Molecular Genetics and Metabolism 101 (2010) 383–390

Contents lists available at ScienceDirect

Molecular Genetics and Metabolism

journal homepage: www.elsevier.com/locate/ymgme

ELMO1 variants and susceptibility to diabetic nephropathy in American Indians

Robert L. Hanson [a], Meredith P. Millis [b], Naomi J. Young [b], Sayuko Kobes [a], Robert G. Nelson [a], William C. Knowler [a], Johanna K. DiStefano [b,*]

[a] Diabetes Epidemiology and Clinical Research Section, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 1550 East Indian School Road, Phoenix, AZ 85014, USA

[b] Translational Genomics Research Institute, Diabetes, Cardiovascular and Metabolic Diseases Division, 445 North Fifth Street, Phoenix, AZ 85004, USA

ARTICLE INFO

Article history:
Received 6 July 2010
Received in revised form 12 August 2010
Accepted 12 August 2010
Available online 19 August 2010

Keywords:
Association analysis
Candidate gene study
End stage renal disease
Genetic susceptibility
Pima Indians

ABSTRACT

Variants in the engulfment and cell motility 1 gene, *ELMO1*, have previously been associated with kidney disease attributed to type 2 diabetes. The Pima Indians of Arizona have high rates of diabetic nephropathy, which is strongly dependent on genetic determinants; thus, we sought to investigate the role of *ELMO1* polymorphisms in mediating susceptibility to this disease in this population. Genotype distributions were compared among 141 individuals with nephropathy and 416 individuals without heavy proteinuria in a family study of 257 sibships, and 107 cases with diabetic ESRD and 108 controls with long duration diabetes and no nephropathy. We sequenced 17.4 kb of *ELMO1* and identified 19 variants. We genotyped 12 markers, excluding those in 100% genotypic concordance with other variants or with a minor allele frequency <0.05, plus 21 additional markers showing association with ESRD in earlier studies. In the family study, the strongest evidence for association was with rs1345365 (odds ratio [OR] = 2.42 per copy of A allele [1.35–4.32]; $P = 0.001$) and rs10951509 (OR = 2.42 per copy of A allele [1.31–4.48]; $P = 0.002$), both of which are located in intron 13 and are in strong pairwise linkage disequilibrium ($r^2 = 0.97$). These associations were in the opposite direction from those observed in African Americans, which suggests that the relationship between diabetic kidney disease and *ELMO1* variation may involve as yet undiscovered functional variants or complex interactions with other biological variables.

1. Introduction

Diabetic nephropathy is the most common cause of end-stage renal disease (ESRD) in the United States and underlies a substantial proportion of morbidity and mortality associated with diabetes. Regardless of diabetes type, duration of the disease is one of the strongest determinants of diabetic nephropathy [1–3]. In addition to duration of diabetes, other risk factors include hyperglycemia, hypertension, and hyperlipidemia [4,5]; however, strong evidence supports a role for genetic factors in mediating susceptibility to diabetic nephropathy and it is likely that the combination of environmental exposures and genetic load determine individual risk for development or progression of the disease [6–8].

Variants in the gene encoding the engulfment and cell motility 1 protein (*ELMO1*) are associated with kidney disease attributed to type 2 diabetes in Japanese [9] and African American individuals [10]. Variants in *ELMO1* are also associated with kidney disease in European Americans with type 1 diabetes [11,12]. However, the particular nephropathy-associated variants have differed across populations, and the contribution of this gene to the development of diabetic nephropathy remains unclear. In the present study, we evaluated the role of *ELMO1* variants in diabetic nephropathy in American Indians. This population is at high risk for diabetic nephropathy, and the occurrence of the disease is strongly familial [13,14]. The goals of this study were twofold: 1) to identify additional *ELMO1* polymorphisms and test them for association with kidney disease in American Indian individuals with T2D, and 2) to assess in this population the association between diabetic kidney disease and *ELMO1* variants previously observed in other populations.

