Multiple Superoxide Dismutase 1/Splicing Factor Serine Alanine 15 Variants Are Associated With the Development and Progression of Diabetic Nephropathy

The Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Genetics Study

Hussam Al-Kateb,[1] Andrew P. Boright,[2] Lucia Mirea,[3,4] Xinlei Xie,[3] Rinku Sutradhar,[3,5] Alireza Mowjoodi,[1] Bhupinder Bharaj,[1] Michelle Liu,[1] Jean M. Bucksa,[6] Valerie L. Arends,[6] Michael W. Steffes,[6] Patricia A. Cleary,[7] Wanjie Sun,[7] John M. Lachin,[7] Paul S. Thorner,[8,9] Michael Ho,[8] Amy Jayne McKnight,[10] A. Peter Maxwell,[10] David A. Savage,[10] Kenneth K. Kidd,[11] Judith R. Kidd,[11] William C. Speed,[11] Trevor J. Orchard,[12] Rachel G. Miller,[12] Lei Sun,[1,4] Shelley B. Bull,[3,4] Andrew D. Paterson,[1,4] and the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Research Group*

BACKGROUND—Despite familial clustering of nephropathy and retinopathy severity in type 1 diabetes, few gene variants have been consistently associated with these outcomes.

RESEARCH DESIGN AND METHODS—We performed an individual-based genetic association study with time to renal and retinal outcomes in 1,362 white probands with type 1 diabetes from the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) study. Specifically, we genotyped 1,411 SNPs that capture common variations in 212 candidate genes for long-term complications and analyzed them for association with the time from DCCT baseline to event for renal and retinal outcomes using multivariate Cox proportion hazards models. To address multiple testing and assist interpretation of the results, false discovery rate q values were calculated separately for each outcome.

RESULTS—We observed association between rs17880135 in the 3′ region of superoxide dismutase 1 (SOD1) and the incidence of both severe nephropathy (hazard ratio [HR] 2.62 [95% CI 1.64–4.18], $P = 5.6 \times 10^{-5}$, $q = 0.06$) and persistent microalbuminuria (1.82 [1.29–2.57], $P = 6.4 \times 10^{-4}$, $q = 0.46$). Sequencing and fine-mapping identified additional SOD1 variants, including rs202446, rs9974610, and rs204732, which were also associated ($P < 10^{-3}$) with persistent microalbuminuria, whereas rs17880135 and rs17881180 were similarly associated with the development of severe nephropathy. Attempts to replicate the findings in three cross-sectional case-control studies produced equivocal results. We observed no striking differences between risk genotypes in serum SOD activity, serum SOD1 mass, or SOD1 mRNA expression in lymphoblastoid cell lines.

CONCLUSIONS—Multiple variations in SOD1 are significantly associated with persistent microalbuminuria and severe nephropathy in the DCCT/EDIC study. ***Diabetes*** **57:218–228, 2008**

From the [1]Program in Genetics and Genome Biology, Hospital for Sick Children, Toronto, Ontario, Canada; the [2]Department of Medicine, University Health Network, University of Toronto, Toronto, Ontario, Canada; the [3]Samuel Lunenfeld Research Institute of Mount Sinai Hospital, Prosserman Centre for Health Research, Toronto, Ontario, Canada; the [4]Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada; the [5]Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada; the [6]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota; [7]The Biostatistics Center, The George Washington University, Rockville, Maryland; the [8]Division of Pathology, Hospital for Sick Children, Toronto, Ontario, Canada; the [9]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada; the [10]Nephrology Research Group, Queen's University of Belfast, Belfast, Northern Ireland, U.K.; the [11]Department of Genetics, Yale University School of Medicine, New Haven, Connecticut; and the [12]University of Pittsburgh, Pittsburgh, Pennsylvania.

Address correspondence and reprint requests to Dr. Andrew Paterson, Program in Genetics and Genome Biology, The Hospital for Sick Children, TMDT Building East Tower, Rm. 15-707, 101 College St., Toronto, Ontario, Canada M5G 1L7. E-mail: andrew.paterson@utoronto.ca.

Received for publication 31 July 2007 and accepted in revised form 30 September 2007. Published ahead of print at http://diabetes.diabetesjournals.org on October 2007. DOI: 10.2337/db07-1059. DCCT clinical trial reg. no. NCT00360815, EDIC clinical trial reg. no. NCT00360893, clinicaltrials.gov.

H.A.-K. and A.P.B. contributed equally to this work, and L.M. and X.X. contributed equally to this work.

*A complete list of investigators and members of the research group appears in *N Engl J Med* 353:2643–2653, 2005.

Diabetic nephropathy resulting in renal failure occurs in ~15% of those with type 1 diabetes, and it clusters in families (1–5). Furthermore, there is strong concordance for the severity and patterns of glomerular lesions among type 1 diabetic sibpairs (6), consistent with genetic effects. The sole linkage study reported for nephropathy in type 1 diabetes found significant evidence for linkage to *Homo sapiens* (HSA) 3q (7). Genome-wide linkage studies performed for diabetic nephropathy in type 2 diabetes have revealed suggestive evidence for a susceptibility region on HSA 7q (8,9). A candidate region linkage study of diabetic nephropathy in type 2 diabetes identified a locus on HSA 18q and association of a variant in CNDP1 (10,11). A genome-wide association study of 81,315 SNPs in Japanese type 2 diabetic nephropathy case subjects detected association of SNPs in SLC12A3 and ELMO1 with diabetic nephropathy (12,13), whereas low-resolution genome-wide microsatellite marker association of diabetic nephropathy in type 1 diabetes using pooled DNA samples identified a locus on HSA 10 (14). A family-based association study of 382 markers in 115 candidate genes using 72 probands

with both type 1 diabetes and end-stage renal disease (ESRD) highlighted several loci (15). Most recently, a genome-wide association study using the Affymetrix 100K assay identified a locus at PVT1 for nephropathy in type 2 diabetes (16). In addition, many candidate gene association studies have been performed for nephropathy in type 1 diabetes, but the results have often been inconsistent (17,18), which may be explained by small sample size and low power, with only one variant typically being tested, by differences in phenotypic definition, or by population stratification.

For diabetic retinopathy, there are reports of clustering of severity in both type 1 and type 2 diabetes (4,19), and recent genome-wide linkage studies of retinopathy in type 2 diabetes (20,21) have highlighted certain chromosomal regions, but candidate gene association studies have generally not produced consistent evidence.

In this report, we describe associations between common variations in 212 candidate genes and the time to renal and retinal outcomes in subjects who participated in The Diabetes Control and Complications Trial (DCCT) and were followed up in the Epidemiology of Diabetes Interventions and Complications (EDIC) study.

RESEARCH DESIGN AND METHODS

The DCCT was a multicenter randomized clinical trial to compare intensive and conventional insulin therapy on the development and progression of early vascular and neurological complications of type 1 diabetes (22,23). We studied 1,369 white probands (out of 1,441 DCCT subjects) and 2,966 of their first-degree relatives (details are provided in online appendix 26 and the online appendix text [available at http://dx.doi.org/10.2337/db07-1059]). To confirm findings from the DCCT/EDIC probands, we identified three other studies of diabetic nephropathy in type 1 diabetes (Table 5); details of the characteristics are in the online appendix text.

Retinal and renal outcomes in DCCT/EDIC

Outcomes were defined as time in years from DCCT baseline until the event (detailed in the online appendix text). Renal outcomes included data up to EDIC year 8 (2001): Persistent microalbuminuria was defined as the time to two consecutive albumin excretion rates (AERs) >30 mg/24 h (>20.8 μg/min); severe nephropathy was the time to AER >300 mg/24 h (>208 μg/min) with prior persistent microalbuminuria or ESRD. For retinal outcomes, data up to EDIC year 10 (2003) were used: ≥3 step progression was defined as the time to ≥3 step progression from DCCT baseline using Early Treatment of Diabetic Retinopathy Study (ETDRS) scoring; severe retinopathy was the time to ETDRS level 53/<53 or greater or scatter laser treatment; clinically significant macular edema (CSME) was the time to macular edema or focal laser treatment.

Molecular methods

Gene and SNP selection and Illumina genotyping. Candidate genes were selected based on their involvement in the pathophysiology of diabetic nephropathy and/or diabetic retinopathy. We considered biological evidence, including gene function, evidence for differential expression in diabetes complications, prior genetic association studies of diabetes complications and their risk factors, and genes in pathways related to these genes. For genes sequenced by SeattleSNPs (24), tagSNPs were selected using LDselect (24), requiring pairwise r^2 cutoff >0.64 and minor allele frequency (MAF) >0.05 in either the European American or PDR90 samples (25) (http://egp.gs.washington.edu). Otherwise, we used Haploview pairwise Tagger (26) to select SNPs with an r^2 cutoff >0.8 in the 30 CEU trios (CEPH samples of European descent) from release 16 (phase I) of the HapMap project (http://www.hapmap.org), including 5 kb flanking either side of the gene (27,28). Selected SNPs underwent bioinformatic evaluation to score their suitability for genotyping via the Illumina BeadArray GoldenGate custom assay (29). SNPs that did not pass criteria were replaced by other tagSNPs when possible or were excluded. In total, ~2,750 SNPs were evaluated bioinformatically, and 1,536 were genotyped in all DCCT/EDIC probands with available DNA (n = 1,418). Additional SNPs were genotyped using Taqman technology (see online appendix text).

