Brief Genetics Report

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. McCarthy[1,2,3]

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 971R 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 971R 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 [1]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, U.K.; the [2]Wellcome Trust Centre for Human Genetics, Oxford, U.K.; the [3]Imperial College Faculty of Medicine, London, U.K.; the [4]Centre for Molecular Genetics, Peninsular Medical School, Exeter, U.K.; the [5]School of Clinical Medical Sciences, University of Newcastle, Newcastle, U.K.; the [6]Centre of Diabetes and Metabolic Medicine, Bart's and The London, Queen Mary's School of Medicine and Dentistry, London, U.K.; the [7]Department of Diabetes and Endocrinology, Norfolk and Norwich University Hospital NHS Trust, Norwich, U.K.; and the [8]Center 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.

Abnormalities in both insulin secretion and action contribute to the development of type 2 diabetes (1). In the search for genetic variants underlying the familial aggregation of defective insulin action that contributes to type 2 diabetes susceptibility (2), there has been an understandable focus on proximal components of the insulin signaling pathway, since functional variation in these molecules could readily explain the widespread disruption of postreceptor signaling that characterizes this condition.

Insulin receptor substrate (IRS)-1 plays a key role in proximal insulin signaling and has been extensively studied in this respect. While many sequence variants within the *IRS1* gene have been identified, the main focus has been on two relatively infrequent nonsynonymous variants, P512A (sometimes called P513A, rs1801276) and G971R (G972R, rs1801278), following an initial study that suggested that the combined prevalence of these variants was increased in type 2 diabetes (3). More than 25 studies, featuring diverse sizes and study designs and using samples recruited from many populations, have subsequently pursued these associations (4). Most attention has been concentrated on the G971R variant, which is more common than P512A (albeit with a minor allele frequency around 5%), and has stronger evidence supporting direct consequences on gene product function (5–7). Individual studies have reached apparently conflicting conclusions (4), although this diversity may be partly explained by questions of study design (notably small sample size) and differences in sample ascertainment and potential ethnic heterogeneity.

Combined analysis can help to address such confounding factors and provide more robust estimates of effect size: a recent meta-analysis of 27 association studies that had been published prior to 2000 (a total of 3,408 case and 5,419 control subjects) indicated that carriage of the 971R allele conferred a significant increase in disease risk (summary risk ratio 1.25 [95% CI 1.05–1.48]) (4). Interestingly, this association was entirely attributable to studies in which cases had been recruited from hospital-based clinics rather than population cohorts. This may reflect the likelihood that selective ascertainment results in samples

TABLE 1
Clinical characteristics of the populations studied

Sample group	Sample type	*n*	Age at diagnosis (cases) or last examination (control subjects) (years)	BMI (kg/m^2)	Waist-to-hip ratio	Percent male
Warren 2 probands	Case	565	55.4 ± 8.4	27.7 ± 4.2 (male), 30.0 ± 5.6 (female)	0.96 ± 0.07 (male), 0.87 ± 0.06 (female)	54.3
U.K. random (ECACC)	Control	347	38.5 ± 8.2*	NA	NA	53.1
Young-onset type 2 diabetes	Case	406	39.6 ± 6.5	30.5 ± 5.0 (male), 33.5 ± 8.1 (female)	0.97 ± 0.06 (male), 0.89 ± 0.08 (female)	57.1
Exeter Family Study	Control	910	31.6 ± 5.4	26.7 ± 3.8 (male), 27.9 ± 4.8 (female)†	0.88 ± 0.06 (male), NA (female)†	49.9

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.

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 values, Kruskal-Wallis test). Combined analysis (Cochran-Mantel-Haenszel) was similarly nonsignificant (P512A $P = 0.90$, G971R $P = 0.50$). The estimate of the common odds ratio (OR) for the association between 971R allele carriage and type 2 diabetes was 1.11 (95% CI 0.86–1.44, $P = 0.44$) and for 512A 0.94 (0.55–1.59, $P = 0.90$). The two SNPs were not found to be in linkage disequilibrium (LD) in any of the sample groups (pairwise $r^2 < 0.01$), and no significant haplotypic association with type 2 diabetes was identified in either case-control comparison (Warren 2 probands versus ECACC, $P = 0.99$; young-onset type 2 diabetic subjects versus the Exeter Family Study, $P = 0.18$) or the combined dataset ($P = 0.49$).

To examine the impact of these findings in light of the recent meta-analysis (4), we combined information from the 27 analyses included in the previous report with the data from the present study and from a further recent report of two G971R association studies that also identified no significant association with type 2 diabetes (12). (Systematic literature review identified no other relevant studies.) There was no evidence for heterogeneity of the ORs across the 31 studies ($P = 0.14$). Not surprisingly, given the overlap in the CIs for the OR estimate in the present study (0.86–1.44) and that from the previous meta-analysis (1.05–1.48), these estimates were not significantly different. However, inclusion of the four new case-control comparisons reduces the overall OR associated with the 971R variant and decreases, but does not

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

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

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 971R 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 >99% 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 overestimated 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 (9,19). 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 thermosequenase 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 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 PSQ96MA platform according to the manufacturer's instructions. Primer sequences and PCR conditions are available from the authors upon request. Overall genotyping success rates were 89 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

