Liver Pyruvate Kinase Polymorphisms Are Associated With Type 2 Diabetes in Northern European Caucasians

Hua Wang,1 Winston Chu,1 Swapan K. Das,1 Qianfang Ren,1 Sandra J. Hasstedt,2 and Steven C. Elbein1

Pyruvate kinase is a key glycolytic enzyme. Isoforms that are expressed in the red cell, liver, pancreatic β-cells, small intestine, and proximal renal tubule are encoded by the 12 exons of the PKLR gene, which maps to chromosome 1q23. We hypothesized that common variants of the PKLR gene could account for the linkage of diabetes to this region. We screened the promoter regions, exons and surrounding introns, and the 3′ untranslated region for mutations. We identified five single-nucleotide polymorphisms (SNPs), and only one (V506I, exon 11) altered the coding sequence. We tested the five SNPs, a poly-T insertion-deletion polymorphism, and an ATT triplet repeat in 131 unrelated diabetic patients and 118 nondiabetic control subjects. The V506I variant was rare and not associated with type 2 diabetes. The four SNPs and the insertion-deletion polymorphism were associated with diabetes, with a 10% difference between individuals with diabetes and nondiabetic individuals (P = 0.001–0.011, relative risk for minor allele 1.85). The same trend was found for the ATT repeat (P = 0.029). Common variants in the PKLR are associated with increased risk of type 2 diabetes, but because of strong linkage disequilibrium between variants, the actual susceptibility allele may be in a different gene. Diabetes 51:2861–2865, 2002

Type 2 diabetes has a strong genetic component. The risk of diabetes for the siblings of a diabetic individual may be three- to fourfold higher than the general population (1). Nonetheless, identification of the susceptibility loci has been challenging. A single locus has been mapped by linkage and narrowed to the calpain 10 gene on chromosome 2q (2), but this locus does not play a major role in diabetes susceptibility in most Caucasian populations (3). Genome-wide scans from a number of laboratories suggest that multiple genes are likely to contribute to diabetes susceptibility. A type 2 diabetes locus on chromosome 1q21-q23 was identified by our laboratory in Caucasian families (4) and independently in Pima Indian families and discordant sib pairs (5). Replication of this linkage was reported subsequently in Amish families (6), English sib pairs (7), and French families (8), and this region was recently linked to glucose and HbA1c in the Framingham Offspring Study (9).

Pyruvate kinase (ATP:pyruvate phosphotransferase, EC 2.7.1.40) is a ubiquitously expressed enzyme that catalyzes the conversion of phosphoenolpyruvate to pyruvate with the generation of ATP. It comprises four isoforms from two genetic loci. M1 and M2 are alternatively spliced products of the PKM gene, which maps to chromosome 15. The 12-exon PKLR gene maps to chromosome 1q23, well within the CI for linkage of type 2 diabetes, and thus is both a functional and positional candidate for type 2 diabetes susceptibility. The PKLR gene undergoes alternative splicing to generate a red cell isoform (exons 1 and 3–12) and a liver isoform (exons 2 and 3–12). These two isoforms are under the control of different promoters and show different tissue distribution. Homozygosity for mutations of the PK-R gene and the resulting red cell pyruvate kinase deficiency are well-documented causes of hemolytic anemia (OMIM266200), but such mutations are rare. In addition to liver, the PK-L isoform is expressed in pancreatic β-cells, small intestine, and proximal renal tubule (10). The gene is upregulated by glucose, perhaps acting through the carbohydrate response element in the L-type promoter region (10), and is among the downstream targets of hepatocyte nuclear factor (HNF)-1α (11,12). Thus, PK-L shares the tissue distribution of genes known to cause more severe early-onset diabetes (maturity-onset diabetes of the young [MODY] [13]) and is among the genes with decreased expression in patients with MODY-1 and -3 (11,12,14). Furthermore, PK-L activity in response to glucose may be regulated by AMP-activated protein kinase (15), which in turn is also a target of metformin (16).

To test the hypothesis that variation in PKLR contributes to diabetes pathogenesis, we screened for mutations in 16 diabetic individuals who were drawn from families with the strongest evidence for linkage to chromosome 1q21-q23. We then tested for an association of five single-nucleotide polymorphisms (SNPs), an insertion-deletion polymorphism, and a trinucleotide repeat polymorphism with type 2 diabetes in patients with type 2 diabetes and control subjects from the same population.