DNA sequencing. To fine-map the causal variant(s), we sequenced ~19 kb of the superoxide dismutase 1 (SOD1) genomic region in all eight DCCT/EDIC probands homozygous for the rare genotype (C/C) at rs17880135. The region encompassed SOD1 and ~5 kb at both the 5′ and 3′. In addition, we sequenced the cDNA of splicing factor serine alanine 15 (SFRS15, the gene 3′ of SOD1) in seven individuals, selected in a way similar to that of the samples for SOD1 sequencing (online appendix text and online appendix 18).

Functional studies and immunohistochemistry. We quantified SOD1 mRNA expression, serum SOD activity, and SOD1 mass in a subset of probands and nondiabetic relatives selected by rs17880135 genotype. SOD1 mRNA expression was measured in Epstein-Barr virus–transformed lymphocytes from 18 individuals homozygous for the rare allele (8 probands and 10 nondiabetic relatives), 38 heterozygotes (16 probands and 22 nondiabetic relatives), and 38 homozygotes (16 probands and 22 nondiabetic relatives) for the common allele at rs17880135 with matching on relevant covariates (online appendix text). To determine whether serum SOD activity and SOD1 mass differed by rs17880135 genotype, we proceeded in a similar fashion (online appendix text).

Statistical methods

DCCT/EDIC probands. SNP allele and genotype frequencies were calculated and deviation from Hardy-Weinberg equilibrium (HWE) was determined using exact tests (online appendix 3). Details about the association analysis for SNPs with DCCT/EDIC outcomes are provided in the online appendix text (17,30). Briefly, SNP genotypes were tested for association with DCCT baseline design and biological covariates (online appendix 4). Cox proportional hazards analysis of discrete time-to-event outcomes was performed using two models with additive genotype coding (31). The first model was "univariate," including DCCT cohort, treatment, and cohort-treatment interaction as covariates, and stratified by the year of entry into DCCT. Our primary analysis was a multivariate Cox model, including cohort, treatment, cohort-treatment interaction, age of diagnosis, DCCT baseline duration, sex, BMI, mean blood pressure, triglyceride, HDL-C, total cholesterol, baseline smoking, A1C eligibility, time-dependent updated mean A1C, and time-dependent indicators for hypertension diagnosis and/or treatment of hypertension, with stratification by DCCT year of entry (online appendix text and online appendix 5). X chromosome SNPs were analyzed separately in men and women (online appendixes 6–9). SNPs with MAF <1% were analyzed separately (online appendixes 10–12) using a modified Cox model designed for sparse data (32,33), including cohort, treatment, cohort-treatment interaction, and sex as covariates.

Assessment of statistical significance. To address the inherent problem of multiple hypotheses testing, we adopted the false discovery rate (FDR) framework and applied the q value method (34) (http://faculty.washington.edu/jstorey/qvalue/). The q value is a measure of statistical significance in terms of FDR, and a test with q value of 0.5 means that the expected FDR is 0.5 if we declare that test and all other tests with smaller q values significant. To use the available phenotype information, we applied the stratified FDR (35) (http://faculty.washington.edu/jstorey/qvalue/) designed to increase power; q values were hence calculated separately for each outcome.

Population stratification and family-based analysis in DCCT/EDIC. To detect population structure, we used STRUCTURE v2.0 to analyze genotype data for SNPs chosen to have low pairwise linkage disequilibrium from the Illumina assay (36). We first included both white and nonwhite probands (n = 50) to identify white probands whose subgroup classification as determined by STRUCTURE disagreed with their self-reported ethnicity; then we analyzed the white individuals alone to identify subgroups within them. Using either 180 SNPs with the highest heterozygosity or 164 ancestry informative SNPs (37), there was no evidence for outliers, and the white probands could not be separated into subgroups. In addition, to further exclude that the association of SOD1 SNPs with persistent microalbuminuria and severe nephropathy was due solely to population stratification, we genotyped three SOD1 SNPs associated with renal outcomes (Table 3) in all available DCCT/EDIC relatives (online appendix 26) and performed family-based analysis (online appendix 25).

Martingale residuals (38) obtained from the multivariate Cox proportional hazards models without genotype, were used as continuous traits in subsequent family-based association analyses (FBAs; see online appendix 26 for family structures). FBA conditions on the parental genotypes were used to test for departures between the observed offspring genotypes and those expected under Mendelian inheritance (39). Analyses were performed assuming an additive genetic model and used FBAT v1.7.3 to perform single-SNP and haplotype analyses (40–42).

Multi-SNP and multistate analysis. Multi-SNP selection analyses were also performed to determine the most likely etiological variant(s) for each of persistent microalbuminuria and severe nephropathy. These included multiple Cox proportional hazards regression models with forward and backward selection of main effects and two-way interactions of SOD1 SNPs (online appendix text) to capture un-genotyped variants and to avoid the uncertainties of phasing unrelated probands. To examine the association of SOD1/SFRS15 SNPs with both renal outcomes, we assumed a three-state progressive

TABLE 1
Single-marker multivariate results with q value <0.5 in DCCT/EDIC probands for severe nephropathy and persistent microalbuminuria

Gene symbol	SNP	Allele		χ_1^2	P value	HR (95% CI)	q value
		Major	Minor (%)				
Severe nephropathy							
SOD1	rs17880135	A	C (5.7)	16.24	5.6×10^{-5}	2.62 (1.64–4.18)	0.07
CASP3	rs2705897	C	A (26.4)	9.87	1.7×10^{-3}	1.62 (1.20–2.17)	0.46
TRPC6	rs3824935	G	A (9.9)	9.77	1.8×10^{-3}	2.00 (1.29–3.08)	0.46
TGFBR2	rs2276768	G	A (10.8)	9.52	2.0×10^{-3}	0.40 (0.22–0.71)	0.46
HPSE	rs4693614	A	G (23.2)	9.44	2.1×10^{-3}	1.61 (1.19–2.19)	0.46
CYP11B2	rs7844961	G	A (8.4)	8.93	2.8×10^{-3}	1.89 (1.25–2.88)	0.46
COX5A	rs8042694	A	G (28.8)	8.92	2.8×10^{-3}	1.59 (1.17–2.16)	0.46
NPHS1	rs3814995	G	A (32.5)	8.58	3.4×10^{-3}	0.61 (0.44–0.85)	0.46
ATP5G3	rs10497435	A	G (4.3)	8.57	3.4×10^{-3}	2.39 (1.33–4.29)	0.46
FLT4	rs307806	G	A (14.9)	8.07	4.5×10^{-3}	1.66 (1.17–2.37)	0.46
FLT4	rs2279622	G	A (6.4)	8.01	4.7×10^{-3}	1.95 (1.23–3.10)	0.46
UQCRC1	rs11715496	C	A (2.0)	7.88	5.0×10^{-3}	2.63 (1.34–5.16)	0.46
BDKRB2	rs4900312	G	A (20.0)	7.81	5.2×10^{-3}	0.57 (0.39–0.85)	0.46
LIPC	rs1968685	C	G (48.2)	7.82	5.2×10^{-3}	0.65 (0.47–0.88)	0.46
PARP1	rs2027440	A	G (15.8)	7.50	6.2×10^{-3}	1.60 (1.14–2.24)	0.48
PARP1	rs3219065	A	G (15.8)	7.50	6.2×10^{-3}	1.60 (1.14–2.24)	0.48
Persistent microalbuminuria							
SOD1	rs17880135	A	C (5.7)	11.64	6.5×10^{-4}	1.82 (1.29–2.57)	0.48
PON1	rs1157745	C	A (27.6)	10.56	1.2×10^{-3}	0.68 (0.54–0.86)	0.48
PON1	rs3917532	A	T (27.6)	10.56	1.2×10^{-3}	0.68 (0.54–0.86)	0.48
AKR1B1	rs2259458	C	A (30.7)	10.19	1.4×10^{-3}	1.36 (1.13–1.64)	0.48
FLT4	rs307806	G	A (14.9)	6.47	1.1×10^{-2}	1.37 (1.08–1.75)	0.66
FLT4	rs10516142	G	A (3.0)	6.29	1.2×10^{-2}	1.85 (1.15–3.03)	0.66

Sample size for the analysis of persistent microalbuminuria, 1,296; sample size for the analysis of severe nephropathy, 1,362. No SNPs had q values <0.5 for any of the retinal outcomes. The HR is specified for each minor allele using the common homozygote genotype as the baseline. Results for two FLT4 SNPs are included for the persistent microalbuminuria results because SNPs in FLT4 also show evidence for association with severe nephropathy. P values are based on the Wald test.

model, focusing on the rate of transition from normal to persistent microalbuminuria, and from persistent microalbuminuria to severe nephropathy. The multistate model provided separate estimates of the transition intensities between states and separate estimates of the SNP and covariate effects (43–45). Note that our model differs from the primary multivariate model above in that the repeated indicator for hypertension was not included.

Statistical analysis of SOD1 mRNA, SOD activity, and SOD1 mass

Replicate log SOD1 expression values were averaged and compared across genotypes using separate general linear models for probands and relatives, taking into account the matched design. Log SOD activity and log SOD1 mass were analyzed similarly. To address concerns about residual non-normality in the latter, comparisons were repeated in conditional and unconditional logistic regression models that treated the genotype as a categorical response and log activity or log mass as an explanatory covariate (online appendix text).

RESULTS

Outcomes in DCCT/EDIC. The numbers of individuals with each of the renal and retinal events or who were censored are listed in online appendix 24. Twenty-three percent (n = 212) and 8% (n = 98) of the probands developed persistent microalbuminuria and severe nephropathy, respectively, during a follow-up of 13.5 ± 2.6 years (means ± SD) in DCCT/EDIC with 10 ± 2 AER measures. The effects of covariates in the Cox proportional hazards models, without the inclusion of SNPs, are provided for the univariate (online appendix 1) and multivariate models (online appendix 2), with HRs calculated for specific changes in quantitative covariates for the multivariate models (online appendix 23).

