Polymorphisms in Both Promoters of Hepatocyte Nuclear Factor 4-α Are Associated With Type 2 Diabetes in the Amish

Coleen M. Damcott,1 Nicole Hoppman,1 Sandra H. Ott,1 Laurie J. Reinhart,1 Jian Wang,1 Toni I. Pollin,1 Jeffrey R. O’Connell,1 Braxton D. Mitchell,1 and Alan R. Shuldiner1,2

Hepatocyte nuclear factor 4-α (HNF4A) is a transcription factor located on chromosome 20q13 that regulates expression of genes involved in glucose metabolism and homeostasis. Recently, two groups independently identified single nucleotide polymorphism (SNPs) in an alternate upstream promoter (P2) of HNF4A that were associated with type 2 diabetes in Ashkenazi Jews and Finns. We genotyped haplotype-tagging SNPs (htSNPs) across the two promoter regions and the coding region of HNF4A in individuals with type 2 diabetes (n = 137), impaired glucose tolerance (IGT) (n = 139), and normal glucose tolerance (n = 342) from the Amish Family Diabetes Study (AFDS) to test for association with type 2 diabetes. In the P1 promoter region, we observed a significant association between the A allele of the GG genotype (57.8 vs. 62.9 years, P = 0.011). In the P2 promoter, the htSNP rs1884614 showed borderine association with both type 2 diabetes (OR 1.40, P = 0.09) and the combined type 2 diabetes/IGT trait (1.35, P = 0.07). In an expanded set of 698 nondiabetic AFDS subjects, we found association between rs1884614 and glucose area under the curve during an oral glucose tolerance test (additive model, P = 0.022; dominant model, P = 0.010). The results of this study provide evidence that variants in both the P1 and P2 promoters of HNF4A increase risk for typical type 2 diabetes. Diabetes 53:3337–3341, 2004

Hepatocyte nuclear factor 4-α (HNF4A) is a transcription factor that is expressed in several tissues, including liver and pancreas, where it regulates expression of genes involved in glucose homeostasis and glucagon secretion, respectively. Relatively rare mutations in HNF4A have been identified that cause maturity-onset diabetes of the young type 1 (rev. in 5), a dominantly inherited, early-onset form of type 2 diabetes characterized by impaired glucose-induced insulin secretion due to pancreatic β-cell dysfunction (6–9). HNF4A expression patterns are complex as a result of alternative splicing and transcription from two different promoters, the proximal P1 promoter and the P2 promoter, which lies ∼45 kb upstream of the P1 promoter (10–13).

The 12 coding exons of HNF4A span ∼29 kb on chromosome 20q13, a region of overlapping linkage to type 2 diabetes in several Caucasian (14–19) and Asian (20,21) populations. Recently, through fine-mapping efforts in this region of chromosome 20q, two groups concurrently identified single nucleotide polymorphisms (SNPs) in the P2 and P1 promoter regions and coding exons of HNF4A that are associated with type 2 diabetes in the Ashkenazi Jews (22) and Finns (FUSION 1 [Finland-United States Investigation of NIDDM Genetics 1]) (23). Silander et al. (23) identified 10 SNPs across a 64-kb region spanning the P2 and P1 promoter regions and exons 1–3 of HNF4A that were associated with type 2 diabetes in the Ashkenazi cohort, the SNPs closer to the P1 promoter and coding exons of HNF4A were not associated with type 2 diabetes (22); however, four SNPs spanning a ∼10-kb region surrounding the P2 promoter were associated with type 2 diabetes (rs4810424, rs1884613, rs1884614, and rs2144908). These four SNPs are located in a 177-kb region of strong linkage disequilibrium (LD), including the 45-kb gap separating HNF4A from P2 (22,23). In addition to the observed association with type 2 diabetes, these P2 SNPs appeared to explain a significant portion of the linkage to chromosome 20q12-q13 observed in both the Ashkenazi Jews and Finns. The replicating evidence pre-
rs2425640, in the P2 and P1 promoters, respectively, that were associated with type 2 diabetes and glucose traits.

The NGT control subjects selected were control subjects with normal glucose tolerance (NGT), individuals with impaired glucose tolerance (IGT), and 342 which included 137 subjects with type 2 diabetes, 139 enrolled in the Amish Family Diabetes Study (AFDS), to capture most or all of the variation across the gene with haplotype blocks defined in the Ashkenazi Jews (22) and Finns (23).

