Published in final edited form as:

Am J Nephrol. 2015 ; 42(2): 99–106. doi:10.1159/000439448.

Re-sequencing of the *APOL1-APOL4* and *MYH9* gene regions in African Americans does not identify additional risks for CKD progression

Gregory A. Hawkins^{1,2}, David J. Friedman³, Lingyi Lu^{1,4}, David R. McWilliams^{1,4}, Jeff W. Chou^{1,4}, Satria Sajuthi^{1,4}, Jasmin Divers^{1,4}, Rulan Parekh⁵, Man Li⁶, Giulio Genovese⁷, Martin R. Pollak³, Pamela J. Hicks⁸, Donald W. Bowden^{2,8}, Lijun Ma^{1,9}, Barry I. Freedman^{1,9,*}, and Carl D. Langefeld^{1,2,*}

¹Center for Public Health Genomics, Wake Forest School of Medicine, Winston-Salem, NC

²Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC

³Department of Internal Medicine, Division of Nephrology, Beth Israel Deaconess Medical Center, Boston MA

⁴Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem, NC

⁵Nephrology, Hospital for Sick Children, Toronto, CA

⁶Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

⁷Broad Institute, Boston MA

⁸Department of Biochemistry and Center for Diabetes Research, Wake Forest School of Medicine, Winston-Salem, NC

⁹Department of Internal Medicine, Section on Nephrology, Wake Forest School of Medicine, Winston-Salem, NC

Abstract

Background—*APOL1* G1 and G2 nephropathy risk variants are associated with non-diabetic end-stage kidney disease (ESKD) in African Americans (AAs) in an autosomal recessive pattern. Additional risk and protective genetic variants may be present near the *APOL1* loci since earlier age ESKD is observed in some AAs with one *APOL1* renal-risk variant and because the adjacent gene *MYH9* is associated with nephropathy in populations lacking G1 and G2 variants.

Methods—Re-sequencing was performed across a ~275 kb region encompassing the *APOL1-APOL4* and *MYH9* genes in 154 AA cases with non-diabetic ESKD and 38 controls without nephropathy who were heterozygous for a single *APOL1* G1 or G2 risk variant.

Correspondence: Barry I. Freedman, M.D., Section on Nephrology, Wake Forest School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157-1053, U.S.A., Phone: 336-716-6192, Fax: 336-716-4318, bfreedma@wakehealth.edu.

*denotes equal contributors

Results—Sequencing identified 3246 non-coding single nucleotide polymorphisms (SNPs), 55 coding SNPs, and 246 insertion/deletions (InDels). No new coding variations were identified. Eleven variants, including a rare *APOL3* Gln⁵⁸Ter null variant (rs11089781), were genotyped in a replication panel of 1571 AA ESKD cases and 1334 controls. After adjusting for *APOL1* G1 and G2 risk effects, these variations were not significantly associated with ESKD. In subjects with <2 *APOL1* G1 and/or G2 alleles (849 cases; 1139 controls), the *APOL3* null variant was nominally associated with ESKD (recessive model, OR 1.81; p=0.026); however, analysis in 807 AA cases and 634 controls from the Family Investigation of Nephropathy and Diabetes (FIND) did not replicate this association.

Conclusion—Additional common variants in the *APOL1-APOL4-MYH9* region do not contribute significantly to ESKD risk beyond the *APOL1* G1 and G2 alleles.

Keywords

African Americans; *APOL1*; kidney disease; FSGS; genetics; DNA sequencing

Introduction

Five common etiologies of end-stage kidney disease (ESKD) are associated with the G1 and G2 coding variants in the apolipoprotein L1 gene (*APOL1*) on chromosome 22q13 [1–3]. This observation has altered our understanding of the pathogenesis of non-diabetic nephropathy in populations with recent African ancestry. *APOL1*-associated nephropathies are inherited in an autosomal recessive pattern; risk is present in those inheriting two copies of the G1 and/or the G2 nephropathy variants.[1] Despite this inheritance pattern, individuals with a single *APOL1* nephropathy variant initiate renal replacement therapy or develop nephropathy at earlier ages than those without *APOL1* nephropathy variants[1;4–6]. In addition, the majority of African Americans (AAs) who inherit two *APOL1* nephropathy variants will not develop ESKD [1;7;8].

These findings suggest that additional factors beyond the G1 and G2 variants modulate risk for development of focal segmental glomerulosclerosis (FSGS), HIV-associated nephropathy (HIVAN), focal global glomerulosclerosis with interstitial and vascular changes (FGGS or hypertension-attributed nephropathy), sickle cell nephropathy, and severe lupus nephritis, diseases comprising the *APOL1*-associated disease spectrum[1;3;4;8–12]. Variation in neighboring genes may contribute to risk, based upon the association of markers in the adjacent non-muscle myosin heavy chain 9 gene (*MYH9*) with ESKD in European and Asian ancestry populations. These ancestral groups generally lack *APOL1* G1 and G2 nephropathy variants[13–17].

