Genetic Analysis of *HNF4A* Polymorphisms in Caucasian-American Type 2 Diabetes

Allison M. Bagwell,[1,2] Jennifer L. Bento,[2,3] Josyf C. Mychaleckyj,[2,4,5] Barry I. Freedman,[3] Carl D. Langefeld,[5] and Donald W. Bowden[1,2,3,4]

Hepatocyte nuclear factor 4α (*HNF4A*), the gene for the maturity-onset diabetes of the young type 1 monogenic form of type 2 diabetes, is within the type 2 diabetes–linked region on chromosome 20q12-q13.1 and, consequently, is a positional candidate gene for type 2 diabetes in the general population. Previous studies have identified only a few rare coding mutations. However, recent studies suggest that single nucleotide polymorphisms (SNPs) located near the P2 (β-cell) promoter of *HNF4A* are associated with diabetes susceptibility. In this study, we evaluated 23 SNPs spanning 111 kb including the *HNF4A* gene for association with type 2 diabetes in a collection of Caucasian type 2 diabetic patients with end-stage renal disease ($n = 300$) and control subjects ($n = 310$). None of the individual SNPs were associated with type 2 diabetes in this collection of case subjects (P values ranging from 0.06 to 0.99). However, haplotype analysis identifies significant differences between haplotype frequencies in type 2 diabetic case and control subjects ($P = 0.013$ to $P < 0.001$), with two uncommon "risk" haplotypes (2.4 and 2.2% of chromosomes) and two uncommon "protective" haplotypes (7.1 and 5.0% of chromosomes) accounting for the evidence of association. Our results suggest that type 2 diabetes linked to 20q12–13 is a heterogeneous disease in which different populations may have different type 2 diabetes susceptibility loci.*Diabetes* 54: 1185–1190, 2005

Several studies have provided evidence for linkage of type 2 diabetes to the long arm of chromosome 20 in Caucasians (1–5) and Asians (6,7), suggesting that one or more type 2 diabetes susceptibility loci are located on chromosome 20q. Hepatocyte nuclear factor 4α (*HNF4A*) is a candidate gene within this linked region because mutations in *HNF4A* have been implicated in the maturity-onset diabetes of the young type 1 subtype of type 2 diabetes, a monogenic form of type 2 diabetes characterized by defective insulin secretion (8). *HNF4A* is a member of the steroid/thyroid hormone receptor superfamily of transcription factors (9), and it interacts with regulatory elements in promoters and enhancers of genes involved in cholesterol, fatty acid, and glucose metabolism (10). Thirteen exons have been identified in *HNF4A*, and alternative splicing of these exons results in at least nine isoforms of the gene. The transcription of three of these isoforms is driven by an alternate promoter known as P2, which is located ~45.5 kb upstream of the P1 promoter (11,12). Recent studies suggest that although both promoters function in pancreatic β-cells (13), it is the P2 promoter that primarily drives transcription in these cells (11,12). In previous studies (14,15), we surveyed the coding region of *HNF4A* and conserved enhancer-like elements in the distal *HNF4A* promoter, finding no evidence for significant association to type 2 diabetes in the general population. Recently, several groups have evaluated single nucleotide polymorphisms (SNPs) in the P2 region for association with type 2 diabetes. These reports suggest that four SNPs in a 10.7-kb region encompassing the P2 promoter are significantly associated with type 2 diabetes disease status in subjects from Finnish type 2 diabetic families (16) and Ashkenazi Jewish type 2 diabetic families (17).

RESEARCH DESIGN AND METHODS

The samples evaluated in this study consisted of a collection of 300 unrelated Caucasian type 2 diabetic patients with end-stage renal disease (ESRD), and a corresponding collection of 310 randomly ascertained unrelated Caucasian subjects with no known history of diabetes. Both sample collections were recruited simultaneously. Ascertainment and recruitment criteria of these collections have been described previously in detail (18–23). Case subjects were recruited on the basis of a diagnosis of type 2 diabetes and ESRD. These case subjects were not derived from the family collection that previously showed evidence of linkage to 20q (1) but were recruited using identical clinical criteria. The type 2 diabetic ESRD subjects were 52% female and had a mean age at diagnosis of diabetes of 46.5 ± 12.8 years, mean BMI at recruitment of 28.5 ± 7.0 kg/m^2, mean maximum reported BMI of 36.1 ± 8.3 kg/m^2, mean duration of diabetes >15 years, and mean HbA_{1c} of 8.6%. Control subjects were 67% female, had a mean age of 45.8 years, and a mean BMI (from self-reported height and weight) of 25.7 kg/m^2.

