Effect of Genetic African Ancestry on eGFR and Kidney Disease

Miriam S. Udler,*†‡ Girish N. Nadkarni,*‡§ Gillian Belbin,†‡ Vaneet Lotay,‡ Christina Wyatt,*§ Omri Gottesman,*‡ Erwin P. Bottinger,*‡§ Eimear E. Kenny,†‡‖¶ and Inga Peter†‡¶

Departments of *Medicine and †Genetics and Genomic Sciences, ‡The Charles Bronfman Institute for Personalized Medicine, §Division of Nephrology, ‖The Center for Statistical Genetics, and ¶The Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, New York

ABSTRACT

Self-reported ancestry, genetically determined ancestry, and *APOL1* polymorphisms are associated with variation in kidney function and related disease risk, but the relative importance of these factors remains unclear. We estimated the global proportion of African ancestry for 9048 individuals at Mount Sinai Medical Center in Manhattan (3189 African Americans, 1721 European Americans, and 4138 Hispanic/Latino Americans by self-report) using genome-wide genotype data. CKD-EPI eGFR and genotypes of three *APOL1* coding variants were available. In admixed African Americans and Hispanic/Latino Americans, serum creatinine values increased as African ancestry increased (per 10% increase in African ancestry, creatinine values increased 1% in African Americans and 0.9% in Hispanic/Latino Americans; $P \leq 1x10^{-7}$). eGFR was likewise significantly associated with African genetic ancestry in both populations. In contrast, *APOL1* risk haplotypes were significantly associated with CKD, eGFR<45 ml/min per 1.73 m^2, and ESRD, with effects increasing with worsening disease states and the contribution of genetic African ancestry decreasing in parallel. Using genetic ancestry in the eGFR equation to reclassify patients as black on the basis of ≥50% African ancestry resulted in higher eGFR for 14.7% of Hispanic/Latino Americans and lower eGFR for 4.1% of African Americans, affecting CKD staging in 4.3% and 1% of participants, respectively. Reclassified individuals had electrolyte values consistent with their newly assigned CKD stage. In summary, proportion of African ancestry was significantly associated with normal-range creatinine and eGFR, whereas *APOL1* risk haplotypes drove the associations with CKD. Recalculation of eGFR on the basis of genetic ancestry affected CKD staging and warrants additional investigation.

J Am Soc Nephrol 26: 1682–1692, 2015. doi: 10.1681/ASN.2014050474

It is estimated that 6.7% of the United States population have CKD, which is defined as eGFR<60 ml/min per 1.73 m^2 and/or evidence of kidney damage.[1] Values of serum creatinine and eGFR and rates of CKD and ESRD vary significantly by ancestry and ethnicity in the United States.[1,2] Paradoxically, whereas eGFR is found to be, on average, higher in African Americans (AAs) than European Americans (EAs),[3] AAs also have 4-fold higher rates of ESRD than EAs.[1] The cause of these differences is thought to be multifactorial, including environmental and genetic contributions.[4] Recently, variants in the Apolipoprotein L1 (*APOL1*) locus, which are found in individuals with African ancestry, were shown to confer a 10-fold increased risk of hypertensive ESRD, 17-fold increased risk of FSGS, and 29-fold increased risk of HIV-associated nephropathy.[5–7] *APOL1* variants were also associated with higher rates of CKD progression in AAs, which was defined by eGFR slope or doubling of serum creatinine,[8] and predicted younger age of dialysis initiation

Received May 14, 2014. Accepted September 2, 2014.

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

Correspondence: Dr. Miriam S. Udler, Massachusetts General Hospital, Thier 11-1101, 50 Blossom Street, Boston MA 02114. Email: mudler@mgh.harvard.edu

in nondiabetic ESRD AAs and Hispanic/Latino Americans (H/LAs).[9,10] These findings strongly support a genetic component to population differences in the prevalence and incidence of CKD.

A potential issue in studies of CKD in ancestrally and ethnically diverse populations is that the most routinely used measurement of kidney function, eGFR, is calculated using equations that incorporate the patient's ancestry as a binary coding of black or not black. The two most commonly used and extensively validated equations are the Modification of Diet in Renal Disease (MDRD)[11] and Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI),[12] and both incorporate into the estimates serum creatinine, age, sex, and ancestry. The study populations used in deriving and validating the CKD-EPI equation were more ethnically diverse than those used for MDRD; however, they only included 10%–30% AAs and 2%–5% H/LAs.[12] Furthermore, classifying a person as black or not black is prone to error and can vastly oversimplify ancestral and ethnic identity,[13] particularly in mixed ancestry (or admixed) populations, such as AAs and H/LAs. Using genome-scale genetic data, it is possible to accurately determine an individual's genetic ancestry, which is the proportion of an individual's genome that is ancestral to continental source populations. In the case of AAs, genetic ancestry is often modeled as a proportion of African and European ancestry,[14,15] and H/LAs are modeled with African, European, and Native American ancestry.[16–18] In this study, we were interested in exploring how genetically determined African ancestry may affect estimates of eGFR and the related disease risk.

This analysis was performed using data from the Bio*Me* Biobank Program of The Charles Bronfman Institute for Personalized Medicine at Mount Sinai Medical Center, which is a repository of genetic data linked to participants' electronic medical records (EMRs). This hospital-based population included 9048 unrelated adult participants for whom laboratory data on eGFR and genome-wide single-nucleotide polymorphisms (SNPs) were available made up of 3189 AAs, 1721 EAs, and 4138 H/LAs. Such a diverse population constitutes an ideal setting to investigate the contribution of genetically determined ancestry and self-reported ancestry on kidney function estimates in normal and disease states. Because *APOL1* variants have recently been identified as important genetic contributors to the increased risk of CKD seen in AAs, we were also interested in determining how much these variants accounted for the variability in genetic risk in our diverse cohort.

RESULTS

Study Population

Global ancestry proportions derived from genome-scale genetic data[19,20] were available for 10,998 biobank participants made up of 3764 AAs, 2040 EAs, and 5194 H/LAs. After quality control, 10,320 individuals (3550 AAs, 2040 EAs, and 4730 H/LAs) were confirmed to be unrelated (Concise Methods), and 9048 of these individuals had laboratory data available to calculate eGFR (3189 AAs, 1721 EAs, and 4138 H/LAs). Clinical characteristics of 9048 study participants included in our analyses are shown in Table 1.

AAs were, on average, younger than participants from the other ethnic groups. Compared with EAs and H/LAs, AAs had significantly higher body mass index (BMI), systolic BP (SBP), and diastolic BP (DBP) after adjustment for age and sex ($P<0.001$) (Table 1). AAs also had higher prevalence of diabetes than EAs with adjustment for age and sex, but it was not significantly different from H/LAs. Participants who self-identified as AAs were, on average, 82.0% (range=0.6%–100%) African by genetic ancestry, whereas H/LAs were 28.6% African (range=0%–100%) and EAs were 2.9% African (range=0%–26.2%) (Table 1). Distributions of percentages of African, European, and Native American ancestries in each population are given in Supplemental Figure 1.

APOL1 Imputation

In total, 4106 biobank participants (3764 AAs, 19 EAs, and 323 H/LAs) were custom genotyped for three *APOL1* variants (G1a rs73885319, G1b rs60910145, and G2 rs71785313) previously shown to be associated with ESRD.[5,7] The genotypes were also imputed using the 1000 Genomes Project phase 1 release (December of 2013) as a reference panel[21] in these individuals as well as an additional 2019 EAs and 4811 H/LAs for whom genome-wide Illumina OmniExpress Exome array SNP data were available but custom *APOL1* genotyping results were not available (Concise Methods). Imputed genotypes were compared with typed genotypes in 4106 individuals for whom both data were available, with >99% concordance in each racial group (Supplemental Table 1). This finding is consistent with previous work showing accurate imputation of common alleles (minor allele frequency>5%).[22–24]

In AAs, the frequencies of *APOL1* risk alleles were 22% for G1 rs73885319, 21% for G1 rs60910145, and 13% for G2 rs71785313. The frequency of carrying two risk alleles (necessary to possess risk in the assumed recessive model) at any of the variants (G2, G1, or G1/G2) was 14.5% in AAs, 2.0% in H/LAs, and 0.05% in EAs.