Genotype data. Of the 1,536 SNPs attempted on the Illumina chip, no genotypes were generated for 86 SNPs, leaving 1,441 autosomal and 9 X chromosome SNPs successfully genotyped. Thirty-seven autosomal and two X chromosome SNPs were nonpolymorphic in white DCCT/EDIC probands. Excepting these SNPs, genotype data were 99.95% complete with only six samples having no genotypes. Sixty-two autosomal SNPs had MAF <1% (online appendix 10), leaving 1,342 autosomal and 7 X chromosome SNPs with MAF >1% (online appendix 3). One proband was discrepant between reported and genotyped sex. Based on 18 duplicate samples ($n \sim$ 52,000 genotypes), genotype agreement was 100%. Marked deviation from HWE (online appendixes 3 and 6) was observed for both rs12431885 (triallelic) and rs2069579 in BDKRB2 and the triallelic rs3091244 (CRP) for which we only genotyped alleles A and G.

Association of Illumina SNPs with outcomes. From the multivariate association analysis with outcomes, there were four SNPs with $q < 0.5$ for persistent microalbuminuria, 16 for severe nephropathy, and none for ≥3 step progression, severe retinopathy, or CSME (Table 1; online appendixes 5, 9, and 12). Of interest, most positive SNPs showed detectable associations with either persistent microalbuminuria or severe nephropathy, but not both. Two exceptions were fms-related tyrosine kinase 4 (FLT4; also known as vascular endothelial growth factor receptor 3) and SOD1. For FLT4, rs307806 was associated with both persistent microalbuminuria and severe nephropathy, whereas two other SNPs in this gene were separately associated with persistent microalbuminuria and severe nephropathy. Also of note, three SNPs in CNDP1 (online appendix 5) had $P < 0.05$ for persistent microalbuminuria,

but their linkage disequilibrium with CNDP1 variants associated with diabetic nephropathy is unknown (11). In this report, we focus on SOD1 because rs17880135 had the lowest q value for any outcome (severe nephropathy, $q = 0.06$), and the same SNP also showed the most significant association with persistent microalbuminuria ($q = 0.48$).

DNA sequencing and SOD1 fine-mapping in DCCT/EDIC probands. In the fine-mapping stage, we followed up the association of rs17880135 by examining a total of 11 SNPs in SOD1/SFRS15 (4 on the Illumina chip and 7 genotyped using Taqman). In designing the Ilumina chip, we had selected five tagSNPs with MAF >5% in SOD1 from the SeattleSNPs project PDR90 sample (24), which identified 80 SNPs including 55 represented in 19 bins and an additional 25 single SNPs bins (online appendix 21). Many of the bins contained SNPs with low MAF. Of the five tag SNPs selected, four passed the Illumina bioinformatic design criteria and were genotyped. Rs17880135 lies 3′ of SOD1 and is in the same linkage disequilibrium bin in SeattleSNPs as rs17881180 in intron 1 and was therefore genotyped in all DCCT/EDIC probands. As expected, rs17881180 was in strong linkage disequilibrium with rs17880135 ($r^2 = 0.912$; Table 2) and had similar associations with renal outcomes (Table 3).

We sequenced 19 kb across the genomic region of SOD1 in eight individuals who were homozygous for the rare (high-risk) genotype at rs17880135 (online appendix 19) and found no coding variations. Of note, all eight individuals were homozygous for the rare allele at rs202446 (46), and 4 of 16 chromosomes carried the rare allele at rs17881203 (which had MAF = 5% in HapMap) and were therefore genotyped in DCCT/EDIC probands (Tables 2 and 3). No variants were found in the cDNA of SFRS15 compared with the reference sequence.

Based on the sequencing, SeattleSNPs, and HapMap, we genotyped six additional SNPs in SOD1/SFRS15—selected predominantly to capture linkage disequilibrium bins across SOD1 (online appendixes 21 and 13–15; Tables 2 and 3). Of note, 5 of the 11 SOD1/SFRS15 SNPs (rs17880135, rs17881180, rs202446, rs9974610, and rs204732) were associated with persistent microalbuminuria (with $P < 10^{-3}$), whereas rs17880135 and rs17881180 were most strongly associated with severe nephropathy (Table 3). We found no evidence for significant nonadditivity for the association of SOD1/SFRS15 SNPs with renal outcomes (online appendix 22). There was no significant association ($P < 0.01$) between any of the SOD1 SNPs and any of the retinal outcomes (online appendixes 5 and 15).

Genotyping SOD1 SNPs in HapMap. To determine linkage disequilibrium between rs17880135 and rs17881180 and other SNPs genotyped by HapMap in the SOD1 region, we genotyped them in the CEU HapMap samples (27,28). We found that none of the SNPs within 100 kb SOD1 genotyped by HapMap are in linkage disequilibrium ($r^2 > 0.5$) with any of the five SNPs associated with persistent microalbuminuria (online appendixes 28 and 29).

Multi-SNP analyses and multistate models. Because some of the SNPs in SOD1 that are associated with persistent microalbuminuria and severe nephropathy are in strong linkage disequilibrium with each other (Table 2), we used multi-SNP regression analyses to attempt to identify the most statistically important SNP(s). For persistent microalbuminuria, either rs202446 (backward selection) or rs204732 (forward selection) remained in the model, suggesting that either of them or other variant(s) in strong linkage disequilibrium with both of them account

TABLE 2
Linkage disequilibrium between SOD1 SNPs in DCCT/EDIC probands

Physical location	31,940,240	31,950,512	31,954,158	31,954,859	31,956,763	31,959,353	31,960,736	31,962,033	31,963,874	31,964,004	32,014,714
Location*	5′	5′	Intron 1	Intron 1	Intron 1	Intron 2	Intron 3	Intron 4	3′	3′ SFRS15	Intron 1 SFRS15
rs no.	rs9974610	rs202446	rs17881180	rs17884536	rs17880998	rs17878806	rs17880753	rs17880196	rs17880135	rs17881203	rs204732
rs9974610		0.835	0.979	1.000	1.000	0.833	0.890	0.970	0.981	0.930	0.835
rs202446	0.595		0.971	1.000	1.000	0.752	0.877	0.922	0.974	0.937	0.904
rs17881180	0.246	0.305		1.000	1.000	0.906	1.000	0.936	0.977	0.918	0.962
rs17884536	0.009	0.007	0.002		1.000	0.970	0.980	1.000	1.000	1.000	1.000
rs17880998	0.025	0.020	0.006	0.343		0.981	0.977	1.000	1.000	1.000	0.955
rs17878806	0.123	0.080	0.037	0.045	0.136		0.937	0.965	0.917	0.910	0.795
rs17880753	0.007	0.005	0.002	0.941	0.322	0.041		1.000	1.000	1.000	0.870
rs17880196	0.095	0.068	0.023	0.019	0.055	0.526	0.019		1.000	1.000	0.828
rs17880135	0.265	0.331	0.912	0.002	0.007	0.041	0.002	0.028		0.949	0.974
rs17881203	0.062	0.083	0.237	0.001	0.002	0.011	0.001	0.008	0.243		0.937
rs204732	0.595	0.752	0.286	0.007	0.020	0.096	0.005	0.059	0.307	0.077	

Linkage disequilibrium between SOD1 and SFRS15 SNPs in white DCCT/EDIC probands. The values above and below the diagonal are D′ and r^2, respectively. According to dbSNP build 125 (http://www.ncbi.nlm.nih.gov/SNP), the following are identical: rs17884536 = rs6650814; rs17880998 = rs4998557; rs17880196 = rs1041740; rs17880753 = rs2234694; rs17878806 = rs9967983. Physical location refers to build 36.2 of the human genome. *All SNPs are in SOD1 apart from those indicated, which are in SFRS15.

TABLE 3
Single-marker multivariate results for time to renal outcomes for SNPs in SOD1 and SFRS15 in DCCT/EDIC probands

SNP	Allele		Persistent microalbuminuria				Severe nephropathy			
	Major	Minor (%)	χ_1^2	*P* value	HR (95% CI)	*q* value	χ_1^2	*P* value	HR (95% CI)	*q* value
rs9974610	A	G (17.6)	16.47	4.9×10^{-5}	1.59 (1.27–1.96)	NA	2.00	0.16	1.28 (0.91–1.82)	NA
rs202446	C	A (14.3)	17.15	3.5×10^{-5}	1.64 (1.29–2.08)	NA	7.64	0.0057	1.65 (1.16–2.35)	NA
rs17881180	G	A (5.2)	15.18	9.8×10^{-5}	2.04 (1.43–2.92)	NA	20.5	5.8×10^{-6}	3.03 (1.88–4.89)	NA
rs17884536	C	A (3.8)	0.37	0.55	1.15 (0.73–1.81)	0.96	0.93	0.34	0.67 (0.29–1.51)	0.88
rs17880998	G	A (10.4)	0.92	0.34	1.15 (0.86–1.54)	0.96	0.11	0.75	1.07 (0.68–1.72)	0.88
rs17878806	A	T (45.5)	0.50	0.48	0.93 (0.78–1.14)	NA	0.92	0.34	0.86 (0.65–1.15)	NA
rs17880753	T	G (3.7)	0.80	0.37	1.23 (0.79–1.91)	NA	0.48	0.49	0.75 (0.33–1.68)	NA
rs17880196	G	A (32)	1.69	0.19	0.88 (0.72–1.06)	0.96	1.42	0.23	0.83 (0.61–1.12)	0.88
rs17880135	A	C (5.7)	11.64	6.4×10^{-4}	1.82 (1.29–2.57)	0.48	16.2	5.6×10^{-5}	2.62 (1.64–4.18)	0.07
rs17881203	A	G (1.5)	1.73	0.19	1.56 (0.80–3.00)	NA	4.19	0.040	2.38 (1.04–5.26)	NA
rs204732	C	T (15.4)	16.48	4.9×10^{-5}	1.59 (1.28–2.00)	NA	3.62	0.057	1.39 (0.99–1.96)	NA

