Published in final edited form as:

Hum Genet. 2009 August ; 126(2): 265–275. doi:10.1007/s00439-009-0667-0.

The Influence of Carnosinase Gene Polymorphisms on Diabetic Nephropathy Risk in African Americans

Caitrin W. McDonough^{1,2,4}, Pamela J. Hicks^{1,2,3}, Lingyi Lu⁶, Carl D. Langefeld⁶, Barry I. Freedman⁵, and Donald W. Bowden^{1,2,3,5}

¹Center for Diabetes Research, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157; USA

²Center for Human Genomics, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157; USA

³Department of Biochemistry, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157; USA

⁴Graduate Program in Molecular Medicine, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157; USA

⁵Department of Internal Medicine, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157; USA

⁶Public Health Sciences, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157; USA

Abstract

Four genome wide linkage scans for diabetic nephropathy have mapped susceptibility loci to chromosome 18q22.3-23 in the region of the carnosinase genes, *CNDP1* and *CNDP2*. *CNDP1* has been associated with diabetic nephropathy in Europeans and European Americans, but not African Americans. Individuals homozygous for a five tri-nucleotide repeat allele (5L; D18S880) are protected from diabetic nephropathy. We identified 64 variants after sequencing the exons, promoter, and 3' UTR of *CNDP1* and *CNDP2* in African American and European American DNA samples. After scanning 44 of these variants, extensive genotyping of 12 SNPs and D18S880 was performed in 1025 African American cases with type 2 diabetes (DM)-associated end-stage renal disease (ESRD) and 1064 African American non-diabetic non-nephropathy controls to assess whether the carnosinase genes influence risk for DM-ESRD in African Americans. Evidence of association with DM-ESRD was seen with 2 SNPs: rs6566810 and rs4892247; 3 two-marker haplotypes: rs6566810 and rs17089362, rs17089362 and rs890336, and rs890334 and rs12717111 (global empirical $p=0.0034$, 0.0275 , and 0.0002 respectively) and 3 three-marker haplotypes: rs6566810, rs17089362, and rs890336; rs890335, rs890334, and rs12717111; and rs890334, rs12717111, and D18S880 (global empirical $p=0.0074$, $1.5E-05$, and 0.0032 respectively). The risk haplotypes (rs6566810, rs17089362 [A,T] and rs6566810, rs17089362, rs890336 [A,T,C]) were most strongly associated with DM-ESRD among African Americans in the non 5L-5L group. Variants in the carnosinase genes appear to contribute to diabetic nephropathy susceptibility in African Americans. Protection from diabetic nephropathy afforded by 5L-5L homozygosity in *CNDP1* may be masked by the effects of additional risk haplotypes in *CNDP1* and *CNDP2*.

Keywords

African Americans; carnosinase; End Stage Renal Disease; DNA polymorphisms; type 2 diabetes; nephropathy

Introduction

As the number of individuals with type 2 diabetes mellitus (DM) has more than doubled in the past twenty-five years, the number with diabetic complications has also grown. In the United States, DM is the leading cause of end-stage renal disease (ESRD), accounting for 44.8% of incident cases (U.S. Renal Data System 2008). Familial clustering of diabetes-associated ESRD suggests a genetic component exists (Freedman et al. 1995; Quinn et al. 1996; Seaquist et al. 1989). Genome-wide linkage scans for diabetic nephropathy (DN) have provided consistent evidence for a DN susceptibility locus at chromosome 18q22.3-23 in Turkish and Pima Indian families (Vardarli et al. 2002), 18q21.1-23 in African American families (Bowden et al. 2004), and 18q22.3 in European American and American Indian families (Iyenger et al. 2007).

The carnosinase genes (carnosinase dipeptidase 1, *CNDP1*, and carnosinase 2, *CNDP2*) are located on chromosome 18 at 18q22.3. *CNDP1* encodes a secreted serum carnosinase that degrades carnosine specifically, while *CNDP2* encodes a nonspecific dipeptidase (Teufel et al. 2003). Carnosine (β -alanyl-L-histidine) has been shown to play a role as a free oxygen radical scavenger, having anti-glycating effects, inhibiting formation of advanced glycation end products (AGEs) (Hipkiss et al. 1998; Boldyrev 2000), and has been reported to influence glucose metabolism (Sauerhofer et al. 2007). D18S880, a microsatellite marker encoding a tri-nucleotide repeat in the signal peptide sequence of exon 2 of *CNDP1*, has previously been shown to be associated with DN in Europeans (Janssen et al. 2005) and European Americans (EA) (Freedman et al. 2007). These consistent observations in European-derived populations suggest a significant role for *CNDP1* with risk for DN, but there have not been detailed studies in other ethnic groups, specifically African Americans. This is of interest since we previously reported evidence for linkage to DN on chromosome 18q in African American (AA) families (Bowden et al. 2004), suggesting a possible role for *CNDP1* in AA DN.

Materials and Methods

Subjects

This study was conducted under Institutional Review Board approval from Wake Forest University School of Medicine and adhered to the tenets of the Declaration of Helsinki. Identification, clinical characteristics, and recruitment of AA and EA patients and controls have been previously described (Yu et al. 1996). Briefly, 1025 unrelated AA patients with type 2 DM were recruited from dialysis facilities. Individuals with a history of ketoacidosis, or who developed diabetes prior to the age of 25 years and received continuous insulin therapy since diagnosis were assumed to have type 1 diabetes and were excluded. Type 2 DM was diagnosed in African Americans who reported developing DM after the age of 25 years and who did not receive only insulin therapy since diagnosis. Cases had DM diagnosed at least 5 years before initiating renal replacement therapy, background or greater diabetic retinopathy, and/or ≥ 100 mg/dl proteinuria on urinalysis in the absence of other causes of nephropathy. Another 1064 unrelated AA controls and 39 unrelated EA controls without a current diagnosis of diabetes or renal disease were recruited from the community and internal medicine clinics. All DM-ESRD cases and non-diabetic controls were born in North Carolina, South Carolina, Georgia, Tennessee, or Virginia. DNA extraction was performed using the PureGene system (Gentra Systems, Minneapolis, MN). DNA was also obtained from 44 Yoruba Nigerians from the

National Institute of General Medical Sciences (NIGMS) Human Variation Collection (Coriell Cell Repositories, Camden, NJ).

D18S880 genotyping

The tri-nucleotide repeat polymorphism, D18S880, in exon 2 of *CNDP1* was genotyped by fragment length analysis on an ABI Prism DNA Analyzer 3700 (Applied Biosystems Incorporated, Foster City, CA) in a manner similar to that described by Janssen *et al.* (Janssen *et al.* 2005). Fragment length was determined using ABI Prism GeneMapper software v3.0 (Applied Biosystems, Inc.). 56 duplicate samples were run for molecular quality control purposes. Of the replicate pairs, 100% were concordant.