In an effort to detect novel risk and protective variants on chromosome 22q13 that impact the risk for nephropathy, Next Generation DNA sequencing (NGS) was performed on a ~275 kb region encompassing the *APOL1-APOL4* and *MYH9* genes using capture enrichment. The primary goal was to identify genetic variations that may contribute to ESKD risk in cases who were heterozygous for *APOL1* G1 or G2 alleles, but which may be absent in subjects who were heterozygous for the *APOL1* G1 or G2 alleles and lacked nephropathy.

Materials and Methods

Study Populations

One hundred and ninety-two self-described subjects of African descent born in North Carolina, South Carolina, Virginia, Georgia, or Tennessee formed the Next Generation Sequencing (NGS) discovery panel (Table 1). Cases had ESKD attributed to FSGS, FGGS, HIVAN, chronic glomerulonephritis (without a kidney biopsy), hypertension, or unknown causes and were recruited from dialysis clinics. Patients with ESKD ascribed to diabetes mellitus, urologic disease, surgical nephrectomy, IgA nephropathy, polycystic kidney disease, membranous glomerulonephritis, membranoproliferative glomerulonephritis and other etiologies of ESKD were excluded. Controls had a serum creatinine concentration <1.5 mg/dl (men) or <1.3 mg/dl (women), with a urine albumin:creatinine ratio <30 mg/g. All 192 subjects (154 cases and 38 controls) in this discovery panel were heterozygous for either the G1 or G2 *APOL1* risk allele with equal proportions of each risk allele in cases and controls.

The complete Wake Forest study population included in the full genotyping set consisted of a total of 1,571 subjects of African descent having non-diabetic ESKD and 1,342 controls without kidney disease (Table 1). Because the 192 participants included in the sequencing discovery panel were a subset derived from the complete genotyping set and constituted a significant number of heterozygotes, these samples were included in the complete analysis in order to limit possible type 1 and type 2 errors that could have occurred due to missing heterozygotes. The complete analysis set of cases met identical disease criteria as NGS study cases. All controls in the complete analysis set denied knowledge of kidney disease and met the serum creatinine concentration cut-offs employed in the NGS analysis; albeit, some subjects lacked serum creatinine concentrations and the majority lacked measures of albuminuria. Therefore, it is possible that a small number of controls had mild or undiagnosed nephropathy, an effect that would reduce power for confirming genetic associations. The Family Investigation of Nephropathy and Diabetes (FIND) [18,19] served as a second replication population, consisting of 807 non-diabetic ESKD cases and 634 non-nephropathy controls with complete phenotypic data (Table 1).

Next Generation Sequencing

For re-sequencing, a 275 kb region of Chr22:36529953-36804892 (hg19) was designated as the capture region. The custom enrichment probes were designed and synthesized using TargetSeq™ system (Life Technologies; Grand Island, New York). Repetitive and low complexity sequences were removed from the enrichment design using RepeatMasker (www.repeatmasker.org). Genomic DNA was sheared to an average of 180 bp using a Covaris S220 focused ultrasonicator and DNA libraries constructed using the SOLiD Fragment Library kit on an automated SOLiD Library Builder (Life Technologies; New York). Samples were barcoded to allow multiplex sequencing. The fragment libraries were enriched and amplified according to the TargetSeq protocol and then sequenced on an ABI 5500XL at 50–75 bp read lengths.

Bioinformatic and Statistical Analysis

Data generated on the ABI SOLiD (.xseq) were analyzed by first pass using LifeScope (Grand Island, New York) Genomic Analysis Software (v 2.5) to generate BAM files against the reference human genome (GRCh37/hg19). A BED file generated for the capture enrichment design was used to filter fragments aligning to the region Chr22:36529953-36804892. BAM files were used for secondary alignment and variant calling. Variants were called using GATK 'best practices'. Multiple-sample variant calling was done using GATK UnifiedGenotyper (v2.6) after de-duplication (PicardTools) and realignment around indels (GATK IndelRealigner).

A logistic regression model was computed to test for association between an individual single nucleotide polymorphism (SNP) and ESKD. In the NGS discovery cohort, the logistic model was adjusted for age and gender as implemented in SNPGWA (www.phs.wakehealth.edu). By design, the discovery cohort had one copy of either the *APOL1* G1 or G2 risk variant. In the validation analysis, comprised of all discovery and additional replication samples, the logistic regression model was computed adjusting for age, gender, global ancestry, and *APOL1* G1/G2 risk.