Genetic African Ancestry and Creatinine

Consistent with previous studies,[2,3] AAs, on average, had higher serum creatinine values than H/LAs and EAs, even after adjustment for age, sex, SBP, DBP, diabetes, smoking, and BMI ($P<0.001$ for both group comparisons) (Table 1). These interpopulation differences were attenuated with adjustment for genetic African ancestry ($P>0.05$) (Figure 1A, Table 2), suggesting that genetic factors and/or population-level environmental factors may account for observed differences in creatinine levels.

In both the admixed AAs and H/LAs, there was a significant trend of increasing creatinine as percentage of African ancestry increased (Figure 1A). This trend was even more significant

Table 1. Characteristics of Biobank study population

Characteristics	Self-Reported Ancestral/Ethnic Group		
	AA	EA	H/LA
N (total=9048)	3189	1721	4138
Age (yr), mean (SEM)	52.7 (0.26)	68.1 (0.22)[a]	55.0 (0.25)[a]
Women (%)	64.8	47.8[a]	63.0
Diabetes (%)	30.1	16.1[b]	32.4
Hemoglobin A1c (%), mean[c]	6.37	5.89[a]	6.44
BMI (kg/m^2), mean (SEM)[c]	30.8 (0.14)	27.2 (0.13)[b]	29.7 (0.10)[b]
SBP (mmHg), mean (SEM)[c]	129.1 (0.25)	126.3 (0.34)[b]	126.2 (0.22)[b]
DBP (mmHg), mean (SEM)[c]	75.9 (0.15)	73.8 (0.19)[b]	72.9 (0.12)[b]
Serum creatinine (mg/dl), mean[d,e]	1.02	0.99[a]	0.92[a]
Serum creatinine (mg/dl), mean adjusted[d,e,f]	1.02	0.89[a]	0.92[a]
eGFR (ml/min per 1.73 m^2) stages (%)			
1: eGFR≥90	45.0	15.1	37.9
2: eGFR=60–89	39.8	59.7	43.9
3: eGFR=30–59	11.4	22.4	14.8
3a: eGFR=45–59	8.0	15.2	10.7
3b: eGFR=30–44	3.4	7.2	4.1
4: eGFR characteristics=15–29	1.4	1.9	1.6
5: eGFR<15	2.3	1.0	1.8
eGFR, mean (SEM)[d]	85.9 (27.5)	70.4 (18.9)	80.9 (24.1)
eGFR, mean adjusted[d,f] (SEM)	86.4 (15.6)	70.5 (8.2)[a]	81.6 (15.4)[a]
CKD (%)[g]	15.0 (0.6)	16.4 (0.9)	15.1 (0.6)
ESRD (%)[h]	3.4	1.6[a]	2.3[a]
ESRD adjusted (%)[f,h]	4.6	1.9[a]	2.9[a]
Genetic ancestry (%), mean (range)			
African	82.0 (0.6–100)	2.9 (0–26.2)	28.6 (0–100)
European	15.8 (0–96.4)	94.5 (68.3–97.3)	54.1 (0–96.5)
Native American	2.3 (0–68.3)	2.6 (0–28.3)	17.3 (0–98.6)
APOL1 risk haplotype frequency (%)	14.5	0.05	2.0

[a] $P<0.001$ compared with AAs.
[b] $P<0.001$ compared with AAs with adjustment for age and sex.
[c] Mean of all yearly medians.
[d] Most recent yearly median value.
[e] Exponent of natural log creatinine.
[f] Adjusted for age, sex, average SBP, average DBP, diabetes status, BMI, and smoking.
[g] CKD defined using algorithm as defined in Concise Methods.
[h] ESRD is defined as eGFR<15 ml/min per 1.73 m^2 or patient on hemodialysis.

when creatinine levels were restricted to normal-range values (<1.3 mg/dl in men and <1.1 mg/dl in women as defined by Peralta *et al.*[25]), with each 10% increase in African ancestry associated with a 1% increase in creatinine in AAs and a 0.9% increase in creatinine in H/LAs ($P\leq 1\times10^{-7}$) (Table 2). When using a threshold for elevated creatinine (≥1.3 mg/dl in men and ≥1.1 mg/dl in women) and considering it as a binary trait,[26] elevated creatinine was significantly associated with percentage of African ancestry in ancestry-pooled individuals using a fully adjusted model (odds ratio [OR], 1.11; 95% confidence interval [95% CI], 1.07 to 1.16; $P<9\times10^{-8}$) (Table 2).

Carrying two *APOL1* risk variants was significantly associated with higher serum creatinine level in all individuals stratified by population group independent of genetic African ancestry in an adjusted model ($P=2\times10^{-10}$) (Table 2). When this analysis was restricted to only diabetic individuals, the association was attenuated ($P=0.01$) (Supplemental Table 2). Additionally, the association between *APOL1* and serum creatinine disappeared in analyses restricted to normal-range creatinine (<1.3 mg/dl in men and <1.1 mg/dl in women), suggesting that it was driven by individuals with kidney disease (Table 2, Supplemental Table 2).

Genetic Ancestry and eGFR

eGFR estimates using the CKD-EPI equation with AAs classified as black and H/LAs and EAs classified as not black were significantly higher in AAs than in H/LAs or EAs in the adjusted model ($P<0.001$ for both comparisons) (Table 2). Unlike the analyses with creatinine, adjustment for genetic African ancestry did not attenuate the interpopulation differences in eGFR between AAs and H/LAs (Figure 1B).

In AAs and H/LAs, there was a significant trend of decreasing eGFR as percentage of African ancestry increased (adjusted $\beta=-1.62$, $P=3\times10^{-8}$ in AAs; $\beta=-0.87$, $P=4\times10^{-8}$ in H/LAs; both per 10% increase in African ancestry) (Table 2), which was only partially attenuated in those with eGFR>90 ml/min per 1.73 m^2 (Table 2).