Sequence analysis and genotyping

The exons of *CNDP2* and *CNDP1* were amplified using intronic primers; 1000bp of the promoter regions and 3' UTRs of *CNDP2* and *CNDP1* were also amplified. Primer sequences are available on request. All products were sequenced in 3 AAs and 3 EAs with two copies of the 5L allele, and in 3 AAs and 3 EAs with one or no copies of the 5L allele using BigDye Terminator v. 1.1 Cycle Sequencing Kits (Applied Biosystems Inc., Foster City, CA) and an Applied Biosystems 3730xl DNA Analyzer (Applied Biosystems Inc., Foster City, CA). Sequences were viewed and compared using Sequencher software version 4.6 (Gene Codes Corporation, Ann Arbor, MI). SNPs identified from sequencing were genotyped using the MassARRAY system from Sequenom, Inc. (Sequenom, San Diego, CA) (Buetow *et al.* 2001) in 380 AA case subjects and 364 AA control subjects. SNPs *CNDP2* rs6566810, rs17089362, rs890336, rs15165, rs890335, rs890334, rs12717111, rs12717112; *CNDP1* rs4892247, rs11659237, rs4891564, and rs2887 demonstrated initial evidence of association in these 380 cases and 364 controls and were then genotyped in the full set of 1025 AA DM-ESRD cases and 1064 AA non DM-ESRD controls. Primer sequences are available on request. Genotyping success rates were >92.1% in case subjects and >92.5 % in control subjects. Concordance among blind duplicate samples was 100%. SNPs rs6566810, rs17089362, rs890335, 890334, and rs12717111 were sequenced in 26 AA controls and 22 AA DM-ESRD cases to verify the accuracy of the genotype calls. The region was PCR amplified and sequenced using BigDye Terminator v. 1.1 Cycle Sequencing Kits (Applied Biosystems Inc, Foster City, CA) and an Applied Biosystems 3730xl DNA Analyzer (Applied Biosystems Inc, Foster City, CA). Sequences were viewed and compared using Sequencher software version 4.6 (Gene Codes Corporation, Ann Arbor, MI). Primer sequences are available on request.

Genotyping for admixture analyses

Seventy Ancestral Informative Markers (AIMs) were genotyped in 1025 AA case subjects, 1064 AA control subjects, 44 Yoruba Nigerians, and 39 EAs using Illumina's Custom Genotyping Service or the MassARRAY system (Sequenom, San Diego, CA) (Buetow *et al.* 2001).

Statistical analyses

The percentage of individuals homozygous for the protective 5 leucine repeat at D18S880 (5L-5L, *CNDP1* Mannheim) was determined in all subjects. The allelic and genotypic frequencies were compared between the case and control groups using the statistical analysis program CLUMP (Sham and Curtis 1995). CLUMP uses a Monte Carlo approach to assess significant departure of observed values in a contingency table from the expected values.

SNPs were tested for departure from Hardy Weinberg proportions using an exact test of HWE proportions for the combined group of cases and controls, and then for cases only, and controls only (Wigginton *et al.* 2005). Those SNPs out of HWE were noted, but still included for the

genotypic analysis. Haplotype block structure was established using Haploview 4.1 (Barrett et al. 2005), defining blocks using the method from Gabriel *et al.* (Gabriel et al. 2002).

Unadjusted measures of linkage disequilibrium (LD) and association were assessed using the software SNPGWA (Harley et al. 2008; <http://www.phs.wfubmc.edu>). SNPGWA computes LD statistics, D' and r^2 , for each pair of tandem SNPs. SNPGWA also performs multiple tests of association including the overall 2-degree of freedom test (genotype), dominant model, recessive model, additive model (Cochran-Armitage trend test), and the corresponding lack of fit to the additive model. Odds ratios, 95% confidence intervals, and p-values were computed for each model of association. Haplotypic association was also calculated for sets of tandem SNPs using the expectation-maximization (EM) algorithm. Tests were first performed using 1,000 permutations. Permutations were increased to 10,000-1,000,000 where association tests indicated possible significance (empirical p-value < 0.10). Individual two-marker and three-marker haplotype frequencies, p-values, odds ratios, and 95% confidence intervals were computed using the program Dandelion (<http://www.phs.wfubmc.edu>). Dandelion is unable to calculate odds ratios when the haplotype frequency is less than or equal to 0.001 in cases and/or controls. SNPs rs12717112 and rs4891564 were removed from haplotype analysis due to inconsistency with HWE in the combined population and low MAF respectively.

Ancestral allele frequencies were estimated from the results of the AIMs genotyped in the Yoruba Nigerians and the European Americans. Individual ancestral proportions were generated for each subject using FRAPPE (Tang et al. 2005), an EM algorithm, under a two-population model. The individual ancestral estimates were used as covariates in the association analyses with the program SNPADDMIX, a module of SNPGWA. SNPADDMIX uses logistic regression tests of additive, dominant, recessive, and two degree of freedom genetic models in order to include adjustments for individual estimates of African ancestry.

D18S880 was converted to a biallelic marker for analysis in Haploview, SNPGWA, SNPADDMIX, and Dandelion. The following alleles were classified as the major allele: 4L, and 5L; while, any other allele (6L, 7L, and 8L) was classified as the minor allele. The study population was also stratified by their D18S880 genotype: subjects with the 4L-4L, 4L-5L, and 5L-5L genotypes were in the 5L-5L only group; whereas, subjects with one or no copies of the 5L allele were in the non 5L-5L group.

Results

All 12 exons of the *CNDP2* gene and all 12 exons of the *CNDP1* gene were sequenced using intronic primers in 6 AA and 6 EA DNAs enriched for the 5L-5L genotype. Additionally, 1000 base pairs (bp) of the promoter region, and the 3'UTR were sequenced in the same individuals. Sixty four variants were identified: 13 coding and 51 non-coding. Results are shown in Table 1.

Forty four variants identified from sequencing and the previously associated marker D18S880 were initially tested for association in 380 AA DM-ESRD cases and 364 AA non-diabetic non-nephropathy controls (Table 2). There were two *CNDP2* SNPs that were monomorphic in AAs: 1358 G>A (R399K), and rs3206931 and four *CNDP1* SNPs that were monomorphic in AAs: IVS4-122 G>A, rs4891562, IVS9+5 G>A, and IVS9+63 G>A. Four *CNDP2* SNPs showed nominal evidence of association under the 2 degree of freedom (2 df) test (p-values 0.002 to 0.074; Table 2). There was also evidence of haplotypic association across *CNDP2*. Six 2-marker haplotypes and eight 3-marker haplotypes were nominally associated (p-values 0.001 to 0.069; Table 2). In *CNDP1*, there was evidence of association with three SNPs (p-values 0.028 to 0.064; Table 2) and nominal evidence of association with three 2-marker haplotypes

and one 3-marker haplotype (p-values 0.057 to 0.077; Table 2). The associations observed in *CNDP1* were at the 3' end of the gene spanning intron 9 to the 3' UTR.

CNDP1 and *CNDP2* polymorphisms demonstrating nominal evidence of association with DM-ESRD from this initial analysis in 380 DM-ESRD cases and 364 controls were carried forward for detailed analysis in a larger African American sample composed of 1025 AA cases with DM-ESRD and 1064 AA non-diabetic non-nephropathy controls. Characteristics of cases and controls are summarized in Table 3. Cases were more often female (62%) than controls (54%). Although the control group was younger, controls were older (mean age 51.4 years) than the age at DM diagnosis in DM-ESRD cases (mean age at DM 41.1 years). The estimated mean proportion of African ancestry was 0.823 in cases and 0.797 in controls.