Genotyping

Two SNPs in *APOL1*, the G1 nephropathy risk allele (rs73885319; rs60910145) and an insertion/deletion polymorphism for the G2 risk allele (rs71785313), were genotyped in subjects using a custom assay designed in the Center for Genomics and Personalized Medicine Research at Wake Forest School of Medicine on the Sequenom platform (San Diego, California) [20]. Of the 270 blind duplicates that were genotyped, only 3 were found to be discordant for a 99% concordance rate. The *APOL1* G1 and G2 genotype calls were visually inspected for quality control.

SNPs identified as being associated with ESKD (Supplementary Table 1) in the first pass association analysis in the 192 subject NGS discovery panel and one SNP in *APOL3* (rs11089781) encoding the *APOL3* null allele (Gln⁵⁸Ter) were genotyped using the Sequenom platform, as described above. Ninety blind duplicates were genotyped and 2 were discordant (98% concordance rate). Admixture estimates were computed based on the frequencies of 106 bi-allelic ancestry informative markers distributed throughout the genome. The maximum likelihood approach as coded in the package Frequentist estimation of individual ancestry proportion (FRAPPE) was used to obtain the proportion of African and European ancestry in each individual [21]. Genotype data at these markers were obtained from 44 HapMap Yoruba individuals (YRI) and 39 European American controls as anchors and provided starting values for the Expectation-Maximization algorithm in FRAPPE.

Results

Deep Sequencing Results

Table 1 contains demographic characteristics for the NGS (discovery) and the full sample of Wake Forest ESKD cases and non-nephropathy controls that were used for replication. In

the NGS discovery study, our goal was to identify new genetic variants that may contribute to ESKD risk in subjects with ESKD that had inherited only a single copy of either the G1 or G2 *APOL1* risk variant (recessive inheritance pattern). A matched set of subjects without nephropathy and a single copy of either the G1 or G2 *APOL1* risk variants served as the comparison group. For subsequent validation by genotyping the full sample of ESKD cases and non-nephropathy controls, a mixture of genotypes representative of these populations was included. Cases with ESKD generally initiated renal replacement therapy in the late fourth or fifth decade of life, and were compared to control subjects of similar age. There were disproportionately greater numbers of females in the NGS control sample relative to NGS cases and the full case/control sample. Body mass index (BMI) was also higher in the non-nephropathy controls, likely reflecting poorer nutrition and inflammation observed with advanced nephropathy.

The mean sequencing depth across the 275 kb region was >100X for the 192 subjects in the NGS discovery panel. The median fraction of mapped reads was 94%. SNP calling using GATK resulted in 3300 SNPs of which 69.8 % were found in dbSNP. Fifty-five coding changes were identified across all five genes, none of which were novel. The transition/transversion ratio was 2.04. One hundred twenty SNPs previously genotyped were detected with Cohen's kappa 0.83, representing high agreement. Two hundred and forty six insertion-deletions (InDels) were also identified of which 60% were found in dbSNP. The *APOL1* G1 risk alleles (rs73885319; rs60910145) and G2 risk allele insertion/deletion (rs71785313) were used as internal controls to validate re-sequencing results in the NGS discovery panel. *APOL1* G1 and G2 genotypes were 100% concordant with previous genotyping in all 192 individuals comprising the NGS discovery sample.

Association Analysis

After adjusting for age and gender and compensating for linkage disequilibrium (LD), 10 variants from the top 25 hits in the dominant and additive models were chosen for genotyping in the complete Wake Forest (validation) sample (Supplementary Table S1). Of these 10 variants, two were found to be monomorphic in the validation set; both of these SNPs had been identified as novel polymorphisms in the deep sequencing analysis. After adjusting for the effects of the *APOL1* G1 and G2 risk alleles and without multiple comparison adjustment, only the *APOL3* SNP rs132672 approached statistical significance ($p=0.053$) in the complete validation set under a recessive model.

A second analysis focused on coding variants that had been identified in the *APOL1-APOL4-MYH9* capture region. A summary of coding variants identified in each gene in the discovery panel cases and controls is shown (Table 2). No novel coding variations were identified in the *APOL1-APOL4-MYH9* genes. Based on allele frequency, a single *APOL3* SNP (rs11089781) encoding the *APOL3* nonsense allele (Gln⁵⁸Ter) was associated with ESKD in the NGS discovery panel ($p=0.02$) with 30% of cases and 15% of controls having the null allele. This SNP was genotyped and found to be associated with ESKD in the full case-control validation sample after adjustment for *APOL1* G1 and G2 compound risk (Table 3, recessive model, odds ratio 1.81; $p=0.026$). The recessive mode of inheritance was consistent with expectations for a null allele in a highly inducible gene. No other coding

variants had strong evidence for association with ESKD. However, further genotyping of rs11089781 in 807 FIND non-diabetic ESKD cases and 634 FIND non-nephropathy controls did not support association (data not shown).