NA, not calculated because that SNP was not genotyped on the initial Illumina assay. According to dbSNP build 125, the following are identical: rs17884536 = rs6650814; rs17880998 = rs4998557; rs17880196 = rs1041740; rs17880753 = rs2234694; rs17878806 = rs9967983. The HR is specified for each minor allele using the common homozygote genotype as the baseline. The *P* value is from the Wald test. Sample sizes for this analysis are provided in online appendixes 5 and 15.

for this association. For severe nephropathy, only rs17881180 remained after selection, but because of the strong linkage disequilibrium with rs17880135, we cannot distinguish between them. Examination of all possible pairwise interactions between SNPs provided no compelling evidence for interactions for either persistent microalbuminuria or severe nephropathy (online appendix 16).

Initially, we analyzed persistent microalbuminuria and severe nephropathy as separate traits; however, our definition of severe nephropathy required that individuals first progress through persistent microalbuminuria. To compare the association of each of the SOD1 SNPs with the different stages of disease, we examined multistate models. rs17880135 and rs17881180 were significantly associated with both persistent microalbuminuria and severe nephropathy and with the progression from persistent microalbuminuria to severe nephropathy ($P < 0.05$; Table 4). In contrast, rs204732 was associated with persistent microalbuminuria but not with the subsequent development of severe nephropathy (heterogeneity $P = 0.044$).

Family-based association in DCCT/EDIC. Although there was no significant preferential allele transmission with severe nephropathy detected for any of the three SNPs genotyped in the relatives (rs17880135, rs17881180, and rs202446), rs202446 was associated with persistent microalbuminuria ($P = 0.036$) in the same direction as the individual-based analysis, consistent with haplotype analysis (data not shown). Of note, the family-level test for severe nephropathy with rs17880135 and rs17881180 suffered from low statistical power due to MAF = 5% and subsequently only 152–163 informative families contributing to the test statistic.

Functional studies. Generally, we observed no significant difference ($P < 0.05$) by genotype at rs17880135 of SOD1 expression in EBV-transformed lymphoblasts, serum SOD1 mass, or SOD activity (Fig. 1), although there was a borderline higher SOD1 mass in probands with A/C compared with A/A genotypes. This effect was not observed in the nondiabetic relatives. Findings for rs17881180 were identical, due to nearly complete linkage disequilibrium, and results for the other three SOD1 SNPs associated with persistent microalbuminuria were nonsignificant (online appendixes 30 and 31).

Other studies. We attempted to confirm the association of SOD1 SNPs with diabetic nephropathy in type 1 diabetes in three cross-sectional studies (Table 5). None of the

TABLE 4
Multistate analyses of SOD1 SNPs

SNP	From normal to PM			From PM to SN			From normal to PM vs. from PM to SN	From normal to SN		
	χ^2	*P* value	HR (95% CI)	χ^2	*P* value	HR (95% CI)	*P* value	χ^2	*P* value	HR (95% CI)
rs9974610	16.48	4.9×10^{-5}	1.56 (1.25–1.92)	0.37	0.55	1.11 (0.79–1.54)	0.087	1.47	0.23	1.22 (0.88–1.69)
rs202446	16.89	3.9×10^{-5}	1.61 (1.28–2.03)	2.49	0.11	1.33 (0.93–1.89)	0.36	4.47	0.035	1.44 (1.03–2.03)
rs17881180	13.70	2.1×10^{-4}	1.93 (1.36–2.73)	6.60	0.01	1.85 (1.16–2.95)	0.89	12.63	3.8×10^{-4}	2.25 (1.44–3.52)
rs17884536	0.15	0.7	1.09 (0.70–1.69)	3.10	0.078	0.47 (0.21–1.09)	0.082	3.52	0.061	0.45 (0.20–1.04)
rs17880998	0.89	0.35	1.15 (0.87–1.51)	0.04	0.84	0.95 (0.59–1.54)	0.51	0.53	0.46	0.84 (0.53–1.33)
rs17878806	0.46	0.5	0.93 (0.78–1.14)	1.06	0.3	0.85 (0.64–1.15)	0.61	4.08	0.043	0.74 (0.55–0.99)
rs17880753	0.42	0.52	1.15 (0.75–1.77)	2.53	0.11	0.51 (0.23–1.17)	0.087	2.22	0.14	0.54 (0.24–1.21)
rs17880196	1.72	0.19	0.88 (0.72–1.06)	1.01	0.32	0.85 (0.62–1.16)	0.86	3.37	0.066	0.75 (0.55–1.02)
rs17880135	10.38	0.0013	1.74 (1.24–2.43)	4.59	0.032	1.65 (1.04–2.60)	0.86	10.51	0.0012	2.07 (1.33–3.22)
rs17881203	1.68	0.2	1.54 (0.80–2.94)	0.33	0.57	1.30 (0.53–3.12)	0.76	1.09	0.3	1.56 (0.68–3.58)
rs204732	16.34	5.3×10^{-5}	1.56 (1.27–1.96)	0.06	0.81	1.04 (0.75–1.45)	0.044	0.90	0.34	1.16 (0.85–1.61)

For details, see legend to Table 2 and online appendix text. Note that this model includes all the covariates in the multivariate model (see RESEARCH DESIGN AND METHODS and online appendix 2) except for the updated indicator for hypertension. *P* values are based on the Wald test. PM, persistent microalbuminuria; SN, severe nephropathy.