Table 4 contains allele and genotype frequencies at the D18S880 marker. In addition to the previously reported alleles for D18S880 (5L, 6L, and 7L by Janssen *et al.* and Freedman *et al.*) a 4L repeat and an 8L repeat were infrequently observed in AAs. As previously reported, significant differences in allele frequency were not seen in AA cases with DM-ESRD versus controls (p=0.641; Table 4A). No significant difference in frequency of the 5L-5L genotype was detected between groups (p=0.842; Table 4B).

In addition to D18S880, 12 SNPs were successfully genotyped in the 1025 AA DM-ESRD cases and 1064 AA controls. Two of these SNPs (rs15165 and rs12717112) were inconsistent with Hardy Weinberg Equilibrium (HWE) proportions (p=0.0046 and 0.0049, respectively) in cases and SNP rs6566810 was inconsistent in controls (p=0.0055) (Supplementary Table 1). In the combined population, rs12717112 was inconsistent with expected HWE proportions (p=0.0008; Supplementary Table 1). Minor allele frequencies (MAF), and genotype frequencies and counts for each SNP are shown in Supplementary Table 1. One SNP, rs4891564 had MAFs below 5% in both cases and controls (Supplementary Table 1).

Admixture-adjusted genotypic results for SNPs associated with DM-ESRD under the 2 df test are shown in Supplementary Table 2. Evidence for association with DM-ESRD was observed with one *CNDP2* SNP rs6566810 and one *CNDP1* SNP rs4892247 (p=0.0049 and 0.0345 respectively; Supplementary Table 2). Results are further detailed in Table 5. SNP rs6566810 was associated with DM-ESRD in the entire population (p=0.0015 recessive, odds ratio [OR] (95% CI) 2.29 (1.37-3.81)); as was rs4892247 (p=0.0146 recessive, OR (95%CI) 0.73 (0.57-0.94)). SNP rs6566810 was out of HWE in the control population (p=0.0055; Supplementary Table 1). The association tests performed are robust to HWE departures, given correct genotyping. In order to ensure that this slight departure was not due to an error in genotyping, genotype calls were confirmed by direct DNA sequencing of 26 controls and 22 cases. Genotypes from DNA sequencing were 100% concordant with those from the Sequenom MassArray system.

When the population was stratified by genotype at the D18S880 marker, there was no evidence of association with rs6566810 and DM-ESRD in the 5L-5L group (p=0.1156, 2 df test; Table 5). However, rs6566810 remained associated in the non 5L-5L group (p=0.0496, 2 df test) and was associated in the recessive model (p=0.0152, OR (95% CI) 2.14 (1.16-3.96); Table 5). SNP rs4892247, was trending to association in the 5L-5L group (p=0.0587, 2 df test) and was significantly associated under the dominant model (p=0.0181, OR (95% CI) 0.70 (0.53-0.94)). In the non 5L-5L group, rs4892247 was associated with DM-ESRD (p=0.0485 2 df test; and p=0.0166 recessive, OR (95% CI) 0.69 (0.50-0.93); Table 5).

Haplotype analyses of all two-marker and three-marker haplotypes are shown in Supplementary Table 3. In all subjects, there were three two-marker haplotypes that were significantly associated with DM-ESRD: rs6566810 and rs17089362, rs17089362 and rs890336, and rs890334 and rs12717111 (global empirical p=0.0034, 0.0275, and 0.0002

respectively; Table 6). Additionally, there were three three-marker haplotypes that were significantly associated with DM-ESRD in all subjects: rs6566810, rs17089362, and rs890336; rs890335, rs890334, and rs12717111; and rs890334, rs12717111, and D18S880 (global empirical $p=0.0074$, $1.5E-05$, and 0.0032 respectively; Table 6). In the non 5L-5L homozygotes, two two-marker haplotypes were associated: rs6566810 and rs17089362, and rs890334 and rs12717111, (global empirical $p=0.0424$, and 0.0071 respectively; Table 6), and two three-marker haplotypes were associated: rs890335, rs890334, and rs12717111; and rs890334, rs12717111, and D18S880 (global empirical $p=0.0012$, and 0.0222 respectively; Table 6). In contrast, when limited to 5L-5L homozygotes no significant evidence of haplotypic association was seen (Table 6).

The individual haplotype frequencies for all subjects, the 5L-5L only group, and the non 5L-5L group are shown in Table 7. For two of the associated two-marker haplotypes (rs6566810, rs17089362; and rs890334, rs12717111) and all of the associated three-marker haplotypes the most associated individual haplotypes were rare, occurring in 1% or less of the total population. In order to confirm that these haplotypes did exist and were not due to genotyping error, SNPs rs6566810, rs17089362, rs890335, rs890334, and rs12717111 were sequenced in 26 controls and 22 cases. All genotypes obtained from DNA sequencing exactly matched those from the Sequenom MassArray system.

In all subjects, two-marker haplotypes rs6566810, rs17089362 [A,T] and rs17089362, rs890336 [T,C] were associated with risk (p -values, OR (95% CI) = 0.008 , 3.21 (1.67 - 6.17) and 0.038 , 1.37 (1.11 - 1.68) respectively; Table 7A). Three-marker haplotype rs6566810, rs17089362, rs890335 [A,T,C]; was also associated with risk (p -values, OR (95% CI) = 0.0004 , 9.11 (3.12 - 26.59); Table 7A). However, two-marker haplotype rs890334, rs12717111 [A,A] and three-marker haplotypes rs890335, rs890334, rs12717111 [A,A,A] and rs890334, rs12717111, D18S880 [A,A,5L-5L] were protective (p -values, OR (95% CI) = 0.006 , 0.16 (0.05 - 0.47); 0.002 , N/A and 0.0155 , N/A respectively; Table 7A).

Upon stratifying the population by D18S880 genotype, only one risk haplotype remained associated in the 5L-5L only group (rs6566810, rs17089362, rs890336 [A,T,C] p -value, OR (95% CI) = 0.041 , 4.93 (1.44 - 16.79); Table 7B). While the protective haplotypes were not associated, they all trended towards association (p -values = 0.084 - 0.092 ; Table 7B). In the non 5L-5L group two risk haplotypes remained associated: rs6566810, rs17089362 [A,T] and rs6566810, rs17089362, rs890336 [A,T,C] (p -values, OR (95% CI) = 0.021 4.07 (1.61 - 10.30), and 0.004 N/A respectively, Table 7C). There was also evidence of association with the three protective haplotypes in the non 5L-5L group: rs890334, rs12717111 [A,A]; rs890335, rs890334, rs12717111 [A,A,A]; and rs890334, rs12717111, D18S880 [A,A,5L-5L] (p -values, OR (95% CI) = 0.028 , 0.17 (0.05 - 0.63); 0.011 , N/A; 0.039 , N/A respectively, Table 7C).