Discussion

This study is the first to describe NGS based re-sequencing of the ~275 kb region encompassing the *APOL1-4-MYH9* region. Based on our strategy of sequencing affected and unaffected subjects heterozygous for either the *APOL1* G1 or G2 risk allele, our data suggest that no additional common genetic variations in this ~275 kb contribute significantly to ESKD risk beyond the *APOL1* G1 and G2 alleles. Although we did identify the common null variant rs11089781 (Gln>Ter) as having a potential recessive effect on ESKD risk within our primary screening population, the inability to reproduce this association in the independent FIND cohort suggests that the initially observed association could be due to a false positive result, a cohort-specific observation, or different diagnoses in FIND participants [18,19]. Given the frequency and the potential functionality of this null variant, analysis in other African ancestry populations with nephropathy is warranted, as are tests to determine how this *APOL3* null variation functions in ESKD.

The data in this study are consistent with a recent report showing that re-sequencing of all of the *APOL1* exons in 1437 subject of African and European ancestry did not identify any new common or rare coding variants in *APOL1* that contribute to additional ESKD risk [22]. Of the 23 coding variants identified in that study, we were able to identify 15 of the same variants. Of the remaining 8 coding variants, our study was not powered (N=192) to detect these rare variants (frequency <1% in all cases) as has been the case in other re-sequencing studies [23]. We did, however, identify one additional rare variant (rs143845266) not identified in the *APOL1* re-sequencing study. Because the *APOL1* re-sequencing study focused only on exons, we are unable to compare our NGS sequencing data obtained in the non-coding regions (introns and promoter region) and we could not compare our InDel sequencing data since no comprehensive InDel sequencing data was reported in the *APOL1* focused re-sequencing study. However, combining our analysis with this *APOL1* focused study, the cumulative data support that additional common coding variations in *APOL1* are highly unlikely to contribute significantly to additional nephropathy risk.

The present analysis utilized a capture bait strategy to enrich for target regions and sequencing read lengths of 50–75 bp, thus our data were not ideal for properly evaluating copy number variations or large structural variations in the *APOL1-4-MYH9* region. There is evidence that the *APOL1* region could be a hotspot for genetic recombination based on the mapping of crossover recruiting motifs within surrounding the *APOL1* gene [24]. Additionally, there are low divergent sequence regions within *APOL4* and *APOL1*, which could be evidence of a recombination event or a large sequence inversion involving these two genes. A recent analysis of 123 subjects with kidney disease and 255 control subjects found that in the subjects with nephropathy, there was a ~4% *APOL1* duplication rate compared to a *APOL1* duplication rate of only ~0.8% in the controls [25]. Thus it is possible that in a small subset of subjects heterozygous for the *APOL1* G1 or G2 allele, individuals could carry two copies of a risk allele. Whether the second copy is expressed or functional

isn't known, but it does suggest that in some subjects a second copy of an *APOL1* G1 or G2 allele could set up an imbalance in gene expression where the dose of the mutant copies of *APOL1* protein could be higher in subjects with an *APOL1* CNV.

Given that additional common genetic variants in the *APOL1-4-MYH9* region that contribute significantly to ESKD risk apparently do not exist, the possibility remains that altered gene expression or epigenetics could also be in play. In the case of *APOL1*, at least 4 splice variants are identified in genomic databases (UCSC Genome Browser: genome.ucsc.edu). One report identified that differential splicing out of exon 4 regulates the toxicity of the G0 (non-G1 or non-G2) *APOL1* allele [26]. Based on a review of *APOL1* exon 4, only 4 extremely rare coding variants have been identified in this exon. Neither this study nor our detailed *APOL1* re-sequencing study identified these 4 rare variants, suggesting that these rare variants are most likely novel germline variants, and that this exonic region is under conservation pressure, critical for proper function of *APOL1*.

In summary, the *APOL1* G1 and G2 renal risk variants are important determinants of ESKD susceptibility. Based on the present analyses and other reports, these two variants appear to be responsible for the bulk of *APOL1* derived risk for progressive non-diabetic nephropathy. While additional coding changes occur in the genes *APOL1*, *APOL2*, *APOL3*, and *APOL4*, none of the variants identified in this study significantly contributed to nephropathy risk. Although this doesn't exclude the possibility that these coding variants may affect gene function, and thus be minor contributors to disease risk, the data support other uncharacterized genetic or epigenetic factors as likely responsible for additional ESKD risk residing in the chromosome 22q13 region.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by NIH grants RO1 DK071891, DK084149, DK070657-01, and O1-DK-57304-01. Computational resources provided, in part, by the Wake Forest Health Science Center for Public Health Genomics

References

1. Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzuere P, Freedman BI, Bowden DW, Langefeld CD, Oleksyk TK, Uscinski Knob AL, Bernhardt 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*. 2010; 329:841–845. [PubMed: 20647424]
2. Palmer ND, Freedman BI. APOL1 and progression of nondiabetic nephropathy. *J Am Soc Nephrol*. 2013; 24:1344–1346. [PubMed: 23813212]
3. 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*. 2010; 128:345–350. [PubMed: 20635188]
4. Freedman BI, Langefeld CD, Andringa KK, Croker JA, Williams AH, Garner NE, Birmingham DJ, Hebert LA, Hicks PJ, Segal MS, Edberg JC, Brown EE, Alarcon GS, Costenbader KH, Comeau