RESEARCH DESIGN AND METHODS

Subjects. All studies were conducted on individuals of Northern European ancestry ascertained in Utah, as described in detail elsewhere (4). Families
TABLE 1
Characteristics of study population

<table>
<thead>
<tr>
<th></th>
<th>Type 2 diabetic family members</th>
<th>Type 2 diabetic population sample</th>
<th>Nondiabetic control subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>70</td>
<td>61</td>
<td>118</td>
</tr>
<tr>
<td>age (years)</td>
<td>62.6 ± 10.5</td>
<td>62.3 ± 12.3</td>
<td>55.6 ± 12.8</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>31.5 ± 5.8</td>
<td>27.6 ± 18.9</td>
<td>28.0 ± 6.4</td>
</tr>
</tbody>
</table>

Data are means ± SE.

were ascertained for at least two diabetic siblings, and only unrelated diabetic family members were included in the case-control study (Table 1). Not all families represented in the case-control study were included in the linkage studies or the family-based association analyses. Most control subjects were spouses of family members who had normal oral glucose tolerance tests. Additional diabetic individuals whose families were not sampled were ascertained from the same population for documented diabetes and a family history of diabetes. All subjects provided written informed consent under a protocol approved by the University of Utah Institutional Review Board. Mutation detection was performed in 16 unrelated individuals with type 2 diabetes who were selected from independent families with the strongest evidence for linkage to chromosome 1q21-q23 (4). Family-based association studies used members of 63 families that have been described in detail elsewhere (4,17,18). Genotypic data were available for up to 698 individuals, including 283 affected members of 63 families that have been described in detail elsewhere (4). Family-based association studies used linkage disequilibrium data for the common alleles in three pools (214 bp plus 217 bp, 220 bp, and 223 bp) against all rare alleles (208, 211, 226, 229, and 232 bp) using a 2 × 4 contingency table. Linkage disequilibrium was calculated using the expectation maximization algorithm in the 2LD program (22) for each pairwise combination of two allele variants. Case and control data were pooled for linkage disequilibrium calculations, and individuals were excluded if they were not successfully typed at both markers in the pair. In addition to testing for allelic association, we tested positively associated SNPs further by testing for an association of any genotype using 2 × 3 χ² analysis and for dominant and recessive models using 2 × 2 χ² analysis. Family-based associations were tested for increased transmission over the expected transmission of each allele from parents to affected offspring of 0.5 using maximum likelihood methods, as we have described previously (3). The effect of PKLR variation on quantitative traits was primarily tested in nondiabetic family members (glucose and insulin) or all family members (BMI and lipids). All skewed measures, including insulin, glucose, BMI, and lipid measures, were log transformed to normality and adjusted for age and sex before analysis. An effect of each SNP on the quantitative traits was tested using measured genotype analysis, as described in detail elsewhere (3).

RESULTS
Among the 12 exons and two promoter regions, we detected five SNPs (Table 1 and Fig. 1). Only PKLR3 in exon 11 changed the amino acid sequence at codon 506 (referenced to the PK-R message) from valine to isoleucine. The two exon 12 SNPs were either synonymous (PKLR4 and Arg569Arg) or in the untranslated region (PKLR5). All other SNPs were intronic and did not alter known regulatory or splicing regions. Of the SNPs identified, all but the nonsynonymous SNP in exon 11 were also reported in the dbSNP database. Of the 13 SNPs proposed in the public database for this gene, we confirmed 4 (Table 2), failed to detect 4 (Table 2), and the remaining 5 fell outside of our amplification primers and thus were not examined.