There was evidence of high linkage disequilibrium (LD) in both carnosinase genes (Supplementary Table 3). Four SNPs in *CNPD2* were in high LD with pairwise D' values of 0.994 ($r^2=0.345$) between rs15165 and rs890335, $D'=0.996$ ($r^2=0.980$) between rs890335 and rs890334, and $D'=0.946$ ($r^2=0.888$) between rs890334 and rs12717111 (Supplementary Table 3). Similarly, three SNPs in *CNDPI* were in high LD with pairwise D' values of 0.935 ($r^2=0.419$) between SNPs rs4892247 and rs11659237, and $D'=0.997$ ($r^2=0.868$) between rs11659237 and rs2887 (Supplementary Table 3). D18S880, in exon 2 of *CNDPI*, was in low LD with the two nearest markers rs12717111 ($D'=0.235$; $r^2=0.006$) and rs4892247 ($D'=0.054$; $r^2=0.003$) (Supplementary Table 3).

Discussion

Genome wide scans for linkage have demonstrated that a gene(s) contributing to DN maps to chromosome 18q22.3-23 in Turkish, Pima Indian, EA, and American Indian families (Vardarli et al. 2002; Iyengar et al 2007) and 18q21.1-23 in AA families (Bowden et al. 2004). The carnosinase genes, *CNDP1* and *CNDP2*, reside on chromosome 18 at 18q22.3. *CNDP1* was associated with DN in a small European population (Janssen et al. 2005) and a larger EA population (Freedman et al. 2007). Janssen *et al.* found the shortest allelic form observed in European-derived populations with 5 leucine repeats (5L; the *CNDP1* Mannheim allele) significantly more often in diabetic controls without DN, compared to cases with DN. The observed association was only seen in those individuals homozygous for the *CNDP1* Mannheim allele. The 5L-5L genotype was also associated with lower serum carnosinase concentrations (Janssen et al. 2005). Freedman *et al.* observed that the 5L-5L genotype was significantly more frequent in non-diabetic controls than subjects with DN-ESRD. The number of leucine repeats in this signal peptide was associated with secretion of serum carnosinase in Cos-7 transfected cells. There was a significant increase in the level of secreted carnosinase in cells containing constructs with six, seven, and eight leucine repeats compared to those with four and five leucine repeats (Riedel et al. 2007). Recently, Krolewski *et al.* claimed to have excluded the carnosinase genes and D18S880 as a cause of type 1 DN (Wanic et al. 2008). While association was not observed with D18S880 frequencies between three study groups (diabetic subjects with normoalbuminuria, diabetic cases with proteinuria, and diabetic cases with ESRD), there was modest association observed in the promoter region of *CNDP1* when comparing normoalbuminuric controls versus ESRD cases (3 promoter SNPs had $p=0.01$) (Wanic et al. 2008).

We further investigated *CNDP1* and the adjacent *CNDP2* gene by genotyping 12 additional SNPs in 1025 AA type 2 DM-ESRD cases and 1064 AA non-diabetic controls. Evidence of association was detected with two SNPs under the 2 df test in all subjects. SNP rs6566810, which was associated with risk, remained associated in the non 5L-5L group but not in the 5L-5L only group. SNP rs4892247, which was associated with protection against DM-ESRD, also remained associated in the non 5L-5L group and was trending towards association in the 5L-5L only group under the 2 df test. Genotypic associations were seen after adjusting for African ancestry.

Significant evidence of association was also seen with three two-marker haplotypes and three three-marker haplotypes; however, most of the associated haplotypes were rare. We saw association with both risk and protective haplotypes. Three risk haplotypes; rs6566810, rs17089362 [A,T]; rs17089362, rs890336 [T,C] and rs6566810, rs17089362, rs890336 [A,T,C] were associated in all subjects. After splitting the population by 5L-5L status, two of the risk haplotypes (rs6566810, rs17089362 [A,T] and rs6566810, rs17089362, rs890336 [A,T,C]) remained associated in the non 5L-5L group; whereas only one risk haplotype remained associated in the 5L-5L only group (rs6566810, rs17089362, rs890336 [A,T,C]). There were also three protective haplotypes associated in all subjects: rs890334, rs12717111 [A,A]; rs890335, rs890334, rs12717111 [A,A,A]; and rs890334, rs12717111, D18S880 [A,A, 5L-5L]. All three of these protective haplotypes remained associated in the non 5L-5L group, and were trending towards association in the 5L-5L only group.

D18S880 was not associated with DM-ESRD in our African-American population under both allelic and genotypic models. However, when the population was stratified by number of leucine repeats at D18S880, the risk allele and haplotypes often remain associated in the non 5L-5L group, and not in the 5L-5L only group. With the protective allele and haplotypes there is still evidence of association in the non 5L-5L group, and weaker evidence of association (or a trend) in the 5L-5L only group. This suggests that the risk haplotypes may be masked by the

protection afforded by the 5L-5L genotype and that additional SNPs and haplotypes could be protective. However, as previously reported, D18S880 is not contained in any haplotype block (Freedman et al. 2007). In addition, D18S880 was not in LD with any other marker genotyped in our population (Supplementary Table 3). The identification of risk haplotypes for DM-ESRD that may counteract the protection afforded by 5L-5L homozygosity is a potentially important observation. This finding could partially explain why African American 5L-5L homozygotes do not appear to be protected from DM-ESRD, in contrast to Europeans and EAs.

Our study design prevents us from distinguishing whether the observed associations were with type 2 DM, diabetic nephropathy, or both. However, the high rates of nephropathy in the diabetic AA population make it difficult to recruit large numbers of diabetic subjects lacking nephropathy after 10 years diabetes duration. Recruitment of a large African American type 2 diabetes non-nephropathy control group will be an important step in confirming that our results are associated with susceptibility to diabetic nephropathy and not diabetes *per se*. The previous association with D18S880 in Europeans and EAs was seen with DN and not diabetes (Janssen et al. 2005; Freedman et al. 2007). *CNDP1*, which encodes a serum secreted dipeptidase (Teufel et al. 2003), is present in human kidney and expressed at significantly greater levels in podocytes and renal parietal epithelial cells in subjects with DN (Janssen et al. 2005). *CNDP2*, which encodes a non-specific dipeptidase, is ubiquitously expressed. However, there is stronger expression in human kidney and liver (Teufel et al. 2003). Therefore, variants in *CNDP2* may also play a role in DN.

This study represents a comprehensive evaluation of both carnosinase genes, *CNDP1* and *CNDP2*, in a large African American population with type 2 diabetes associated ESRD. While we did not detect association with D18S880, we did find association with variants in *CNDP2* and other variants in *CNDP1* with DM-ESRD. It was necessary to split the AA study group into potentially protected 5L-5L homozygotes and less protected non 5L-5L homozygotes in order to detect additional SNPs and haplotypes that may confer risk for diabetic nephropathy in this ethnic group. We conclude that both *CNDP1* and *CNDP2* play roles in DN susceptibility in the high risk African American population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by NIH grants RO1 DK066358 (DWB), RO1 DK053591 (DWB), RO1 DK070941 (BIF), and in part by the General Clinical Research Center of the Wake Forest University School of Medicine grant M01 RR007122.