From the [1]Molecular Genetics Program, Wake Forest University School of Medicine, Winston-Salem, North Carolina; the [2]Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina; the [3]Department of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, North Carolina; the [4]Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina; and the [5]Department of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina.

Address correspondence and reprint requests to Donald W. Bowden, Department of Biochemistry, Wake Forest University School of Medicine, Medical Center Blvd., Winston-Salem, NC 27157. E-mail: dbowden@wfubmc.edu.

Received for publication 15 June 2004 and accepted in revised form 22 November 2004.

Additional information for this article can be found in an online appendix at http://diabetes.diabetesjournals.org.

ESRD, end-stage renal disease; HNF, hepatocyte nuclear factor; LD, linkage disequilibrium; SNP, single nucleotide polymorphism.

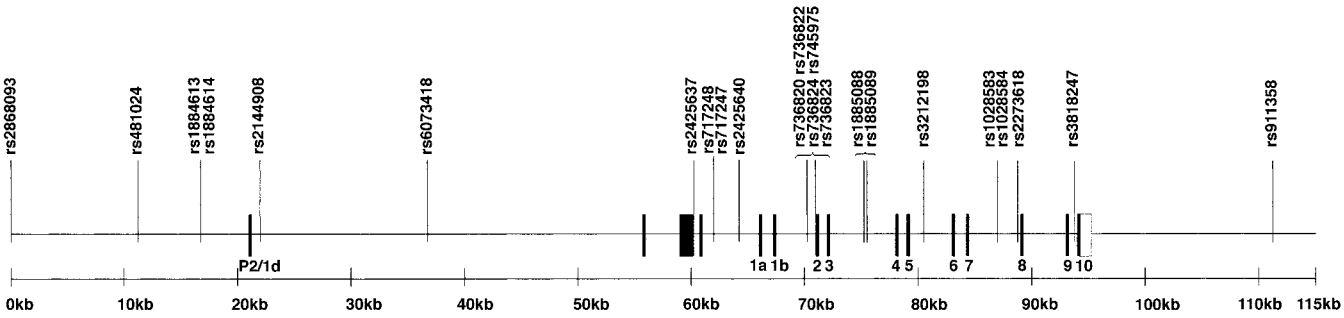

FIG. 1. Genomic map of the *HNF4A* gene with the locations of the 23 genotyped SNPs. The labeled shaded regions are exons, numbered 1–10. The unlabeled shaded regions indicate evolutionarily conserved regions, as identified previously (15). The ruler along the bottom represents the relative location and spacing of SNPs in kilobases within the 111.23-kb region containing *HNF4A*.

SNP selection and genotyping. Sixteen SNPs used in this study were selected from the dbSNP public database (rs2868093, rs6073418, rs717248, rs717247, rs736820, rs736822, rs736824, rs745975, rs736823, rs1885088, rs1885089, rs3212198, rs1028583, rs1028584, rs2273618, and rs911358) and were mapped to unique locations in and around the *HNF4A* gene. Seven additional dbSNPs (rs4810424, rs1884613, rs1884614, rs2144908, rs2425637, rs2425640, and rs3818247) located in and around the *HNF4A* gene were selected on the basis of previous reports of association with type 2 diabetes in Finnish (16) and/or Ashkenazi (17) populations. SNP genotyping was performed on a Sequenom MassArray Genotyping System (Sequenom, San Diego, CA) as previously described (24).

Statistical analysis. Pearson's test of homogeneity of proportions was applied to analyze allele frequency differences between diabetic and nondiabetic subjects. The SNP-Analysis software package (http://www.fhcrc.org/labs/Kruglyak/Downloads) was used to evaluate SNPs for Hardy-Weinberg equilibrium and to calculate pairwise linkage disequilibrium (LD) statistics. Within the regions of strongest LD, HAPLO.SCORE (25 and http://www.mayo.edu/statgen/) and Dandelion (26) were used to estimate haplotypes and to test for association of these haplotypes within the case-control population.

Direct sequencing. The coding regions encompassed by specific haplotypes were sequenced to assess whether coding mutations contribute to the observed haplotype associations. Specifically, the P2 promoter region and alternate exon 1d were sequenced in 27 Caucasian type 2 diabetic case subjects with either of the risk haplotypes and in an equal number of Caucasian control subjects with neither of the risk haplotypes. Exon 4 was sequenced in 16 Caucasian control subjects bearing either of the protective haplotypes and in an equal number of Caucasian type 2 diabetic case subjects with neither of the protective haplotypes. In addition, SNP rs1884613 and the surrounding genomic region were sequenced in 96 Caucasian type 2 diabetic case subjects to confirm individual genotypes assigned by the Sequenom system at this particular SNP.