ME, Criswell LA, Harley JB, James JA, Kamen DL, Lim SS, Merrill JT, Sivits KL, Niewold TB, Patel NM, Petri M, Ramsey-Goldman R, Reville JD, Salmon JE, Tsao BP, Gibson KL, Byers JR, Vinnikova AK, Lea JP, Julian BA, Kimberly RP. End-stage renal disease in African Americans with lupus nephritis is associated with APOL1. *Arthritis Rheumatol*. 2014; 66:390–396. [PubMed: 24504811]

5. Kanji Z, Powe CE, Wenger JB, Huang C, Ankers E, Sullivan DA, Collerone G, Powe NR, Tonelli M, Bhan I, Bernhardt 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*. 2011; 22:2091–2097. [PubMed: 21997398]
6. 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*. 2012; 27:1498–1505. [PubMed: 22357707]
7. Freedman BI, Skorecki K. Gene-gene and gene-environment interactions in apolipoprotein L1 gene-associated nephropathy. *Clin J Am Soc Nephrol*. 2014; 9:2006–2013. [PubMed: 24903390]
8. 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*. 2011; 22:2129–2137. [PubMed: 21997394]
9. Ashley-Koch AE, Okocha EC, Garrett ME, Soldano K, De Castro LM, Jonassaint JC, Orringer EP, Eckman JR, Telen MJ. MYH9 and APOL1 are both associated with sickle cell disease nephropathy. *Br J Haematol*. 2011; 155:386–394. [PubMed: 21910715]
10. Larsen CP, Beggs ML, Saeed M, Walker PD. Apolipoprotein L1 risk variants associate with systemic lupus erythematosus-associated collapsing glomerulopathy. *J Am Soc Nephrol*. 2013; 24:722–725. [PubMed: 23520206]
11. Lipkowitz MS, Freedman BI, Langefeld CD, Comeau ME, Bowden DW, Kao WH, Astor BC, Bottinger EP, Iyengar SK, Klotman PE, Freedman RG, Zhang W, Parekh RS, Choi MJ, Nelson GW, Winkler CA, Kopp JB. Apolipoprotein L1 gene variants associate with hypertension-attributed nephropathy and the rate of kidney function decline in African Americans. *Kidney Int*. 2013; 83:114–120. [PubMed: 22832513]
12. 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. APOL1 risk variants, race, and progression of chronic kidney disease. *N Engl J Med*. 2013; 369:2183–2196. [PubMed: 24206458]
13. Cheng W, Zhou X, Zhu L, Shi S, Lv J, Liu L, Zhang H. Polymorphisms in the nonmuscle myosin heavy chain 9 gene (MYH9) are associated with the progression of IgA nephropathy in Chinese. *Nephrol Dial Transplant*. 2011; 26:2544–2549. [PubMed: 21245129]
14. Cooke JN, Bostrom MA, Hicks PJ, Ng MC, Hellwege JN, Comeau ME, Divers J, Langefeld CD, Freedman BI, Bowden DW. Polymorphisms in MYH9 are associated with diabetic nephropathy in European Americans. *Nephrol Dial Transplant*. 2012; 27:1505–1511. [PubMed: 21968013]
15. Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, Oleksyk T, McKenzie LM, Kajiyama H, Ahuja TS, Berns JS, Briggs W, Cho ME, Dart RA, Kimmel PL, Korbet SM, Michel DM, Mokrzycki MH, Schelling JR, Simon E, Trachtman H, Vlahov D, Winkler CA. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet*. 2008; 40:1175–1184. [PubMed: 18794856]
16. O'Seaghdha CM, Parekh RS, Hwang SJ, Li M, Kottgen A, Coresh J, Yang Q, Fox CS, Kao WH. The MYH9/APOL1 region and chronic kidney disease in European-Americans. *Hum Mol Genet*. 2011; 20:2450–2456. [PubMed: 21429915]
17. Pattaro C, Kottgen A, Teumer A, Garnaas M, Boger CA, Fuchsberger C, Olden M, Chen MH, Tin A, Taliun D, Li M, Gao X, Gorski M, Yang Q, Hundertmark C, Foster MC, O'Seaghdha CM, Glazer N, Isaacs A, Liu CT, Smith AV, O'Connell JR, Struchalin M, Tanaka T, Li G, Johnson AD, Gierman HJ, Feitosa M, Hwang SJ, Atkinson EJ, Lohman K, Cornelis MC, Johansson A, Tonjes A, Dehghan A, Chouraki V, Holliday EG, Sorice R, Kutalik Z, Lehtimäki T, Esko T, Deshmukh

