linkage disequilibrium in HapMap CEU (r<sup>2</sup>=0.96, D'=1.0). The published SNP is rs7805747.

<sup>c</sup>This SNP showed significant evidence for population stratification.
<table>
<thead>
<tr>
<th>SNP</th>
<th>Position (Build 36)</th>
<th>Variant type</th>
<th>Effect Allele</th>
<th>Other Allele</th>
<th>Arizona</th>
<th>The Dakotas</th>
<th>Oklahoma</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Freq</td>
<td>β (SE)</td>
<td>P</td>
<td>n</td>
<td>Freq</td>
</tr>
<tr>
<td>rs4821472</td>
<td>34977906</td>
<td>Flanking 5'UTR</td>
<td>G</td>
<td>A</td>
<td>0.05</td>
<td>−0.02 (0.20)</td>
<td>0.93</td>
</tr>
<tr>
<td>rs9610467</td>
<td>34979520</td>
<td>Intron</td>
<td>A</td>
<td>G</td>
<td>0.13</td>
<td>0.14 (0.11)</td>
<td>0.22</td>
</tr>
<tr>
<td>rs7284919</td>
<td>34982110</td>
<td>Intron</td>
<td>G</td>
<td>A</td>
<td>0.11</td>
<td>0.18 (0.12)</td>
<td>0.13</td>
</tr>
<tr>
<td>rs136148</td>
<td>34982877</td>
<td>Intron</td>
<td>G</td>
<td>A</td>
<td>0.07</td>
<td>−0.14 (0.16)</td>
<td>0.39</td>
</tr>
<tr>
<td>rs4820224</td>
<td>34983221</td>
<td>Intron</td>
<td>A</td>
<td>G</td>
<td>0.05</td>
<td>−0.05 (0.19)</td>
<td>0.77</td>
</tr>
<tr>
<td>rs136159a</td>
<td>34986969</td>
<td>Intron</td>
<td>A</td>
<td>G</td>
<td>0.03</td>
<td>−0.28 (0.21)</td>
<td>0.18</td>
</tr>
<tr>
<td>rs713929a</td>
<td>34987542</td>
<td>Intron</td>
<td>A</td>
<td>G</td>
<td>0.03</td>
<td>−0.28 (0.21)</td>
<td>0.17</td>
</tr>
<tr>
<td>rs2239785a</td>
<td>34991276</td>
<td>missense</td>
<td>G</td>
<td>A</td>
<td>0.04</td>
<td>−0.35 (0.19)</td>
<td>0.07</td>
</tr>
<tr>
<td>rs136174a</td>
<td>34991482</td>
<td>synonymous</td>
<td>C</td>
<td>A</td>
<td>0.03</td>
<td>−0.28 (0.21)</td>
<td>0.18</td>
</tr>
<tr>
<td>rs136175a</td>
<td>34991512</td>
<td>missense</td>
<td>G</td>
<td>A</td>
<td>0.03</td>
<td>−0.28 (0.21)</td>
<td>0.18</td>
</tr>
<tr>
<td>rs136176a</td>
<td>34991592</td>
<td>missense</td>
<td>G</td>
<td>A</td>
<td>0.03</td>
<td>−0.28 (0.21)</td>
<td>0.17</td>
</tr>
<tr>
<td>rs136177a</td>
<td>34991788</td>
<td>synonymous</td>
<td>G</td>
<td>A</td>
<td>0.03</td>
<td>−0.29 (0.20)</td>
<td>0.16</td>
</tr>
<tr>
<td>rs3886200a</td>
<td>34993573</td>
<td>Flanking 3'UTR</td>
<td>G</td>
<td>A</td>
<td>0.12</td>
<td>0.17 (0.12)</td>
<td>0.14</td>
</tr>
<tr>
<td>rs2012928</td>
<td>34993948</td>
<td>Flanking 3'UTR</td>
<td>G</td>
<td>A</td>
<td>0.10</td>
<td>0.28 (0.13)</td>
<td>0.03</td>
</tr>
<tr>
<td>rs136183</td>
<td>34996271</td>
<td>Flanking 3'UTR</td>
<td>A</td>
<td>G</td>
<td>0.12</td>
<td>0.20 (0.12)</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Effect estimates (β) are the change in estimated GFR residuals per each copy of the effect allele (Materials and Methods). SNP, single nucleotide polymorphism; Freq, effect allele frequency (minor allele); n, number; UTR, untranslated region.