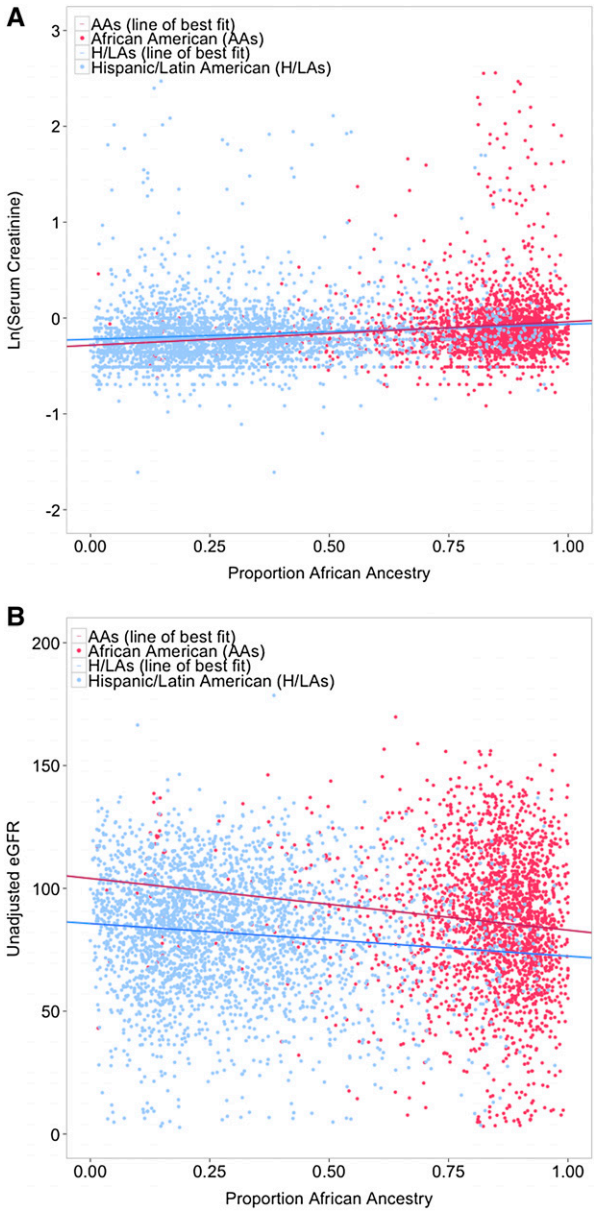

Figure 1. Adjustment for proportion genetic African ancestry attenuates differences in serum creatinine between AAs and H/LAs, but not differences in eGFR. As shown in A, the natural log of serum creatinine values in H/LAs and AAs increase with proportion of genetic African ancestry (unadjusted β=0.02 units on log scale per 10% genetic African ancestry; $P=2\times10^{-5}$ in AAs; unadjusted β=0.01 units on log scale per 10% genetic African ancestry; $P=8\times10^{-5}$ in H/LAs). After adjusting for proportion of genetic African ancestry in a full model, there was no longer a significant difference in average creatinine levels between these populations ($P>0.05$). B shows that eGFR levels decrease with increased proportion of genetic African ancestry in both AAs and H/LAs (unadjusted $\beta=-1.69$ ml/min per 1.73 m^2 per 10% genetic African ancestry; $P=2\times10^{-7}$ in AAs; unadjusted $\beta=-1.26$ ml/min per 1.73 m^2 per 10% genetic African ancestry; $P=7\times10^{-12}$ in H/LAs). However, unlike the analyses with creatinine, adjustment for genetic African ancestry did not attenuate the interpopulation differences in eGFR between AAs and H/LAs ($P<10^{-10}$ in adjusted model).

With the addition of *APOL1* risk variants to the model, the effect of genetic African ancestry on eGFR remained significant in both AA and H/LA populations and seemed to be driven by individuals who do not have diabetes (Table 2, Supplemental Table 2). Additionally, the *APOL1* risk haplotypes were no longer significantly associated with eGFR in those with eGFR>90 ml/min per 1.73 m^2 (Table 2, Supplemental Table 2).

Genetic Ancestry and CKD/ESRD

CKD was defined on the basis of data in the patients' EMRs using a validated algorithm, such that either two values of eGFR<60 ml/min per 1.73 m^2 were recorded ≥3 months apart or one value of eGFR<60 ml/min per 1.73 m^2 was recorded in addition to use of particular International Classification of Diseases, Ninth Revision (ICD-9) codes or chart documentation (Concise Methods). When adjusted for age, sex, diabetes, SBP, and DBP, AAs had higher rates of CKD than non-AAs (OR, 1.50; 95% CI, 1.24 to 1.82; $P<0.001$ compared with EAs; OR, 1.18; 95% CI, 1.02 to 1.37; $P=0.03$ compared with H/LAs; data not shown). These intergroup differences in prevalence of CKD were attenuated with adjustment for genetic African ancestry and *APOL1* risk haplotypes; these two factors were also independently associated with CKD risk (genetic African ancestry: OR, 1.08; 95% CI, 1.04 to 1.17; $P=1\times10^{-4}$; *APOL1* risk haplotype: OR, 1.57; 95% CI, 1.19 to 2.05; $P=0.001$) (Table 2).

AAs also had higher rates of ESRD (defined as eGFR<15 ml/min per 1.73 m^2 or on dialysis) even after adjustment for traditional CKD risk factors (Table 1). Adjusted rates of ESRD were no longer significantly higher in AAs when also adjusted for percentage of African ancestry (Table 2). With addition of the *APOL1* variants to the model, African ancestry was no longer significantly associated with ESRD (Table 2), suggesting that *APOL1* variants account for the majority of increased ESRD risk in individuals of African ancestry. Once again, the associations with *APOL1* haplotypes and ESRD seemed to be driven by individuals who do not have diabetes (OR, 4.81; 95% CI, 2.57 to 8.98; $P=9\times10^{-7}$); however, we also observed a study-wide borderline association in patients with diabetes (OR, 2.81; 95% CI, 1.54 to 5.10; $P=0.001$) (Supplemental Table 2).

Recalculated eGFR on the Basis of Proportion of Genetic African Ancestry

GFR estimates using CKD-EPI and MDRD equations incorporate ancestry and ethnicity as a binary variable: black or not black. The CKD-EPI equation increases the GFR of black individuals by a factor of 1.159. It is unclear how well these binary categories, which largely rely on perceived or self-reported ancestry, account for the actual underlying environmental and genetic differences in GFR, and the relationship between serum creatinine and GFR between populations.[27] When individuals in our study were partitioned by percentage of genetic African ancestry (<25%, 25%–50%, 50%–75%, and ≥75%), a significant increase in serum creatinine levels was noted only in those with ≥50% genetic African ancestry compared with

Table 2. Contributions of genetic ancestry, *APOL1*, and self-reported ancestry to normal renal function and disease states in full model*

Characteristics	Genetic African Ancestry (per 10%)		*APOL1* Risk Haplotype		Comparison of AA Versus (*P* Value)	
	β or OR (95% CI)	*P* Value	β or OR (95% CI)	*P* Value	EA	H/LA
AAs						
Log creatinine[a]	0.02 (0.008 to 0.03)	1×10^{-4}	0.08 (0.04 to 0.12)	1×10^{-4}		
Log creatinine[a] restricted[b]	0.01 (0.007 to 0.016)	3×10^{-7}	0.0003 (−0.02 to 0.02)	0.98		
eGFR (ml/min per 1.73 m^2)	−1.62 (−2.19 to −1.04)	3×10^{-8}	−2.69 (−5.12 to −0.25)	0.03		
eGFR>90 ml/min per 1.73 m^2	−1.07 (−1.52 to −0.62)	4×10^{-6}	−0.94 (−3.04 to 1.15)	0.36		
Elevated creatinine[c]	1.14 (1.06 to 1.23)	0.001	1.27 (0.98 to 1.66)	0.07		
CKD	1.19 (1.09 to 1.30)	1×10^{-4}	1.32 (0.98 to 1.79)	0.07		
eGFR<45 ml/min per 1.73 m^2	1.07 (0.96 to 1.20)	0.22	1.69 (1.15 to 2.48)	0.007		
ESRD (eGFR<15 ml/min per 1.73 m^2)	1.09 (0.93 to 1.28)	0.31	2.69 (1.66 to 4.35)	5×10^{-5}		
H/LAs						
Log creatinine[a]	0.009 (0.003 to 0.01)	0.001	0.26 (0.18 to 0.35)	2×10^{-9}		
Log creatinine[a] restricted[b]	0.009 (0.005 to 0.01)	7×10^{-8}	0.01 (−0.05 to 0.06)	0.82		
eGFR (ml/min per 1.73 m^2)	−0.87 (−1.18 to −0.56)	4×10^{-8}	−11.33 (−16.14 to −6.51)	4×10^{-6}		
eGFR>90 ml/min per 1.73 m^2	−0.34 (−0.60 to −0.09)	0.008	−2.32 (−6.49 to 1.85)	0.27		
Elevated creatinine[c]	1.09 (1.03 to 1.14)	0.001	3.46 (1.86 to 6.43)	9×10^{-5}		
CKD	1.03 (0.98 to 1.09)	0.20	3.02 (1.57 to 5.81)	0.001		
eGFR<45 ml/min per 1.73 m^2	1.05 (0.99 to 1.12)	0.13	4.07 (1.92 to 8.65)	2×10^{-4}		
ESRD (eGFR<15 ml/min per 1.73 m^2)	0.89 (0.79 to 1.01)	0.07	11.06 (4.47 to 27.35)	2×10^{-7}		
All individuals						
Log creatinine[a]	0.01 (0.008 to 0.02)	2×10^{-8}	0.11 (0.07 to 0.14)	2×10^{-5}	0.59	0.12
Log creatinine[a] restricted[b]	0.01 (0.007 to 0.01)	1×10^{-14}	0.002 (−0.17 to 0.02)	0.83	0.02	0.19
eGFR (ml/min per 1.73 m^2)	−1.17 (−1.44 to −0.90)	4×10^{-17}	−4.23 (−6.20 to −2.26)	2×10^{-5}	<0.001	<0.001
eGFR>90 ml/min per 1.73 m^2	−0.64 (−0.87 to −0.40)	4×10^{-8}	−1.27 (−2.89 to 0.35)	0.13	<0.001	<0.001
Elevated creatinine[c]	1.11 (1.07 to 1.16)	9×10^{-8}	1.49 (1.17 to 1.91)	0.001	0.24	0.49
CKD	1.08 (1.04 to 1.13)	1×10^{-4}	1.57 (1.19 to 2.05)	0.001	0.12	0.02
eGFR<45 ml/min per 1.73 m^2	1.08 (1.02 to 1.14)	0.008	2.01 (1.43 to 2.83)	7×10^{-5}	0.02	0.13
ESRD (eGFR<15 ml/min per 1.73 m^2)	0.99 (0.90 to 1.08)	0.77	3.54 (2.30 to 5.41)	8×10^{-5}	0.09	0.16