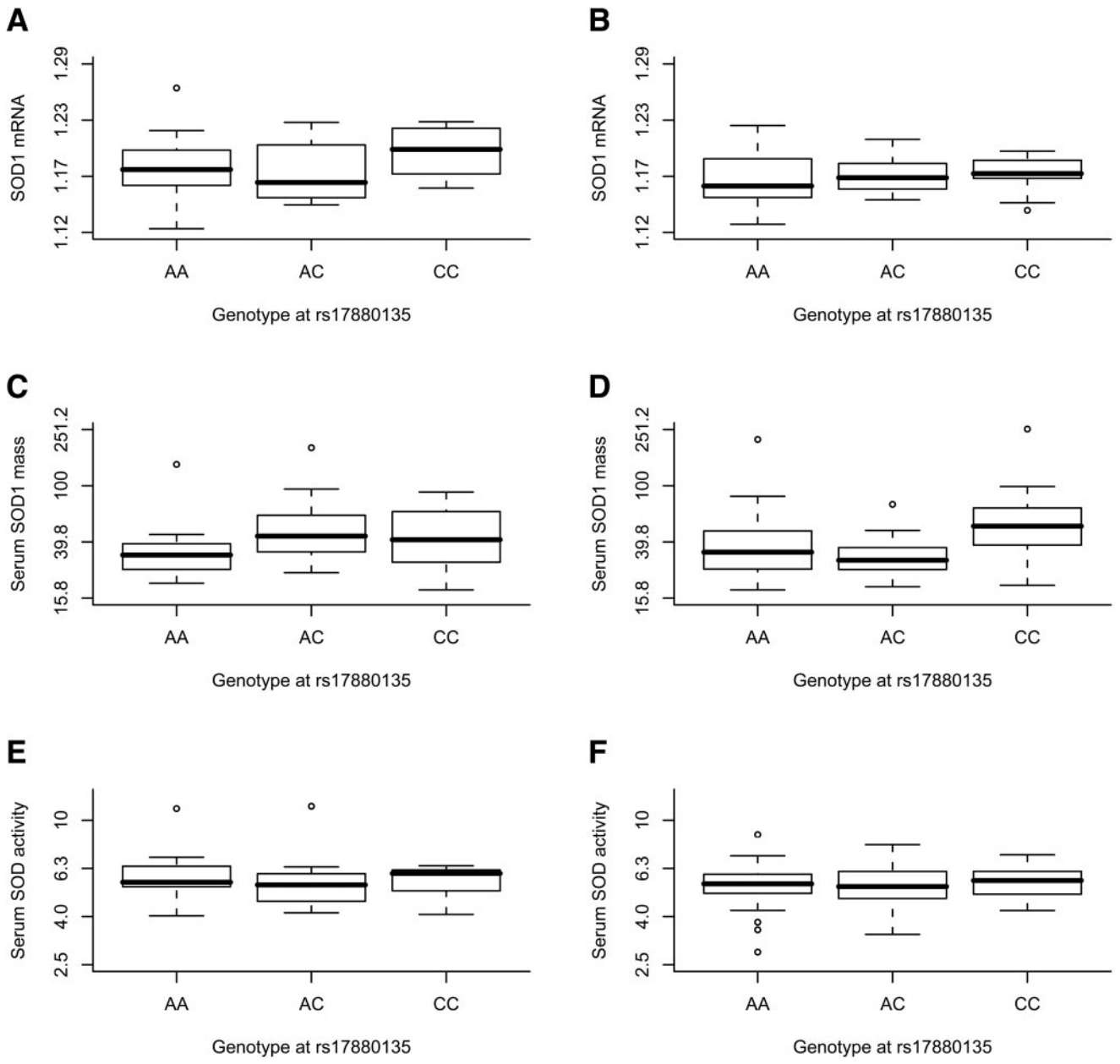

FIG. 1. Log SOD1 gene expression in EBV cells (cycle threshold number compared with b-actin; *A* and *B*), log serum SOD1 mass (ng/ml) (*C* and *D*), and log serum SOD activity (units/ml) (*E* and *F*) by genotype at rs17880135, separately for DCCT/EDIC probands (*A*, *C*, and *E*) and nondiabetic relatives (*B*, *D*, and *F*). The box and whisker plots have lines at the median, boxes at the inter-quartile values, whiskers at the nearest value not beyond 1.5 times the inter-quartile range, and lines beyond the whiskers are outliers. There are no significant differences by genotype for SOD1 expression for the probands ($P = 0.42$) or relatives ($P = 0.60$). For SOD1 mass, there was no difference between those homozygous for the minor allele (C/C) and those homozygous for the major allele (A/A) ($P = 0.23$ and 0.12, probands and relatives, respectively), but there was a borderline result for A/C vs. A/A ($P = 0.08$) in probands, but not relatives ($P = 0.16$). For SOD activity, there were no significant differences by genotype in either relatives or probands ($P > 0.6$). The *P* values reported for mRNA are for multiple linear regression models with additive genotype coding and adjustment for the matching variables sex, age, A1C, and AER or cystatin for the proband and relatives, respectively. The *P* values for SOD1 mass and SOD activity are from unconditional logistic regression with adjustment for the matching variables. The results for the same individuals by genotype at rs17881180, rs202446, rs9974610, and rs204732 are in online appendixes 30 and 31.

SOD1 SNPs were significantly ($P < 0.05$) associated in any of the studies, but the 95% CIs were wide, suggesting modest power to detect association. Although each study had different definitions of diabetic nephropathy (online appendix text), meta-analysis did not detect significant between-study heterogeneity. The pooled 95% CI could not exclude an odds ratio as large as 1.84.

DISCUSSION

SOD1 (Copper-zinc super oxide dismutase) is an antioxidant enzyme present in the cytosol, nuclei, peroxisomes, and inner mitochondrial space. Its primary function is to lower intracellular concentrations of superoxide anion, a reactive oxygen species. There is ample evidence from animal models that oxidative stress is involved in the pathogenesis of diabetic nephropathy (47–49). In a rat model of type 1 diabetes, renal mitochondria produced significantly increased quantities of superoxide anion and showed evidence of oxidative damage (50). Moreover, mRNA expression of SOD1 is increased in the kidneys of diabetic rats and is normalized upon the administration of insulin and return of metabolic control (51). The role of

TABLE 5
Univariate association of SOD1 SNPs in other studies

Study	SNP Genotype	Case subjects			Control subjects			Cochran-Armitage *P*	Odds ratio (95% CI)
		m/m	m/M	M/M	m/m	m/M	M/M		
GoKinD	Rs17800135	1(0.1%)	98(11.9%)	727(88.0%)	2(0.2%)	98(10.9%)	802(88.9%)	0.61	1.08(0.81–1.44)
	rs17881180	1(0.1%)	93(11.2%)	737(88.7%)	1(0.1%)	94(10.3%)	812(89.5%)	0.58	1.09(0.81–1.46)
	rs202446	14(1.7%)	187(22.5%)	629(75.8%)	17(1.8%)	221(24.3%)	670(73.8%)	0.35	0.91(0.75–1.11)
	rs204732	16(1.9%)	211(25.5%)	599(72.5%)	28(3.1%)	221(24.4%)	655(72.5%)	0.62	0.95(0.79–1.15)
Ireland	rs17880135	1(0.4%)	38(14.2%)	228(85.4%)	2(0.5%)	53(12.0%)	386(87.5%)	0.46	1.18(0.77–1.79)
Pittsburgh Epidemiology of Diabetes Complications (cross-sectional analysis)†									
Microalbuminuria	rs17800135	1(1.0%)	9(9.2%)	88(89.8%)	0(0.0%)	29(12.0%)	213(88.0%)	0.85	0.94(0.46–1.89)
Microalbuminuria	rs17801180	0(0.0%)	10(10.1%)	89(89.9%)	0(0.0%)	26(10.6%)	219(89.4%)	0.89	0.95(0.44–2.04)
Severe nephropathy	rs17800135	2(1.8%)	13(11.8%)	95(86.3%)	0(0.0%)	29(12.0%)	213(88.0%)	0.39	1.31(0.71–2.42)
Severe nephropathy	rs17801180	0(0.0%)	13(11.8%)	97(88.2%)	0(0.0%)	26(10.6%)	219(89.4%)	0.74	1.13(0.56–2.29)
Meta-analysis*	rs17880135								1.13(0.69–1.84)
Cross-sectional analysis of DCCT/EDIC									
PM	rs17880135	3(1.0%)	39(12.5%)	271(86.6%)	5(0.5%)	90(9.2%)	879(90.3%)	0.059	1.40(0.98–2.00)
PM	rs202446	13(4.3%)	79(26.0%)	212(69.7%)	14(1.5%)	231(24.5%)	697(74.0%)	0.033	1.31(1.02–1.68)
SN	rs17880135	2(1.7%)	20(17.4%)	93(80.9%)	5(0.5%)	90(9.2%)	879(90.3%)	0.0014	2.05(1.30–3.24)
SN	rs202446	7(6.3%)	29(26.1%)	75(67.6%)	14(1.5%)	231(24.5%)	697(74.0%)	0.023	1.51(1.05–2.17)

The inclusion/exclusion criteria for case and control subjects for each study is provided in the online appendix text. The association of SNP genotypes with diabetic nephropathy was performed using the Cochran-Armitage exact trend test, and association was estimated by odds ratios using logistic regression with additive coding of the number of minor alleles as a covariate. Meta-analysis was conducted by linear regression of the log odds ratio estimates on study indicators, with inverse variance weighting, in fixed and random effects models. M, major allele; m, minor allele; PM, persistent microalbuminuria; SN, severe nephropathy. *Case definition from each study differed. †Overt nephropathy and ESRD were combined from Pittsburgh Epidemiology of Diabetes Complications.

SOD1 in diabetic nephropathy is supported in murine models of both type 1 and type 2 diabetes. When SOD1 is overexpressed, there is a reduction in renal cell injury, including albuminuria (52,53), leading to the conclusion that increases in cellular SOD1 activity attenuate diabetic renal injury. Consistent with this, when SOD1 is knocked out, diabetic renal disease develops faster (54). Figure 2 shows the distribution of SOD1 protein in human kidney, with the strongest staining in podocytes and parietal cells, with minimal expression in mesangial cells. The somatic mutation frequency has been shown to be reduced by overexpression of SOD1 (55). Mutations in SOD1 are present in ~10% of patients with familial amyotrophic lateral sclerosis (56), with most mutations nonsynonymous and genetically dominant (57,58). Although there are multiple transcripts of SOD1, the functional relevance of most is unclear (59–61). There is no known functional effect of the SOD1/SFRS15 SNPs that are significantly associated with persistent microalbuminuria and severe nephropathy, and little is known about SFRS15 (62).

Evidence for association of SOD1 in DCCT/EDIC. We initially tested 1,411 common variations in 212 candidate genes for nephropathy and retinopathy in the DCCT/EDIC probands using time-to-event analysis, including all appropriate covariates. The most striking finding was for rs17800135 in SOD1 with severe nephropathy and persistent microalbuminuria, and we proceeded to fine-map this by DNA sequencing and genotyping seven additional SNPs. We found multiple SOD1 SNPs to be associated with the development of each of persistent microalbuminuria and severe nephropathy, and evidence for differential association of rs204732 with persistent microalbuminuria and severe nephropathy from multistate models. Allele and haplotype frequencies of four SOD1 SNPs vary across global populations (online appendix 20; Fig. 3) and are important information for power calculations for replication studies in other populations. Specifically, the rare (risk) alleles at rs17800135 and rs17881180 were present solely in the European and Indian population samples and also showed variation in frequency within Europe. The argument that the association of SOD1 SNPs with diabetic nephropathy is solely due to population stratification is mitigated by the observed association from family-based analyses of rs202446 with persistent microalbuminuria.

We observed a HR of 2.62 (95% CI 1.64–4.18; $P = 5.6 \times 10^{-5}$) for the association of rs17880135 with time to severe nephropathy. This SNP was chosen by selecting the smallest *P* value after testing 1,411 SNPs for association, leading to upward bias in the magnitude of the genetic effect estimate (i.e., HR) (63,64), a phenomenon known as the Beavis effect or Winner's curse. To obtain more realistic estimates of the effect size expected in a replication sample, we calculated bootstrap estimates based on repeated sample splitting in which approximately two-thirds of the individuals were used to identify the SNP with the smallest *P* value and the HR for that SNP was estimated in the remaining individuals (65–68), yielding a weighted bootstrap estimate of 1.65 for the HR for rs17880135. To detect a HR of 1.65 in a study of similar design to DCCT/EDIC, assuming MAF = 5%, 8% of probands with the outcome, and power of >80%, a study sample size of 4,375 individuals (350 with diabetic nephropathy) would be required for $\alpha = 0.05$ (6,250 with 500 case subjects for $\alpha = 0.01$). This suggests that large sample sizes are needed to achieve replication studies with sufficient power. The point estimates for population attributable risk in for the time-to-event analysis in DCCT/EDIC for rs17881180 with severe nephropathy was 13.8%, whereas for rs202446 with persistent microalbuminuria, it was 14.2% (69).

