

# High-Density Haplotype Structure and Association Testing of the Insulin-Degrading Enzyme (*IDE*) Gene With Type 2 Diabetes in 4,206 People

Jose C. Florez,<sup>1,2,3,4</sup> Steven Wiltshire,<sup>5</sup> Christina M. Agapakis,<sup>1,3</sup> Noël P. Burt,<sup>3</sup> Paul I.W. de Bakker,<sup>1,3,6</sup> Peter Almgren,<sup>7</sup> Kristina Bengtsson Boström,<sup>8</sup> Tiinamaija Tuomi,<sup>9</sup> Daniel Gaudet,<sup>10</sup> Mark J. Daly,<sup>2,3</sup> Joel N. Hirschhorn,<sup>3,6,11</sup> Mark I. McCarthy,<sup>5,12</sup> David Altshuler,<sup>1,2,3,4,6</sup> and Leif Groop<sup>7</sup>

The insulin-degrading enzyme is responsible for the intracellular proteolysis of insulin. Its gene *IDE* is located on chromosome 10, in an area with suggestive linkage to type 2 diabetes and related phenotypes. Due to the impact of genetic variants of this gene in rodents and the function of its protein product, it has been proposed as a candidate gene for type 2 diabetes. Various groups have explored the role of the common genetic variation of *IDE* on insulin resistance and reported associations of various single nucleotide polymorphisms (SNPs) and haplotypes on both type 2 diabetes and glycemic traits. We sought to characterize the haplotype structure of *IDE* in detail and replicate the association of common variants with type 2 diabetes, fasting insulin, fasting glucose, and insulin resistance. We assessed linkage disequilibrium, selected single-marker and multimarker tags, and genotyped these markers in several case-control and family-based samples totalling 4,206 Caucasian individuals. We observed no statistically significant evidence of association between single-marker or multimarker tests in *IDE* and type 2 diabetes. Nominally significant differences in quantitative

traits are consistent with statistical noise. We conclude that common genetic variation at *IDE* is unlikely to confer clinically significant risk of type 2 diabetes in Caucasians. *Diabetes* 55:128–135, 2006

The insulin-degrading enzyme (IDE or insulysin, EC 3.4.24.56) is a ~110-kDa member of the M16 family of zinc-metalloendopeptidases involved in the proteolysis of various amyloidogenic peptides such as insulin, glucagon, amyloid  $\beta$ -protein, amylin, and atrial natriuretic protein (rev. in 1). Due to its role in intracellular insulin degradation, it has been postulated as a possible contributor to insulin resistance in humans.

Several lines of evidence implicate the *IDE* gene in the pathogenesis of type 2 diabetes. First, IDE activity is reduced by 30% in the diabetic Goto-Kakizaki rat model, and two amino acid substitutions in the *Ide* gene have been shown to be responsible for postprandial hyperglycemia in congenic strains (2). Second, *Ide*-null mice display hyperinsulinemia and glucose intolerance (3). Third, the human *IDE* gene is located on chromosome 10q23-24, within 4–30 Mb from markers that have shown suggestive linkage for type 2 diabetes and related phenotypes in a number of whole-genome linkage studies (4–7). Fourth, several association studies have reported nominal association of *IDE* single nucleotide polymorphisms (SNPs) or haplotypes with type 2 diabetes and/or glycemic traits (8–10).

In type 2 diabetes, robust reproducible associations continue to be documented for the P12A polymorphism in the peroxisome proliferator-activated receptor  $\gamma$  (encoded by *PPARG*) and the E23K polymorphism in *KCNJ11*, which encodes the ATP-sensitive  $K^+$  channel Kir6.2 (recently rev. in 11). We aim to provide similar levels of evidence for other intriguing reports of association in high-likelihood candidate genes; as previously noted (12), when adequately powered samples are studied, true genetic associations are often replicated, adding to the statistical significance of a reported association. We therefore set out to perform a comprehensive test of common variation in the *IDE* gene for association with type 2 diabetes and related traits in a large set of both case-control and family-based samples.