*Adjusted for age, sex, diabetes status, SBP, DBP, BMI, and smoking.
[a]Natural log of creatinine used in model.
[b]Creatinine<1.3 mg/dl in men and <1.1 mg/dl in women.
[c]Creatinine≥1.3 mg/dl in men and ≥1.1 mg/dl in women.

those with <25% (Figure 2A). To investigate the role of genetic ancestry on GFR measures, we recategorized individuals, choosing individuals with ≥50% genetic African ancestry to define black individuals and those with <50% genetic African ancestry to define not black individuals. For AAs who were found by genetic ancestry to be <50% African, we divided their current eGFR by 1.159 to remove the correction factor for black. In a similar fashion, for H/LAs who were found to be ≥50% African, we increased their eGFR by a factor of 1.159. No EAs in our study had ≥50% African ancestry, and thus, their eGFRs were not adjusted.

Using these criteria, 131 self-reported AAs (4.1%) and 609 H/LAs (14.7%) had eGFRs that were recalculated. With the adjusted eGFR measures, 33 AAs (1%) were reclassified to a higher CKD stage, and 180 H/LAs (4.3%) were reclassified to a lower (less severe) CKD stage (Table 3). Additionally, with this eGFR adjustment, there was no longer a significant difference between eGFR values in AAs and H/LAs in the same category of percentage of African ancestry (Figure 2B).

We then investigated whether individuals who were reclassified to a new CKD stage on the basis of genetic African ancestry had metabolic characteristics that were more similar to their new staging than their old staging, which would support the reclassification. In the ancestry-pooled and stratified analyses, average levels of serum bicarbonate and potassium significantly varied with CKD stage ($P_{trend}<0.001$ in adjusted and unadjusted regression analyses for both electrolytes) (Supplemental Table 3). We focused on electrolyte values in the H/LA population, because considerably more H/LA individuals than AA or EA individuals were reclassified to a new stage using genetic ancestry (Table 3). Serum levels of yearly median bicarbonate and potassium were evaluated in H/LAs who were reclassified to a new CKD stage, and these values were compared with yearly median values of other individuals

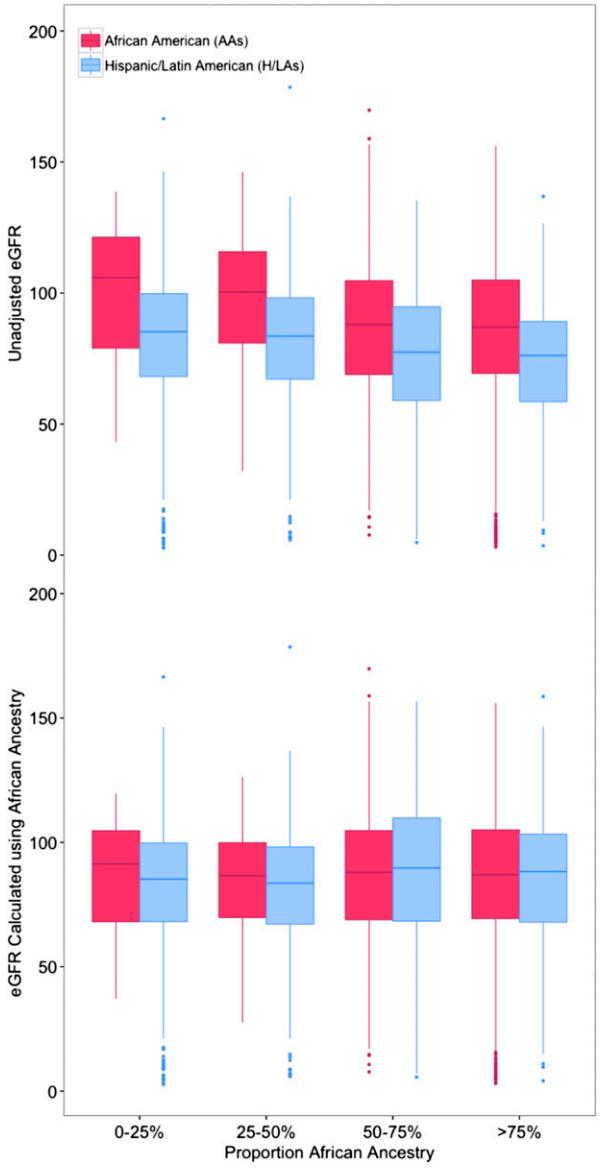

Figure 2. GFR in AAs and H/As calculated using the CKD-EPI equation is compared with a modification of the equation incorporating genetic African ancestry. Unadjusted eGFR (shown at the top of the figure) is higher on average in AAs than H/LAs, and this is seen in each category of proportion genetic African ancestry. The eGFR was then recalculated to incorporate genetic African ancestry, as follows. For self-reported AAs who were found by genetic ancestry to be <50% African, their current eGFR was divided by 1.159 to remove the black correction factor. In a similar fashion, for H/LAs who were found to be ≥50% African, their GFR was increased by a factor 1.159. With eGFR recalculated using genetic African ancestry, there was no longer a significant difference between eGFR values in AAs and H/LAs by category of percentage of African ancestry.

in the original and the new CKD stage assignments (Supplemental Table 3). Because average potassium and bicarbonate levels did not differ significantly between HA/LAs in CKD stages 1 and 2 but were significantly different between those with CKD stages 2 and 3 (Supplemental Table 3), we focused on 52 H/LAs who were reclassified from CKD stage 3 to CKD stage 2 and had measured bicarbonate and potassium levels. Their mean bicarbonate level was 26.9 (95% CI, 26.2 to 27.6), which was higher than the mean bicarbonate of 466 HL/As with their original CKD stage 3 (mean=26.0; 95% CI, 25.8 to 26.3; P=0.02) and not statistically different from 1529 H/LAs with their newly assigned CKD stage 2 (mean=26.6; 95% CI, 26.5 to 26.8; P=0.78) adjusting for age and sex (Supplemental Table 3). Likewise, potassium levels in these 52 reclassified HL/As were lower than H/LAs with their original CKD stage 3 (mean=4.1; 95% CI, 4.0 to 4.2 [n=466]; mean=4.3; 95% CI, 4.3 to 4.4; P=0.004 using log-transformed potassium values) and not significantly different from the H/LAs with their newly assigned CKD stage 2 (mean=4.2; 95% CI, 4.2 to 4.2 [n=1529]; P=0.10) adjusting for age and sex (Supplemental Table 3). Thus the average serum bicarbonate and potassium levels of the reclassified individuals were more similar to their new classification group than their old classification group, supporting their new reclassification.