2. Materials and methods

2.1. Subjects

All subjects in this analysis are participants in a longitudinal study of type 2 diabetes and its complications conducted in the Gila River Indian Community since 1965 [15]. In this study, community members were invited to biennial examinations, which included a 75 g oral glucose tolerance test, used to diagnose diabetes, and measurement of urinary protein–creatinine and (since 1982) albumin–creatinine ratio. Some of the individuals in the present analysis participated previously in a genome-wide linkage scan for diabetic nephropathy [13]. A description of the study group in whom sequence variants were genotyped has also been published [16]. Genotyping was performed in the family-based

* Corresponding author. Translational Genomics Research Institute, 445 North Fifth Street, Phoenix, AZ 85004, USA. Fax: +1 602 343 8844.
E-mail address: jdistefano@tgen.org (J.K. DiStefano).

1096-7192/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2010.08.014

study of individuals who participated in the original linkage scan, and in a case-control study of diabetic ESRD. The family-based sample was comprised of 141 individuals with nephropathy (i.e., urinary protein-to-creatinine ratio >500 mg/g) and 416 individuals without nephropathy (protein creatinine ratio<500 mg/g) [13]. Among the affected individuals 62% were women, the mean (±SD) age was 51.0±11.5 years and mean duration of diabetes was 17.9±7.9 years, while among the unaffected individuals 65% were women, the mean age was 42.2±11.9 years and mean duration of diabetes was 8.0±7.2 years. The case-control study consisted of 107 cases with diabetic ESRD (68% women, mean age=55.9±8.9 years, mean duration=20.4±7.1 years) and 108 control subjects with diabetes duration >10 years and a maximum urinary albumin–creatinine ratio <300 mg/g observed in the longitudinal study (56% women, mean age=58.8±9.7 years, mean duration=20.7±5.5 years); none of these individuals was a first-degree relative of another in the sample. There were 71 individuals who were part of both the case-control and family-based studies. The study was approved by the Institutional Review Boards of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the Translational Genomics Research Institute. All subjects provided written informed consent prior to participation in the study.

2.2. SNP detection

We sequenced 17.4 kb of *ELMO1*, including all exons, exon–intron boundaries, and 3 kb of the 5′ regulatory region, in 36 Pima Indians (18 ESRD cases, 18 diabetic controls) to identify genetic variants. DNA was amplified in a final reaction volume of 10 μl using 60 ng genomic DNA, 10× standard PCR buffer containing 1.5 mM $MgCl_2$ (Applied Biosystems; Foster City, CA), 800 μM dNTPs, 0.4 μM oligonucleotide primers, and 0.4 U AmpliTaq Gold (Applied Biosystems). PCR cycling conditions consisted of an initial denaturation at 96 °C for 7 min, followed by 35 cycles of 96 °C for 20 s, 57 °C for 30 s, and 72 °C for 45 s, ending with a final elongation step at 72 °C for 5 min. Following amplification, AMPure magnetic bead technology (Beckman Coulter Genomics; Danvers, MA) was used to remove unconsumed dNTPs and oligonucleotide primers. Amplicons were bidirectionally sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems; Foster City, CA) and 35 cycles of 96 °C for 10 s, 55 °C for 10 s, and 60 °C for 4 m and subsequently purified with CleanSEQ magnetic bead technology (Beckman Coulter Genomics). Sequences were generated by the AB3730*xl* DNA Analyzer platform (Applied Biosystems) and analyzed using Mutation Surveyor software v2.61 (SoftGenetics; State College, PA). Information on all sequencing primers is available upon request.

In addition to variants identified by sequencing, we also selected SNPs that showed association with kidney disease in Japanese [9] and African American [10] individuals with T2D. These included 8 SNPs from the study conducted in Japanese individuals: rs1541727, rs4723596, rs11983698, rs4723593, rs7799004, rs1558688, rs3807163, and rs7804092 [9], and 13 SNPs from the study conducted in African American individuals: rs9969311, rs2717972, rs6462740, rs2058730, rs2160430, rs7782590, rs10951509, rs1981740, rs1345365, rs6979467, rs6462733, rs6967682, and rs6462731 [10].