From the <sup>1</sup>Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts; the <sup>2</sup>Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts; the <sup>3</sup>Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts; the <sup>4</sup>Department of Medicine, Harvard Medical School, Boston, Massachusetts; the <sup>5</sup>Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, U.K.; the <sup>6</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts; the <sup>7</sup>Department of Endocrinology, University Hospital MAS, Lund University, Malmö, Sweden; the <sup>8</sup>Department of Clinical Science, University Hospital MAS, Lund University, Malmö, Sweden; the <sup>9</sup>Department of Medicine, Helsinki University Central Hospital; Folkhalsan Genetic Institute, Folkhalsan Research Center, and Research Program for Molecular Medicine, University of Helsinki, Helsinki, Finland; the <sup>10</sup>Chicoutimi Hospital, University of Montreal Community Genomic Center, Quebec, Canada; the <sup>11</sup>Divisions of Genetics and Endocrinology, Children's Hospital, Boston, Massachusetts; and the <sup>12</sup>Oxford Centre for Diabetes, Endocrinology, and Metabolism, University of Oxford, Oxford, U.K.

Address correspondence and reprint requests to Jose C. Florez, Diabetes Unit, Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114. E-mail: jcflorez@partners.org.

Received for publication 26 July 2005 and accepted in revised form 4 October 2005.

D.A. and L.G. jointly supervised this project.

L.G. has been a member of an advisory board for and has received consulting fees from Aventis-Sanofi, Bristol-Myers Squibb, Kowa, and Roche.

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

AUC, area under the curve; CEPH, Centre d'Etude du Polymorphisme Humain; HOMA-IR, homeostasis model assessment of insulin resistance; IDE, insulin-degrading enzyme; SNP, single nucleotide polymorphism.

© 2006 by the American Diabetes Association.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

TABLE 1  
Clinical characteristics of patient samples

Sample	Sex (M/F)	Age (years)	BMI (kg/m <sup>2</sup> )	Fasting plasma glucose (mmol/l)	HbA <sub>1c</sub> (%)* or plasma glucose at 2-h OGTT (mmol/l)†
Scandinavian trios					
Affected probands	168/153	39 ± 9	27 ± 5	7.2 ± 2.6	8.5 ± 2.9†
Parents	236/236				
Sibships					
Diabetes/severe impaired glucose tolerance sib	280/329	65 ± 10	29 ± 5	9.3 ± 3.3	14.3 ± 5.6†
Normal glucose tolerance sib	275/305	62 ± 10	26 ± 3	5.4 ± 0.4	6.0 ± 1.1†
Scandinavia C/C					
Diabetes/severe impaired glucose tolerance	252/219	60 ± 10	28 ± 5	9.8 ± 3.4	15.0 ± 5.3†
Normal glucose tolerance	254/217	60 ± 10	27 ± 4	6.2 ± 1.8	6.8 ± 2.8†
Sweden C/C					
Diabetes/severe impaired glucose tolerance	267/247	66 ± 12	28 ± 4	9.6 ± 2.9	6.5 ± 1.5*
Normal glucose tolerance	267/247	66 ± 12	28 ± 4	5.5 ± 0.7	ND
Canada C/C					
Diabetes	70/57	53 ± 8	29 ± 5	6.4 ± 1.8	12.8 ± 2.1†
Normal glucose tolerance	70/57	52 ± 8	29 ± 4	5.1 ± 0.6	6.1 ± 1.1†

Data are means ± SD. Plasma glucose was measured at baseline (fasting) and 2 h after an oral glucose tolerance test (OGTT). C/C, case/control. ND, not determined.

## RESEARCH DESIGN AND METHODS

**Clinical samples.** The diabetic samples are presented in Table 1 and have been described elsewhere (13,14). Briefly, they comprise 321 Scandinavian trios; 1,189 Scandinavian siblings discordant for type 2 diabetes; a Scandinavian case-control sample totaling 942 subjects individually matched for age, BMI, and geographic region; a case-control sample from Sweden totaling 1,028 subjects who were individually matched for sex, age, and BMI; and an individually matched case-control sample totaling 254 subjects from the Saguenay Lac-St. Jean region in Quebec, Canada. We note that these samples have been validated by the replication of the two most widely reproduced associations in type 2 diabetes, *PPARG* P12A (13) and *KCNJ11* E23K (15) and by the overlap with other groups' findings in the promoter region of the hepatocyte nuclear factor 4 $\alpha$  (14).