DISCUSSION

In a large ancestrally diverse cohort, we investigated the associations of (*1*) self-reported ancestry, (*2*) ancestry inferred from genetic data, and (*3*) *APOL1* genotype with various components of kidney function. Specifically, the proportion of genetically derived African ancestry was significantly associated with normal-range variation of creatinine and eGFR, whereas *APOL1* risk haplotypes were associated with CKD and ESRD.

Our results are consistent with previous reports that African ancestry is positively associated with elevated serum creatinine levels in young AA men.[26] We were able to extend these findings to men and women of a broader age range and those who self-reported as H/LA. In our study, the H/LA population exhibited an average of 29% (range=0%–100%) African ancestry, which reflects the predominantly (approximately 80%) Caribbean origin of the H/LA participants.[28] Higher levels of serum creatinine in individuals of African ancestry compared with other ethnic groups have been attributed to larger muscle mass and differences in tubular creatinine secretion.[2,27,29] Our findings strongly support that genetic ancestry is predictive of physiologic differences in both serum creatinine and eGFR.

The MDRD and CKD-EPI equations are the most widely used indices of kidney function; however, it is debated whether eGFR accurately estimates kidney function in admixed populations, such as AAs and H/LAs.[27] Given the large variation in genetically derived African ancestry, especially among H/LAs, we reclassified AAs and H/LAs as black and not black on the basis of an individual's proportion of African ancestry (≥50% versus <50%) rather than self-report and recalculated eGFR. This resulted in lower eGFR in 4.1% of AAs and higher eGFR in 14.7% of H/LAs, with reclassification to a higher CKD stage

Table 3. Reclassification of CKD stage on the basis of genetic ancestry versus self-reported ancestry

CKD Stage Using Self-Reported Ancestry	CKD Stage Using Genetic Ancestry[a]					
	1	2	3a	3b	4	5
AAs[b]						
1	1414	22				
2		1260	10			
3a			252	3		
3b				108	1	
4					46	0
5						73
HAs[c]						
1	1567					
2	111	1705				
3a		57	386			
3b			18	152		
4				10	58	
5					2	72

[a]GFR calculated on the basis of genetic ancestry as follows: regardless of self-reported ancestry, if <50% African ancestry, use CKD-EPI[9] for not black, and if >50% African, use CKD-EPI for black.
[b]Total number reclassified: 36 (1%).
[c]Total number reclassified: 198 (4.8%).

in 1.0% of AAs and a lower CKD stage in 4.3% of H/LAs. The reclassification of individuals to a different CKD stage was supported by their potassium and bicarbonate levels matching more closely with their newly assigned group than their old group. The cutoff of 50% African ancestry to define black was on the basis of analysis of differences in serum creatinine in individuals partitioned by percentage of African ancestry (<25%, 25%–50%, 50%–75%, and >75%) and observation that a significant difference in creatinine level was only present in individuals with ≥50% African ancestry compared with those with <25% African ancestry. However, future studies, including a gold standard measure of GFR, will have the benefit of including genetic ancestry as a continuous variable (because they will not be limited by the structure of existing equations) and will be able to determine whether use of genetic ancestry improves estimates of GFR. Because ancestry estimates are increasingly within reach (through biobanks, clinical sequencing, or direct-to-consumer genetic testing), this information could ultimately be incorporated into clinical decision support applications delivered through EMRs similarly to existing personalized care approaches, such as genetic prediction of drug response variability.[30,31]

If our findings are confirmed with studies using gold standard measurements of GFR, the clinical implications of improved GFR estimation with incorporation of genetic ancestry would be quite significant. Current practice guidelines by Kidney Disease Outcome Quality Initiative recommend additional clinical investigations beginning at eGFR<60 ml/min per 1.73 m^2, which include anemia evaluation, nutritional workup, assessment of bone disease and calcium and phosphorous metabolism, and monitoring of patient wellbeing.[32] Incorporation of genetic African ancestry in our cohort led to 0.3% of AAs being newly classified as eGFR<60 ml/min per 1.73 m^2 and 1.4% of H/LAs no longer being classified as such, which could have measurable public health consequences for the estimated 6.7% of the United States population with CKD.[1]

In our patient population, *APOL1* risk haplotypes were significantly associated with CKD, eGFR<45 ml/min per 1.73 m^2, and ESRD in AAs and H/LAs. We also observed significant associations between *APOL1* and serum creatinine as well as eGFR, which seemed to be driven by disease-range values. Additionally, in analyses in patients who do not have diabetes as well as patients with diabetes, *APOL1* risk haplotypes were associated with CKD and ESRD; in patients with diabetes, the study-wise borderline significance of these associations (P=0.05 for CKD; P=0.001 for ESRD) may be affected by limited power, with only one third of study participants classified as diabetic. There have been differing results in studies as to whether the risk effect of *APOL1* extends to diabetic kidney disease,[8,33–36] which may also, in part, be attributable to differences in study power as well as prevalence of *APOL1*-associated intermediate phenotypes in the diabetic populations.

Because *APOL1* variants were genotyped in only a subset of our population and the *APOL1* locus is not represented on the Illumina OmniExpress Exome array, to increase the power of our study, we imputed the *APOL1* variants using genome-wide genotype data in subjects for whom genotyping data were not available. Our results indicated >99% concordance rates for the subset of samples for which both typed and imputed genotype data were generated. This exercise shows that the *APOL1* haplotypes can be successfully imputed and further analyzed in conjunction with other clinical traits in large cohorts with existing genome-wide genotype data. The success of imputation may vary on the basis of the types of genotyping arrays, reference panels, imputation software, and the number of individuals with African ancestry, for whom these haplotypes are common and can be more accurately imputed.[22–24]

The strengths of our study include the largest and most ethnically diverse cohort to date to address the effect of genetic ancestry on eGFR. Our subjects were enrolled in a biobank setting, where longitudinal phenotypic data contained within the EMR can be extracted, allowing adjustment for potential confounders.[29] Importantly, this study's discoveries, after they are validated, could be introduced into clinical research protocols and the EMR to augment clinical care. Moreover, we used a large genome-wide genotype panel to quantify the genetic ancestry for each participant. Previous studies mostly relied on self-reported ancestry, country of origin, or a small set of ancestry informative markers.[25,26,37]

This study's limitations include the lack of a gold standard GFR measure to assess the improvement in accuracy of eGFR with recalculation using African genetic ancestry; however, we were primarily interested to see whether using genetic ancestry would significantly affect eGFR estimates. Although eGFR values changed in 18.8% of participants and CKD staging changed in 5.3% of participants, it is unknown whether recalculated estimates would increase the sensitivity and specificity of CKD diagnosis. Furthermore, although the EMR offers access to a broad range of phenotypes for many individuals, the depth of the phenotyping is limited. For example, we had limited data on albuminuria and cause of kidney disease. There is also a potential bias introduced by using a clinical cohort where laboratory testing was performed on the basis of clinical appropriateness in contrast to standardized cohort studies, where data are available for all individuals. We tried to minimize this bias in our analysis of electrolytes by looking at commonly measured electrolytes, which would be checked in individuals routinely at all stages of CKD. Our findings were also possibly confounded by socioeconomic and environmental factors (*e.g.*, diet, lifestyle, and access to care) that were not accounted for in our study. Nevertheless, we adjusted for self-reported ancestry, which in the United States is strongly associated with socioeconomic status,[4] as well as established risk factors, such as diabetes and hypertension, which are likely to be affected by similar environmental exposures.