All samples were sequenced with both forward and reverse primers using the BigDye Terminator v.1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. The sequencing reactions were then analyzed with the Applied Biosystems 3730xl DNA sequencer (Applied Biosystems, Foster City, CA).

RESULTS

In this study, we evaluated 23 SNPs covering a 111-kb genomic region spanning distal to the P2 promoter through the *HNF4A* coding sequence (Fig. 1). All SNPs were located in noncoding regions except rs736823 (in exon 2), resulting in a H98Y amino acid change. The average spacing between SNPs was 5 kb, with inter-SNP distances ranging from 90 bp to 23.5 kb. Included in this evaluation were the four SNPs showing significant association in both the Finnish (16) and Ashkenazi Jewish (17) populations (rs4810424, rs1884613, rs1884614, and rs2144908). In addition, the SNPs we evaluated included two SNPs showing association in the Finnish population only (rs2425637 and rs2425640) and two SNPs showing association in the Ashkenazi Jewish population only (rs1028583 and rs3818247).

The 23 SNPs were genotyped in a collection of Caucasian type 2 diabetic ESRD case subjects ($n = 300$) and control subjects ($n = 310$), and single SNP association analysis was performed. In the control population, rs2425637 ($P = 0.009$) and rs2425640 ($P = 7.6 \times 10^{-9}$) did not conform to Hardy-Weinberg equilibrium. Rs2425640 is one of the SNPs for which association with type 2 diabetes in the Ashkenazi Jewish population was observed. Single SNP association analysis is summarized in Table 1, which shows the SNP identifier, the relevant alleles for each SNP, frequencies in case and control subjects, and *P* values from the test for allelic association. This analysis revealed little or no evidence of association between these 23 SNPs and type 2 diabetes or even a suggestive trend (*P* values between 0.06 and 0.99). The lowest *P* value is with rs717248, which was not an associated SNP in the Finnish and Ashkenazi studies.

Even if no individual SNPs in *HNF4A* are associated with disease status, it may be possible to identify *HNF4A* haplotypes that are significantly associated with increased

TABLE 1
Association analysis for *HNF4A* SNPs in Caucasian type 2 diabetic ESRD subjects and Caucasian control subjects

SNP	Alleles (minor/major)	Frequency in patients ($n = 300$)	Frequency in control subjects ($n = 310$)	*P*
rs2868093	T/C	0.134/0.866	0.147/0.853	0.51
rs4810424	C/G	0.18/0.82	0.164/0.836	0.49
rs1884613	G/C	0.171/0.829	0.172/0.828	0.98
rs1884614	T/C	0.176/0.824	0.168/0.832	0.74
rs2144908	A/G	0.17/0.83	0.181/0.819	0.61
rs6073418	T/C	0.353/0.647	0.356/0.644	0.93
rs2425637	G/T	0.471/0.529	0.471/0.529	0.99
rs717248	G/A	0.027/0.973	0.048/0.952	0.06
rs717247	G/A	0.288/0.712	0.304/0.696	0.53
rs2425640	A/G	0.22/0.78	0.204/0.796	0.51
rs736820	A/G	0.364/0.636	0.356/0.644	0.75
rs736822	T/A	0.235/0.765	0.235/0.765	0.99
rs736824	G/A	0.441/0.559	0.438/0.562	0.91
rs745975	A/G	0.236/0.764	0.228/0.772	0.76
rs736823	A/G	0.07/0.93	0.069/0.931	0.94
rs1885088	A/G	0.207/0.793	0.218/0.782	0.66
rs1885089	T/C	0.207/0.793	0.226/0.774	0.42
rs3212198	C/T	0.437/0.563	0.43/0.57	0.82
rs1028583	T/G	0.421/0.579	0.412/0.588	0.75
rs1028584	A/C	0.421/0.579	0.405/0.595	0.57
rs2273618	T/C	0.412/0.588	0.4/0.6	0.67
rs3818247	T/G	0.391/0.609	0.359/0.641	0.24
rs911358	T/A	0.259/0.741	0.218/0.782	0.1

