Genome-Wide Scans for Diabetic Nephropathy and Albuminuria in Multi-Ethnic Populations: The Family Investigation of Nephropathy and Diabetes

Received for publication 16 August 2006 and accepted in revised form 20 February 2007.


§ Members’ affiliations are listed in the appendix.

*FIND-Genetic Analysis and Data Coordinating Center; Department of Epidemiology & Biostatistics, Case Western Reserve University, Cleveland, Ohio.

Short Title: Diabetic Nephropathy Genome Scan

Corresponding Author:
Dr. Sudha Iyengar
Department of Epidemiology and Biostatistics
Case Western Reserve University
Wolstein Research Building, Rm #1315
10900 Euclid Avenue
Cleveland, OH 44106-7281
E-mail: ski@case.edu

Additional information can be seen in an online-only appendix found at http://diabetes.diabetesjournals.org
Abstract

The Family Investigation of Nephropathy and Diabetes (FIND) was initiated to map genes underlying susceptibility to diabetic nephropathy (DN). Eleven centers participated under a single collection protocol to recruit large numbers of diabetic sib pairs concordant and discordant for DN. We report the findings from the first phase genetic analyses in 1,227 participants from 378 pedigrees of European American, African American, Mexican American, and American Indian descent recruited from eight centers. Model-free linkage analyses, using a dichotomous definition for DN in 397 sib pairs, as well as the quantitative trait urine albumin:creatinine ratio (ACR), were performed using the Haseman-Elston linkage test on 404 microsatellite markers. The strongest evidence of linkage to the trait DN was on chromosomes 7q21.3, 10p15.3, 14q23.1 and 18q22.3. In ACR (883 diabetic sib pairs), the strongest linkage signals were on chromosomes 2q14.1, 7q21.1 and 15q26.3. These results confirm regions of linkage to DN on chromosomes 7q, 10p and 18q from prior reports, making it important that genes underlying these peaks be evaluated for their contribution to nephropathy susceptibility. Large family collections consisting of multiple members with diabetes and advanced nephropathy are likely to accelerate the identification of genes causing DN, a life-threatening complication of diabetes mellitus.
Diabetic nephropathy (DN; OMIM #603933, http://www.ncbi.nlm.nih.gov/omim) is a common microvascular complication of type 1 and type 2 diabetes mellitus (DM). Increasing prevalence of DN and end-stage renal disease (ESRD) attributed to diabetes have been observed, particularly in older adults (1-3). In 1999-2003, the estimated prevalence of type 2 DM (T2DM) among all ESRD patients in the U.S. was approximately 29.1%, but among new ESRD cases during the same period was 40.5% (1). Although progression of DN to ESRD and/or development of a cardiovascular complication is common among patients with advanced chronic kidney disease (CKD), aggressive treatment of hyperglycemia, proteinuria and hypertension may slow its progression (4-6).

Major genes underlying susceptibility to DN have yet to be identified, despite multiple candidate gene and genome scan investigations (7;8). Six genome screens have examined linkage for ESRD or nephropathy in small to moderate sized cohorts containing sibling pairs and nuclear families that were recruited at single centers (9-14). Three of these studies were conducted at a single site (11-13). Although these studies did not provide definitive evidence for linkage that met genome-wide criteria, several putative regions of linkage were reported.

Imperatore et al (9) performed a genome-wide survey in 98 diabetic Pima Indian sibling pairs concordant for DN. Suggestive evidence for linkage was observed on chromosomes 3, 7, 9 and 20. A genome scan was performed in 18 large Turkish families that were enriched for the presence of T2DM and DN (10). There was a highly significant LOD score of 6.1 on chromosome 18q22.3-23, between markers D18S469 and D18S58. A genome scan for DN in African Americans identified loci on chromosomes 7p, 12p, 14q, 16p, 18q and 21q using ordered subsets analysis (11). A recent genome scan for albuminuria revealed evidence for linkage to 22q, 5q and 7q in 59 large, predominantly European American pedigrees enriched for members with T2DM (14). Additional evidence for linkage to 21p was observed when these analyses were restricted to DM-affected relative pairs. Consistent evidence for linkage of DN (and non-diabetic nephropathy) has also been observed on chromosomes 10p (15;16), in the syntenic region of the rodent RF1 locus (15-17), and on 3q (11;18).

The Family Investigation of Nephropathy and Diabetes (FIND) consortium was established in 1999 to identify DN susceptibility genes (19). The FIND consisted initially of eight centers, with three centers added to enhance minority recruitment, a Genetic Analysis and Data Coordinating Center (GADCC), and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). The consortium is assembling a database of families, as well as cases and controls, and conducting genome wide scans using microsatellite (MS) loci in this initial set of recruited families from four ethnic groups: European Americans (EA), African Americans (AA), Mexican Americans (MA), and American Indians (AI). FIND has power to detect linkage to DN in and among four ethnic groups, as well as the ability to confirm prior positive linkage findings using independent families recruited from different geographic areas under a uniform protocol and disease definition. We report results of the first FIND MS-based genome scan for the traits DN and urinary albumin-to-creatinine ratio (ACR) in four ethnic groups.