In summary, serum creatinine levels are higher in AAs compared with H/LAs, but this difference disappears with adjustment for percentage of African ancestry. Higher proportion of genetic African ancestry is associated with higher serum creatinine levels and lower eGFR, especially within the normal value ranges, whereas *APOL1* genotype is more predictive of kidney disease. Significant increases in serum creatinine occur in individuals with ≥50% genetic African ancestry compared with <25% genetic African ancestry. Recalculation of eGFR on the basis of a 50% threshold for genetic African ancestry led to reclassification of the CKD stage in 5.3% patients, and serum bicarbonate and potassium values in these individuals were more similar to their newly assigned group, supporting the reclassification. If validated and proven to be feasible, genetic ancestry could be incorporated to enhance GFR estimates in clinical and epidemiologic studies and ultimately, improve patient care.

CONCISE METHODS

Subjects

Study participants were recruited from the Bio*Me* Biobank Program of The Charles Bronfman Institute for Personalized Medicine at Mount Sinai Medical Center from 2007 on. The Bio*Me* Biobank is a consented EMR-linked medical care setting biorepository drawing from a population of over 70,000 inpatient and 800,000 outpatient visits annually. Mount Sinai Medical Center serves the diverse local communities of upper Manhattan, including Central Harlem (86% AA), East Harlem (88% H/LA), and Upper East Side (88% EA/white), with broad health disparities.[38] Bio*Me* populations include 28% AA, 38% H/LA (predominantly of Caribbean origin), 23% EA/white, and 11% East Asian, South Asian, or other ancestry (the final category was not included in this study). At the time of enrollment, participants were asked to describe their family background and ethnicity. For the purposes of this study, participants were classified as AA if they self-reported to be AA (92.5%), African (0.3%), black (2.2%), or non-Hispanic Afro-Caribbean (5%); H/LA if they self-reported to be H/LA (94%), Native American (0.2%), mixed H/LA and African ancestry (2%), mixed H/LA and European ancestry (1.2%), Central American (1%), or South American (1.6%); and EA if they self-reported to be Caucasian or white (67%), Jewish (21.6%), or European (11%). Since 2007, almost 30,000 Mount Sinai patients have enrolled in the Bio*Me* program. Bio*Me* operations are fully integrated in clinical care processes, including direct recruitment from clinical sites' waiting areas and phlebotomy stations by dedicated Biobank recruiters independent of clinical care providers before or after a clinician standard of care visit. Recruitment currently occurs at a broad spectrum of over 30 clinical care sites. This study included 10,320 unrelated ($\pi \leq 5\%$[39]) biobank participants selected for having genome-wide SNP genotyping without prior knowledge of laboratory values available in their EMRs. The Bio*Me* Biobank Program (Institutional Review Board 07–0529) operates under a Mount Sinai Institutional Review Board-approved research protocol. All study participants provided written informed consent.

Phenotypic Data

Phenotypic data were extracted from patients' EMRs. For continuous variables (serum creatinine, potassium, bicarbonate, eGFR, SBP, DBP, BMI, and hemoglobin A1C), all values were extracted from the EMR, and median values were calculated for each year from 2003 to 2012. In our analyses, we used the most recent yearly median creatinine and eGFR values, and the yearly median bicarbonate and potassium values from that same year. The average of all yearly median values was used for SBP, DBP, and BMI. Type 2 diabetes status was determined using the Electronic Medical Records and Genomics (eMERGE) Network type 2 diabetes phenotyping algorithm.[40] eGFR was determined using the CKD-EPI creatinine equation: $141 \times \min(\text{serum creatinine}/k \text{ or } 1)^{\alpha} \times \max(\text{serum creatinine}/k \text{ or } 1)^{1.209} \times 0.993^{\text{age}} \times (1.018 \text{ for women}) \times (1.159 \text{ for black})$, where serum creatinine is in milligrams per deciliter, k is 0.7 for women and 0.9 for men, α is -0.329 for women and -0.411 for men, min is the minimum of serum creatinine/k or 1, and max is the maximum of serum creatinine/k or 1.[9] With this calculation, we incorporated self-reported ancestry. CKD status was derived using data in the patients' EMRs, with CKD defined on the basis of the 2012 Kidney Disease Improving Global Outcomes criteria (eGFR<60 ml/min per 1.73 m^2 for duration≥3 months). Participants were defined as having CKD if they had either two eGFR values≥3 months apart that were <60 ml/min per 1.73 m^2 or one value of eGFR that was <60 ml/min per 1.73 m^2 recorded in addition to a documented diagnosis of CKD in the problem list, ICD-9 code list, or the text of progress notes. This algorithm has been validated by expert reviewers at multiple institutions through

eMERGE and has a positive predictive value of 96% and a negative predictive value of 100%.

Genetic Data

Genetic Ancestry Analysis

Genome-wide SNP genotyping was performed in 10,998 individuals using the Illumina OmniExpress Exome array (n>900,000). SNPs passing genotype calling quality control metrics were further pruned before analysis using the PLINK software[39] as follows: (*1*) SNPs with minor allele frequency>0.01 with a call rate>0.9 were included, (*2*) SNPs in tight linkage disequilibrium (r^2>0.8) were removed, and (*3*) SNPs that are within regions of persistent haplotypes, which are known to confound principal component analysis,[41] were removed (for example, the Human Leukocyte Antigen: chr6: 27,000,000–35,000,000 [NCBI37/hg19]; Lactase gene: chr2: 135,000,000–137,000,000 [NCBI37/hg19]; and a common inversion: chr8: 6,000,000–16,000,000 [NCBI37/hg19]). Finally, we intersected the data with a reference panel representing the putative ancestral diversity in the Bio*Me* cohort, including five populations from the 1000 Genomes Project[21]: Utah residents with ancestry from northern and western Europe (n=85), Yoruba in Ibadan, Nigeria (n=88), Colombians in Medellin, Colombia (n=60), people with Mexican ancestry in Los Angeles, California (n=59), and Puerto Ricans in Puerto Rico (n=55). We also intersected the data with a panel of Native American individuals (n=42) who were genotyped on the Affymetrix 6.0 platform. A total of 99,296 SNPs remained after pruning and merging. Global proportions of EA, AA, or Native American ancestry per individual were determined using the ADMIXTURE algorithm.[19,20] All 10,320 Bio*Me* Biobank participants and panels were included together in a single run with a putative ancestral population number of k=3 and 5-fold crossvalidation (yielding a log likelihood of 0.59156).

APOL1 Genotyping and Imputation

In total, 4106 (3764 AA, 19 EA, and 323 H/LA) Biobank participants were genotyped for three *APOL1* variants (G1a rs73885319, G1b rs60910145, and G2 rs71785313) using Luminex custom genotyping. *APOL1* G1/G2 genotype testing incorporates PCR and multiplex allele specific primer extension with Luminex's proprietary Universal Tag sorting system on the Luminex 100 xMAP platform. To validate this genotyping method, we performed intra- and inter-assay variation studies that included 48 positive control and 10 negative control samples. Sanger sequencing was used to confirm all of genotypes called by the Luminex method. Among 58 representative samples with all four haplotypes on G1 and G2 loci, the Sanger sequencing results completely agreed with the *APOL1* genotyping results.