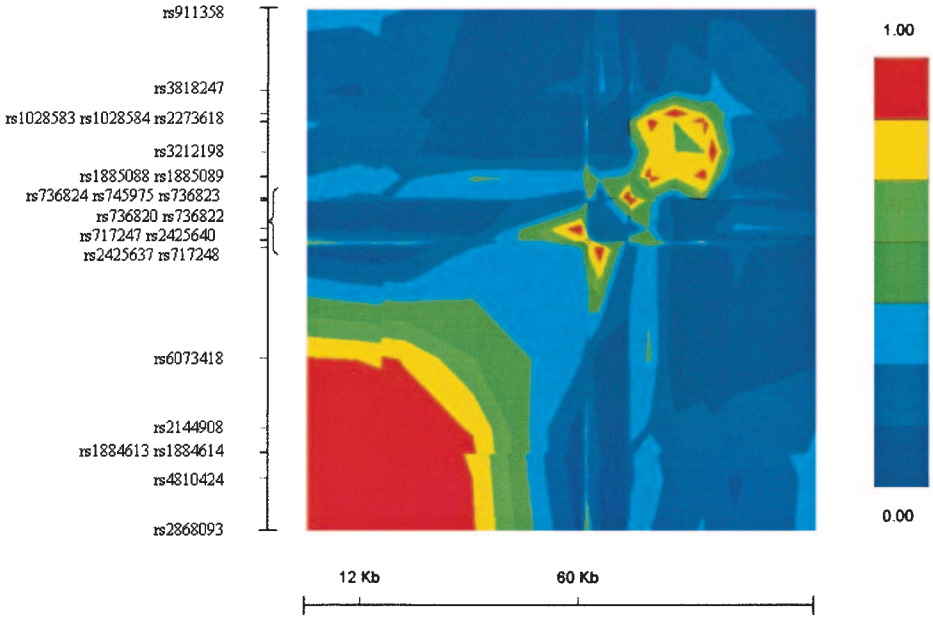

FIG. 2. Marker-to-marker D′ plots for the 23 *HNF4A* SNPs. Inter-SNP D′ values are graphically represented using the Graphical Overview of LD (GOLD) software package, which generates the color-coded plot of the pairwise disequilibrium statistics with horizontal and vertical axes that are scaled according to the inter-SNP distances (31).

or decreased type 2 diabetes risk (27). Therefore, LD and the haplotype structure of *HNF4A* were examined to determine whether specific SNP haplotypes were associated with type 2 diabetes in our population. Inter-SNP D′ values were calculated for the *HNF4A* gene region (Fig. 2). When D′ >0.7 was used to define strong disequilibrium, two six-SNP blocks were identified. The first block, beginning with rs2868093 ~20 kb upstream of P2 and ending with rs6073418 ~17 kb downstream of P2, represents the block of strongest disequilibrium, which spans a region of ~36.8 kb. The second block of more modest disequilibrium, beginning with rs1885088 in intron 3 and ending with rs2273618 in intron 7, covers a 13.5-kb region. Marker to marker D′ values between SNPs in the first block and SNPs in the second block indicate that the two blocks are not in LD with each other.

Each six-SNP disequilibrium block was analyzed using HAPLO.SCORE (25) and Dandelion (26). Haplotype frequencies were estimated, and association analyses were performed with respect to type 2 diabetes in our case-control population. Table 2 summarizes the results of HAPLO.SCORE and Dandelion evaluating the six-SNP haplotypes, in which the specific haplotypes are identified, the frequency of each haplotype is estimated in case and control subjects, a score for each haplotype (Hap score) is calculated, and an empirical *P* value is calculated for the significance of each Hap score. A negative Hap score indicates that the haplotype occurs more frequently in control subjects, whereas a positive Hap score indicates that the haplotype occurs more frequently in type 2 diabetic case subjects. In addition, a global *P* value, which reflects the significance of the difference between type 2 diabetic case and control subjects, is reported.

In the analysis of the six-SNP disequilibrium block (rs2868093, rs4810424, rs1884613, rs1884614, rs2144908, and rs6073418) encompassing the P2 region, the fre-

TABLE 2
Haplotype analysis of six-SNP *HNF4A* haplotypes in type 2 diabetic ESRD subjects vs. control subjects using HAPLO.SCORE (25) and Dandelion (26)

Haplotype	Case subject frequency (%)	Control subject frequency (%)	Hap score	Empirical Hap-specific *P* value	Global simulated *P* value
Six SNPs in strongest LD block (rs2868093, rs4810424, rs1884613, rs1884614, rs2144908, rs6073418)					
TCGTAC	9.6	13.2	−1.88	0.053	
CGCCGT	30.8	35.8	−1.5	0.13	
CCGTAC	2.8	3.3	−0.04	0.95	
CGCCGC	47.8	46.3	0.29	0.78	
TCCTAC	2.4	0	3.28	0.0003	
CGGCGT	2.2	0	3.81	1.00E-06	2.00E-06
Six SNPs in second LD block (rs1885088, rs1885089, rs3212198, rs1028583, rs1028584, rs2273618)					
GCTTAC	0.3	6.8	−4.26	<0.0001	
GCCGCT	0	5.0	−3.92	<0.0001	
GCCGCC	2.2	2.3	−0.01	0.99	
ATTGCC	20.2	20.2	0.01	0.99	
GCTGCC	34.9	29.9	1.32	0.19	
GCCTAT	40.1	32.4	1.97	0.05	<0.0001