1. McCarthy MI, Froguel P: Genetic approaches to the molecular understanding of type 2 diabetes. *Am J Physiol* 283:E217–E225, 2002
2. Bergman RN, Zaccaro DJ, Watanabe RM, Haffner SM, Saad MF, Norris JM, Wagenknecht LE, Hokanson JE, Rotter JI, Rich SS: Minimal model-based insulin sensitivity has greater heritability and a different genetic basis than homeostasis model assessment or fasting insulin. *Diabetes* 52:2168–2174, 2003
3. Almind K, Bjørbaek C, Vestergaard H, Hansen T, Echwald S, Pedersen O: Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin dependent diabetes mellitus. *Lancet* 342:828–832, 1993
4. Jellema A, Zeegers MP, Feskens EJ, Dagnelie PC, Mensink RP: Gly972Arg variant in the insulin receptor substrate-1 gene and association with type 2 diabetes: a metaanalysis of 27 studies. *Diabetologia* 46:990–995, 2003
5. Almind K, Inoue G, Pedersen O, Kahn C: A common amino acid polymorphism in insulin receptor substrate-1 causes impaired insulin signalling: evidence from transfection studies. *J Clin Invest* 97:2569–2575, 1996
6. Hribal ML, Federici M, Porzio O, Lauro D, Borboni P, Accili D, Lauro R, Sesti G: The Gly→ Arg972 amino acid polymorphism in insulin receptor substrate-1 affects glucose metabolism in skeletal muscle cells. *J Clin Endocrinol Metab* 85:2004–2013, 2000
7. Porzio O, Federici M, Hribal ML, Lauro D, Accili D, Lauro R, Borboni P, Sesti G: The Gly972→Arg amino acid polymorphism in IRS-1 impairs insulin secretion in pancreatic beta cells. *J Clin Invest* 104:357–364, 1999
8. Frayling TM, Walker M, McCarthy MI, Evans JC, Ayres S, Allen LI, Ellard S, Lynn S, Turner RC, O'Rahilly S, Hitman GA, Hattersley AT: Parent-offspring trios: a resource to facilitate the identification of type 2 diabetes genes. *Diabetes* 48:2475–2479, 1999
9. Frayling TM, Wiltshire S, Hitman GA, Walker M, Levy JC, Sampson M, Groves CJ, Menzel S, McCarthy MI, Hattersley AT: Young-onset type 2 diabetes families are the major contributors to genetic loci in the Diabetes U.K. Warren 2 genome scan and identify putative novel loci on chromosomes 8q21, 21q22, and 22q11. *Diabetes* 57:1857–1863, 2003
10. Ioannidis JPA, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis JG: Genetic associations in large versus small studies: an empirical assessment. *Lancet* 361:567–571, 2003
11. Keavney B, McKenzie C, Parish S, Palmer A, Clark S, Youngman L, Delepine M, Lathrop M, Peto R, Collins R, the ISIS Collaborators: Large-scale test of hypothesised associations between the angiotensin-converting-enzyme insertion/deletion polymorphism and myocardial infarction in about 5000 cases and 6000 controls. *Lancet* 355:34–442, 2000
12. Van Dam RM, Hoebee B, Seidell JC, Schaap MM, Blaak EE, Feskens EJ: The insulin receptor substrate-1 Gly972Arg polymorphism is not associated with type 2 diabetes mellitus in two population-based studies. *Diabet Med* 21:752–758, 2004
13. Egger M, Davey Smith G, Schneider M, Minder C: Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315:629–934, 1997
14. Risch N, Teng J: The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases I DNA pooling. *Genome Res* 8:1273–1288, 1998
15. Florez JC, Sjögren M, Burtt N, Orho-Melander M, Schayer S, Sun M, Almgren P, Lindblad U, Tuomi T, Gaudet D, Hudson TJ, Daly MJ, Ardlie KG, Hirschhorn JN, Altshuler D, Groop L: Association testing in 9,000 people fails to confirm the association of the insulin receptor substrate-1 G972R polymorphism with type 2 diabetes. *Diabetes* 53:3313–3318, 2004
16. Kovacs P, Hanson RL, Lee YH, Yang X, Kobes S, Permana PA, Bogardus C, Baier LJ: The role of insulin receptor substrate-1 gene (IRS1) in type 2 diabetes in Pima Indians. *Diabetes* 52:3005–3009, 2003
17. 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
18. Withers DJ, Burks DJ, Towery HH, Altamuro SL, Flint CL, White MF: Irs-2 coordinates Igf-1 receptor-mediated beta-cell development and peripheral insulin signalling. *Nat Genet* 23:32–40, 1999
19. Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M, O'Rahilly S, Frayling TM, Bell JI, Lathrop GM, Bennett A, Dhillon R, Fletcher C, Groves CJ, Jones E, Prestwich P, Simecek N, Subba Rao PV, Wishart M, Foxon R, Howell S, Smedley D, Cardon LR, Menzel S, McCarthy MI: A genome-wide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes U.K. Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. *Am J Hum Genet* 69:553–569, 2001
20. Minton J, Hattersley AT, Owen KR, McCarthy MI, Walker M, Barrett T, Frayling TM: Association studies of genetic variation in the WFS1 gene and type 2 diabetes in U.K. populations. *Diabetes* 51:1287–1290, 2002
21. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG: Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 53:79–91, 2002