aSNPs used for haplotype inference.
approximately 1.4% of the variation of eGFR (10). In addition, we were able to detect associations at these loci using a much smaller sample than the European discovery study (~67,000 European ancestry individuals versus ~3000 American Indians). Therefore, our study exemplifies the advantages of using diverse ancestry individuals and/or family data to identify associations with complex diseases.

For the APOL1 locus, we did not identify any associations for either single SNPs or haplotypes with kidney function. The known risk alleles at this locus have been restricted to individuals of African ancestry (16), because they seem to have arisen through positive selection in Africa. APOL1 is a trypanolytic factor in human serum, and the variants associated with kidney disease also confer resistance to Trypanosoma brucei rhodesiense, which is transmitted by the tsetse fly and causes sleeping sickness (41,42). In American Indians, we did not detect the African-derived risk alleles in DNA sequence of APOL1 coding regions (available from whole-exome sequencing data) in 94 SHFS individuals (data not shown), providing additional evidence that these risk variants are only present in Africans. However, we did detect strong LD among the SNPs that we genotyped, suggesting that there may be other selective pressures on this gene region in American Indians. Although negative, our findings do not rule out the possibility that additional variants in the APOL1 gene may influence kidney traits in American Indians.

Our study was limited to GWA published SNPs, and therefore, we cannot exclude the presence in American Indians of one or more potentially functional variants (common or rare) in loci not showing significant associations. In addition, APOL1 SNPs were available only in a subset of SHFS individuals with available genotyping. However, this study is the first study to show generalization of kidney-related loci to American Indians, a population at high risk of CKD. Our findings show the feasibility of genetic studies in this population, even if samples are limited, and provide additional support for a potential biologic relevance of these loci across different ancestries. Future studies on the association of these loci with longitudinal kidney outcomes will require genotyping of the SHS cohort, which has a long follow-up and incident events. Fine mapping of these loci and other known loci and genotyping of American Indian specific at-risk variants identified in sequencing data are needed in this population.

Prior studies have explored epidemiologic factors related to changes in eGFR and/or UACR. Most of the prior population studies have a low prevalence of albuminuria and reduction of kidney function (10,11). Here, we showed significant associations with loci relevant to kidney function and/or damage in a population of American Indians with a high prevalence of kidney disease. Therefore, our study extends the genetic findings for kidney function and UACR to American Indians, a population group that has been consistently underrepresented in genomics investigations, despite having a high prevalence of CKD.

Acknowledgments

The Strong Heart Study is supported by National Heart, Lung, and Blood Institute Grants U01 HL65520, U01 HL41642, U01 HL41652, U01 HL41654, and U01 HL65521. Funding for genotyping was obtained through American Heart Association Grant 0675001N. This study was conducted in part by facility improvements funded by National Center for Research Resources Grant C06 RR013556.

Disclosures

None.

References


knowledge identifies six novel genes related to kidney function. 


*Science* 339: 841–845, 2010


*Hum Mol Genet* 127: 295–310, 2017


*Nat Genet* 41: 315–323, 2009


42. Kronenberg F: APOL1 variants and kidney disease. There is no such thing as a free lunch. *Nephrol Dial Transplant* 26: 775–778, 2011

**Received:** February 25, 2013 **Accepted:** August 5, 2013

Published online ahead of print. Publication date available at www.jasn.org.

This article contains supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.02300213/-/DCSupplemental.

See related editorial, “Generalizability of Genetic Findings Related to Kidney Function and Albuminuria,” on pages 8–11.