H, Ulivi S, Chu AY, Murgia F, Trompet S, Imboden M, Kollerits B, Pistis G, Harris TB, Launer LJ, Aspelund T, Eiriksdottir G, Mitchell BD, Boerwinkle E, Schmidt H, Cavalieri M, Rao M, Hu FB, Demirkan A, Oostra BA, de AM, Turner ST, Ding J, Andrews JS, Freedman BI, Koenig W, Illig T, Doring A, Wichmann HE, Kolcic I, Zemunik T, Boban M, Minelli C, Wheeler HE, Igl W, Zaboli G, Wild SH, Wright AF, Campbell H, Ellinghaus D, Nothlings U, Jacobs G, Biffar R, Endlich K, Ernst F, Homuth G, Kroemer HK, Nauck M, Stracke S, Volker U, Volzke H, Kovacs P, Stumvoll M, Magi R, Hofman A, Uitterlinden AG, Rivadeneira F, Aulchenko YS, Polasek O, Hastie N, Vitart V, Helmer C, Wang JJ, Ruggiero D, Bergmann S, Kahonen M, Viikari J, Nikopensius T, Province M, Ketkar S, Colhoun H, Doney A, Robino A, Giulianini F, Kramer BK, Portas L, Ford I, Buckley BM, Adam M, Thun GA, Paulweber B, Haun M, Sala C, Metzger M, Mitchell P, Ciullo M, Kim SK, Vollenweider P, Raitakari O, Metspalu A, Palmer C, Gasparini P, Pirastu M, Jukema JW, Probst-Hensch NM, Kronenberg F, Toniolo D, Gudnason V, Shuldiner AR, Coresh J, Schmidt R, Ferrucci L, Siscovick DS, van Duijn CM, Borecki I, Kardia SL, Liu Y, Curhan GC, Rudan I, Gyllenstein U, Wilson JF, Franke A, Pramstaller PP, Rettig R, Prokopenko I, Witteman JC, Hayward C, Ridker P, Parsa A, Bochud M, Heid IM, Goessling W, Chasman DI, Kao WH, Fox CS. Genome-wide association and functional follow-up reveals new loci for kidney function. *PLoS Genet.* 2012; 8:e1002584. [PubMed: 22479191]

18. Iyengar SK, Abboud HE, Goddard KA, Saad MF, Adler SG, Arar NH, Bowden DW, Duggirala R, Elston RC, Hanson RL, Ipp E, Kao WH, Kimmel PL, Klag MJ, Knowler WC, Meoni LA, Nelson RG, Nicholas SB, Pahl MV, Parekh RS, Quade SR, Rich SS, Rotter JI, Scavini M, Schelling JR, Sedor JR, Sehgal AR, Shah VO, Smith MW, Taylor KD, Winkler CA, Zager PG, Freedman BI. Genome-wide scans for diabetic nephropathy and albuminuria in multiethnic populations: the family investigation of nephropathy and diabetes (FIND). *Diabetes.* 2007; 56:1577–1585. [PubMed: 17363742]
19. Knowler WC, Coresh J, Elston RC, Freedman BI, Iyengar SK, Kimmel PL, Olson JM, Laetke R, Sedor JR, Seldin MF. The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. *J Diabetes Complications.* 2005; 19:1–9. [PubMed: 15642484]
20. Freedman BI, Langefeld CD, Turner J, Nunez M, High KP, Spainhour M, Hicks PJ, Bowden DW, Reeves-Daniel AM, Murea M, Rocco MV, Divers J. Association of APOL1 variants with mild kidney disease in the first-degree relatives of African American patients with non-diabetic end-stage renal disease. *Kidney Int.* 2012; 82:805–811. [PubMed: 22695330]
21. Tang H, Peng J, Wang P, Risch NJ. Estimation of individual admixture: analytical and study design considerations. *Genet Epidemiol.* 2005; 28:289–301. [PubMed: 15712363]
22. Limou S, Nelson GW, Lecordier L, An P, O'hUigin CS, David VA, Binns-Roemer EA, Guiblet WM, Oleksyk TK, Pays E, Kopp JB, Winkler CA. Sequencing rare and common APOL1 coding variants to determine kidney disease risk. *Kidney Int.* 2015
23. Flannick J, Thorleifsson G, Beer NL, Jacobs SB, Grarup N, Burt NP, Mahajan A, Fuchsberger C, Atzmon G, Benediktsson R, Blangero J, Bowden DW, Brandslund I, Brosnan J, Burslem F, Chambers J, Cho YS, Christensen C, Douglas DA, Duggirala R, Dymek Z, Farjoun Y, Fennell T, Fontanillas P, Forsén T, Gabriel S, Glaser B, Gudbjartsson DF, Hanis C, Hansen T, Hreidarsson AB, Hveem K, Ingelsson E, Isomaa B, Johansson S, Jørgensen T, Jørgensen ME, Kathiresan S, Kong A, Kooner J, Kravic J, Laakso M, Lee JY, Lind L, Lindgren CM, Linneberg A, Masson G, Meitinger T, Mohlke KL, Molven A, Morris AP, Potluri S, Rauramaa R, Ribel-Madsen R, Richard AM, Rolph T, Salomaa V, Segrè AV, Skärstrand H, Steinthorsdottir V, Stringham HM, Sulem P, Tai ES, Teo YY, Teslovich T, Thorsteinsdottir U, Trimmer JK, Tuomi T, Tuomilehto J, Vaziri-Sani F, Voight BF, Wilson JG, Boehnke M, McCarthy MI, Njølstad PR, Pedersen O, Groop L, Cox DR, Stefansson K, Altshuler D. Go-T2D Consortium; T2D-GENES Consortium. *Nat Genet.* 2014; 46:357–63. [PubMed: 24584071]
24. Genovese G, Friedman DJ, Pollak MR. APOL1 variants and kidney disease in people of recent African ancestry. *Nat Rev Nephrol.* 2013; 9:240–244. [PubMed: 23438974]
25. Ruchi R, Genovese G, Lee J, Charoonratana VT, Bernhardt AJ, Alper SL, Kopp JB, Thadhani R, Friedman DJ, Pollak MR. Copy Number Variation at the APOL1 Locus. *PLoS One.* 2015; 10:e0125410. [PubMed: 25933006]
26. Khatua AK, Cheatham AM, Kruzel ED, Singhal PC, Skorecki K, Popik W. Exon 4-encoded sequence is a major determinant of cytotoxicity of apolipoprotein L1. *Am J Physiol Cell Physiol.* 2015; 309:C22–C37. [PubMed: 25924622]