quencies of haplotypes in case and control subjects were strongly statistically different (overall $P = 2 \times 10^{-6}$) (Table 2). This highly significant difference was not due to the difference in frequencies between the three common haplotypes that were estimated. Two of these haplotypes, TCGTAC and CGCCGT, are more common in the control population. The third haplotype, CGCCGC, is more common in the case population, but these differences are not statistically significant. The evidence for association with diabetes is due to substantial differences in frequencies in case and control subjects ($P < 0.001$) for two uncommon haplotypes, TCCTAC (2.4% of case subjects and 0% of control subjects) and CGGCGT (2.2% of case subjects and 0% of control subjects). A single SNP, rs1884613, differentiates between the uncommon risk and common, but not risk-associated, haplotypes (TC**C**TAC risk versus TC**G**TAC and CG**G**CGT risk versus CG**C**CGT).

Because significant risk haplotype associations were observed in the absence of single SNP associations, we investigated whether nearby functional or coding SNPs were contributing to the observed haplotype associations. We sequenced the P2 promoter and alternate exon 1d (~774 bp) in all Caucasian type 2 diabetic case subjects having either the TCCTAC (11 heterozygotes and 2 homozygotes) or CGGCGT (14 heterozygotes) risk haplotype. In addition, we sequenced an equal number of Caucasian control subjects. Our sequence analysis identified no mutations or polymorphisms in the P2/exon 1d region, suggesting that in our type 2 diabetic population there are no polymorphisms in the nearby functional or coding sequences that are in LD with the type 2 diabetes risk haplotype (i.e., there is no obvious sequence difference that may be the true causal variant).

Because of the unusual observation that a single SNP differentiates between risk and nonrisk haplotypes in this first LD block, we confirmed the genotyping results by sequencing rs1884613 and the surrounding genomic sequence (~540 bp) in 96 Caucasian type 2 diabetic case subjects in an effort to confirm the genotypes assigned by the Sequenom method. Concordance of 100% was observed with the genotypes determined using the Sequenom method (data not shown).

In the analysis of the second six-SNP disequilibrium block (rs1885088, rs1885089, rs3212198, rs1028583, rs1028584, and rs2273618) encompassing the region from intron 3 to intron 7, the haplotype frequencies in case and control subjects were statistically different (overall $P < 0.0001$) (Table 2). Two of the common haplotypes (GC CTAT and GCTGCC) occur more frequently in the case population, but only the GCCTAT haplotype is marginally significant ($P = 0.05$) for association with type 2 diabetes, whereas the third common haplotype (ATTGCC) occurs at equal frequencies in the case and control populations. Similar to the P2 disequilibrium block, evidence for association in this six-SNP block is due to significant differences in frequencies in case and control subjects for two uncommon haplotypes, GCCGCT (6.80% of control subjects and 0.34% of case subjects) and GCTTAC (5.02% of control subjects and 0% of case subjects), which are significantly more prevalent in the control population than in the case subjects ($P < 0.001$).

Recently, a rare loss-of-function mutation in exon 4 of *HNF4A* has been identified, and this mutation was found to be associated with type 2 diabetes in Japanese subjects (28). Because our uncommon protective haplotypes in the second LD block encompass exon 4, we scanned our populations for the T130I mutation by sequencing exon 4 in Caucasian control subjects having either the GCCGCT (19 heterozygotes and 3 homozygotes) or GCTTAC (18 heterozygotes and 7 homozygotes) protective haplotype. In addition, we sequenced an equal number of Caucasian type 2 diabetic case subjects bearing neither of the protective haplotypes. Our sequencing data identified no T130I mutations or any other coding mutations in exon 4 (data not shown). These results are not surprising because the T130I mutation is associated with an increased risk for type 2 diabetes, and therefore it would be surprising to observe this allele in control individuals bearing a protective haplotype.

DISCUSSION

Recent studies suggest that four SNPs near the P2 promoter of *HNF4A* are significantly associated with diabetes status and that evidence for linkage at 20q13 could be attributed to the families carrying a risk allele (16,17). Independently, we evaluated 23 SNPs spanning 111.23 kb of the *HNF4A* gene region in a case-control association study. Included in this analysis were the four SNPs in the P2 promoter region found to be associated in both Finnish and Ashkenazi samples (rs4810424, rs1884613, rs1884614, and rs2144908), three SNPs found to be associated in Finnish samples only (rs2425637, rs2425640, and rs1885088), and two SNPs found to be associated in Ashkenazi samples only (rs1028583 and rs3818247). In contrast to the results of Love-Gregory et al. (17) and Silander et al. (16), we observed little or no evidence of association with single SNPs (Table 1).