**Quantitative trait measurements.** A 75-g oral glucose tolerance test was performed in a subset of the control Scandinavian subjects ( $n = 857$ , 435 male). Plasma glucose was measured by a glucose oxidase method on a Beckman Glucose Analyzer (Beckman Instruments, Fullerton, CA). Fasting insulin was measured by radioimmunoassay. Insulin resistance by homeostasis model assessment (HOMA-IR) was calculated as in Matthews et al. (16). Insulin area under the curve (AUC) during the oral glucose tolerance test was calculated by the trapezoidal method. Because of nonnormality, fasting insulin, HOMA-IR, insulin AUC, and BMI were logarithmically transformed.

**Genotyping.** DNA from all samples underwent whole-genome amplification with the protocol developed by Molecular Staging (17). We have recently confirmed the robust fidelity and genome representation of this technology using direct sequencing and high-density oligonucleotide arrays probing >10,000 SNP alleles (18). Genotyping was performed by primer extension of multiplex products with detection by matrix-assisted laser desorption ionization-time of flight mass spectroscopy using a Sequenom platform as previously described (15). Our genotyping success rate was 94.6% and our consensus rate was 99.3%, based on 15,487 (17%) duplicate genotypes. The detailed list of SNPs and their genotyping assays can be found in supplementary Table 1 in the online appendix (available at <http://diabetes.diabetesjournals.org>). Genotype counts for the various samples tested in this study can be found in supplementary Table 2 in the online appendix; both Tables are also posted on our website ([http://genetics.mgh.harvard.edu/AltshulerWeb/publicationdata/Florez\\_IDE.html](http://genetics.mgh.harvard.edu/AltshulerWeb/publicationdata/Florez_IDE.html)).

**Haplotype structure.** To evaluate the haplotype structure of the *IDE* gene at high density, we first downloaded data for 35 SNPs genotyped in the 120 chromosomes from a multigenerational panel of Centre d'Etude du Polymorphisme Humain (CEPH) Caucasian pedigrees by the HapMap project (19). We targeted a segment that would begin ~20 kb upstream of the *IDE* transcription start site and end ~10 kb downstream from the end of the 3' untranslated region and expanded this region in both directions until we noted decay of linkage disequilibrium, as evidenced by the end of a haplotype block (defined as in Gabriel et al. [20]). We then selected 23 other SNPs previously examined for association with type 2 diabetes and related traits in this region (8–10) and genotyped them in the same panel. Finally, 185 additional SNPs were added at

regularly spaced intervals to refine areas of low SNP density or clarify the extent of linkage disequilibrium in the CEPH reference sample. The position of all SNPs in each category is displayed in Fig. 1.

These SNPs span 224 kb, from ~59 kb upstream of the transcription start site to ~44 kb downstream of the end of the *IDE* 3' untranslated region. Three SNPs failed assay design, 32 of the SNPs attempted were technical failures (<75% genotyping percentage), 6 failed Hardy-Weinberg equilibrium ( $P < 0.01$ ), 1 had more than two Mendel errors, and 80 were found to be monomorphic in this population; thus, the final set comprised 121 (50%) working SNPs that were used to evaluate patterns of linkage disequilibrium. The average spacing between these 121 SNPs is 1.8 kb. Case and control chromosomes were phased together using the expectation-maximization algorithm of Excoffier and Slatkin (21) modified to process larger data files using the partition-ligation approach as previously described (15).

**Tag SNP selection.** To correlate our findings with those of the literature, we chose to genotype all working, previously associated SNPs in our disease panels. We selected additional tag SNPs from the CEPH reference sample by a newly developed algorithm named Tagger (available at <http://www.broad.mit.edu/mpg/tagger/>) (22), which selects tag SNPs to construct single-marker and multimarker (haplotype) tests to capture alleles of interest based on the computed correlation  $r^2$  between them (in this case, set at  $r^2 > 0.8$  with a logarithm of odds score >2.0). This method also allows the user to "force in" markers of interest, such that the additional tags selected capture those alleles not already captured by the forced-in set. This procedure resulted in 21 single-marker and 13 multimarker tags, which were carried forward in the disease samples (Fig. 1; supplementary Table 3 of the online appendix). Together, these 34 tags constitute the specific tests of association between the trait and themselves (as well as the other variants captured by them). The small number of tests minimizes both the genotyping burden and the number of statistical comparisons. We did not attempt to capture rare variants (minor allele frequency <5%) because our sample size did not have enough power to detect modest genetic effects at low allele frequencies (see below).