**Research Design and Methods**

**Study Population**

The FIND family study design was described previously (19) and the sample described herein was collected from the eight
original FIND centers during the first half of the recruitment period (2001-2003). In brief, families of probands with DN having a diabetic sib with or without nephropathy were recruited. Living parents and other relatives (i.e. avuncular, cousin, half-sibling, and grandparental affected pairs) were also recruited, when available. FIND is predominantly a sib pair study (>90% have no relatives other than sib pairs). Informed consent was obtained from every subject and approval for recruitment was secured from the institutional review board at each center, including the coordinating center, and a certificate of confidentiality was filed at the National Institutes of Health prior to the start of enrollment.

Phenotypic evaluation
Participants were interviewed regarding prior diagnoses and treatment of kidney disease and diabetes. Medical information was recorded on standardized questionnaires, family history was collected using a standardized instrument and information was obtained from medical record review. Participants without ESRD contributed urine for quantification of ACR and protein:creatinine ratio (PCR), and blood for cell line creation, DNA extraction, hemoglobin A1c (HbA1c), serum creatinine (SC), blood urea nitrogen (BUN) and glucose concentrations; while those with ESRD provided only blood for cell lines, DNA extraction and HbA1c. Blood was obtained in hemodialysis patients prior to initiation of treatment and before heparin administration. Blood and urine samples were shipped to the Laboratory of Genomic Diversity (LGD) for centralized processing. Samples for assay were shipped from LGD to the central clinical laboratory (Penn Medical Laboratories, MEDSTAR). De-identified clinical, demographic and assay information were entered into a centralized database housed at the GADCC. Body mass index (BMI) was calculated using current weight and height recorded on medical questionnaires.

Definitions
Diabetes Participants were considered to have DM if currently or previously treated with insulin and/or oral hypoglycemic medicines. Subjects reporting DM but not treated with these medicines, and those without a history of DM, had HbA1c and fasting plasma glucose concentration measured at study entry. HbA1C concentrations above 6.0% were considered suggestive of DM and fasting plasma glucose and/or oral glucose tolerance testing was then performed to confirm the diagnosis. American Diabetes Association 1997 criteria (20) were used to define diabetes in participants not previously known to have DM. In previously diagnosed individuals, the date of diagnosis of DM was obtained from the medical history, with confirmatory medical record review, when possible. Subjects with either type 1, type 2, or other types of DM were eligible.

Nephropathy Subjects were considered to have overt proteinuria in the presence of a historical 24 hour urine collection with $\geq 500$ mg protein/24h or $\geq 300$ mg albumin/24h, random PCR $> 0.5$ g/g, or random ACR $\geq 0.3$ g/g. ESRD was defined as the need for chronic renal replacement therapy with either dialysis or renal transplantation. Whether an individual was considered affected or unaffected in the analysis depended on the participant category (see Eligibility criteria below).

Diabetic retinopathy The diagnosis required medical record documentation of an ophthalmologic exam demonstrating microaneurysms, proliferative diabetic retinopathy or macular edema. Alternatively, the subject may have had a history of retinal laser surgery (photocoagulation) for diabetic retinopathy.

Eligibility criteria for probands
Probands met above criteria for DM and had DN defined by one of the following:
(1) kidney biopsy revealing DN in the presence of overt proteinuria defined by the following: a) Nodular and/or diffuse increases in mesangial matrix accumulation and b) Thickened glomerular basement membranes and/or arteriolar hyalinization and c) Absence of mesangial immunoglobulin or paraprotein deposits by immunoflorescence microscopy, absence of amyloid deposits by Congo Red staining or electron microscopy, absence of electron dense deposits within the glomerular basement membrane or glomerular capillary subendothelial space.; (2) ESRD attributed to DN based upon either: (a) onset of DM ≥ 5 years prior to renal replacement therapy with diabetic retinopathy; (b) onset of DM ≥ 5 years prior to renal replacement therapy with historic 24 hour urine protein ≥ 3 gm (PCR ≥ 3.0 g/g); or (c) diabetic retinopathy with historic 24 hour urine protein > 3 gm (PCR > 3.0 g/g); (3) CKD (non-ESRD) attributed to DN based upon either diabetic retinopathy with historic 24 hour urine protein ≥ 1 gram (PCR ≥ 1.0 g/g); or 24 hour urine protein excretion > 3 gm (PCR > 3.0 g/g) after diabetes duration > 10 years.

Eligibility criteria for family members

Entry of a proband with DN into the FIND Family protocol required participation of either two living parents (regardless of the presence or absence of diabetes or nephropathy), or at least one full diabetic sibling classified as either DN-affected or lacking DN. Enrollment of a DN sib required one of the following: (1) kidney biopsy consistent with DN (regardless of the degree of proteinuria as defined by the biopsy criteria above); (2) urine albumin excretion ≥ 30 mg/24 hour (ACR ≥ 0.03 g/g) regardless of DM duration; (3) SC concentration ≥ 1.6 mg/dl in men, ≥ 1.4 mg/dl in women, or ESRD.