Our LD and haplotype structure indicate that there are two distinct disequilibrium blocks separated by recombination within this gene region that are each defined by six SNPs. This LD structure broadly corresponds to that reported by Love-Gregory et al. (17). We used this LD structure to guide the haplotype analysis. Our analysis indicates that in the LD block encompassing the P2 promoter region (Table 2), there are two rare, completely mismatching haplotypes (TCCTAC and CGGCGT) that are significantly associated with type 2 diabetes risk. These risk-associated haplotypes differ from the more common, nonrisk haplotypes at a single SNP locus, rs1884613. However, rs1884613 showed no trend toward association in our single SNP analysis. Direct sequencing confirmed that the individual genotypes assigned at this particular locus were correct, indicating that the rare risk haplotypes were not estimated due to genotyping error. In addition to genotyping error, an alternative possibility is that the haplotype-estimating programs (HAPLO.SCORE and Dandelion) do not correctly assign haplotypes in this circumstance of unusual, uncommon haplotypes. Extensive exploration of this possibility has led us to conclude that haplotype assignments are correct (data not shown). A summary of the haplotype assignments for each individual with the uncommon risk or uncommon protective haplotypes is shown in supplemental Table 1 (available at http://diabetes.diabetesjournals.org). It is noteworthy that the probability assigned to each

haplotype call is >70% for all haplotypes and that the average probability of correct assignment is 88% for the P2 haplotypes and 98% for the intron 3 to intron 7 haplotypes. In addition, we have observed that haplotype-specific *P* values for uncommon haplotypes may not be robustly estimated using HAPLO.SCORE (D. Schaid, personal communication), and the reported Hap scores may better reflect the contributions of specific haplotypes to overall differences between case and control subjects. Focusing on the Hap scores alone does not alter our interpretation of the haplotype analysis.

Alternate exon 1d and the P2 promoter region were sequenced in all individuals carrying either of these risk haplotypes. No mutations or polymorphisms were identified, which suggests that the observed associations are not driven by nearby functional or coding mutations. Because the two risk haplotypes are completely mismatching and because they differ from the common nonrisk haplotypes at a single locus that is not individually associated with an increased risk, we hypothesize that unique combinations of SNP alleles may be driving the observed associations and that the phenotype is due to multiple variants working in combination.

In the LD block covering the genomic region from intron 3 to intron 7 (Table 2), there are two rare haplotypes (GCTTAC and GCCGCT) that are significantly associated with a decreased risk for type 2 diabetes and a single common risk haplotype GCCTAT that is marginally significant for type 2 diabetes risk ($P = 0.05$). Because the T130I loss-of-function mutation in exon 4 was recently identified (28), we sequenced exon 4 in all individuals carrying either of these protective haplotypes. The mutation was not present in any of the control samples bearing the protective haplotypes. Because no nearby coding mutations have been identified that can account for these haplotypic associations, a combination of SNPs may be required to confer protection. Baroso et al. (27) recently suggested the possibility of protective haplotypes in the coding region of *HNF4A*; however, SNPs comprising such protective haplotypes were not genotyped in this study.

Whereas we used the LD structure of the gene region to guide our haplotype analysis, Silander et al. (16) estimated haplotype frequencies using six SNPs spanning two distinct LD blocks and regions of recombination between blocks. Because our analysis included four of these six SNPs (rs2144908, rs2425637, rs2425640, and rs1885088), we estimated haplotypes using these four SNPs. Although our analysis identified the same haplotypes as those reported by Silander et al. (16), none of the haplotypes were significantly associated (*P* values ranging from 0.16 to 0.99; overall empirical $P = 0.73$) (data not shown).

HNF4A SNPs do not seem to play the major role in type 2 diabetes susceptibility in our study population that has been observed in studies of Finnish (16) and Ashkenazi Jewish (17) families. It is interesting to note that we observed stronger evidence for association to type 2 diabetes susceptibility with the nearby PTPN1 gene on 20q13 (29, 30). The results of *HNF4A* haplotype analysis suggest, however, that there is some type 2 diabetes risk associated with *HNF4A* in our population. It should be informative to test these haplotype effects in other populations. The risk haplotypes in the P2 region were only observed in case subjects, suggesting that they may have a substantial impact on *HNF4A* gene expression or message stability. Similar conclusions can be drawn about the *HNF4A* protective haplotypes. It should be noted that failure to replicate the evidence reported by Silander et al. (16) and Love-Gregory et al. (17) could be due to lack of statistical power. Our case-control population has 80% power to detect $P \leq 0.05$ for an odds ratio of 1.5. If the true odds ratio for *HNF4A* is lower, this study may not have the power to detect a significant association.