As expected by these studies suggests that SNPs near the P2 promoter of HNF4A increase susceptibility to type 2 diabetes.

Although no evidence for linkage to type 2 diabetes was detected on chromosome 20q12-q13 in our genome-wide scan (average marker density = 9.7 cM) in the Old Order Amish (logarithm of odds = 0.00 between markers D20S107 and D20S119, which are ~6 cM apart) (24), we tested whether SNPs in HNF4A and its promoters were associated with type 2 diabetes in the Amish. We selected six haplotype-tagging SNPs (htSNPs) spanning the P2 and P1 promoters and the HNF4A coding region from the LD blocks defined in the Ashkenazi Jews (22) and Finns (23). Given that the Amish are a young founder population, we hypothesized that haplotype blocks would be as large as or larger than those in the other populations, thus allowing us to capture most or all of the variation across the gene with these SNPs. These SNPs were genotyped in 618 individuals enrolled in the Amish Family Diabetes Study (AFDS), which included 137 subjects with type 2 diabetes, 139 individuals with impaired glucose tolerance (IGT), and 342 control subjects with normal glucose tolerance (NGT). The NGT control subjects selected were ≥38 years of age in order to increase the probability of their capacity for diabetes resistance. Table 1 summarizes the allele frequencies in individuals with type 2 diabetes, IGT, and NGT and the results of genotypic association analysis for each SNP. All SNPs conformed to Hardy-Weinberg expectations. For rs2425640, one of the SNPs located in the P1 promoter region, the frequency of the A allele at the rs1884614 SNP was lower in control subjects with NGT than in both the type 2 diabetic group (genotypic OR 1.40, \( P = 0.09 \)) and the combined type 2 diabetic/IGT group (genotypic OR 1.35, \( P = 0.07 \)), although these differences did not achieve statistical significance, as observed in the Ashkenazi Jews and Finns.

We genotyped one (rs1884614) of the four P2 promoter SNPs reported to be associated with type 2 diabetes and in near-perfect LD with each other in both the Ashkenazi Jewish and Finnish populations. In the Amish, the frequency of the A allele at the rs1884614 SNP was lower in control subjects with NGT than in both the type 2 diabetic group (genotypic OR 1.40, \( P = 0.09 \)) and the combined type 2 diabetic/IGT group (genotypic OR 1.35, \( P = 0.07 \)), although these differences did not achieve statistical significance, as observed in the Ashkenazi Jews and Finns. None of the other SNPs in the P1 promoter region or the coding region were associated with type 2 diabetes in the Amish, including the other SNPs observed to be associated with type 2 diabetes in the Finns (rs2425637 and rs3212183) and Ashkenazi Jews (rs3818247). Haplotype analysis revealed that only those haplotypes containing the rs2425640 A allele and rs1884614 A allele were associated with increased type 2 diabetes prevalence (results not shown).

Table 2 shows the pairwise LD (\( D' \) and \( r^2 \)) among the genotyped SNPs in the Amish. The haplotype block structure in the Amish appears very similar to that reported in Finns and Ashkenazi Jews, suggesting that the SNP density we chose is adequate for the detection of most of the common variation in HNF4A. As shown in the Ashkenazi Jews and the Finns, the SNP representing the P2 haplotype block that was genotyped in the Amish (rs1884614) was

<table>
<thead>
<tr>
<th>SNP location (kb)†</th>
<th>SNP name</th>
<th>Major/minor allele</th>
<th>Type 2 diabetes ( (n = 137) )</th>
<th>IGT ( (n = 139) )</th>
<th>NGT ( (n = 342) )</th>
<th>Diabetes vs. NGT*</th>
<th>Diabetes + IGT vs. NGT*</th>
</tr>
</thead>
<tbody>
<tr>
<td>−3.926</td>
<td>rs1884614</td>
<td>G/A</td>
<td>0.208</td>
<td>0.180</td>
<td>0.139</td>
<td>1.40</td>
<td>0.09</td>
</tr>
<tr>
<td>39.604</td>
<td>rs2425637</td>
<td>G/T</td>
<td>0.380</td>
<td>0.368</td>
<td>0.372</td>
<td>1.00</td>
<td>0.09</td>
</tr>
<tr>
<td>43.592</td>
<td>rs2425640</td>
<td>G/A</td>
<td>0.414</td>
<td>0.363</td>
<td>0.363</td>
<td>1.60</td>
<td><strong>0.03</strong></td>
</tr>
<tr>
<td>50.693</td>
<td>rs3212183</td>
<td>T/C</td>
<td>0.377</td>
<td>0.374</td>
<td>0.368</td>
<td>0.97</td>
<td>0.08</td>
</tr>
<tr>
<td>66.316</td>
<td>rs1028583</td>
<td>G/T</td>
<td>0.411</td>
<td>0.368</td>
<td>0.382</td>
<td>0.83</td>
<td>0.05</td>
</tr>
<tr>
<td>73.035</td>
<td>rs3818247</td>
<td>G/T</td>
<td>0.255</td>
<td>0.281</td>
<td>0.265</td>
<td>0.93</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*P values are based on genotype frequencies, and ORs reflect the odds of disease associated with having two copies of the minor allele versus the odds of disease associated with having two copies of the major allele and were adjusted for age, sex, and pedigree structure. Reported \( P \) values are not adjusted for multiple comparisons. \( P \) values < 0.05 are shown in bold.