To test the role of PKLR in type 2 diabetes, we typed SNPs PKLR1–5 (Table 2) in 118 unaffected individuals and 131 diabetic individuals, all of Northern European ancestry. Characteristics of the study population are shown in Table 1. Although the control population was significantly younger than the diabetic population, their age was similar to the age of onset of the diabetic subjects (data not shown). We also typed two known polymorphisms that lay outside of our screening regions: the two allele, 9-bp

---

**FIG. 1.** Map of the PKLR gene and variants. The approximate location of the 12 exons of PKLR is shown. Names of variants correspond to Table 5 in the online appendix. Black bars show location of coding sequence, and gray bars show the noncoding regions. Exons 1 and 2 are alternatively spliced; exon 1 but not exon 2 is present in the PK-R gene, whereas the PK-L gene begins with exon 2.
insertion-deletion polymorphism comprising a poly-T region in intron 10 (PKLR-T) and an ATT microsatellite in intron 11 (PKLR-ATT). As shown in Table 2, the only nonsynonymous SNP, PKLR3 (V506I), was uncommon and equally frequent in case and control subjects. In contrast, PKLR1, PKLR2, PKLR4, PKLR5, and PKLR-T all showed significant allelic association associated with type 2 diabetes ($P = 0.01–0.001$) (Table 2). In each case, the difference in allele frequency between case and control subjects was $\sim 10\%$, and the minor allele frequencies were similar (Table 2). The relative risk for diabetes with the minor allele was 1.86 (CI 1.22–2.65). The microsatellite polymorphism had nine alleles, five of which were equally frequent in case and control subjects and were relatively rare (online appendix, Table 4). Alleles 214 and 217 were more common in diabetic subjects than in control subjects. When we pooled the less common alleles to test the association in a 2 × 4 contingency table (see RESEARCH DESIGN AND METHODS), we again found evidence for an association of pooled alleles with type 2 diabetes ($P = 0.029$).

Given the similar allele frequencies and associations of PKLR1, PKLR2, PKLR4, PKLR5, and PKLR-T, we tested for pairwise linkage disequilibrium between each allele polymorphism. The V506I variant was in moderate disequilibrium with the other three variants (D’ = 0.7, $P = 0.003$). In contrast, each of the remaining SNPs and the insertion-deletion variant were in strong and essentially complete disequilibrium, with D’ $\sim 0.95$ ($P < 0.0001$, $\chi^2 > 400$). Thus, only two haplotypes were seen for SNPs 1, 2, 4, and 5 and the T repeat insertion-deletion variant: CCAC-insertion and TTCT-deletion. Analysis of genotypic data showed that all five variants were significant in 2 × 3 contingency tables ($P < 0.05$). Analysis of individual genotypes were most significant with a dominant-at risk” haplotype (CCAC-insertion) or, equivalently, a recessive-protective major haplotype.

We typed SNPs 4 and 5 and the insertion-deletion polymorphism in all family members. We observed only two major haplotypes: 98.4% of all haplotypes were either AC-insertion or CT-deletion, and 1.6% were apparent recombinant haplotypes. Consequently, haplotype data provided no additional information for this gene. No excess transmission was detected for either haplotype using a transmission disequilibrium test ($\tau = 0.53$, $P = 0.5$). Furthermore, we found no evidence that either haplotype was associated with fasting or postchallenge glucose, lipid measures, or several indexes of insulin secretion based on the oral glucose tolerance test, as previously described (23). We also found no evidence that PKLR haplotypes were associated with altered insulin secretion (acute insulin response to glucose or disposition index) or insulin sensitivity in 126 members of the families who had undergone intravenous glucose tolerance tests (24) (data not shown).

DISCUSSION

Pyruvate kinase is a key glycolytic enzyme, and altered expression could be expected to alter glucose metabolism and energy production. Most of the mutations described in PKLR alter splicing or amino acids in the common exons 3–12 and thus would be expected to affect both L and R isoforms. Nonetheless, neither homozygous nor heterozygous carriers have been reported to have abnormalities in glucose metabolism. Furthermore, most missense mutations cause mild disease. PK-L transcription in the pancreas is clearly under the control of HNF-1$\alpha$ (15,25) and is increased by glucose and other carbohydrates (10,15), possibly acting through AMP kinase. However, recent data suggest that most pyruvate kinase activity in the pancreas may come from the M2 isoform of the PKM gene (12). Thus, whether the HNF-1$\alpha$ regulation of PKLR activity contributes to the $\beta$-cell defects observed in subjects with MODY1 or MODY3 is unclear.