Because the *APOL1* SNPs were not assayed on the chip, we used the genome-wide Illumina OmniExpress array SNP data and the 1000 Genomes Project phase 1 panel to impute the three *APOL1* genotypes as implemented in SHAPEIT[42] and IMPUTE2[43] without knowledge of custom *APOL1* genotyping results. Imputed genotypes were compared with typed genotypes in 4106 individuals (3764 AA, 19 EA, and 323 H/LA) for whom both were available, and there was >99% concordance in all groups (Supplemental Table 1).

Statistical Methods

The intergroup differences in clinical characteristics were assessed with logistic regression for binary traits and linear regression for continuous traits. Variables that were not normally distributed (serum creatinine and serum potassium) were log-transformed before analysis. For the trait hemoglobin A1c, which had a distribution more severely departing from normal, even after log transformation, intergroup differences were assessed with the Wilcoxon rank-sum test.

The relationships between continuous variables and ancestral/ethnic groups, genetic ancestry, and *APOL1* variants were assessed using linear regression with adjustment of age, sex, and percentage of African ancestry. Additional covariates included BMI, SBP, DBP, presence of diabetes, and smoking status (ever or never). The relationships between dichotomous variables and ancestral/ethnic groups, genetic ancestry, and *APOL1* variants were assessed with logistic regression to allow adjustment for additional covariates.

Associations with *APOL1* variants and covariates were assessed with the recessive model (carriage of two copies of any variant versus zero or one copy) using the Fisher exact test and logistic regression with adjustment for the covariates.

All analyses were performed using STATA/IC, version 11. The study-wide threshold for statistical significance was set at 0.001 as a conservative estimate on the basis of multiple testing.

ACKNOWLEDGMENTS

We thank the participants in the Bio*Me* Biobank for their invaluable contribution to biomedical research. We also are grateful to the reviewers and to Dr. Noah Zaitlen at University of California, San Francisco for helpful insights that improved this manuscript.

The Bio*Me* Biobank program is supported by The Andrea and Charles Bronfman Philantropies. O.G., E.P.B., E.E.K. and I.P. were supported, in part, by eMERGE Network Grant U01-HG006380 (to E.P.B.). The Electronic Medical Records and Genomics Network was initiated and funded by the National Human Genome Research Institute through Grants U01-HG006828 (to Cincinnati Children's Hospital Medical Center/Harvard), U01-HG006830 (to Children's Hospital of Philadelphia), U01-HG006389 (to Essentia Institute of Rural Health), U01-HG006382 (to Geisinger Clinic), U01-HG006375 (to Group Health Cooperative), U01-HG006379 (to Mayo Clinic), U01-HG006380 (to Icahn School of Medicine at Mount Sinai), U01-HG006388 (to Northwestern University), U01-HG006378 (to Vanderbilt University), and U01-HG006385 (to Vanderbilt University serving as the Coordinating Center).

DISCLOSURES

None.

REFERENCES

1. United States Renal Data System: 2013 Annual Data Report, 2013. Available at: http://www.usrds.org/adr.aspx. Accessed May 14, 2014