Table 1

	Wake Forest Discovery Samples			Wake Forest Complete Sample Set *			FIND Replication		
	ESKD Cases (N=154)	Controls (N=38)	Total (N=192)	ESKD Cases (N=1427)	Controls (N=1304)	Total (N=2731)	ESKD Cases (N=394)	Controls (N=660)	Total (N=1054)
Female (%)	46.8	89.5	55.2	43.7	51.7	47.5	33.0	61.8	51.0
Age (years)	55.9±13.7	45.9±12.1	54.0±14.0	53.7±14.3	49.2±12.2	51.6±13.5	47.5±11.7	46.6±11.8	47.0±11.8
Duration ESKD (years)	4.7±5.1	-	-	5.6±5.7	-	-	-	-	-
BMI (kg/m ²)	27.0±6.3	32.5±8.7	28.2±7.2	27.5±7.1	29.6±7.2	28.5±7.3	-	-	-
BUN (mg/dL)	-	10.7±3.4	-	-	13.7±5.2	-	-	-	-
Serum creatinine (mg/dL)	-	0.8±0.1	-	-	1.0±0.4	-	-	-	-

* includes the 202 samples from the Discovery set

Table 2

	Chr 22 position	rs ID	Reference allele	Amino Acid Change	Cases MAF	Controls MAF	
APOL1	36657740	rs41297245	G	Gly ¹¹² Arg	0.02	0.025	
	36661330	rs2239785	G	Glu ¹⁶⁶ Lys	0.27	0.32	
	36661409	rs116136671	A	Asn ¹⁹² Ser	0.044	0.01	
	36661536	rs136174	C	Ala ²³⁴ Ala	0.96	0.96	
	36661566	rs136175	G	Met ²⁴⁴ Ile	0.96	0.96	
	36661646	rs136176	G	Arg ²⁷¹ Lys	0.96	0.96	
	36661674	rs73885316	C	Asn ²⁸⁰ Lys	0.03	0.01	
	36661691	rs73403889	G	Gly ²⁸⁶ Asp	0.003	-	
	36661791	rs141788376	C	Arg ³¹⁹ Arg	0.003	-	
	36661796	rs150588135	G	Arg ³²¹ Gln	0.003	-	
	36661810	rs139590386	A	Ile ³²⁶ Val	0.003	-	
	36661842	rs136177	G	Arg ³⁶⁶ Arg	0.05	0.04	
APOL2	36661891	rs16996616	G	Asp ³⁵³ Asn	0.08	0.04	
	36661906	rs73885319	A	Ser ³⁵⁸ Gly	0.25	0.28	
	36661940	rs143845266	T	Val ³⁶⁹ Glu	-	0.01	
	36662034	rs60910145	T	Ile ⁴⁰⁹ Met	0.25	0.28	
	36623671	rs73885303	G	Ala ²⁶⁵ Thr	0.006	0.01	
	36623902	rs148745051	G	Gly ¹⁸⁸ Ser	0.003	-	
	36623920	rs7285167	C	Arg ¹⁸² Cys	0.63	0.61	
	36623980	rs15125817	C	Ala ¹⁶² Val	0.006	-	
	36624231	rs118097350	C	Arg ⁷⁸¹ Lys	0.01	-	
	APOL3	36537431	rs141045937	C	Ala ³⁴² Ala	0.003	-
		36537500	rs132618	A	Gly ³¹⁹ Gly	0.13	0.16
		36537650	rs147303125	A	Val ²⁶⁹ Val	0.003	-
36537725		rs61731692	G	Tyr ²⁴⁴ Tyr	0.08	0.06	
36537740		rs138851629	G	Ile ²³⁹ Ile	0.003	-	