Diabetic sibs were classified as having DN (forming a DN concordant sib pair) if they had elevated urine albumin excretion (> 300 mg/24 hours or ACR > 0.3 g/g) or ESRD attributed to DN. Diabetic sibs were classified as unaffected by nephropathy (forming a DN discordant sib pair) if they had diabetes duration ≥ 10 years with normal SC concentration (male < 1.6 mg/dl; female < 1.4 mg/dl) and normal urine albumin excretion (< 30 mg/24 hours or ACR < 0.03 g/g) without historical evidence of kidney disease. Diabetic sibs who lacked ESRD or elevated SC concentrations and had spot urine ACR values between 0.03 and 0.3 g/g (in the microalbuminuric range) were included in the ACR quantitative trait genome scan, but not the dichotomous DN trait analysis.

Genotyping and genetic analytic methods

Genotyping and linkage analyses were conducted on 883 full sib pairs with DM from 378 families. DNA was extracted from either lymphoblastoid cell lines or buffy coats at the GADCC and shipped to the Center for Inherited Disease Research (CIDR) for genotyping. CIDR genotyped 404 markers on 22 autosomes and 2 sex chromosomes using a marker set based on the Marshfield Genetics version 8 screening set from Research Genetics, with an average marker spacing of 9 cM. Mendelian inconsistencies were identified using the MARKERINFO program from the Statistical Analysis for Genetic Epidemiology S.A.G.E. (21) software package. All programs are part of the S.A.G.E. suite of programs, unless otherwise specified. Errors in relationship specification were identified through the use of all markers with the program RELTEST. Ninety eight individuals in 54 full sibships were reclassified as unrelated and 65 individuals in 54 full sibships were reclassified as half-sibs. We tested for deviation from Hardy-Weinberg proportions separately for each ethnicity and no significant departure was observed at a 1% significance level. The number of Mendelian inconsistencies blanked due to genotyping error in consistent pedigrees was 260 (0.05%).
We evaluated information in four ethnic groups containing diverse genomes. Maximum likelihood estimation was used, as implemented in the program FREQ, to estimate the marker allele frequencies separately in each ethnic group. Multipoint identity by descent (IBD) allele sharing estimates were computed separately within each of the four ethnic groups using the program GENIBD. For the IBD calculations, 378 pedigrees with 1,227 individuals, 1,337 full sib pairs (of whom 883 were DM concordant), 147 half-sibling pairs, 226 parent-offspring pairs, 97 avuncular pairs and 28 cousin pairs were used.

Linkage analyses were performed in two ways. First, DN was examined as a binary variable (affected versus unaffected, excluding individuals with microalbuminuria), and as quantitative traits (urine ACR and PCR). DN was dichotomized based upon affection status (affected [with both DM and DN], versus unaffected [without DN after DM duration ≥ 10 years]), thus modeling affection status as a function of DM with or without kidney disease. This stringent definition of DN was based on a random urine ACR ≥ 0.3 g/g and the analyses were based on the binary outcome of presence or absence of DN.

Urine ACR and PCR were evaluated in linkage analysis as quantitative traits. Because urine assays are difficult to interpret in individuals with ESRD, we evaluated linkage to ACR and PCR by either excluding or including ESRD/transplant participants in two separate models. The model that included ESRD participants fitted each ESRD individual with an ACR of 3.0 g/g (or total PCR of 3.5 g/g). Individuals with chronic renal failure not yet on dialysis whose measured ACR was > 3.0 g/g or PCR was > 3.5 (N=83) were set to (winsorized) values of 3.0 and 3.5, respectively, to ensure that all values of ACR and PCR were set on a similar scale of measurement. However, the linkage analysis results obtained without winsorizing these values did not materially alter the results (data not shown).

For binary and quantitative traits, the Haseman-Elston (HE) regression linkage test (22), as extended for sibships (23), available in SIBPAL, was performed separately for each ethnicity using the multipoint IBD sharing estimates. SIBPAL performs linear regression based modeling of sib-pair traits as a function of marker allele IBD sharing. Under the latest version of this method, the weighted combination of squared trait difference and squared mean-corrected trait sum was used, further adjusted for the non-independence of sib pairs and the non-independence of squared trait sums and differences (23). For the binary trait, sex was added as a covariate in the HE regression as a 0,1 variable and included as the sum, which assumes that the effect of being a male-female pair is half-way between the effects of the two same-sex pairs. Quantitative traits were adjusted for sex and age at diabetes diagnosis prior to linkage analysis. The residuals were used as the trait values in the HE regression. For regions suggestive of linkage, asymptotic p-values were validated by obtaining empirical p-values in SIBPAL.

For the binary trait, “mean” tests and “proportion” tests were performed using SIBPAL, separately for concordant affected pairs (pairs who had DN and DM), discordant pairs (one member of this pair had DN and DM, while the other had long standing DM but no DN), and concordant unaffected pairs (individuals with long standing DM but no nephropathy). These tests determine if affected pairs share more alleles IBD and discordant pairs share fewer alleles IBD, a pattern expected at a true disease susceptibility locus.

A separate linkage analysis was performed for each ethnicity and we combined p-values across ethnicity using a method proposed by Fisher (24). Fisher’s
method written as $-2\sum_{i=1}^{4}\log_{e}(p_{i})$, where $p_{i}$ is the p-value for the $i^{th}$ ethnicity, compared to a $\chi^{2}$. Fisher’s method was used because it was not desirable to treat all the ethnic groups as a single sample, with the attendant allele frequencies and demographic differences. This method enables us to combine information on all ethnic groups after accounting for group specific differences.