References

- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263–5. [PubMed: 15297300]
- Boldyrev AA. Problems and perspectives in studying the biological role of carnosine. *Biochemistry (Mosc)* 2000;65:751–6. [PubMed: 10951091]
- Bowden DW, Colicigno CJ, Langefeld CD, Sale MM, Williams A, Anderson PJ, Rich SS, Freedman BI. A genome scan for diabetic nephropathy in African Americans. *Kidney Int* 2004;66:1517–26. [PubMed: 15458446]
- Buetow KH, Edmonson M, MacDonald R, Clifford R, Yip P, Kelley J, Little DP, Strausberg R, Koester H, Cantor CR, Braun A. High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser

desorption/ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci U S A* 2001;98:581–4. [PubMed: 11136232]

- Freedman BI, Hicks PJ, Sale MM, Pierson ED, Langefeld CD, Rich SS, Xu J, McDonough C, Janssen B, Yard B, van der Woude FJ, Bowden DW. A leucine repeat in the carnosinase gene CNDP1 is associated with diabetic end-stage renal disease in European Americans. *Nephrol Dial Transplant* 2007;22:1131–5. [PubMed: 17205963]
- Freedman BI, Tuttle AB, Spray BJ. Familial predisposition to nephropathy in African-Americans with non-insulin-dependent diabetes mellitus. *Am J Kidney Dis* 1995;25:710–3. [PubMed: 7747724]
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science* 2002;296:2225–9. [PubMed: 12029063]
- Harley JB, Alarcón-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL, Tsao BP, Vyse TJ, Langefeld CD, Nath SK, Guthridge JM, Cobb BL, Mirel DB, Marion MC, Williams AH, Divers J, Wang W, Frank SG, Namjou B, Gabriel SB, Lee AT, Gregersen PK, Behrens TW, Taylor KE, Fernando M, Zidovetzki R, Gaffney PM, Edberg JC, Rioux JD, Ojwang JO, James JA, Merrill JT, Gilkeson GS, Seldin MF, Yin H, Baechler EC, Li QZ, Wakeland EK, Bruner GR, Kaufman KM, Kelly JA. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci. *Nat Genet* 2008;40:204–10. [PubMed: 18204446]
- Hipkiss AR, Preston JE, Himsworth DT, Worthington VC, Keown M, Michaelis J, Lawrence J, Mateen A, Allende L, Eagles PA, Abbott NJ. Pluripotent protective effects of carnosine, a naturally occurring dipeptide. *Ann N Y Acad Sci* 1998;854:37–53. [PubMed: 9928418]
- Iyengar SK, Abboud HE, Goddard KA, Saad MF, Adler SG, Arar NH, Bowden DW, Duggirala R, Elston RC, Hanson RL, Ipp E, Kao WH, Kimmel PL, Klag MJ, Knowler WC, Meoni LA, Nelson RG, Nicholas SB, Pahl MV, Parekh RS, Quade SR, Rich SS, Rotter JI, Scavini M, Schelling JR, Sedor JR, Sehgal AR, Shah VO, Smith MW, Taylor KD, Winkler CA, Zager PG, Freedman BI. Genome-wide scans for diabetic nephropathy and albuminuria in multiethnic populations: the family investigation of nephropathy and diabetes (FIND). *Diabetes* 2007;56:1577–85. [PubMed: 17363742]
- 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, Zirje M, Rondeau E, Mathieson P, Saleem MA, Meyer J, Köppel 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* 2005;54:2320–7. [PubMed: 16046297]
- Quinn M, Angelico MC, Warram JH, Krolewski AS. Familial factors determine the development of diabetic nephropathy in patients with IDDM. *Diabetologia* 1996;39:940–5. [PubMed: 8858216]
- Riedl E, Koeppl H, Brinkkoetter P, Sternik P, Steinbeisser H, Sauerhoefer S, Janssen B, van der Woude FJ, Yard BA. A CTG polymorphism in the CNDP1 gene determines the secretion of serum carnosinase in Cos-7 transfected cells. *Diabetes* 2007;56:2410–3. [PubMed: 17601991]
- Sauerhöfer S, Yuan G, Braun GS, Deinzer M, Neumaier M, Gretz N, Floege J, Kriz W, van der Woude F, Moeller MJ. L-carnosine, a substrate of carnosinase-1, influences glucose metabolism. *Diabetes* 2007;56:2425–32. [PubMed: 17601992]
- Sequist ER, Goetz FC, Rich S, Barbosa J. Familial clustering of diabetic kidney disease. Evidence for genetic susceptibility to diabetic nephropathy. *N Engl J Med* 1989;320:1161–5. [PubMed: 2710189]
- Sham PC, Curtis D. Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Ann Hum Genet* 1995;59:97–105. [PubMed: 7762987]
- Tang H, Peng J, Wang P, Risch NJ. Estimation of individual admixture: analytical and study design considerations. *Genet Epidemiol* 2005;28:289–301. [PubMed: 15712363]
- Teufel M, Saudek V, Ledig JP, Bernhardt A, Boularand S, Carreau A, Cairns NJ, Carter C, Cowley DJ, Duverger D, Ganzhorn AJ, Guenet C, Heintzelmann B, Laucher V, Sauvage C, Smirnova T. Sequence identification and characterization of human carnosinase and a closely related non-specific dipeptidase. *J Biol Chem* 2003;278:6521–31. [PubMed: 12473676]
- US Renal Data System. USRDS 2008 Annual Data Report: Atlas of End-Stage Renal Disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; Bethesda, MD: 2008.

- 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* 2002;62:2176–83. [PubMed: 12427143]
- Wanic K, Placha G, Dunn J, Smiles A, Warram JH, Krolewski AS. Exclusion of polymorphisms in carnosinase genes (CNDP1 and CNDP2) as cause of diabetic nephropathy in type 1 diabetes mellitus: results of large case - control and follow - up studies. *Diabetes* 2008;57:2547–51. [PubMed: 18753673]
- Wigginton JE, Cutler DJ, Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 2005;76:887–93. [PubMed: 15789306]
- Yu H, Bowden DW, Spray BJ, Rich SS, Freedman BI. Linkage analysis between loci in the renin-angiotensin axis and end-stage renal disease in African Americans. *J Am Soc Nephrol* 1996;7:2559–64. [PubMed: 8989734]

Table 1

Variants identified from sequencing the exons, promoter regions, and 3'UTRs of the Camosinase genes: CNDP1 (Table 1A) and CNDP2 (Table 1B)

A.