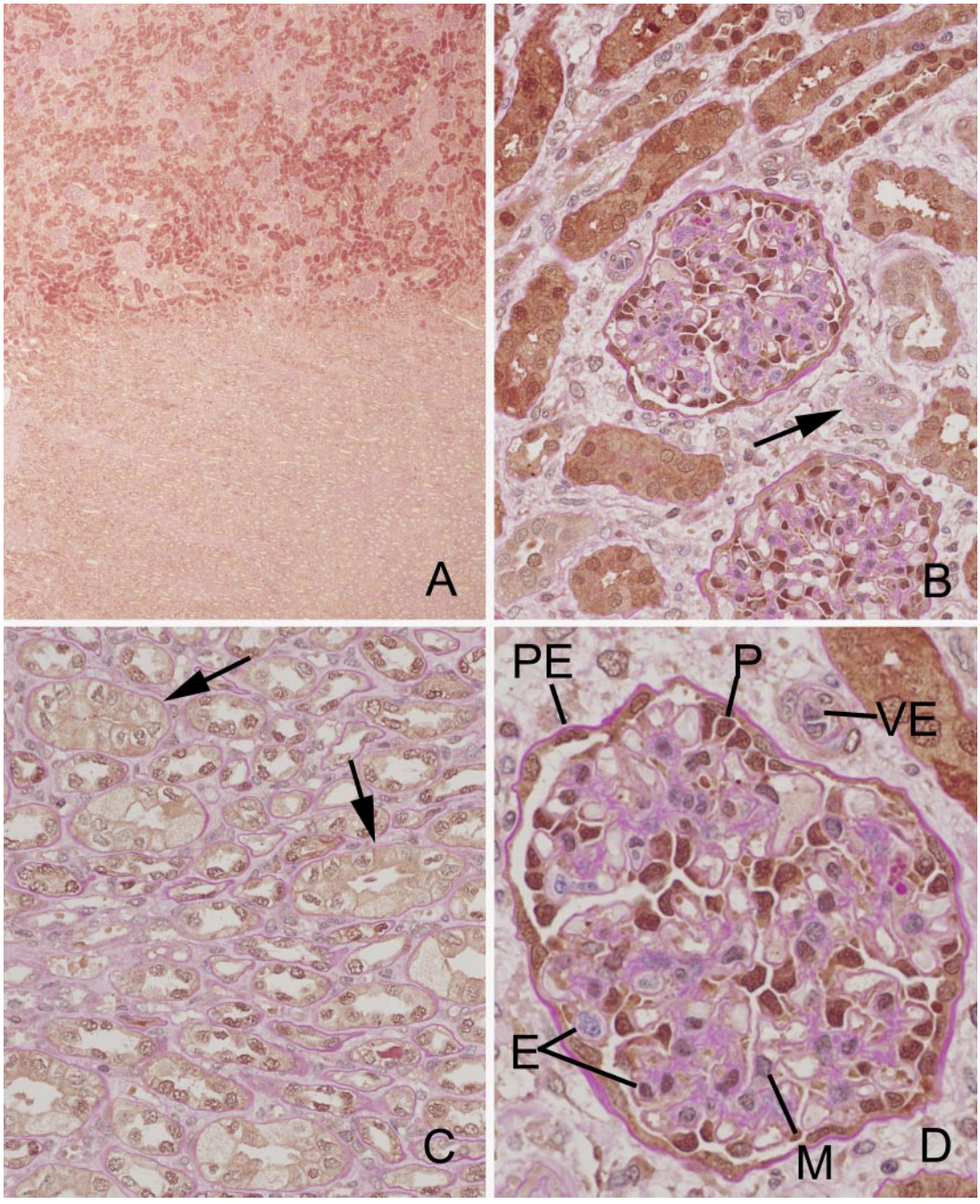

FIG. 2. Localization of SOD1 in normal human kidney by immunohistochemistry. *A*: The expression in cortex (top half) is considerably stronger than in the medulla (bottom half). *B*: In the cortex, SOD1 is expressed in glomeruli and proximal and distal tubules in both cytoplasm and nuclei. Vessels (arrow) show only weak expression in endothelial cells. *C*: In the medulla, there is weak expression in the loops of Henle and in collecting ducts (arrows), both nuclear and cytoplasmic. *D*: Within glomeruli, the strongest expressers of SOD1 are podocytes (P) and parietal epithelial cells (PE), whereas glomerular endothelial cells (E) show variable expression. Mesangial cells (M) are uniformly negative, and vascular endothelial cells (VE) show weak expression. (Please see http://dx.doi.org/10.2337/db07-1059 for a high-quality digital representation of this figure.)

Evidence for association of SOD1 in other studies. Analyses of three cross-sectional studies with a total of 1,301 diabetic nephropathy case subjects and 1,585 control subjects failed to provide supportive evidence for the association of specific SOD1 SNPs with diabetic nephropathy. However, these studies were of modest size and cross-sectional in nature. Furthermore, the case definitions and characteristics of diabetic nephropathy varied considerably among studies. For example, in DCCT/EDIC, only 10% of severe nephropathy case subjects had developed ESRD by EDIC year 8, whereas 57 and 65% of case subjects in the Irish and GoKinD met this criterion, respectively. Therefore, lack of confirmation, assuming that the DCCT/EDIC result is a true positive, may be due to different phenotypes among these studies.

In addition to the Beavis effect and different phenotypes, failure to replicate the SOD1 association in the cross-sectional studies may be in part due to differences in study design. We performed a cross-sectional analysis of DCCT/EDIC based on the repeated AER measures to compare

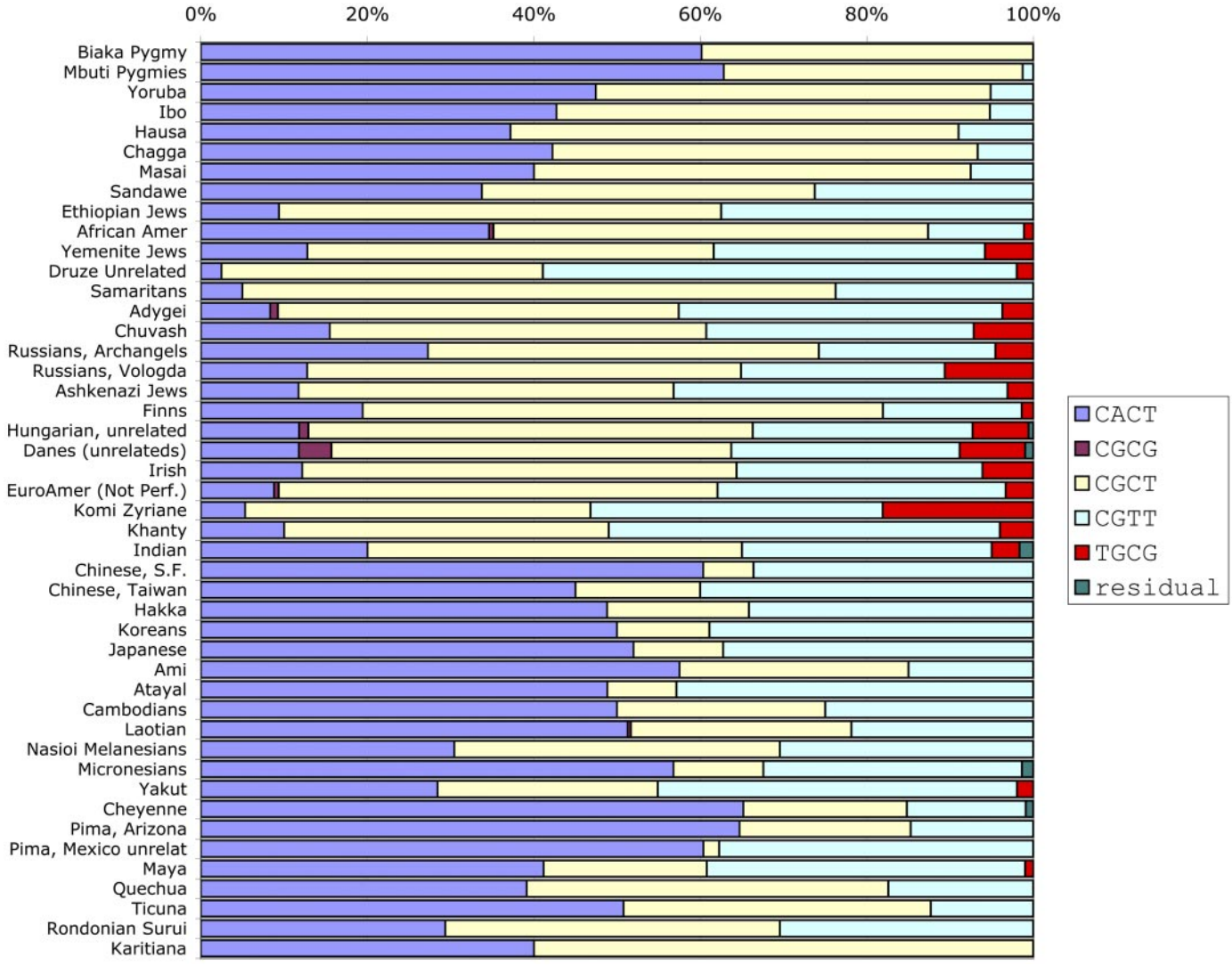

FIG. 3. Distribution of haplotype frequencies across the SOD1 locus in a worldwide study. Stacked bars represent haplotype frequencies for the four SOD1 SNPs in order: rs17881180-rs4998557-rs1041740-rs17880135. Individual marker allele frequencies are provided in online appendix 20. The respective ALFRED site (http://alfred.med.yale.edu/) UIDs are SI003835T, SI003836U, SI003837V, and SI003838W; the four-site haplotype UID is SI006518U. Exact allele frequencies for each site and population and for the haplotypes in each population and descriptions of the populations, specific sample of each, and sample sizes can be found in ALFRED and online appendix 20.