Molecular genetic quality control (QC) analyses were performed by comparing a forensic marker panel for DNA samples collected from buffy coat and cell line sources in a 5% random sample of FIND participants. A total of 2.5% of the 400 FIND participants tested were found to have mis-matched DNA from the buffy coat and cell line sources, equating to an approximate 2.5% overall error rate in the genome scans. Clinical QC was performed with an independent auditing agency site visiting each center.

Results
Eight centers contributed samples for this microsatellite (MS) genome scan. Fourteen percent of families were EA (54 pedigrees), 25% were AA (96 pedigrees), 52% were MA (196 pedigrees), and 9% were AI (32 pedigrees). A one-way analysis of variance (ANOVA) was used to compare ethnic-specific variation in proband characteristics. Differences between the groups included gender and BMI ($p < 0.017$ and 0.014 respectively, on the ANOVA with 3 degrees of freedom). AA had the lowest male:female ratio and highest BMI. The difference in BMI paralleled the nationally reported trends of higher BMI among AAs (25). Remaining clinical characteristics were similar between probands in the four ethnic groups.

We also compared the clinical characteristics of the participants based on presence/absence of kidney disease and their ascertainment status (Table 1). The study included 349 probands, of which 81% were on dialysis or had a renal transplant. The remaining probands had severe proteinuria with reduced glomerular filtration rate (GFR) indicating a high likelihood of rapid progression to renal replacement therapy. Fifty three of the 390 relatives in this report (13.6%) also had ESRD. The age ranges of the probands and their affected relatives were similar; however, the DN affected relatives were younger than the DM without nephropathy relatives. Glycemic control assessed by HbA1C was poor in all three groups, with probands and DM without nephropathy siblings having slightly better control than DN affected relatives. Diabetes duration was longest in the probands (mean 23 ± 8.4 years), differing from other affected relatives who had shorter diabetes duration (mean 16 ± 10 years). As anticipated, the probands and relatives had different biochemical parameters (GFR, SC, BUN, ACR and PCR).

DN was evaluated as a binary trait using the HE regression. The linkage analysis used full sib pairs who were either concordantly affected, discordant, or concordantly unaffected for DN. There were 49 EA, 80 AA, 225 MA, and 43 AI sibling pairs (Table 2). After pooling p-values using Fisher’s method, the most significant linkage signals ($p \leq 0.003$) were detected on chromosomes 7q21.3, 10p15.3, 14q23.1, and 18q22.3. The entire genome scan in all ethnicities is shown in Figure 1. The analogous plots for the two other methods of pooling results for ethnicities are given in supplementary material.

We next evaluated which ethnic group(s) contributed to the linkage signals on chromosomes 7, 10, 14 and 18. AA families had a suggestive linkage peak on chromosome 7 at 106 cM ($p \leq 0.000022$; Figure 2). The small sample of AI families showed a significant linkage peak on chromosome 10 at 0 cM ($p \leq 0.000022$;
Figure 2) and chromosome 14 at 55 cM ($p=7.23 \times 10^{-5}$; Figure 2). Finally, the small sample of EA families showed evidence for linkage on chromosome 18 at 116 cM ($p=2.17 \times 10^{-3}$; Figure 2). Details of the peaks are provided in Table 3.

The mean allele sharing at the linkage peaks are provided in Table 4. For the peaks on chromosomes 7, 10, 14, and 18 there was significantly decreased allele sharing ($p \leq 0.03$) among the discordant sibling pairs for the AA, AI, and EA ethnicities, respectively, indicating that the discordant pairs proved most of the evidence for linkage. However, significant excess sharing was also observed among the ASPs on chromosome 7 in AA ($p = 0.0248$), and chromosome 10 in AI ($p=0.0498$), and suggestive excess sharing in AIs on chromosome 14 ($p = 0.1733$) and EAs on chromosome 18 ($p=0.1959$). The loci that best follow the anticipated patterns of allele sharing are located on chromosomes 7 and 10.

We conducted a quantitative trait linkage analysis of ACR (Table 2). There were 942 individuals making up a total of 883 sibling pairs used in the analysis. Values of urine ACR were imputed for ESRD participants as 3.0 g/g (as described in "Genotyping and genetic analytic methods", above). Median GFR and ACR (IQ range) for the 942 subjects used in the analysis was 79.7 (56.0-103.7) and 0.05 (0.01-0.38), respectively. For urine ACR, suggestive evidence for linkage was observed on chromosomes 2q14.1 in AI families, 7q21.1 in EA families, and 15q26.3 in AA families (Figure 3 and Table 5). The linkage results for PCR confirmed the results for ACR on 2p and 7q, with smaller p-values observed for PCR (not shown).

**Discussion**

This report contains the first phase genome scan results for DN and for the quantitative traits ACR and PCR in the multi-ethnic FIND. FIND families contain large numbers of sibling pairs concordant for DM, and either concordant or discordant for the presence of DN. The current linkage analysis for the discrete trait DN was performed using 397 informative full sib pairs, and 883 sib pairs were used for the quantitative trait urine ACR. Due to the stringent criteria for the dichotomous trait (and thus the smaller sample size), this group has less power to detect linkage than the analysis of the quantitative trait (see supplementary material). This interim report used a MS marker-based genome scan. In contrast, the final FIND genome scan will be performed using a dense map of single nucleotide polymorphisms (SNPs) with a spacing of 0.64 cM and will include all the FIND families recruited through 2005.