Of the five variants that we identified after extensive screening of the PKLR gene, only PKLR3 (V506I) altered the coding sequence, but it was not associated with type 2 diabetes. None of the remaining variants alter a known coding or regulatory region. Nonetheless, all five two allele variants were associated with type 2 diabetes, with $\sim 10\%$ difference in allele frequency between case and control subjects and resulting in a relative risk estimated at 1.8 (CI 1.2–2.8). This trend was confirmed by the analysis of the ATT microsatellite. However, we did not confirm this association in analysis of full families using a modification of the transmission disequilibrium test. Three explanations are possible for this paradox. The initial case-control
observation might represent a type 1 error. Because the six polymorphisms are in strong linkage disequilibrium, this error might be observed for any variant typed in this region. Alternatively, population stratification can cause discrepancies between case-control and family-based association studies, but no stratification is anticipated in this fairly uniform well-matched population. We believe the most likely explanation is the reduced power of the TDT test relative to case-control designs (26), particularly when a relatively small number of large families is tested.

We also failed to detect a physiological consequence of the PKLR variants in nondiabetic family members who had undergone coreal phenotype typing, including alterations in fasting glucose, fasting insulin, or measures of insulin sensitivity and secretion. In contrast, we were able to show an increased fasting glucose with calpain 10 polymorphisms (3). This may reflect the influence of the current haplotypes on a phenotype that we did not measure, such as hepatic glucose production.

Because all five associated variants were in strong mutual disequilibrium, we cannot determine which variant, if any, represents the true susceptibility locus. The responsible variant may lie in another noncoding region of the gene or in another nearby gene. This region is characterized by extensive and nearly complete linkage disequilibrium that extends at least 70 kb centromeric and an unknown distance telomeric, with multiple genes included in the region (27). Furthermore, increasing data support a role for at least some intron variants in altered gene expression and disease susceptibility (2,28). Such effects could be tissue specific, and it is difficult to determine whether the primary role of PKLR is in the pancreas or liver.

Can the association that we report here account for the previously reported linkage? This region of chromosome 1 is particularly gene rich, and multiple genes in this region may be contributing to the evidence for linkage in this region. For example, in Crohn’s disease, a clear disease-causing mutation failed to explain all of the evidence for linkage to a single locus (29). Methods to determine whether an association can account for the evidence for linkage have been proposed by Horikawa et al. (2) and more recently by Sun et al. (30), but neither method is easily applied to the multiplex families on which our linkage was based. In the present study, we found no clear differences in risk or allele frequency between those diabetic individuals who belonged to multiplex families and those who were recruited without other family members (data not shown). Based on the strength of our linkage result and the relatively modest increase in risk seen with the minor alleles of SNPs 1, 2, 4, and 5 and the intron 10 insertion-deletion polymorphism, we would not expect this finding to completely explain the linkage in this region. Given the strong linkage disequilibrium among these four variants, functional studies will be required to identify the causative susceptibility variant.

In summary, we provide evidence that variants in the PKLR gene are associated with an increased diabetes risk, with a relative risk of ~1.8. The increase in risk of PKLR alleles for type 2 diabetes estimated by our study is approximately in the same range (2) or greater than (31) the risk attributed to other susceptibility genes. These findings suggest a model of multiple susceptibility genes, each contributing modest risk. Although this gene is expressed in both liver and pancreas, the means by which these variants increase risk is currently uncertain. Both biological studies of PKLR expression and extensive association studies to determine the extent of linkage disequilibrium 5’ and 3’ to the gene will be needed to define this susceptibility locus.

ACKNOWLEDGMENTS

This work was supported by grant DK39311 from the National Institutes of Health/National Institute of Diabetes and Digestive and Kidney Diseases and by the Department of Veterans Affairs. Subject ascertainment was supported in part by a GENNID Family Center Acquisition grant from the American Diabetes Association. Clinical studies were supported by grants from the National Institutes of Health/National Center for Research Resources to the General Clinical Research Centers of the University of Utah (M01RR03655) and the University of Arkansas for Medical Sciences (M01RR14288).

REFERENCES

12. Shih DQ, Screenan S, Munoz KN, Phillipson L, Pontoglio M, Yaniv M, Polonsky KS, Stoffel M: Loss of HNF-1α function in mice leads to abnormal