**Statistical analysis.** Power calculations were performed with the program of Purcell et al. (23) (available at <http://pngu.mgh.harvard.edu/~purcell/gpc/>). To examine the association of SNPs and haplotypes with type 2 diabetes, we used simple  $\chi^2$  analysis in the case-control samples, the transmission disequilibrium test (24) in the diabetic trios, and the discordant allele test (25) in the sibpairs; the first two have been implemented in Haploview (available at <http://www.broad.mit.edu/mpg/haploview/>) for both single-marker and multimarker association testing (26). Possible ambiguity in haplotype assignments due to incomplete data were accounted for by estimating haplotype frequencies in the tests of disease association (supplementary Table 3 of the online appendix); data completeness was ~97.9% after phasing. Each multimarker haplotype test was assessed against all other possibilities at the corresponding locus. Results from the various samples (and where relevant, from available previously published genotypes) were combined by Mantel-Haenszel meta-analysis of the odds ratios (12). Homogeneity of odds ratios among study samples was tested using a Pearson  $\chi^2$  goodness of fit as previously described (12).



**FIG. 1.** Linkage disequilibrium (LD) plot across the *IDE* locus. The horizontal black line depicts the 224-kb DNA segment of chromosome 10q23-24 analyzed in our CEPH sample. The locations of 243 SNPs are indicated above the black line: the 42 SNPs that failed are depicted in black, the 80 monomorphic SNPs are in green, the 35 HapMap SNPs are in red, and the 121 working SNPs are in blue. Their rs designation, genotyping percentage (scale, 75–100%) and minor allele frequency are shown immediately below the black line. A linkage disequilibrium plot is depicted in the bottom part of the figure based on the measure  $D'$ : each square represents the magnitude of linkage disequilibrium for a single pair of markers, with red color indicating linkage disequilibrium that is strong ( $D' > 0.8$ ) and statistically significant (logarithm of odds  $> 2.0$ ). The haplotypes spanning this block are shown above the linkage disequilibrium plot, with the thickness of the blue line indicating their frequency in the CEPH reference sample (figure prepared using the program LocusView v2.0, T. Petryshen, A. Kirby, M. Ainscow, unpublished software).

**Quantitative trait comparisons.** Only nondiabetic subjects were included in our analyses of diabetes-related phenotypes. Their chromosomes were phased as above, and individuals were sorted by their diploid genotypes at each locus; each most likely inferred multimarker haplotype test was compared against all other possibilities at the corresponding loci. Fasting plasma glucose, fasting insulin levels, and HOMA-IR were compared by ANOVA across the three genotypic groups for each marker. To perform direct comparisons with the findings of Karamohamed et al. (9) and Gu et al. (10), haplotypes were reconstructed from the SNPs based on the linkage disequilibrium structure in the reference panel. We corrected our best nominal  $P$  value for the multiple hypotheses examined by performing 1,000 permutations and obtaining an empirical  $P$  value based on the number of times our best  $F$  statistic was exceeded in the resampling procedure.

**RESULTS**

**Characterization of common sequence variation at *IDE*.** Consistent with previous studies performed at lower SNP density (8–10,27–29), the *IDE* region is characterized

by a high degree of linkage disequilibrium and reduced haplotype diversity (Fig. 1). From the initial set of 121 working SNPs, the 106 common variants with minor allele frequency  $> 5\%$  can be well captured with  $r^2 > 0.8$  by a set of 22 tag SNPs combined in 22 single-marker and 13 multimarker tests. One SNP (rs6583826) did not genotype well in the disease samples; nevertheless, dropping this tag SNP from the tag set still allowed us to capture the remaining SNPs with minor allele frequency  $> 5\%$  with a minimum  $r^2 > 0.8$ , with the sole exception of rs6583826 itself, which was captured at  $r^2 = 0.738$ . We therefore carried these 21 tag SNPs forward in our disease samples (Table 2).