Suggestive evidence for linkage to the dichotomous trait of DN was observed on chromosomes 7q21.3, 10p15.3, 14q23.1, and 18q22.3. The linkage peak at 102.0 cM on chromosome 7q was driven predominantly by the AA families, although there was strong evidence for linkage in ethnicity-combined FIND families. DN was also linked to 10p, 14q and 18q in all FIND families; however the AI families made major contributions to the chromosome 10p and 14q peaks, and both EAs and AIs contributed to the evidence for linkage to 18q.

When ACR was evaluated as a quantitative trait, suggestive evidence for linkage was observed on chromosomes 2q, 7q and 15q. Again, the relative ethnic contributions to each peak differed, with the relatively small number of AI families predominantly contributing to the chromosome 10p and 14q peaks, and both EAs and AIs contributed to the evidence for linkage to 18q.

These results should be considered in the context of other genome scans that have evaluated albuminuria, GFR, overt nephropathy and ESRD in diabetic and non-diabetic families. While the majority of AA families in this report were recruited by Wake
Forest University School of Medicine, the FIND families were different from those contained in the previously published diabetic ESRD genome scan from that institution (11-13). Several Pima families reported previously (9) were included in this FIND analysis, although they are a small component of total participants in this report. The FIND 7q, 10p and 18q DN linkage regions replicate peaks previously observed in type 2 DN and type 2 diabetic ESRD in Turkish, AA and Pima families. The novel FIND peaks on 2, 14 and 15 likely reflect the inclusion of additional ethnic groups, or possibly different FIND inclusion criteria for DN. The FIND inclusion criteria for probands required the presence of severe DN, either ESRD or proteinuria > 1 g/day, which is likely to progress to ESRD.

It is possible that the susceptibility to albuminuria and to CKD with reduced GFR may not share many genetic determinants and that the genes regulating renal function may differ from those controlling proteinuria (26). However, longitudinal studies will ultimately be required to reach this conclusion. The FIND DN linkage result on 10p replicates linkage to diabetic and non-diabetic ESRD previously observed in reports evaluating AA families (15;16). Ewens et al. (27) and McKnight et al. (17) used association based methods in type 1 diabetic populations of modest size, and observed associations with the neuropilin 1 and the D10S1435 marker, respectively. This genomic stretch is also orthologous to the RF-5 locus of the rat reported by Brown et al (28). Therefore, it is possible that a general “renal failure” susceptibility gene exists on 10p, a gene promoting renal failure in the presence of hyperglycemia, as well as other systemic insults including high blood pressure. The chromosome 18q peak has been observed previously in DN in several different ethnic groups (10;11), and it is reportedly due to polymorphisms in the carnosinase gene (CNDP1) (29;30). Individuals homozygous for the 5 leucine repeat (CNDP1 Mannheim allele) were at reduced risk for development of DN. CNDP1 and other positional candidate genes under the 18q peak should be evaluated for their role in susceptibility to DN, because multiple independent studies show convergence for linkage results in this region.

The pathogenesis of type 1 versus type 2 diabetic nephropathy and their respective genetic bases remain unknown. Although, common genes for these disorders have been postulated, the loci that were linked with diabetic albuminuria and nephropathy in the Joslin Diabetes Clinic population (type 1 diabetes) are clearly different from those identified in this, and other, type 2 diabetic nephropathy cohorts. Therefore, it remains possible that different genetic loci are implicated in these two forms of DN.

Issues concerning heterogeneity of the DN disease phenotype, typically expressed as either loss of GFR (e.g., presence of CKD or ESRD) versus albuminuria per se, abound. It is clear that the “classic” DN phenotype typically encompasses both loss of GFR in concert with excess albuminuria. Therefore, the FIND attempted to define a severe DN phenotype for affected probands, that would be present in those with diabetic nephropathy as their cause for ESRD (or advanced CKD) and also those at high risk for rapid progression to diabetic ESRD (as defined by proteinuria > 1 gram/day). The association between high level proteinuria and rapid loss of GFR in DN is well established (31;32). However, there is heterogeneity of disease states when referring to diabetic subjects with milder levels of albuminuria (e.g., microalbuminuria), and comparing them to subjects with diabetic ESRD. This relates to the fact that microalbuminuria may regress to normal levels and is independently associated with risk for CV events far more strongly than it is to development of DN. This
heterogeneity led the FIND investigators to perform independent genome scans for: (1) the amalgamated dichotomous phenotype of DN, (2) the quantitative trait of urine ACR, and (3) the quantitative trait of GFR.