2. Gallagher D, Visser M, De Meersman RE, Sepulveda D, Baumgartner RN, Pierson RN, Harris T, Heymsfield SB: Appendicular skeletal muscle mass: Effects of age, gender, and ethnicity. *J Appl Physiol (1985)* 83: 229–239, 1997
3. Stevens LA, Claybon MA, Schmid CH, Chen J, Horio M, Imai E, Nelson RG, Van Deventer M, Wang HY, Zuo L, Zhang YL, Levey AS: Evaluation of the Chronic Kidney Disease Epidemiology Collaboration equation for estimating the glomerular filtration rate in multiple ethnicities. *Kidney Int* 79: 555–562, 2011
4. Patzer RE, McClellan WM: Influence of race, ethnicity and socioeconomic status on kidney disease. *Nat Rev Nephrol* 8: 533–541, 2012
5. Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI, Bowden DW, Langefeld CD, Oleksyk TK, Uscinski Knob AL, Bernhardy AJ, Hicks PJ, Nelson GW, Vanhollebeke B, Winkler CA, Kopp JB, Pays E, Pollak MR: Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 329: 841–845, 2010
6. Kopp JB, Nelson GW, Sampath K, Johnson RC, Genovese G, An P, Friedman D, Briggs W, Dart R, Korbet S, Mokrzycki MH, Kimmel PL, Limou S, Ahuja TS, Berns JS, Fryc J, Simon EE, Smith MC, Trachtman H, Michel DM, Schelling JR, Vlahov D, Pollak M, Winkler CA: APOL1 genetic variants in focal segmental glomerulosclerosis and HIV-associated nephropathy. *J Am Soc Nephrol* 22: 2129–2137, 2011
7. Tzur S, Rosset S, Shemer R, Yudkovsky G, Selig S, Tarekegn A, Bekele E, Bradman N, Wasser WG, Behar DM, Skorecki K: Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum Genet* 128: 345–350, 2010
8. Parsa A, Kao WH, Xie D, Astor BC, Li M, Hsu CY, Feldman HI, Parekh RS, Kusek JW, Greene TH, Fink JC, Anderson AH, Choi MJ, Wright JT Jr., Lash JP, Freedman BI, Ojo A, Winkler CA, Raj DS, Kopp JB, He J, Jensvold NG, Tao K, Lipkowitz MS, Appel LJ; AASK Study InvestigatorsCRIC Study Investigators: APOL1 risk variants, race, and progression of chronic kidney disease. *N Engl J Med* 369: 2183–2196, 2013
9. Tzur S, Rosset S, Skorecki K, Wasser WG: APOL1 allelic variants are associated with lower age of dialysis initiation and thereby increased dialysis vintage in African and Hispanic Americans with non-diabetic end-stage kidney disease. *Nephrol Dial Transplant* 27: 1498–1505, 2012
10. Kanji Z, Powe CE, Wenger JB, Huang C, Ankers E, Sullivan DA, Collerone G, Powe NR, Tonelli M, Bhan I, Bernhardy AJ, Dibartolo S, Friedman D, Genovese G, Pollak MR, Thadhani R: Genetic variation in APOL1 associates with younger age at hemodialysis initiation. *J Am Soc Nephrol* 22: 2091–2097, 2011
11. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D; Modification of Diet in Renal Disease Study Group: A more accurate method to estimate glomerular filtration rate from serum creatinine: A new prediction equation. *Ann Intern Med* 130: 461–470, 1999
12. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J; CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration): A new equation to estimate glomerular filtration rate. *Ann Intern Med* 150: 604–612, 2009
13. Caulfield T, Fullerton SM, Ali-Khan SE, Arbour L, Burchard EG, Cooper RS, Hardy BJ, Harry S, Hyde-Lay R, Kahn J, Kittles R, Koenig BA, Lee SS, Malinowski M, Ravitsky V, Sankar P, Scherer SW, Seguin B, Shickle D, Suarez-Kurtz G, Daar AS: Race and ancestry in biomedical research: Exploring the challenges. *Genome Med* 1: 8, 2009
14. Bryc K, Auton A, Nelson MR, Oksenberg JR, Hauser SL, Williams S, Froment A, Bodo JM, Wambebe C, Tishkoff SA, Bustamante CD: Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc Natl Acad Sci U S A* 107: 786–791, 2010
15. Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I, Beaty TH, Mathias R, Reich D, Myers S: Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet* 5: e1000519, 2009
16. Tang H, Coram M, Wang P, Zhu X, Risch N: Reconstructing genetic ancestry blocks in admixed individuals. *Am J Hum Genet* 79: 1–12, 2006
17. Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, Hammer M, Bustamante CD, Ostrer H: Colloquium paper: Genome-wide patterns of population structure and admixture among Hispanic/Latino populations. *Proc Natl Acad Sci U S A* 107[Suppl 2]: 8954–8961, 2010
18. Gravel S, Zakharia F, Moreno-Estrada A, Byrnes JK, Muzzio M, Rodriguez-Flores JL, Kenny EE, Gignoux CR, Maples BK, Guiblet W, Dutil J, Via M, Sandoval K, Bedoya G, Oleksyk TK, Ruiz-Linares A, Burchard EG, Martinez-Cruzado JC, Bustamante CD; 1000 Genomes Project: Reconstructing Native American migrations from whole-genome and whole-exome data. *PLoS Genet* 9: e1004023, 2013
19. Alexander DH, Novembre J, Lange K: Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19: 1655–1664, 2009
20. Zhou H, Alexander D, Lange K: A quasi-Newton acceleration for high-dimensional optimization algorithms. *Stat Comput* 21: 261–273, 2011
21. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA; 1000 Genomes Project Consortium: An integrated map of genetic variation from 1,092 human genomes. *Nature* 491: 56–65, 2012
22. Radmanesh F, Devan WJ, Anderson CD, Rosand J, Falcone GJ: Accuracy of imputation to infer unobserved APOE epsilon alleles in genome-wide genotyping data [published online ahead of print January 22, 2014]. *Eur J Hum Genet* 10.1038/ejhg.2013.308
23. Nelson SC, Doheny KF, Pugh EW, Romm JM, Ling H, Laurie CA, Browning SR, Weir BS, Laurie CC: Imputation-based genomic coverage assessments of current human genotyping arrays. *G3 (Bethesda)* 3: 1795–1807, 2013
24. Howie B, Marchini J, Stephens M: Genotype imputation with thousands of genomes. *G3 (Bethesda)* 1: 457–470, 2011
25. Peralta CA, Li Y, Wassel C, Choudhry S, Palmas W, Seldin MF, Risch N, Siscovick D, Arnett D, Psaty B, Shlipak MG: Differences in albuminuria between Hispanics and whites: An evaluation by genetic ancestry and country of origin: The multi-ethnic study of atherosclerosis. *Circ Cardiovasc Genet* 3: 240–247, 2010
26. Peralta CA, Risch N, Lin F, Shlipak MG, Reiner A, Ziv E, Tang H, Siscovick D, Bibbins-Domingo K: The Association of African Ancestry and elevated creatinine in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Nephrol* 31: 202–208, 2010
27. Delanaye P, Mariat C, Maillard N, Krzesinski JM, Cavalier E: Are the creatinine-based equations accurate to estimate glomerular filtration rate in African American populations? *Clin J Am Soc Nephrol* 6: 906–912, 2011
28. Moreno-Estrada A, Gravel S, Zakharia F, McCauley JL, Byrnes JK, Gignoux CR, Ortiz-Tello PA, Martínez RJ, Hedges DJ, Morris RW, Eng C, Sandoval K, Acevedo-Acevedo S, Norman PJ, Layrisse Z, Parham P, Martínez-Cruzado JC, Burchard EG, Cuccaro ML, Martin ER, Bustamante CD: Reconstructing the population genetic history of the Caribbean. *PLoS Genet* 9: e1003925, 2013
29. Goldwasser P, Aboul-Magd A, Maru M: Race and creatinine excretion in chronic renal insufficiency. *Am J Kidney Dis* 30: 16–22, 1997
30. Gottesman O, Scott SA, Ellis SB, Overby CL, Ludtke A, Hulot JS, Hall J, Chatani K, Myers K, Kannry JL, Bottinger EP: The CLIPMERGE PGx Program: Clinical implementation of personalized medicine through electronic health records and genomics-pharmacogenomics. *Clin Pharmacol Ther* 94: 214–217, 2013
31. Pulley JM, Denny JC, Peterson JF, Bernard GR, Vnencak-Jones CL, Ramirez AH, Delaney JT, Bowton E, Brothers K, Johnson K, Crawford DC, Schildcrout J, Masys DR, Dilks HH, Wilke RA, Clayton EW, Shultz E, Laposata M, McPherson J, Jirjis JN, Roden DM: Operational implementation of prospective genotyping for personalized medicine: The design of the Vanderbilt PREDICT project. *Clin Pharmacol Ther* 92: 87–95, 2012
32. National Kidney Foundation: Clinical Practice Guidelines for Nutrition in Chronic Renal Failure, 2000. Available at: http://www.kidney.org/professionals/kdoqi/guidelines/doqi_nut.html. Accessed May 14, 2014

33. Friedman DJ, Kozlitina J, Genovese G, Jog P, Pollak MR: Population-based risk assessment of APOL1 on renal disease. *J Am Soc Nephrol* 22: 2098–2105, 2011
34. Foster MC, Coresh J, Fornage M, Astor BC, Grams M, Franceschini N, Boerwinkle E, Parekh RS, Kao WH: APOL1 variants associate with increased risk of CKD among African Americans. *J Am Soc Nephrol* 24: 1484–1491, 2013
35. Freedman BI, Hicks PJ, Bostrom MA, Comeau ME, Divers J, Bleyer AJ, Kopp JB, Winkler CA, Nelson GW, Langefeld CD, Bowden DW: Non-muscle myosin heavy chain 9 gene MYH9 associations in African Americans with clinically diagnosed type 2 diabetes mellitus-associated ESRD. *Nephrol Dial Transplant* 24: 3366–3371, 2009
36. Kao WH, Klag MJ, Meoni LA, Reich D, Berthier-Schaad Y, Li M, Coresh J, Patterson N, Tandon A, Powe NR, Fink NE, Sadler JH, Weir MR, Abboud HE, Adler SG, Divers J, Iyengar SK, Freedman BI, Kimmel PL, Knowler WC, Kohn OF, Kramp K, Leehey DJ, Nicholas SB, Pahl MV, Schelling JR, Sedor JR, Thornley-Brown D, Winkler CA, Smith MW, Parekh RS; Family Investigation of Nephropathy and Diabetes Research Group: MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nat Genet* 40: 1185–1192, 2008
37. Stehman-Breen CO, Gillen D, Steffes M, Jacobs DR Jr., Lewis CE, Kiefe CI, Siscovick D: Racial differences in early-onset renal disease among young adults: The coronary artery risk development in young adults (CARDIA) study. *J Am Soc Nephrol* 14: 2352–2357, 2003
38. Tayo BO, Teil M, Tong L, Qin H, Khitrov G, Zhang W, Song Q, Gottesman O, Zhu X, Pereira AC, Cooper RS, Bottinger EP: Genetic background of patients from a university medical center in Manhattan: Implications for personalized medicine. *PLoS ONE* 6: e19166, 2011
39. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575, 2007
40. Kho AN, Hayes MG, Rasmussen-Torvik L, Pacheco JA, Thompson WK, Armstrong LL, Denny JC, Peissig PL, Miller AW, Wei WQ, Bielinski SJ, Chute CG, Leibson CL, Jarvik GP, Crosslin DR, Carlson CS, Newton KM, Wolf WA, Chisholm RL, Lowe WL: Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. *J Am Med Inform Assoc* 19: 212–218, 2012
41. Patterson N, Price AL, Reich D: Population structure and eigenanalysis. *PLoS Genet* 2: e190, 2006
42. Delaneau O, Zagury JF, Marchini J: Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 10: 5–6, 2013
43. Howie BN, Donnelly P, Marchini J: A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 5: e1000529, 2009

This article contains supplemental material online at http://jasn.asnjournals.org/lookup/suppl/doi:10.1681/ASN.2014050474/-/DCSupplemental.