Chr. 18 Position	Marker	Gene	Location	Change	AA Frequency	CA Frequency
70313555	-1,022 (TAAA)8>(TAAA)9	CNDP2	Promoter	TAAA Repeat	4 of 5 (0.80)	4 of 6 (0.67)
70313676	rs11873377	CNDP2	Promoter	C → T	6 of 10 (0.60)	12 of 12 (1.0)
70314094	-483 G>A	CNDP2	Promoter	G → A	1 of 10 (0.10)	0 of 12 (0.00)
70314140	rs7237740	CNDP2	Promoter	C → T	4 of 12 (0.33)	9 of 12 (0.75)
70314224	rs58558076	CNDP2	Promoter	C → T	4 of 12 (0.33)	0 of 12 (0.00)
70314390	rs12964619	CNDP2	Promoter	G → A	1 of 12 (0.08)	2 of 12 (0.17)
70314439	rs6566810	CNDP2	Promoter	T → A	1 of 12 (0.08)	1 of 12 (0.08)
70314572	rs6566811	CNDP2	Promoter	G → A	4 of 12 (0.33)	8 of 12 (0.67)
70318104	rs3764509	CNDP2	5' UTR	C → G	7 of 12 (0.58)	0 of 12 (0.00)
70319588	rs2303463 (P35P)	CNDP2	Exon 3	C → T	2 of 12 (0.17)	0 of 12 (0.00)
70321162	IVS3+1,475 A>T	CNDP2	Intron 3	A → T	1 of 12 (0.08)	0 of 12 (0.00)
70327063	rs2278161 (Y126H)	CNDP2	Exon 5	A → G	1 of 12 (0.08)	0 of 12 (0.00)
70329110	rs34916325 (R180Q)	CNDP2	Exon 6	G → A	4 of 12 (0.33)	0 of 12 (0.00)
70329141	rs2278159 (Y190Y)	CNDP2	Exon 6	A → G	1 of 12 (0.08)	0 of 12 (0.00)
70329180	rs2278158	CNDP2	Intron 6	C → A	1 of 12 (0.08)	0 of 12 (0.00)
70331703	rs2241511	CNDP2	Intron 7	T → C	6 of 12 (0.50)	8 of 12 (0.67)
70331719	rs2241510	CNDP2	Intron 7	C → T	2 of 12 (0.17)	0 of 12 (0.00)
70331723	rs8090744	CNDP2	Intron 7	G → A	5 of 12 (0.42)	0 of 12 (0.00)
70331951	IVS8+17 G>A	CNDP2	Intron 8	G → A	2 of 12 (0.17)	0 of 12 (0.00)
70334481	1104 G>A (L314L)	CNDP2	Exon 9	G → A	1 of 12 (0.08)	0 of 12 (0.00)
70336618	rs17089362	CNDP2	Intron 9	G → T	2 of 6 (0.33)	0 of 2 (0.00)
70336841	1358 G>A (R399K)	CNDP2	Exon 10	G → A	2 of 12 (0.17)	2 of 12 (0.17)
70336952	rs890337	CNDP2	Intron 10	T → C	1 of 4 (0.25)	0 of 0 (n/a)
70337384	rs2241508	CNDP2	Intron 11	T → C	3 of 12 (0.25)	12 of 12 (1.0)
70338592	rs890336	CNDP2	3' UTR	G → A	3 of 10 (0.30)	12 of 12 (1.0)
70339208	rs3206931	CNDP2	3' UTR	C → T	2 of 10 (0.20)	2 of 10 (0.20)
70339231	rs15165	CNDP2	3' UTR	G → A	5 of 10 (0.50)	10 of 10 (1.0)
70339232	rs890335	CNDP2	3' UTR	C → T	5 of 10 (0.50)	0 of 10 (0.00)

A.

Chr. 18 Position	Marker	Gene	Location	Change	AA Frequency	CA Frequency
70339327	rs7577	CNDP2	3' UTR	A → G	0 of 10 (0.00)	2 of 10 (0.20)
70339351	rs890334	CNDP2	Downstream	C → T	5 of 10 (0.50)	0 of 10 (0.00)
70339436	+100 A>—	CNDP2	Downstream	A Deletion	3 of 6 (0.50)	6 of 6 (1.0)
70340230	rs12717111	CNDP2	Downstream	A → C	5 of 12 (0.42)	0 of 10 (0.00)
70340290	rs12717112	CNDP2	Downstream	C → T	5 of 12 (0.42)	0 of 10 (0.00)

B.

Chr. 18 Position	Marker	Gene	Location	Change	AA Frequency	CA Frequency
70351839	rs2346062	CNDP1	Promoter	A → G	6 of 12 (0.50)	12 of 12 (1.0)
70351839	rs2346063	CNDP1	Promoter	C → G	4 of 12 (0.33)	10 of 12 (0.83)
70351856	rs2081548	CNDP1	Promoter	C → T	5 of 12 (0.42)	0 of 12 (0.00)
70352752	80 T>C	CNDP1	5' UTR	T → C	1 of 12 (0.08)	0 of 12 (0.00)
70352927	IVS1+21 ->A	CNDP1	Intron 1	A insertion	1 of 6 (0.17)	0 of 6 (0.00)
70377435	rs12605490	CNDP1	Intron 2	A → T	7 of 12 (0.58)	10 of 12 (0.83)
70377484	rs12605520	CNDP1	Intron 2	C → T	1 of 12 (0.08)	6 of 12 (0.50)
70377519	rs12607796	CNDP1	Intron 2	T → C	7 of 12 (0.58)	10 of 12 (0.83)
70377558	rs9807352 (A58A)	CNDP1	Exon 3	C → A	2 of 12 (0.17)	1 of 12 (0.08)
70377561	rs11538622 (I59I)	CNDP1	Exon 3	C → A	1 of 12 (0.08)	0 of 12 (0.00)
70378877	rs9955580	CNDP1	Intron 3	C → T	1 of 12 (0.08)	0 of 12 (0.00)
70378949	IVS4-122 G>A	CNDP1	Intron 3	G → A	5 of 12 (0.42)	6 of 12 (0.50)
70378995	IVS4-76 C>T	CNDP1	Intron 3	C → T	1 of 12 (0.08)	0 of 12 (0.00)
70379104	rs4263028 (V113I)	CNDP1	Exon 4	A → G	11 of 12 (0.92)	12 of 12 (1.0)
70379249	rs4329999	CNDP1	Intron 4	A → G	10 of 10 (1.0)	9 of 12 (0.75)
70380147	rs9319909	CNDP1	Intron 4	C → G	3 of 12 (0.25)	2 of 12 (0.17)
70380192	rs9967490	CNDP1	Intron 4	A → G	6 of 12 (0.50)	8 of 12 (0.67)
70385317	rs4891562	CNDP1	Intron 5	C → A	6 of 12 (0.50)	6 of 12 (0.50)
70385615	rs12960862 (Y241Y)	CNDP1	Exon 6	T → C	6 of 12 (0.50)	6 of 12 (0.50)
70389452	1019 C>A (H270N)	CNDP1	Exon 7	C → A	3 of 12 (0.25)	1 of 12 (0.08)
70395188	rs62621182 (D316N)	CNDP1	Exon 8	G → A	3 of 12 (0.25)	0 of 8 (0.00)
70396547	IVS9+5 G>A	CNDP1	Intron 9	G → A	3 of 12 (0.25)	0 of 12 (0.00)

A.