In summary, the initial preliminary phase of the FIND linkage analysis detected evidence for linkage to DN on chromosomes 7q21.3, 10p15.3, 14q23.1, and 18q22.3, and to albuminuria on 2q14.1, 7q21.1 and 15q26.3. The first phase FIND results replicate the evidence for linkage to DN on chromosomes 7q, 10p and 18q in diverse ethnic groups, as well as on 10p in diabetic and non-diabetic ESRD. The final phase FIND genome scan with a SNP-based marker set will be performed in approximately 5000 individuals in more than 1200 families. Genomic regions revealing significant and consistent evidence for linkage to DN in large family-based studies, such as the FIND, should be fine mapped in order to identify DN susceptibility genes. Our strategy for prioritizing ultra fine mapping of linkage regions entails ranking the regions by the best p-value and their allele sharing characteristics. These genes have the potential to improve our understanding of the pathogenesis of DN and hold great promise for the identification of novel therapeutic strategies.

Acknowledgements
We kindly thank all FIND participants. This study was supported by the research grant U01DK57292-05 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and, in part, by the Intramural Research Program of the NIDDK. This work was supported by the National Center for Research Resources for the General Clinical Research Center (GCRC) grants: Case Western Reserve University M01-RR-000080; Wake Forest University M01-RR-07122; Harbor-UCLA Medical Center M01-RR-00425; College of Medicine - University of California Irvine M01-RR-00827-29; University of New Mexico HSC M01-RR-00997; Frederic C. Bartter M01-RR-01346. Genotyping was performed by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to the Johns Hopkins University; contract number N01-HG-65403. The results of this analysis were obtained by using the S.A.G.E. package of genetic epidemiology software, which is supported by a U.S. Public Health Service Resource Grant (RR03655) from the National Center for Research Resources.

Appendix: Members of the Family Investigation of Nephropathy and Diabetes Research Group
Genetic Analysis and Data Coordinating Center, Case Western Reserve University, Cleveland, Ohio: SK Iyengar*, RC Elston¶, KAB Goddard**, JM Olson** (deceased), S Ialacci#, C Fondran, A Horvath, G Jun, K Kramp, SRE Quade, M Slaughter, E Zaletel. Participating Investigation Centers:

External Advisory Committee: D Warnock (chair), R Chakraborty, GM Dunston, SJ O’Brien (ad hoc), R Spielman.* Principal Investigator † Former Principal Investigator ** Co-investigator # Program Coordinator § University of California, Davis, CA= † University of California, Irvine, CA ‡ Study Chair

Electronic- Database Information Online Mendelian Inheritance In Man (OMIM), http://www.ncbi.nlm.nih.gov/omim (For DN)
Reference List


Table 1. Clinical characteristics of the genotyped individuals stratified by proband and DN status

<table>
<thead>
<tr>
<th></th>
<th>DN Probands</th>
<th>DN Relatives</th>
<th>DM Without Nephropathy Relatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>349</td>
<td>390</td>
<td>147</td>
</tr>
<tr>
<td>Male, N (%)</td>
<td>169 (46.0)</td>
<td>180 (46.2)</td>
<td>43 (29.3)</td>
</tr>
<tr>
<td>Age, years (sd)</td>
<td>57 (10.7)</td>
<td>56 (11.4)</td>
<td>59 (10.0)</td>
</tr>
<tr>
<td>BMI, kg/m² (sd)</td>
<td>30 (7.2)</td>
<td>32 (8.6)</td>
<td>32 (7.3)</td>
</tr>
<tr>
<td>Ethnicity, N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>European American</td>
<td>50 (14.3)</td>
<td>42 (10.8)</td>
<td>27 (18.4)</td>
</tr>
<tr>
<td>African American</td>
<td>90 (25.8)</td>
<td>95 (24.4)</td>
<td>33 (22.4)</td>
</tr>
<tr>
<td>Mexican American</td>
<td>179 (51.3)</td>
<td>210 (53.8)</td>
<td>80 (54.4)</td>
</tr>
<tr>
<td>American Indian</td>
<td>30 (8.6)</td>
<td>43 (11.0)</td>
<td>7 (4.8)</td>
</tr>
<tr>
<td>ESRD, N (%)</td>
<td>282* (80.8)</td>
<td>53 (13.6)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age diagnosed, years (sd)</td>
<td>34 (12.3)</td>
<td>41 (13.1)</td>
<td>42 (11.5)</td>
</tr>
<tr>
<td>Duration, years (sd)</td>
<td>23 (8.4)</td>
<td>16 (10.1)</td>
<td>17 (7.0)</td>
</tr>
<tr>
<td>Biochemistry†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycosylated hemoglobin, % (sd)</td>
<td>7.8 (1.8)</td>
<td>8.6 (2.4)</td>
<td>7.8 (2.0)</td>
</tr>
<tr>
<td>Serum creatinine, mg/dl (sd)</td>
<td>8.58 (3.02)</td>
<td>2.6 (3.2)</td>
<td>0.99 (.16)</td>
</tr>
<tr>
<td>Blood Urea Nitrogen mg/dl, (sd)</td>
<td>72.9 (17.2)</td>
<td>29.9 (25.5)</td>
<td>15.7 (4.9)</td>
</tr>
<tr>
<td>Urine PCR, g/g (sd)</td>
<td>3.29 (.69)</td>
<td>1.38 (1.4)</td>
<td>0.17 (.26)</td>
</tr>
<tr>
<td>Urine ACR, g/g (sd)</td>
<td>2.75 (.7)</td>
<td>0.93 (1.2)</td>
<td>0.01 (.01)</td>
</tr>
<tr>
<td>GFR, ml/min/1.73m² (sd)</td>
<td>10.6 (15.0)</td>
<td>68.5 (42.0)</td>
<td>85.6 (24.5)</td>
</tr>
<tr>
<td>GFR, ml/min/1.73m² (median)</td>
<td>3 (3-3)</td>
<td>0.1 (0.05-0.8)</td>
<td>0.009 (0.006-0.01)</td>
</tr>
<tr>
<td>GFR, ml/min/1.73m² (Q25-Q75)</td>
<td>5 (5-5)</td>
<td>76.2 (47.4-98.0)</td>
<td>89.8 (73.9-105.1)</td>
</tr>
</tbody>
</table>