Table 2 shows the pairwise LD (\( D' \) and \( r^2 \)) among the genotyped SNPs in the Amish. The haplotype block structure in the Amish appears very similar to that reported in Finns and Ashkenazi Jews, suggesting that the SNP density we chose is adequate for the detection of most of the common variation in HNF4A. As shown in the Ashkenazi Jews and the Finns, the SNP representing the P2 haplotype block that was genotyped in the Amish (rs1884614) was

<table>
<thead>
<tr>
<th>rs1884614</th>
<th>rs2425637</th>
<th>rs2425640</th>
<th>rs3212183</th>
<th>rs1028583</th>
<th>rs3818247</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1884614</td>
<td>—</td>
<td>0.40</td>
<td>0.12</td>
<td>0.13</td>
<td>0.38</td>
</tr>
<tr>
<td>rs2425637</td>
<td>0.05</td>
<td>—</td>
<td>0.81</td>
<td>0.83</td>
<td>0.06</td>
</tr>
<tr>
<td>rs2425640</td>
<td>0</td>
<td>0.28</td>
<td>—</td>
<td>0.05</td>
<td>0.33</td>
</tr>
<tr>
<td>rs3212183</td>
<td>0.01</td>
<td>0.65</td>
<td>0.17</td>
<td>—</td>
<td>0.12</td>
</tr>
<tr>
<td>rs1028583</td>
<td>0.04</td>
<td>0</td>
<td>0.09</td>
<td>0.01</td>
<td>—</td>
</tr>
<tr>
<td>rs3818247</td>
<td>0.13</td>
<td>0.01</td>
<td>0.07</td>
<td>0.01</td>
<td>0.27</td>
</tr>
</tbody>
</table>

*Values in the upper right represent \( D' \), while values in the bottom left represent \( r^2 \). Shown in bold are the two SNPs, rs1884614 and rs2425640, in the P2 and P1 promoters, respectively, that were associated with type 2 diabetes and glucose traits.
clearly not in LD with rs2425640 in the P1 promoter or with the other HNF4A SNPs.

In addition to the case/control analysis, we genotyped rs1884614 and rs2425640 in an additional 217 nondiabetic Amish subjects to create an expanded set of 698 nondiabetic individuals (NGT [n = 568] and IGT [n = 130]) and performed an association analysis with diabetes-related quantitative traits. Figure 1 shows the mean plasma glucose levels at 30-min intervals during a 3-h oral glucose tolerance test (OGTT) according to genotype at the rs1884614 SNP. Carriers of the A “risk” allele exhibited higher total glucose area under the curve during the OGTT (additive model, \( P = 0.022 \); dominant model, \( P = 0.010 \)). There was no association with total insulin area under the curve during the OGTT (additive model, \( P = 0.919 \); dominant model, \( P = 0.919 \)).