Chr. 18 Position	Marker	Gene	Location	Change	AA Frequency	CA Frequency
70396605	IVS9+63 G>A	CNDP1	Intron 9	G → A	3 of 12 (0.25)	0 of 12 (0.00)
70398258	rs4892247	CNDP1	Intron 9	T → C	2 of 8 (0.25)	6 of 12 (0.50)
70398426	1459 C>T (H416H)	CNDP1	Exon 10	C → T	1 of 10 (0.10)	0 of 10 (0.00)
70401742	rs11659237	CNDP1	Intron 10	C → T	6 of 12 (0.50)	5 of 12 (0.42)
70402912	1869 T>C	CNDP1	3' UTR	T → C	1 of 12 (0.08)	0 of 12 (0.00)
70403066	rs4891564	CNDP1	3' UTR	C → A	1 of 12 (0.08)	1 of 12 (0.08)
70403117	rs2887	CNDP1	3' UTR	C → T	4 of 12 (0.33)	6 of 12 (0.50)
70403419	+180 T>—	CNDP1	Downstream	T Deletion	6 of 6 (1.0)	6 of 6 (1.0)
70403785	+546 —>A(A)	CNDP1	Downstream	A(A) insertion	6 of 6 (1.0)	5 of 6 (0.83)

Frequencies are shown as number out of chromosomes with data for SNPs and number out of samples with data for insertions/deletions.

Table 2

Initial genotyping results in 380 AA DM-ESRD cases and 364 AA controls

Marker	Gene	Location	Cases MAF	Controls MAF	2 Degrees of Freedom P-value	2-marker Haplotype P-value	3-marker Haplotype P-value
rs11873377	CNDP2	Promoter	0.244	0.282	0.218	0.007	0.053
rs7237740	CNDP2	Promoter	0.305	0.339	0.407	0.074	0.087
rs58558076	CNDP2	Promoter	0.208	0.236	0.315	0.369	0.006
rs12964619	CNDP2	Promoter	0.149	0.134	0.426	0.001	0.016
rs6566810	CNDP2	Promoter	0.255	0.181	0.002	0.013	0.011
rs3764509	CNDP2	5' UTR	0.506	0.483	0.460	0.331	0.317
rs2303463	CNDP2	Exon 3	0.240	0.224	0.866	0.544	0.624
rs2278161	CNDP2	Exon 5	0.100	0.099	0.873	0.435	0.451
rs34916325	CNDP2	Exon 6	0.068	0.087	0.068	0.255	0.148
rs2241510	CNDP2	Intron 7	0.122	0.104	0.317*	0.093	0.184
rs8090744	CNDP2	Intron 7	0.168	0.210	0.099	0.086	0.012
IVS8+17 G>A	CNDP2	Intron 8	0.014	0.021	0.319	0.143	0.230
rs17089362	CNDP2	Intron 9	0.090	0.114	0.018	0.448	0.576
1358 G>A	CNDP2	Exon 10	0.000	0.000	–	–	–
rs890337	CNDP2	Intron 10	0.242	0.230	0.318	0.519	0.602
rs2241508	CNDP2	Intron 11	0.276	0.252	0.346	0.758	0.638
rs890336	CNDP2	3' UTR	0.295	0.304	0.626	0.364	0.274
rs3206931	CNDP2	3' UTR	0.000	0.000	–	–	–
rs15165	CNDP2	3' UTR	0.298	0.342	0.017	0.242	0.403
rs890335	CNDP2	3' UTR	0.134	0.161	0.372	0.413	0.638
rs7577	CNDP2	3' UTR	0.093	0.086	0.724	0.543	0.754
rs890334	CNDP2	Downstream	0.136	0.164	0.155	0.360	0.069
rs12717111	CNDP2	Downstream	0.141	0.172	0.338	0.044	0.044
rs12717112	CNDP2	Downstream	0.138	0.160	0.180	0.041	0.034
rs2346063	CNDP1	Promoter	0.355	0.328	0.245	0.525	0.889
rs2081548	CNDP1	Promoter	0.148	0.156	0.137	0.940	0.656
D18S880	CNDP1	Exon 2	0.399	0.385	0.780	0.396	0.148
rs12605490	CNDP1	Intron 2	0.430	0.386	0.301	0.191	0.118

Marker	Gene	Location	Cases MAF	Controls MAF	2 Degrees of Freedom P-value	2-marker Haplotype P-value	3-marker Haplotype P-value
rs12605520	CNDP1	Intron 2	0.167	0.155	0.612	0.323	0.438
rs12607796	CNDP1	Intron 2	0.431	0.401	0.413	0.450	0.476
rs9807352	CNDP1	Exon 3	0.147	0.130	0.445	0.508	0.751
IVS4-122 G>A	CNDP1	Intron 3	0.000	0.000	–	–	–
rs4329999	CNDP1	Exon 4	0.084	0.075	0.798	0.793	0.442
rs9319909	CNDP1	Intron 4	0.316	0.327	0.905	0.343	0.481
rs9967490	CNDP1	Intron 4	0.433	0.459	0.685	0.594	0.720
rs4891562	CNDP1	Intron 5	0.000	0.000	–	–	–
rs12960862	CNDP1	Exon 6	0.328	0.355	0.784	0.769	0.759
1019 C>A	CNDP1	Exon 7	0.207	0.208	0.992	0.514	0.207
rs62621182	CNDP1	Exon 8	0.158	0.129	0.386	0.057	0.059
IVS9+5 G>A	CNDP1	Exon 9	0.000	0.000	–	–	–
IVS9+63 G>A	CNDP1	Exon 9	0.000	0.000	–	–	–
rs4892247	CNDP1	Intron 9	0.354	0.421	0.028	0.077	0.150
rs11659237	CNDP1	Intron 10	0.460	0.401	0.039	0.075	0.178
rs4891564	CNDP1	3' UTR	0.030	0.042	0.392	0.091	–
rs2887	CNDP1	3' UTR	0.491	0.437	0.064	–	–

SNPs highlighted in bold were carried forward in the study. MAF=Minor Allele Frequency. SNPs with a MAF less than 0.05 are denoted in bold, italics. P-values less than 0.08 are also denoted in bold italics.

* Denotes dominant genetic model.

Table 3

Demographic characteristics of the whole African American cohort

	DM-ESRD Cases N=1025		Controls N=1064	
	N [^]	Mean (± SD) or %	N [^]	Mean (± SD) or %
% Female (N)	1025	62% (634)	1062	54% (577)
BMI	989	29.6 (± 6.9)	782	29.6 (± 7.3)
Age at Exam (yrs.)	979	61.7 (± 10.2)	784	51.4 (± 11.4)
Age at T2DM Onset (yrs.)	982	41.1 (± 12.2)	–	–
Age at ESRD Onset (yrs.)	1013	58.2 (± 10.6)	–	–

N[^] Denotes number with data available

Table 4
The allele frequencies (Table 4A) and genotype frequencies (Table 4B) of D18S880 in the African American cohort

A.	Group (N)	Allele Frequencies			
		4L	5L	6L	8L
	Healthy Controls (1048)	0.0239	0.5997	0.3669	0.0086
	DM-ESRD (1013)	0.0276	0.5923	0.3707	0.0089
B.	Group (N)	Genotype Frequencies			
		4L,5L/5L	All Others		
	Healthy Controls (1048)	0.3903	0.6097		
	DM-ESRD (1013)	0.3860	0.6140		

Table 5
Genotypic association for rs6566810 and rs4892247 after adjustment for individual admixture