SD is defined as standard deviation
*Three probands are not on dialysis and have no urine results, but were classified as ESRD for the analyses. †Laboratory values in participants with ESRD (either on dialysis or after renal transplant) were set to: serum creatinine 10 mg/dl, BUN 80 mg/dl, PCR 3.5 g/g, ACR 3 g/g, and GFR 5.0 ml/min/1.73m².
Table 2. Sample size for diabetic nephropathy as dichotomous trait genome scan and Urine ACR as quantitative trait genome scan

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Pedigrees</th>
<th>Dichotomous Trait</th>
<th>Quantitative Trait</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FULL SIB PAIRS</td>
<td>HALF SIB PAIRS</td>
<td>SIB AVUNCULAR</td>
</tr>
<tr>
<td>AA</td>
<td>CA 96</td>
<td>D 51</td>
<td>CU 27</td>
</tr>
<tr>
<td>AI</td>
<td>CA 32</td>
<td>D 27</td>
<td>CU 15</td>
</tr>
<tr>
<td>EA</td>
<td>CA 54</td>
<td>D 17</td>
<td>CU 29</td>
</tr>
<tr>
<td>MA</td>
<td>CA 196</td>
<td>D 122</td>
<td>CU 87</td>
</tr>
<tr>
<td>Total</td>
<td>CA 378</td>
<td>D 217</td>
<td>CU 158</td>
</tr>
</tbody>
</table>

AA= African American; AI= American Indian; EA= European American, MA= Mexican American
CA (concordant affected) = both individuals in the sib pair have DM and DN.
D (discordant) = both individuals in the sib pair have DM, one has DN and the other is unaffected with DN after DM duration of at least 10 years.
CU (concordant unaffected) = both individuals in the sib pair have DM with a duration of at least 10 years, neither has DN.
NA not applicable to quantitative traits
Table 3. Summary of linkage peaks for the dichotomous diabetic nephropathy trait where the nominal p-value reached ≤ 0.003

<table>
<thead>
<tr>
<th>CHROMOSOME</th>
<th>ETHNIC GROUP</th>
<th>ASYMPTOTIC P-VALUE</th>
<th>EMPIRICAL P-VALUE</th>
<th>PEAK LOCATION (cM)</th>
<th>FLANKING MARKERS†</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>AA</td>
<td>1.71x10^-5</td>
<td>6.00x10^-5</td>
<td>7q21.3 (104)</td>
<td>D7S2212-D7S821</td>
</tr>
<tr>
<td>7</td>
<td>ALL*</td>
<td>8.60x10^-5</td>
<td></td>
<td>7q21.3 (104)</td>
<td>D7S2212-D7S821</td>
</tr>
<tr>
<td>10</td>
<td>AI</td>
<td>4.45x10^-5</td>
<td>1.40x10^-4</td>
<td>10p15.3 (0)</td>
<td>D10S1435-D10S189</td>
</tr>
<tr>
<td>10</td>
<td>ALL*</td>
<td>1.29x10^-4</td>
<td></td>
<td>10p15.3 (0)</td>
<td>D10S1435-D10S189</td>
</tr>
<tr>
<td>14</td>
<td>AI</td>
<td>7.23x10^-5</td>
<td>2.00x10^-3</td>
<td>14q23.1 (55)</td>
<td>D14S587-D14S588</td>
</tr>
<tr>
<td>14</td>
<td>ALL*</td>
<td>7.80x10^-4</td>
<td></td>
<td>14q21.1 (55)</td>
<td>D14S587-D14S588</td>
</tr>
<tr>
<td>18</td>
<td>EA</td>
<td>2.17x10^-3</td>
<td>3.15x10^-2</td>
<td>18q22.3 (116)</td>
<td>D18S1371-D18S1390</td>
</tr>
<tr>
<td>18</td>
<td>ALL*</td>
<td>1.61x10^-3</td>
<td></td>
<td>18q22.3 (116)</td>
<td>D18S1371-D18S1390</td>
</tr>
</tbody>
</table>

AA= African American; AI= American Indian; Empirical p-value = Permutation p-value computed for 50,000 replicates; cM= centimorgan
*Linkage evidence for the full sample was determined by post-hoc pooling of the p-values for all the ethnic groups using Fisher’s method, as described in the materials and methods. Ethnic-specific p-values are reported only for those groups that showed significant evidence of linkage.
† The flanking markers represent the nearest marker on either side of the linkage peak where there was a 1 unit drop in the p-value to define the peak.
Table 4. Mean Allele Sharing for Affected and Discordant Sib Pairs at the Peak Locations for Diabetic Nephropathy

