Association Studies of Insulin Receptor Substrate 1 Gene (IRS1) Variants in Type 2 Diabetes Samples Enriched for Family History and Early Age of Onset

Eleftheria Zeggini,1,2 James Parkinson,3 Stephanie Halford,3 Katharine R. Owen,4 Timothy M. Frayling,4 Mark Walker,5 Graham A. Hitman,6 Jonathan C. Levy,1 Mike J. Sampson,7 Edith J.M. Feskens,8 Andrew T. Hattersley,4 and Mark I. McCarthy1,2,3

The gene encoding insulin receptor substrate-1 (IRS1) represents a strong biological candidate for a contributory role in type 2 diabetes susceptibility. Indeed, functional studies have implicated the G971R variant, and a recent meta-analysis of 27 association studies suggested that carriage of G971R was associated with a 25% increase in disease risk. However, this association has not been evaluated in large samples. The present study genotyped the P512A and G971R IRS1 variants in 971 U.K. type 2 diabetic subjects ascertained for strong family history and/or early onset, as well as 1,257 control subjects matched by ethnicity. There was no evidence for association with type 2 diabetes for either variant. (For example, the odds ratio [OR] for carriage of G971R was 1.11 [95% CI 0.86–1.44, \( P = 0.44 \).) An updated meta-analysis (31 studies; 5,104 case and 7,418 control subjects) remained significant for the G971R association (\( P = 0.025 \)), albeit with a diminished OR (1.15 [95% CI 1.02–1.31]). Additional studies of IRS1 variation will be required to obtain a robust estimate of the overall contribution of IRS1 variation to type 2 diabetes susceptibility, but the current study suggests that previous studies have overestimated the magnitude of this effect. Diabetes 53:3319–3322, 2004

From the 1Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, U.K.; the 2Wellcome Trust Centre for Human Genetics, Oxford, U.K.; the 3Imperial College Faculty of Medicine, London, U.K.; the 4Centre for Molecular Genetics, Peninsular Medical School, Exeter, U.K.; the 5School of Clinical Medical Sciences, University of Newcastle, Newcastle, U.K.; the 6Centre of Diabetes and Metabolic Medicine, Bart’s and The London, Queen Mary’s School of Medicine and Dentistry, London, U.K.; the 7Department of Diabetes and Endocrinology, Norfolk and Norwich University Hospital NHS Trust, Norwich, U.K.; and the 8Center of Nutrition and Health, National Institute for Public Health and the Environment, Bilthoven, the Netherlands.

Address correspondence and reprint requests to Mark McCarthy, Robert Turner Chair of Diabetes, Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital Site, Old Road, Headington, Oxford OX3 7LJ, U.K. E-mail: mark.mccarthy@drl.ox.ac.uk.

Received for publication 9 July 2004 and accepted in revised form 24 August 2004.

E.Z. and J.P. contributed equally to this study.

ECACC, European Collection of Cell Cultures; IRS, insulin receptor substrate; LD, linkage disequilibrium; SNP, single nucleotide polymorphism.

© 2004 by the American Diabetes Association.
with more severe disease and/or earlier age of diagnosis and thereby a greater genetic contribution to susceptibility (8,9).

Recent surveys of the performance of complex trait association studies have emphasized the overwhelming importance of sample size in making appropriate inferences (10) and the profound impact that publication bias can have on association meta-analysis (11). For both reasons, it is essential that the predominantly small studies of the IRS1 variants undertaken thus far are informed by the findings of the larger association studies that are now increasingly possible.

The current study has examined the role of the IRS1 P512A and G971R variants in genetic susceptibility to type 2 diabetes using two large, well-characterized case samples from the U.K. These were compared with 1,257 ethnically matched control subjects (see Research Design and Methods and Table 1).

Neither of the IRS1 single nucleotide polymorphisms (SNPs) deviated significantly from Hardy-Weinberg equilibrium in any group studied. The genotypic distributions of the P512A and G971R variants are depicted in Table 2. There was no significant association between IRS1 genotype and type 2 diabetes in either case-control comparison (Warren 2 probands versus European Collection of Cell Cultures [ECACC]; P512A \( P = 1.00 \), G971R \( P = 0.78 \); young-onset type 2 diabetic subjects versus the Exeter Family Study: P512A \( P = 0.74 \), G971R \( P = 0.19 \), exact \( P \)

<table>
<thead>
<tr>
<th>Sample group</th>
<th>Sample type</th>
<th>( n )</th>
<th>Age at diagnosis (cases) or last examination (control subjects)</th>
<th>BMI (kg/m(^2))</th>
<th>Waist-to-hip ratio</th>
<th>Percent male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young-onset type 2 diabetes</td>
<td>Case</td>
<td>406</td>
<td>39.6 ± 6.5</td>
<td>30.5 ± 5.0 (male), 33.5 ± 8.1 (female)</td>
<td>0.97 ± 0.06 (male), 0.89 ± 0.08 (female)</td>
<td>57.1</td>
</tr>
<tr>
<td>Exeter Family Study</td>
<td>Control</td>
<td>910</td>
<td>31.6 ± 5.4</td>
<td>26.7 ± 3.8 (male), 27.9 ± 4.8 (female)†</td>
<td>0.88 ± 0.06 (male), NA (female)†</td>
<td>49.9</td>
</tr>
</tbody>
</table>

Continuous data are means ± SD. *Age information was available for only 34% of subjects in this group. †BMI in female subjects was measured during pregnancy: meaningful measures of waist-to-hip ratio not available.