with our primary time-to-event analysis that included covariates (Table 5). We defined persistent microalbuminuria case subjects as ever having two consecutive AERs >30 mg/day during DCCT/EDIC, severe nephropathy to be two consecutive AERs >30 mg/day and one AER >300 mg/day or ESRD, whereas control subjects had to have 15 years of diabetes and never two consecutive AERs >30. In this cross-sectional analysis, the evidence for association between rs17880135 and persistent microalbuminuria was borderline significant ($P = 0.059$), whereas the association between rs17880135 and severe nephropathy ($P = 0.0014$) remained at a weaker level. However, the latter was 2 orders of magnitude less significant than the results of the primary analysis, and a similar pattern of results was observed for rs202446. Thus, even using this generous cross-sectional definition of renal disease (renal measures collected during the whole of DCCT/EDIC) produces a reduction in the evidence for association. Attempts to replicate the DCCT/EDIC associations in cross-sectional studies may have inherent limitations, particularly if the genetic effect operates to modify the time when diabetic nephropathy develops, rather than just influencing susceptibility.

We did not identify functional differences between the associated SOD1 SNPs, but these studies were limited by the low frequency of some genotypes and by the use of available samples. Specifically lymphoblasts were used for expression studies, and SOD activity and SOD1 mass were measured in serum, whereas renal tissue would have been preferred. In conclusion, using tagSNPs, sequencing, fine-mapping, and detailed statistical analysis, we have found that multiple variants in SOD1 are associated with the time to development of persistent microalbuminuria and subsequent progression to severe nephropathy in the phenotypically rich DCCT/EDIC study. The involvement of SOD1 is consistent with the known molecular mechanisms for the development of diabetic complications (49). Although we were unable to observe confirmation in three other studies, we cannot exclude an effect size similar to that obtained from cross-sectional analysis of the DCCT/EDIC data. This indicates that even the combined sample size of these studies may have been insufficient to reliably detect

the SOD1/SFRS15 association, and that studies of additional populations is warranted.

ACKNOWLEDGMENTS

H.A.-K. has received support from Premier's Research Excellence Award, Government of Ontario. L.M. has received a CIHR Research Scholarship. R.S. has received support from the Samuel Lunenfeld Research Institute Interface Training Program: Applying Genomics to Human Health (CIHR). K.K.K. has received grant P01-GM57672 toward typing the global sample from the National Institute of General Medical Sciences. S.B.B. has received a Canadian Institutes of Health Research (CIHR) Senior Investigator award. A.D.P. holds a Canada Research Chair in the Genetics of Complex Diseases. This work has received National Institute of Diabetes and Digestive and Kidney Diseases Contract N01-DK-6-2204, National Institute of Diabetes and Digestive and Kidney Diseases Grant R01-DK-077510, and Juvenile Diabetes Research Foundation Grant 1-2006-757 and support from the Canadian Network of Centres of Excellence in Mathematics. The GoKinD study sample collection has received support from the Juvenile Diabetes Research Foundation in collaboration with the Joslin Diabetes Center and George Washington University and from the U.S. Centers for Disease Control and Prevention.

Clinical data and DNA from the DCCT/EDIC study will be made available through the National Institute of Diabetes and Digestive and Kidney Diseases repository at https://www.niddkrepository.org/niddk/home.do. Clinical data and DNA from the Genetics of Kidneys in Diabetes (GoKinD) study are available at www.gokind.org.

The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Diabetes and Digestive and Kidney Diseases or the National Institutes of Health.

The contributions to the global sample of many researchers from around the world and thousands of individuals who volunteered are gratefully acknowledged.

REFERENCES

1. Seaquist ER, Goetz FC, Rich S, Barbosa J: Familial clustering of diabetic kidney disease: evidence for genetic susceptibility to diabetic nephropathy. *N Engl J Med* 320:1161–1165, 1989
2. Borch-Johnsen K, Norgaard K, Hommel E, Mathiesen ER, Jensen JS, Deckert T, Parving HH: Is diabetic nephropathy an inherited complication? *Kidney Int* 41:719–722, 1992
3. Quinn M, Angelico MC, Warram JH, Krolewski AS: Familial factors determine the development of diabetic nephropathy in patients with IDDM. *Diabetologia* 39:940–945, 1996
4. The Diabetes Control and Complications Trial Research Group: Clustering of long-term complications in families with diabetes in the Diabetes Control and Complications Trial. *Diabetes* 46:1829–1839, 1997
5. Harjutsalo V, Katoh S, Sarti C, Tajima N, Tuomilehto J: Population-based assessment of familial clustering of diabetic nephropathy in type 1 diabetes. *Diabetes* 53:2449–2454, 2004
6. Fioretto P, Steffes MW, Barbosa J, Rich SS, Miller ME, Mauer M: Is diabetic nephropathy inherited? Studies of glomerular structure in type 1 diabetic sibling pairs. *Diabetes* 48:865–869, 1999
7. Moczulski DK, Rogus JJ, Antonellis A, Warram JH, Krolewski AS: Major susceptibility locus for nephropathy in type 1 diabetes on chromosome 3q: results of novel discordant sib-pair analysis. *Diabetes* 47:1164–1169, 1998
8. Imperatore G, Hanson RL, Pettitt DJ, Kobes S, Bennett PH, Knowler WC, Pima Diabetes Genes Group: Sib-pair linkage analysis for susceptibility genes for microvascular complications among Pima Indians with type 2 diabetes. *Diabetes* 47:821–830, 1998
9. Krolewski AS, Poznik GD, Placha G, Canani L, Dunn J, Walker W, Smiles A, Krolewski B, Fogarty DG, Moczulski D, Araki S, Makita Y, Ng DP, Rogus J, Duggirala R, Rich SS, Warram JH: A genome-wide linkage scan for genes controlling variation in urinary albumin excretion in type II diabetes. *Kidney Int* 69:129–136, 2006
10. Vardarli I, Baier LJ, Hanson RL, Akkoyun I, Fischer C, Rohmeiss P, Basci A, Bartram CR, Van Der Woude FJ, Janssen B: Gene for susceptibility to diabetic nephropathy in type 2 diabetes maps to 18q22.3–23. *Kidney Int* 62:2176–2183, 2002
11. Janssen B, Hohenadel D, Brinkkoetter P, Peters V, Rind N, Fischer C, Rychlik I, Cerna M, Romzova M, de Heer E, Baelde H, Bakker SJ, Zirie M, Rondeau E, Mathieson P, Saleem MA, Meyer J, Koppel H, Sauerhoefer S, Bartram CR, Nawroth P, Hammes HP, Yard BA, Zschocke J, van der Woude FJ: Carnosine as a protective factor in diabetic nephropathy: association with a leucine repeat of the carnosinase gene CNDP1. *Diabetes* 54:2320–2327, 2005
12. Tanaka N, Babazono T, Saito S, Sekine A, Tsunoda T, Haneda M, Tanaka Y, Fujioka T, Kaku K, Kawamori R, Kikkawa R, Iwamoto Y, Nakamura Y, Maeda S: Association of solute carrier family 12 (sodium/chloride) member 3 with diabetic nephropathy, identified by genome-wide analyses of single nucleotide polymorphisms. *Diabetes* 52:2848–2853, 2003
13. Shimazaki A, Kawamura Y, Kanazawa A, Sekine A, Saito S, Tsunoda T, Koya D, Babazono T, Tanaka Y, Matsuda M, Kawai K, Iiizumi T, Imanishi M, Shinosaki T, Yanagimoto T, Ikeda M, Omachi S, Kashiwagi A, Kaku K, Iwamoto Y, Kawamori R, Kikkawa R, Nakajima M, Nakamura Y, Maeda S: Genetic variations in the gene encoding ELMO1 are associated with susceptibility to diabetic nephropathy. *Diabetes* 54:1171–1178, 2005
14. McKnight AJ, Maxwell AP, Sawcer S, Compston A, Setakis E, Patterson CC, Brady HR, Savage DA: A genome-wide DNA microsatellite association screen to identify chromosomal regions harboring candidate genes in diabetic nephropathy. *J Am Soc Nephrol* 17:831–836, 2006
15. Ewens KG, George RA, Sharma K, Ziyadeh FN, Spielman RS: Assessment of 115 candidate genes for diabetic nephropathy by transmission/disequilibrium test. *Diabetes* 54:3305–3318, 2005
16. Hanson RL, Craig DW, Millis MP, Yeatts KA, Kobes S, Pearson JV, Lee AM, Knowler WC, Nelson RG, Wolford JK: Identification of PVT1 as a candidate gene for end-stage renal disease in type 2 diabetes using a pooling-based genome-wide single nucleotide polymorphism association study. *Diabetes* 56:975–983, 2007
17. Boright AP, Paterson AD, Mirea L, Bull SB, Mowjoodi A, Scherer SW, Zinman B: Genetic variation at the ACE gene is associated with persistent microalbuminuria and severe nephropathy in type 1 diabetes: the DCCT/EDIC Genetics Study. *Diabetes* 54:1238–1244, 2005
18. Ng DP, Tai BC, Koh D, Tan KW, Chia KS: Angiotensin-I converting enzyme insertion/deletion polymorphism and its association with diabetic nephropathy: a meta-analysis of studies reported between 1994 and 2004 and comprising 14,727 subjects. *Diabetologia* 48:1008–1016, 2005
19. Hallman DM, Huber JC Jr, Gonzalez VH, Klein BE, Klein R, Hanis CL: Familial aggregation of severity of diabetic retinopathy in Mexican Americans from Starr County, Texas. *Diabetes Care* 28:1163–1168, 2005
20. Hallman DM, Boerwinkle E, Gonzalez VH, Klein BE, Klein R, Hanis CL: A genome-wide linkage scan for diabetic retinopathy susceptibility genes in Mexican Americans with type 2 diabetes from Starr County, Texas. *Diabetes* 56:1167–1173, 2007
21. Looker HC, Nelson RG, Chew E, Klein R, Klein BE, Knowler WC, Hanson RL: Genome-wide linkage analyses to identify loci for diabetic retinopathy. *Diabetes* 56:1160–1166, 2007
22. The DCCT Research Group: The Diabetes Control and Complications Trial: design and methodologic considerations for the feasibility phase. *Diabetes* 35:530–545, 1986
23. The Diabetes Control and Complications Trial Research Group: The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 329:977–986, 1993
24. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA: Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 74:106–120, 2004
25. Collins FS, Brooks LD, Chakravarti A: A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res* 8:1229–1231, 1998
26. Barrett JC, Fry B, Maller J, Daly MJ: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265, 2005
27. Gibbs RA, Belmont JW, Hardenbol P, et al.: The International HapMap Project. *Nature* 426:789–796, 2003
28. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P: A haplotype map of the human genome. *Nature* 437:1299–1320, 2005
29. Oliphant A, Barker DL, Stuelpnagel JR, Chee MS: BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques* Suppl:56–58, 60–61, 2002