<table>
<thead>
<tr>
<th>ETHNIC GROUP</th>
<th>CHROMOSOME</th>
<th>DSP MEAN SHARING*</th>
<th>P-VALUE</th>
<th>ASP MEAN SHARING*</th>
<th>P-VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>7</td>
<td>0.38</td>
<td>0.0013</td>
<td>0.58</td>
<td>0.0248</td>
</tr>
<tr>
<td>AI</td>
<td>10</td>
<td>0.27</td>
<td>0.0007</td>
<td>0.58</td>
<td>0.0498</td>
</tr>
<tr>
<td>AI</td>
<td>14</td>
<td>0.30</td>
<td>0.0287</td>
<td>0.56</td>
<td>0.1733</td>
</tr>
<tr>
<td>EA</td>
<td>18</td>
<td>0.40</td>
<td>0.0308</td>
<td>0.57</td>
<td>0.1959</td>
</tr>
</tbody>
</table>

AA= African American; AI= American Indian; EA= European American; DSP= Discordant Sibling Pairs; ASP= Affected Sibling Pairs

*Under the null hypothesis, sib pairs should share 50% of their genome identical by descent. At a disease locus discordant sibs should share less than 50% and affected sibs should share more than 50%.
Table 5. Summary of linkage peaks for urine ACR where the nominal p-value reached \( \leq 0.006 \)

<table>
<thead>
<tr>
<th>CHROMOSOME</th>
<th>ETHNIC GROUP</th>
<th>ASYMPTOTIC P-VALUE</th>
<th>EMPIRICAL P-VALUE*</th>
<th>PEAK LOCATION (cM)</th>
<th>FLANKING MARKERS†</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>AI</td>
<td>8.24x10^{-5}</td>
<td>0.0015</td>
<td>2q14.1 (124)</td>
<td>D2S410-D2S1328</td>
</tr>
<tr>
<td>2</td>
<td>ALL</td>
<td>5.58x10^{-3}</td>
<td></td>
<td>2q14.1 (124)</td>
<td>D2S410-D2S1328</td>
</tr>
<tr>
<td>7</td>
<td>EA</td>
<td>3.76x10^{-5}</td>
<td>0.0006</td>
<td>7q21.1 (91)</td>
<td>D7S3046-D7S2212</td>
</tr>
<tr>
<td>7</td>
<td>ALL</td>
<td>3.04x10^{-4}</td>
<td></td>
<td>7q21.1 (91)</td>
<td>D7S3046-D7S2212</td>
</tr>
<tr>
<td>15</td>
<td>AA</td>
<td>1.85x10^{-4}</td>
<td>0.0055</td>
<td>15q26.3 (106)</td>
<td>D15S657-D15S642</td>
</tr>
<tr>
<td>15</td>
<td>ALL</td>
<td>2.48x10^{-3}</td>
<td></td>
<td>15q26.3 (106)</td>
<td>D15S657-D15S642</td>
</tr>
</tbody>
</table>

AI= American Indian; EA= European American; AA= African American; cM= centimorgan
*Empirical p-value = Permutation p-value computed for 50,000 replicates
† The flanking markers represent the nearest marker on either side of the linkage peak where there was a 1 unit drop in the p-value to define the peak.
Figure 1. Plot of the genome scan for DN as a dichotomous trait. For each autosome (1-22) the genetic distance along the chromosome is plotted on the X-Axis and the -log_{10} (P-value) is plotted on the Y-axis. On the right side of chromosomes 4, 8, 12, 16 and 20, the corresponding LOD score is plotted on the Y-axis. The dashed lines represent p-values of 0.000022 and 0.000744 which meet the Lander-Kruglyak criterion (33) of significant and suggestive linkage, respectively. LOD scores of 3.7 and 2.1 equate to the Lander-Kruglyak p-values of significant and suggestive linkage, respectively.

Figure 2. Chromosomes 7, 10, 14, and 18 linkage results for the DN as a dichotomous displayed by ethnic group. The genetic distance along the chromosome is plotted on the X-Axis and the –
log_{10}(P-value) is plotted on the Y-axis. The markers that were genotyped are labeled along the top of the graph. On the right side the corresponding LOD score is plotted on the Y-axis. The dashed lines represent p-values of 0.000022 and 0.000744 which meet the Lander-Kruglyak criterion (33) of significant and suggestive linkage, respectively. LOD scores of 3.7 and 2.1 equate to the Lander-Kruglyak p-values of significant and suggestive linkage, respectively.

Figure 3. Plot of the genome scan for urine ACR in all diabetic participants. For each autosome (1-22) the genetic distance along the chromosome is plotted on the X-Axis and the $-\log_{10}(P$-value) is plotted on the Y-axis. On the right side the corresponding LOD score is plotted on the Y-axis. The dashed lines represent p-values of 0.000022 and 0.000744 which meet the Lander-Kruglyak criterion (33) of significant and suggestive linkage, respectively.
Kruglyak criterion (33) of significant and suggestive linkage, respectively. LOD scores of 3.7 and 2.1 equate to the Lander-Kruglyak p-values of significant and suggestive linkage, respectively.