Chr 22 position	rs ID	Reference allele	Amino Acid Change	Cases MAF	Controls MAF
36537763	rs61741884	C	Val ²³² Met	0.006	-
36537798	rs116147257	G	Thr ²²⁰ Ile	0.03	0.10
36537893	rs3827346	A	Asn ¹⁸⁸ Asn	0.02	0.04
36538053	rs6000152	G	Ala ¹³⁵ Val	0.075	0.14
36541536	rs141517147	G	Ala ¹¹² Val	0.003	-
36556729	rs79419411	T	Arg ⁷¹ Gly	0.11	0.15
36556768	rs11089781	G	Gln⁵⁸Ter	0.30	0.15
36556823	rs132653	G	Ser ³⁹ Arg	0.45	0.48
36556882	rs148712729	T	Cys ¹³ Trp	0.003	0.01
APOL4					
36587154	rs111781032	C	Thr ³³⁹ Asn	0.1	0.06
36587202	rs6000172	C	Ser ³²³ Leu	0.74	0.65
36587215	rs61730820	G	Glu ³¹⁹ Lys	0.013	-
36587223	rs6000173	C	Ala ³¹⁶ Glu	0.74	0.65
36587279	rs6000174	T	Leu ²⁹⁷ Leu	0.74	0.65
36587310	rs76473883	G	Arg ²⁸⁷ Gln	0.013	0.04
36587311	rs150226555	C	Arg ²⁸⁷ Trp	-	0.025
36587346	rs61730819	G	Arg ²⁷⁵ His	0.1	0.06
36587486	rs2227167	T	Ser ²²⁸ Ser	0.74	0.65
36587511	rs2227168	G	Arg ²²⁰ His	0.75	0.68
36587682	rs146308862	G	Gly ¹⁶³ Glu	-	0.01
36587704	rs132700	A	Met ¹⁵⁶ Val	0.32	0.44
36587886	rs78582347	G	Met ⁹⁴ Val	0.1	0.06
36587952	rs2227169	G	Glu ⁷² Glu	0.73	0.64
36591380	rs2007468	T	Phe ⁶² Phe	0.87	0.78
36591402	rs140324556	G	Ser ⁵⁵ Thr	0.003	-
36591475	rs77639244	C	Arg ³¹ Cys	0.1	0.06
36598049	rs80587	G	Val ¹² Leu	0.65	0.56
36598058	rs132736	A	Ile ⁹ Val	0.64	0.50
36598081	rs6000181	T	Met ¹ Leu	0.05	0.01

Table 3

	N (case/controls)	MAF(case/control)	Dominant model		Additive model		Recessive model	
			p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95% CI)
All Subjects	1559/1332	0.30/0.22	0.31	1.09(0.92, 1.29)	0.18	1.1(0.96,1.25)	0.2	1.25(0.89,1.76)
Subjects with 0 or 1 <i>APOLI</i> G1 or G2 allele	849/1139	0.21/0.19	0.38	1.09(0.90,1.32)	0.13	1.14(0.96,1.35)	0.022	1.83(1.09,3.08)
Subjects with 0 <i>APOLI</i> G1 or G2 allele	374/562	0.11/0.12	0.24	0.81(0.58,1.15)	0.40	0.87(0.64,1.20)	0.41	1.6(0.52,4.90)
Subjects with 1 <i>APOLI</i> G1 or G2 allele	463/574	0.28/0.26	0.38	1.12(0.87,1.45)	0.15	1.17(0.94,1.46)	0.064	1.75(0.97,3.18)
Subjects with 2 <i>APOLI</i> G1 or G2 allele	722/196	0.40/0.40	0.64	1.08(0.78,1.50)	0.79	1.03(0.82,1.30)	0.92	0.98(0.64,1.50)
Subjects with 1 G1 allele	166/219	0.26/0.21	0.09	1.46(0.95,2.25)	0.084	1.41(0.96,2.08)	0.49	1.55(0.44,5.46)
Subjects with 1 G2 allele	297/355	0.30/0.28	0.75	0.95(0.70,1.31)	0.69	1.06(0.81,1.37)	0.10	1.76(0.90,3.46)

* P-values not adjusted for multiple comparisons

OR, odds ratio