In addition, as noted above, the study population described here consists of case subjects ascertained as unrelated Caucasian Americans with a diagnosis of type 2 diabetes with associated ESRD. The populations studied by Silander et al. (16) (Finnish type 2 diabetic families) and Love-Gregory et al. (17) (type 2 diabetic families of Ashkenazi Jewish heritage) were ascertained in a different manner. Although it seems unlikely, there is the possibility that the *HNF4A* risk alleles observed by Silander et al. (16) and Love-Gregory et al. (17) confer protection from ESRD and would be rare in the ESRD population. Comparison of allele frequencies (data not shown) does not directly suggest this. Overall, the results reported here, in combination with our analysis of PTPN1, and the results of Silander et al. (16) and Love-Gregory et al. (17) are consistent with a model in which multiple genes on 20q12–13 contribute to type 2 diabetes, and different genes affect risk to various degrees in different populations.

ACKNOWLEDGMENTS

D.W.B. has received support from National Institutes of Health Grants R01-DK-56289 and R01-DK-53591.

REFERENCES

1. Bowden DW, Sale M, Howard TD, Qadri A, Spray BJ, Rothschild CB, Akots G, Rich SS, Freedman BI: Linkage of genetic markers on human chromosomes 20 and 12 to NIDDM in Caucasian sib pairs with a history of diabetic nephropathy. *Diabetes* 46:882–886, 1997
2. Ji L, Malecki M, Warram JH, Yang Y, Rich SS, Krolewski AS: New susceptibility locus for NIDDM is localized to human chromosome 20q. *Diabetes* 46:876–881, 1997
3. Zouali H, Hani EH, Philippi A, Vionnet N, Beckmann JS, Demenais F, Froguel P: A susceptibility locus for early-onset non-insulin dependent (type 2) diabetes mellitus maps to chromosome 20q, proximal to the phosphoenolpyruvate carboxykinase gene. *Hum Mol Genet* 6:1401–1408, 1997
4. Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, Balow J Jr, Ally DS, Kohtamaki K, Chines P, Birznieks G, Kaleta HS, Musick A, Te C, Tannenbaum J, Eldridge W, Shapiro S, Martin C, Witt A, So A, Chang J, Shurtleff B, Porter R, Boehnke M, et al.: Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. *Proc Natl Acad Sci U S A* 96:2198–2203, 1999
5. Permutt MA, Wasson JC, Suarez BK, Lin J, Thomas J, Meyer J, Lewitzky S, Rennich JS, Parker A, DuPrat L, Maruti S, Chayen S, Glaser B: A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. *Diabetes* 50:681–685, 2001
6. Luo TH, Zhao Y, Li G, Yuan WT, Zhao JJ, Chen JL, Huang W, Luo M: A genome-wide search for type II diabetes susceptibility genes in Chinese Hans. *Diabetologia* 44:501–506, 2001
7. Mori Y, Otabe S, Dina C, Yasuda K, Populaire C, Lecoeur C, Vatin V, Durand E, Hara K, Okada T, Tobe K, Boutin P, Kadowaki T, Froguel P: Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate Loci on 7p and 11p. *Diabetes* 51:1247–1255, 2002
8. Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI: Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* 384:458–460, 1996