Marker	Genotypic Association							
	2 DF		Dominant		Additive		Recessive	
	P-value		P-value	OR (95% CI)	P-value	OR (95% CI)	P-value	OR (95% CI)
rs6566810	All	0.0049	0.1281	1.16 (0.96-1.39)	0.0165	1.22 (1.04-1.43)	0.0015	2.29 (1.37-3.81)
	5L-5L Only	0.1156	0.1927	1.22 (0.90-1.66)	0.0812	1.27 (0.97-1.66)	0.0595	2.45 (0.96-6.21)
	Non 5L-5L	0.0496	0.3308	1.12 (0.89-1.43)	0.0915	1.19 (0.97-1.46)	0.0152	2.14 (1.16-3.96)
rs4892247	All	0.0345	0.1009	0.86 (0.72-1.03)	0.0160	0.86 (0.75-0.97)	0.0146	0.73 (0.57-0.94)
	5L-5L Only	0.0587	0.0181	0.70 (0.53-0.94)	0.0261	0.78 (0.63-0.97)	0.3111	0.80 (0.52-1.23)
	Non 5L-5L	0.0485	0.7508	0.96 (0.76-1.21)	0.1405	0.89 (0.76-1.04)	0.0166	0.69 (0.50-0.93)

Results are shown in All subjects: 1025 DM-ESRD cases, 1064 controls; 5L-5L Only subjects: 390 DM-ESRD cases and 408 controls; and Non 5L-5L subjects: 623 DM-ESRD cases and 640 controls
P-values less than 0.05 are denoted in bold, italics. OR = odds ratio, 95% CI = 95% confidence interval

Table 6Haplotypic association across *CNDP1* and *CNDP2*

	All	5L-5L Only	Non 5L-5L
	Empirical P-value	Empirical P-value	Empirical P-value
Two-Marker			
rs6566810, rs17089362	<i>0.0034</i>	0.1205	<i>0.0424</i>
rs17089362, rs890336	<i>0.0275</i>	0.2370	0.1191
rs890334, rs12717111	<i>0.0002</i>	0.0585	<i>0.0071</i>
Three-Marker			
rs6566810, rs17089362, rs890336	<i>0.0074</i>	0.2059	0.0838
rs890335, rs890334, rs12717111	<i>0.000015</i>	0.0656	<i>0.0012</i>
rs890334, rs12717111, D18S880	<i>0.0032</i>	0.0578	<i>0.0222</i>

Results are shown in All subjects: 1025 DM-ESRD cases, 1064 controls; 5L-5L Only subjects: 390 DM-ESRD cases and 408 controls; and Non 5L-5L subjects: 623 DM-ESRD cases and 640 controls

P-values less than 0.05 are denoted in bold, italics

Table 7

Frequencies of associated haplotypes in All subjects (Table 7A), 5L-5L Only subjects (Table 7B), and Non 5L-5L subjects (Table 7C)

A.									
		Empirical P-value	Haplotypes	Frequency Cases	Frequency Controls	P-value	Odds Ratio (95%CI)		
Two-Marker									
rs6566810, rs17089362	T,G	0.0034	T,G	0.681	0.720	0.047	0.83 (0.72-0.94)		
	A,T		A,T	0.018	0.005	0.008	3.21 (1.67-6.17)		
rs17089362, rs890336	T,C	0.0275	T,C	0.107	0.081	0.038	1.37 (1.11-1.68)		
rs890334, rs12717111	A,A	0.0002	A,A	0.002	0.012	0.006	0.16 (0.05-0.47)		
Three-Marker									
rs6566810, rs17089362, rs890336	T,G,T	0.00074	T,G,T	0.523	0.556	0.133	0.88 (0.78-0.99)		
	A,T,C		A,T,C	0.015	0.001	0.0004	9.11 (3.12-26.59)		
rs890335, rs890334, rs12717111	A,A,A	0.000015	A,A,A	0.000	0.009	0.002	N/A		
	A,A,C		A,A,C	0.155	0.129	0.095	1.23 (1.04-1.47)		
rs890334, rs12717111, D18S880	A,A,5L-5L	0.0032	A,A,5L-5L	0.0008	0.008	0.0155	N/A		
	A,C,5L-5L		A,C,5L-5L	0.114	0.095	0.155	1.23 (1.00-1.50)		
B.									
		Empirical P-value	Haplotypes	Frequency Cases	Frequency Controls	P-value	Odds Ratio (95%CI)		
Two-Marker									
rs6566810, rs17089362	T,G	0.1205	T,G	0.665	0.716	0.12	0.79 (0.64-0.97)		
	A,T		A,T	0.020	0.008	0.157	2.37 (0.97-5.78)		
rs17089362, rs890336	T,C	0.237	T,C	0.141	0.108	0.149	1.36 (1.01-1.84)		
rs890334, rs12717111	A,A	0.0585	A,A	0.001	0.011	0.092	N/A		
Three-Marker									

A.		Empirical P-value	Haplotypes	Frequency Cases	Frequency Controls	P-value	Odds Ratio (95%CI)
rs6566810, rs17089362, rs890336		0.2059	T,G,T A,T,C	0.514 0.018	0.537 0.003	0.508 0.041	0.91 (0.75-1.11) 4.93 (1.44-16.79)
rs890335, rs890334, rs12717111		0.0656	A,A,A A,A,C	0.000 0.180	0.008 0.157	0.084 0.371	N/A 1.18 (0.91-1.54)
rs890334, rs12717111, D18S880		0.0578	A,A,5L-5L A,C,5L-5L	0.001 0.181	0.011 0.157	0.089 0.372	N/A 1.18 (0.91-1.54)
C.							
		Empirical P-value	Haplotypes	Frequency Cases	Frequency Controls	P-value	Odds Ratio (95%CI)
Two-Marker							
rs6566810, rs17089362		0.0424	T,G A,T	0.691 0.017	0.724 0.004	0.196 0.021	0.85 (0.72-1.01) 4.07 (1.61-10.30)
rs17089362, rs890336		0.1191	T,C	0.085	0.063	0.136	1.38 (1.02-1.86)
rs890334, rs12717111		0.0071	A,A	0.002	0.012	0.028	0.17 (0.05-0.63)
Three-Marker							
rs6566810, rs17089362, rs890336		0.0838	T,G,T A,T,C	0.529 0.014	0.569 0.0003	0.151 0.004	0.85 (0.73-0.99) N/A
rs890335, rs890334, rs12717111		0.0012	A,A,A A,A,C	0.000 0.139	0.010 0.112	0.011 0.157	N/A 1.27 (1.00-1.61)
rs890334, rs12717111, D18S880		0.0222	A,A,5L-5L A,C,5L-5L	0.000 0.075	0.007 0.053	0.039 0.112	N/A 1.44 (1.05-1.99)

All subjects: 1025 DM-ESRD cases, 1064 controls; 5L-5L Only subjects: 390 DM-ESRD cases and 408 controls; and Non 5L-5L subjects: 623 DM-ESRD cases and 640 controls

N/A: Odds ratio not calculated due to low haplotype frequency (≤ 0.001) in cases and/or controls

P-values less than 0.05 are denoted in bold, italics