30. Al-Kateb H, Mirea L, Xie X, Sun L, Liu M, Chen H, Bull SB, Boright AP, Paterson AD: Multiple variants in vascular endothelial growth factor (VEGFA) are risk factors for time to severe retinopathy in type 1 diabetes: the DCCT/EDIC genetics study. *Diabetes* 56:2161–2168, 2007
31. Cox DR: Regression models and life tables (with discussion). *J Royal Stat Soc B* 34:187–220, 1972
32. Heinze G, Schemper M: A solution to the problem of monotone likelihood in Cox regression. *Biometrics* 57:114–119, 2001
33. Heinze G, Ploner M: SAS and SPLUS programs to perform Cox regression without convergence problems. *Comput Methods Programs Biomed* 67: 217–223, 2002
34. Storey JD, Tibshirani R: Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100:9440–9445, 2003
35. Sun L, Craiu RV, Paterson AD, Bull SB: Stratified false discovery control for large-scale hypothesis testing with application to genome-wide association studies. *Genet Epidemiol* 30:519–530, 2006
36. Pritchard JK, Stephens M, Donnelly P: Inference of population structure using multilocus genotype data. *Genetics* 155:945–959, 2000
37. Rosenberg NA, Li LM, Ward R, Pritchard JK: Informativeness of genetic markers for inference of ancestry. *Am J Hum Genet* 73:1402–1422, 2003
38. Therneau TM, Grambsch PM, Fleming TA: Martingale-based residuals for survival models. *Biometrika* 77:147–160, 1990
39. Rabinowitz D, Laird N: A unified approach to adjusting association tests for population admixture with arbitrary pedigree structure and arbitrary missing marker information. *Hum Hered* 50:211–223, 2000
40. Horvath S, Xu X, Laird NM: The family based association test method: strategies for studying general genotype-phenotype associations. *Eur J Hum Genet* 9:301–306, 2001
41. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM: Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 26:61–69, 2004
42. Rakovski CS, Xu X, Lazarus R, Blacker D, Laird NM: A new multimarker test for family-based association studies. *Genet Epidemiol* 31:9–17, 2007
43. Andersen PK, Borgan O, Gill RD, Keiding N: *Statistical Models Based on Counting Processes.* New York, Springer-Verlag, 1993
44. Therneau TM, Grambsch PM: *Modeling Survival Data: Extending the Cox Model.* New York, Springer, 2000
45. Kalbfleisch JD, Prentice RL: *The Statistical Analysis of Failure Time Data.* New York, John Wiley & Sons, 2002
46. Broom WJ, Parton MJ, Vance CA, Russ C, Andersen PM, Hansen V, Leigh PN, Powell JF, Al-Chalabi A, Shaw CE: No association of the SOD1 locus and disease susceptibility or phenotype in sporadic ALS. *Neurology* 63:2419–2422, 2004
47. Evans JL, Goldfine ID, Maddux BA, Grodsky GM: Oxidative stress and stress-activated signaling pathways: a unifying hypothesis of type 2 diabetes. *Endocr Rev* 23:599–622, 2002
48. Koya D, Hayashi K, Kitada M, Kashiwagi A, Kikkawa R, Haneda M: Effects of antioxidants in diabetes-induced oxidative stress in the glomeruli of diabetic rats. *J Am Soc Nephrol* 14:S250–S253, 2003
49. Brownlee M: The pathobiology of diabetic complications: a unifying mechanism. *Diabetes* 54:1615–1625, 2005
50. Rosca MG, Mustata TG, Kinter MT, Ozdemir AM, Kern TS, Szweda LI, Brownlee M, Monnier VM, Weiss MF: Glycation of mitochondrial proteins from diabetic rat kidney is associated with excess superoxide formation. *Am J Physiol Renal Physiol* 289:F420–F430, 2005
51. Sechi LA, Ceriello A, Griffin CA, Catena C, Amstad P, Schambelan M, Bartoli E: Renal antioxidant enzyme mRNA levels are increased in rats with experimental diabetes mellitus. *Diabetologia* 40:23–29, 1997
52. DeRubertis FR, Craven PA, Melhem MF, Salah EM: Attenuation of renal injury in *db/db* mice overexpressing superoxide dismutase: evidence for reduced superoxide-nitric oxide interaction. *Diabetes* 53:762–768, 2004
53. Craven PA, Melhem MF, Phillips SL, DeRubertis FR: Overexpression of Cu2+/Zn2+ superoxide dismutase protects against early diabetic glomerular injury in transgenic mice. *Diabetes* 50:2114–2125, 2001
54. Derubertis FR, Craven PA, Melhem MF: Acceleration of diabetic renal injury in the superoxide dismutase knockout mouse: effects of tempol. *Metabolism* 56:1256–1264, 2007
55. Kunishige M, Hill KA, Riemer AM, Farwell KD, Halangoda A, Heinmoller E, Moore SR, Turner DM, Sommer SS: Mutation frequency is reduced in the cerebellum of Big Blue mice overexpressing a human wild type SOD1 gene. *Mutat Res* 473:139–149, 2001
56. Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng HX, et al.: Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 362:59–62, 1993
57. Gaudette M, Hirano M, Siddique T: Current status of SOD1 mutations in familial amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord* 1:83–89, 2000
58. Selverstone Valentine J, Doucette PA, Zittin Potter S: Copper-zinc superoxide dismutase and amyotrophic lateral sclerosis. *Annu Rev Biochem* 74:563–593, 2005
59. Sherman L, Dafni N, Lieman-Hurwitz J, Groner Y: Nucleotide sequence and expression of human chromosome 21-encoded superoxide dismutase mRNA. *Proc Natl Acad Sci U S A* 80:5465–5469, 1983
60. Kilk A, Laan M, Torp A: Human CuZn superoxide dismutase enzymatic activity in cells is regulated by the length of the mRNA. *FEBS Lett* 362:323–327, 1995
61. Hirano M, Hung WY, Cole N, Azim AC, Deng HX, Siddique T: Multiple transcripts of the human Cu, Zn superoxide dismutase gene. *Biochem Biophys Res Commun* 276:52–56, 2000
62. Hirosawa M, Nagase T, Ishikawa K, Kikuno R, Nomura N, Ohara O: Characterization of cDNA clones selected by the GeneMark analysis from size-fractionated cDNA libraries from human brain. *DNA Res* 6:329–336, 1999
63. Beavis WD: QTL analyses: power, precision, and accuracy. In *Molecular Dissection of Complex Traits.* Paterson AH, Ed. New York, CRC Press, 1998, p. 145–162
64. Goring HH, Terwilliger JD, Blangero J: Large upward bias in estimation of locus-specific effects from genomewide scans. *Am J Hum Genet* 69:1357–1369, 2001
65. Ambroise C, McLachlan GJ: Selection bias in gene extraction on the basis of microarray gene-expression data. *Proc Natl Acad Sci U S A* 99:6562–6566, 2002
66. Sun L, Bull SB: Reduction of selection bias in genomewide studies by resampling. *Genet Epidemiol* 28:352–367, 2005
67. Wu LY, Sun L, Bull SB: Locus-specific heritability estimation via the bootstrap in linkage scans for quantitative trait loci. *Hum Hered* 62:84–96, 2006
68. Efron B, Tibshirani R: Improvements on cross-validation: the 632+ bootstrap method. *J Am Stat Assoc* 92:438–548, 1997
69. Kleinbaum DG, Kupper LL, Morgenstern H: *Epidemiologic Research: Principles and Quantitative Methods*. Belmont, CA, Lifetime Learning Publications, 1982