2.3. SNP genotyping

SNPs were genotyped by either allelic discrimination PCR (AD-PCR) in conjunction with the 7000 Sequence Detection System (Applied Biosystems) or the iPLEX assay (Sequenom, Inc.; La Jolla, CA) according to the manufacturers' protocols. SNPs genotyped using AD-PCR included 37457413, rs1731981, rs1882072, rs1882071, rs13242348, rs2717968, rs17170970, rs1420421, 37139247, rs2072522, rs741301 and rs7785934. All remaining markers (i.e., rs1541727, rs9969311, rs2717972, rs6462740, rs2058730, rs2160430, rs7782590, rs10951509, rs1981740, rs1345365, rs6979467, rs6462733, rs6967682, rs6462731, rs4723596, rs11983698, rs4723593, rs7799004, rs1558688, rs3807163, rs7804092) were genotyped using the iPLEX platform. We designed iPLEX primers and multiplex conditions with the Assay Design v3.1 software (Sequenom). For the iPLEX assay, reaction products were dispensed onto a 384-element SpectroCHIP bioarray (Sequenom) using a MassARRAY Nanodispenser, and analyzed by MassARRAY Workstation v4.0 software.

The observed genotype frequency for each SNP was assessed for deviation from that expected under Hardy–Weinberg equilibrium and encrypted samples were used to assess data quality. Assays were considered successful and genotype data subsequently analyzed if 1) a minimum of 90% of all genotyping calls were obtained, 2) markers did not deviate significantly ($P \leq 0.05$) from Hardy–Weinberg equilibrium, and 3) genotyping error results were <3%.

2.4. Power calculations

Power for association studies depends on the sample size, the level of type I error specified, and the genotype frequencies in cases and controls. Genotype frequencies will depend on population frequency of the allele of interest and its effect on disease liability. For the present power calculations, a functional allele with an additive effect on disease liability was assumed (comprised of a mixture of three normal distributions). By 20 years of duration of diabetes, the cumulative incidence of ESRD in Pimas is ~0.15 and the cumulative incidence of macroalbuminuria is ~0.85, thus, the cases and controls are assumed to represent the upper and lower 15% of the liability distribution, respectively [17]. Individuals in the family-based study were estimated by application of the generalized estimating equation approach to simulated data to represent an effective sample size of 133 cases and 395 controls for a dichotomous trait with prevalence of 24% (i.e., the upper 24% and the lower 76% of the liability distribution). With these assumptions, the minimal effect size detectable with 80% power at $p<0.05$ was calculated using the formulas derived by Hanson et al [18] for various values of the minor allele frequency (MAF). The results, shown in Table 1 indicate that the present case-control study is well-powered to detect an association accounting for ~1.5% of the variance in liability, given the selection parameters, this corresponds to an odds ratio of 1.7–3.8 per copy of the risk allele, depending on its frequency. Similarly, the family study is powered to detect an association accounting for ~2.7% of the variance, corresponding to an odds ratio of 1.5–2.6.

2.5. Statistical analysis

The extent to which observed genotype frequencies for each SNP deviated from that expected under Hardy–Weinberg equilibrium (HWE) was assessed (χ^2 with 1 df); none of the markers varied significantly from HWE. The statistical evidence for association and the strength of the association between genotype and affection status,

Table 1
Minimal detectable effect size at $P<0.05$ with 80% power for case-control and family association studies.

MAF	Case-control study		Family study	
	OR	h^2	OR	h^2
0.05	3.80	0.017	2.57	0.030
0.10	2.54	0.016	1.96	0.029
0.20	1.99	0.015	1.65	0.028
0.30	1.82	0.015	1.55	0.027
0.40	1.75	0.015	1.50	0.027
0.50	1.73	0.015	1.49	0.027

OR represents the odds ratio per copy of the risk allele and h^2 is the proportion of variance in liability to disease explained by the variant.