TABLE 2
P512A and G971R genotype distributions in the sample groups under study

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Warren 2 probands</th>
<th>ECACC</th>
<th>Young-onset type 2 diabetes</th>
<th>Exeter Family Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample type</td>
<td>Case</td>
<td>Control</td>
<td>Case</td>
<td>Control</td>
</tr>
<tr>
<td>512 PP</td>
<td>462 (96.3)</td>
<td>315 (96.3)</td>
<td>363 (96.8)</td>
<td>771 (96.4)</td>
</tr>
<tr>
<td>512 PA</td>
<td>18 (3.7)</td>
<td>12 (3.7)</td>
<td>12 (3.2)</td>
<td>29 (3.6)</td>
</tr>
<tr>
<td>512 AA</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>971 GG</td>
<td>479 (85.8)</td>
<td>296 (85.3)</td>
<td>332 (83.4)</td>
<td>766 (86.2)</td>
</tr>
<tr>
<td>971 GR</td>
<td>78 (14.0)</td>
<td>48 (13.8)</td>
<td>63 (15.8)</td>
<td>119 (13.4)</td>
</tr>
<tr>
<td>971 RR</td>
<td>1 (0.2)</td>
<td>3 (0.9)</td>
<td>3 (0.8)</td>
<td>4 (0.4)</td>
</tr>
</tbody>
</table>

Data are \( n \) (%). P512A: Warren 2 probands versus ECACC, genotype-wise \( P = 1.0 \), allele-wise \( P = 1.0 \); young-onset type 2 diabetes versus Exeter Family Study, genotype-wise \( P = 0.74 \), allele-wise \( P = 0.74 \); combined, genotype-wise \( P = 0.90 \), allele-wise \( P = 0.90 \). G971R: Warren 2 probands versus ECACC, genotype-wise \( P = 0.78 \), allele-wise \( P = 0.65 \); young-onset type 2 diabetes versus Exeter Family Study, genotype-wise \( P = 0.19 \), allele-wise \( P = 0.20 \); combined, genotype-wise \( P = 0.50 \), allele-wise \( P = 0.51 \). All \( P \) values are derived from exact implementations of standard contingency table tests: where three (genotype) classes were available, the Kruskal-Wallis test for genotype trend was used.
abolish, significance at the 5% level. In the augmented meta-analysis, which includes 5,104 case and 7,418 control subjects, carriage of the G971R variant appears to increase risk of type 2 diabetes by ~15% (combined OR 1.15 [1.02–1.31], P = 0.025). The previous meta-analysis found that the G971R association was restricted to those studies with hospital-based case ascertainment. We repeated the meta-analysis on this subset of studies, including the current datasets. There was no heterogeneity of ORs (23 studies; P = 0.28). Again, the magnitude of the overall OR was reduced in the augmented meta-analysis but remained significant (combined OR 1.28 [1.09–1.49], P = 0.002).

Using the regression asymmetry test (13), we found no significant evidence for study selection bias.

The key finding of the present study is that despite a sample size larger than that of any of the studies included in the previous meta-analysis, no evidence of association was seen between either IRS1 variant and type 2 diabetes. Although the present study comprises two separate case-control comparisons, similarities in the ascertainment schemes (e.g., case selection designed to enhance susceptibility–allele frequency), subject characteristics (e.g., ethnicity), and measured ORs support their combination into a single analysis. The impact of the present study is further enhanced by the case selection for strong family history and/or early onset of disease. These maneuvers have been shown to confer a substantial and beneficial effect on power to detect susceptibility effects (9,14). Taking this enhancement into account (explicitly in respect of membership of an affected sibship), the current study has ~85% power to detect (at the 5% level) an effect size equivalent to that found in the previous meta-analysis (OR 1.25) and >90% power to detect the higher effect size seen when analysis was restricted to hospital-based samples (OR 1.43).

Nevertheless, inclusion of these data (together with findings from a recent publication [12]) in an augmented meta-analysis did not abolish the combined evidence that G971R variation influences type 2 diabetes susceptibility. This is not altogether surprising given the relative size of the present study compared with the combined sample size in the previous meta-analysis. Even in the absence of study selection bias, a robust estimate of the true effect size associated with IRS1 variation is most likely to be obtained from larger, well-powered studies (10). In this regard, data from Florez et al. (15), indicating that analysis of >9,000 subjects failed to detect evidence of association with the G971R variant, corroborate our findings and indicate that the effect of this variant is likely to have been underestimated in the previous literature.