9. Sladek FM, Zhong WM, Lai E, Darnell JE Jr.: Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev* 4:2353–2365, 1990
10. Stoffel M, Duncan SA: The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. *Proc Natl Acad Sci U S A* 94:13209–13214, 1997
11. Thomas H, Jaschkowitz K, Bulman M, Frayling TM, Mitchell SM, Roosen S, Lingott-Frieg A, Tack CJ, Ellard S, Ryffel GU, Hattersley AT: A distant upstream promoter of the HNF-4alpha gene connects the transcription factors involved in maturity-onset diabetes of the young. *Hum Mol Genet* 10:2089–2097, 2001
12. Boj SF, Parrizas M, Maestro MA, Ferrer J: A transcription factor regulatory circuit in differentiated pancreatic cells. *Proc Natl Acad Sci U S A* 98:14481–14486, 2001
13. Eeckhoute J, Moerman E, Bouckenooghe T, Lukoviak B, Pattou F, Formstecher P, Kerr-Conte J, Vandewalle B, Laine B: Hepatocyte nuclear factor 4 alpha isoforms originated from the P1 promoter are expressed in human pancreatic beta-cells and exhibit stronger transcriptional potentials than P2 promoter-driven isoforms. *Endocrinology* 144:1686–1694, 2003
14. Price JA, Fossey SC, Sale MM, Brewer CS, Freedman BI, Wuerth JP, Bowden DW: Analysis of the HNF4 alpha gene in Caucasian type II diabetic nephropathic patients. *Diabetologia* 43:364–372, 2000
15. Bagwell AM, Bailly A, Mychaleckyj JC, Freedman BI, Bowden DW: Comparative genomic analysis of the HNF-4alpha transcription factor gene. *Mol Genet Metab* 81:112–121, 2004
16. Silander K, Mohlke KL, Scott LJ, Peck EC, Hollstein P, Skol AD, Jackson AU, Deloukas P, Hunt S, Stavrides G, Chines PS, Erdos MR, Narisu N, Conneely KN, Li C, Fingerlin TE, Dhanjal SK, Valle TT, Bergman RN, Tuomilehto J, Watanabe RM, Boehnke M, Collins FS: Genetic variation near the hepatocyte nuclear factor-4α gene predicts susceptibility to type 2 diabetes. *Diabetes* 53:1141–1149, 2004
17. Love-Gregory LD, Wasson J, Ma J, Jin CH, Glaser B, Suarez BK, Permutt MA: A common polymorphism in the upstream promoter region of the hepatocyte nuclear factor-4α gene on chromosome 20q is associated with type 2 diabetes and appears to contribute to the evidence for linkage in an Ashkenazi Jewish population. *Diabetes* 53:1134–1140, 2004
18. 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* 7:2559–2564, 1996
19. Freedman BI, Yu H, Spray BJ, Rich SS, Rothschild CB, Bowden DW: Genetic linkage analysis of growth factor loci and end-stage renal disease in African Americans. *Kidney Int* 51:819–825, 1997
20. Yu H, Bowden DW, Spray BJ, Rich SS, Freedman BI: Identification of human plasma kallikrein gene polymorphisms and evaluation of their role in end-stage renal disease. *Hypertension* 31:906–911, 1998
21. Yu H, Anderson PJ, Freedman BI, Rich SS, Bowden DW: Genomic structure of the human plasma prekallikrein gene, identification of allelic variants, and analysis in end-stage renal disease. *Genomics* 69:225–234, 2000
22. Freedman BI, Yu H, Anderson PJ, Roh BH, Rich SS, Bowden DW: Genetic analysis of nitric oxide and endothelin in end-stage renal disease. *Nephrol Dial Transplant* 15:1794–1800, 2000
23. Freedman BI, Rich SS, Yu H, Roh BH, Bowden DW: Linkage heterogeneity of end-stage renal disease on human chromosome 10. *Kidney Int* 62:770–774, 2002
24. Bensen JT, Langefeld CD, Hawkins GA, Green LE, Mychaleckyj JC, Brewer CS, Kiger DS, Binford SM, Colicigno CJ, Allred DC, Freedman BI, Bowden DW: Nucleotide variation, haplotype structure, and association with end-stage renal disease of the human interleukin-1 gene cluster. *Genomics* 82:194–217, 2003
25. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA: Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 70:425–434, 2002
26. Green LE, Lange EM, Langefeld CD: Power comparison for phase-known versus phase-unknown haplotype analyses for case-control designs. *Am J Hum Genet* 69:1948a, 2001
27. Barroso I, Luan J, Middelberg RP, Harding AH, Franks PW, Jakes RW, Clayton D, Schafer AJ, O'Rahilly S, Wareham NJ: Candidate gene association study in type 2 diabetes indicates a role for genes involved in beta-cell function as well as insulin action. *PLoS Biol* 1:E20, 2003
28. Zhu Q, Yamagata K, Miura A, Shihara N, Horikawa Y, Takeda J, Miyagawa J, Matsuzawa Y: T130I mutation in HNF-4alpha gene is a loss-of-function mutation in hepatocytes and is associated with late-onset type 2 diabetes mellitus in Japanese subjects. *Diabetologia* 46:567–573, 2003
29. Bento JL, Palmer ND, Mychaleckyj JC, Lange LA, Langefeld CD, Rich SS, Freedman BI, Bowden DW: Association of protein-tyrosine phosphatase1B gene polymorphisms with type 2 diabetes. *Diabetes* 53:3007–3012, 2004
30. Palmer ND, Bento JL, Mychaleckyj JC, Langefeld CD, Campbell JK, Norris JM, Haffner SM, Bergman RN, Bowden DW: Association of protein tyrosine phosphatase 1B gene polymorphisms with measures of glucose homeostasis in Hispanic Americans: the IRAS Family Study. *Diabetes* 53:3013–3019, 2004
31. Abecasis GR, Cookson WO: GOLD: graphical overview of linkage disequilibrium. *Bioinformatics* 16:182–183, 2000