In conclusion, we found that htSNPs in the P1 and P2 regions of HNF4A are associated with type 2 diabetes and diabetes-related traits in the Amish. Rs2425640 in the P1 region was also associated with type 2 diabetes in the Finns; however, contrary to our findings in the Amish, in which the A allele was the at-risk allele for both type 2 diabetes risk and an earlier onset of diabetes, the frequency of the G allele was significantly higher in Finnish subjects with type 2 diabetes. This discrepancy between the two populations may indicate that this SNP is not the functional SNP but is marking an at-risk haplotype that differs between the Amish and Finns. Of note, this SNP was not associated with type 2 diabetes in the Ashkenazi Jews, but others in the region were associated with type 2 diabetes, suggesting again that the SNPs thus far examined may be marking at-risk haplotypes in the different populations. Alternatively these discrepancies between popula-

**FIG. 1.** Mean plasma glucose (A) and insulin (B) levels at 30-min intervals during a 3-h 75-g OGTT according to SNP rs1884614 genotype groups. Carriers of the A “risk” allele exhibited higher total glucose area under the curve during the OGTT (additive model, \( P = 0.022 \); dominant model, \( P = 0.010 \)). There was no association with total insulin area under the curve during the OGTT (additive model, \( P = 0.919 \); dominant model, \( P = 0.919 \)).
tions could represent false-positive or false-negative results. Rs1884614, an hSNP in the P2 region of HNF4A was associated with glucose levels during an OGTT in the Amish and was also associated with type 2 diabetes in both the Ashkenazi Jews and the Finns. In all populations studied to date, the P1 and P2 SNPs reside in different haplotype blocks, suggesting the presence of two independent variants influencing type 2 diabetes risk. This replication across several studies lends further support to the possibility that variation in the P1 and P2 regions of HNF4A, or SNPs in strong LD with these regions, contributes to the pathogenesis of type 2 diabetes. Of note, our genome scan did not provide any evidence for linkage to type 2 diabetes or related traits to this region of chromosome 20 in the Amish (24). This observation is likely due to the relative insensitivity of linkage analysis compared with association analysis and suggests that this allele may also influence type 2 diabetes risk more broadly in other populations. Although HNF4A is the strongest candidate gene for type 2 diabetes in this region, the P2 SNPs reside in a large haplotype block that contains several other known and predicted genes and expressed sequence tags; therefore, the possibility must be considered that the pathogenic SNP(s) may reside in another gene. Further studies in other populations, as well as functional analysis, will be required to further define the role of variation in HNF4A in type 2 diabetes pathogenesis.

RESEARCH DESIGN AND METHODS

The AFDS was initiated in 1995 with the goal of identifying susceptibility genes for type 2 diabetes and related traits in a cohort of individuals from the Old Order Amish population in Lancaster County, Pennsylvania. Details of the AFDS design, recruitment, phenotyping, and pedigree structure have been described previously (25). Briefly, probands with previously diagnosed type 2 diabetes (onset between 35 and 65 years of age) and all first- and second-degree relatives of probands and spouses over the age of 18 were recruited. Phenotypic characterization of study participants included medical and family history, anthropometry, and a 3-h 75-g OGTT with insulin levels. The diagnosis of type 2 diabetes was defined on the basis of the OGTT using criteria of the American Diabetes Association (2-h glucose ≥11.1 mmol/l or fasting glucose ≥7 mmol/l), by current treatment with diabetes medications, or by a previous physician-documented diagnosis of diabetes. IGT was defined by a 2-h OGTT glucose between 7.8 and 11.1 mmol/l or fasting glucose ≤7 mmol/l. NGT was defined by a fasting glucose ≤7.8 mmol/l. The total glucose and insulin areas under the curve during the 3-h OGTT and insulin (fasting and insulin area under the curve during the 3-h OGTT) were estimated according to HNF4A genotypes in an expanded set of nondiabetic AFDS subjects (n = 698). To account for the relatedness among family members, the measured genotype approach was used (29), in which we estimated the likelihood of specific genetic models given the pedigree structure. Parameter estimates were obtained by maximum likelihood methods, and the significance of association was tested by likelihood ratio tests. Within each model, we simultaneously estimated the effects of age and sex. Insulin values were transformed by their natural logarithms (ln) to reduce skewness. Quantitative trait analyses were conducted using the SOLAR software (28).

ACKNOWLEDGMENTS

This work was supported by research grants R01-DK54261, K24-DK02673, U01-DK58026, and K07-CA67960; the University of Maryland General Clinical Research Center Grant M01 RR 16500; the General Clinical Research Centers Program; the National Center for Research Resources (NCRR); the National Institutes of Health; and the Baltimore Veterans Administration Geriatric Research and Education Clinical Center.

We gratefully acknowledge our Amish liaisons and field workers and the extraordinary cooperation and support of the Amish community, without whom these studies would not be possible.

REFERENCES