It is important to consider other potential explanations of the disparity between the current findings and several previous studies. Clinical heterogeneity is one possibility, but the indication from the previous meta-analysis that the consequences of G971R variation are restricted to hospital-based cases (as opposed to those ascertained from population samples) makes this unlikely, since the present study is clearly in the former category. Alternatively, population-specific associations could be the result of population-specific differences in the LD patterns between the targeted SNP and an unidentified etiological variant. There is no evidence that LD relationships within IRS1 are particularly diverse between closely related populations, but a recent study in Pima Indians (in whom the G971R variant was not detected) provided evidence for associations with other IRS1 SNPs (16). However, a recent well-powered study of IRS1 variation in type 2 diabetes that typed variants in codons 235 and 805, but not those in the present study, failed to detect any evidence of association (17).

Importantly, the G971R substitution has been shown to alter IRS-1 function in a variety of cell types (5–7). However, these experimental data are not inconsistent with a failure to detect an association with type 2 diabetes. In the whole organism, redundancy and developmental plasticity may provide compensation for isolated abnormalities of protein function. It is worth noting that Irs-1–deficient mice do not develop diabetes (18).

Ultimately, resolution of whether IRS1 variation truly influences susceptibility to type 2 diabetes will require several large studies on the scale of the present and other current reports, combined with a more comprehensive survey of variation within the gene. With the availability of large collaborative endeavors, information on the local LD structure, and improved genotyping methods, such a study should be feasible in the near future.

**RESEARCH DESIGN AND METHODS**

The present study included analysis of two case and two control groups. Clinical characteristics are provided in Table 1. In the first case-control analysis, we compared unrelated type 2 diabetic probands (n = 565) from the Diabetes U.K. Warren 2 sibpair repository (Warren 2 probands) (19) with 347 random U.K. population control samples from the ECACC (Salisbury, U.K.). For the second comparison, the case sample combines two sets of young-onset type 2 diabetic subjects (n = 406) with almost identical clinical characteristics: offspring from parent-offspring trios (n = 157) ascertained for type 2 diabetes (8) and young-onset (<45 years) type 2 diabetic subjects (n = 249) (20). These cases were compared with 910 unrelated parents from a consecutive birth cohort (the Exeter Family Study), 825 of whom were normoglycemic (20). All case samples are therefore strongly selected for inherited type 2 diabetes through early disease onset and/or positive family history. Other forms of diabetes were excluded using a combination of clinical, immunological, and genetic criteria (9,10). Confirmation of glycemic status in the control populations was limited to fasting plasma glucose measures for the Exeter Family Study samples. All subjects are of exclusively British/Irish origin. Informed consent was obtained from all subjects, and all studies were carried out in accordance with the principles of the Declaration of Helsinki (2000).

**SNP genotyping.** The G971R (rs1801278) SNP was genotyped using the Sequenom MASSARRAY system. Following PCR and dNTP removal with shrimp alkaline phosphatase, the sequencing oligonucleotide was extended using thermostable and the products resin-captured, arrayed, and analyzed in a Bruker Biflex III Mass Spectrometer according to the MASSARRAY protocol (Sequenom, San Diego, CA). The mass spectra were processed by an analysis module provided by the SpectroTypER software. The P512A (rs1801276) SNP was genotyped by Pyrosequencing (Biotage, Uppsala, Sweden). Following PCR, biotinylated amplicons were captured on magnetic beads and pyrosequenced on a PSQ960MA platform according to the manufacturer’s instructions. Primer sequences and PCR conditions are available from the authors upon request. Overall genotyping success rates were 99 and 98.4% for the P512A and G971R SNPs, respectively. Based on duplicate genotyping of ~20% of samples, we estimate the error rate to be <1%.

**Statistical analyses.** Both SNPs were tested for deviation from Hardy-Weinberg equilibrium in case and control subject separately using HelixTree (GoldenHelix, Bozeman, MT). Genotype and allele frequency distributions were compared by standard contingency table methods, and ORs with 95% CIs were calculated in Stata (version 8; Stata, College Station, TX). Genotype trend comparisons were also carried out using the Kruskal-Wallis test and the generalized Cochran-Mantel-Haenszel statistic. As appropriate, exact versions of these tests were implemented using StatXact 6 (Cytel Software, Cambridge, MA). Pairwise measures of LD were calculated and two-point haplotype frequencies estimated using the expectation-maximization algorithm in Helix-
Tree. Haplotypic associations were investigated by haplotype trend regression (21).

**Meta-analysis.** Estimates of combined ORs for meta-analyses were calculated using the fixed-effects Mantel-Haenszel method after testing for homogeneity of ORs using Stata. Meta-analyses included all studies previously reported (4), as well as the present study and one additional recent study (12). Systematic surveys of online publication databases failed to detect any relevant additional published studies of these variants. Evidence for study selection bias was sought using the unweighted regression asymmetry test (13).

**ACKNOWLEDGMENTS**

The work was supported by Diabetes U.K., who were also responsible for funding the collection of many of the samples used. A.T.H. is a Wellcome Trust Research Leave Fellow.

We thank J. Florez, D. Altshuler, and L. Groop for sharing data. We are grateful to Golden Helix for providing us with the HelixTree software package.

**REFERENCES**