**Association study.** To estimate the sample size required to replicate the main single-SNP association with type 2

TABLE 2  
Association study of single-SNP and multimarker haplotypes in *IDE* with type 2 diabetes

Test	Single tag SNP test	Position	Alleles	Minor allele frequency	Odds ratio (95% CI)	<i>P</i>
1	rs2263638	94158777	C/T	0.34	1.02 (0.92–1.13)	0.73
2	rs2209972*	94169008	C/T	0.31	0.98 (0.87–1.09)	0.66
3	rs967878*	94169328	C/A	0.47	1.06 (0.95–1.18)	0.27
4	rs884526*	94170442	G/T	0.31	0.98 (0.88–1.10)	0.76
5	rs12257053	94175650	T/C	0.11	1.11 (0.94–1.32)	0.22
6	rs11186994	94184819	C/T	0.31	0.97 (0.87–1.07)	0.51
7	rs10882063	94189317	G/T	0.43	1.07 (0.96–1.19)	0.24
8	rs2251101*	94201284	T/C	0.23	0.95 (0.85–1.07)	0.43
9	rs4646958	94204339	T/A	0.09	0.93 (0.78–1.10)	0.39
10	rs1887922*	94214145	T/C	0.18	0.97 (0.84–1.12)	0.69
11	rs2275218*	94215277	T/C	0.07	0.97 (0.76–1.24)	0.82
12	rs2250090*	94221786	G/A	0.02	1.14 (0.84–1.54)	0.39
13	rs2249960*	94223100	A/G	0.09	1.05 (0.87–1.27)	0.63
14	rs11187031	94251771	T/G	0.25	1.08 (0.96–1.21)	0.21
15	rs4641376	94275508	C/A	0.18	1.03 (0.90–1.18)	0.64
16	rs10882083	94311953	C/G	0.40	1.01 (0.91–1.12)	0.85
17	rs12355977	94317963	G/C	0.10	1.08 (0.92–1.27)	0.35
18	rs4646953*	94323935	A/G	0.21	1.09 (0.97–1.23)	0.16
19	rs10882086	94355005	T/C	0.10	1.00 (0.84–1.20)	0.97
20	rs12252836	94366308	C/A	0.05	1.02 (0.81–1.29)	0.87
21	rs12260688	94369694	C/T	0.03	0.92 (0.67–1.26)	0.61

  

Test	Multimarker tag SNP test	Haplotype	Frequency	Odds ratio (95% CI)	<i>P</i>
22	rs12260688, rs12252836, rs2275218	C,C,T	0.86	0.98 (0.84–1.14)	0.77
23	rs12252836, rs2251101	A,C	0.04	0.94 (0.74–1.20)	0.64
24	rs4646953, rs1887922	A,T	0.61	1.03 (0.93–1.15)	0.54
25	rs4646953, rs4641376	G,A	0.08	0.94 (0.78–1.14)	0.56
26	rs10882083, rs2251101	G,T	0.35	0.99 (0.89–1.10)	0.91
27	rs10882083, rs1887922	C,T	0.42	1.00 (0.90–1.11)	0.96
28	rs4641376, rs4646958	A,A	0.09	1.10 (0.92–1.32)	0.29
29	rs4641376, rs11187031, rs2275218	C,T,T	0.58	1.00 (0.90–1.11)	0.97
30	rs4641376, rs11187031	C,T	0.65	1.02 (0.91–1.13)	0.76
31	rs2275218, rs4646958	T,T	0.84	0.90 (0.78–1.04)	0.15
32	rs1887922, rs10882063	C,G	0.16	1.03 (0.89–1.18)	0.70
33	rs10882063, rs2209972	T,T	0.29	1.00 (0.90–1.13)	0.94
34	rs11186994, rs2209972	C,C	0.69	0.96 (0.86–1.06)	0.41

Twenty-one tag SNPs (*top panel*) and 13 multimarker haplotypes defined by these SNPs (*bottom panel*) were tested for association with type 2 diabetes in our samples. Results from the various samples were combined by Mantel-Haenszel meta-analysis of the odds ratios. Chromosomal position is according to the NCBI 35 release; alleles and odds ratios of individual SNPs are reported as major versus minor allele. \*SNPs previously associated with type 2 diabetes or related phenotypes (8–10).

diabetes of Groves et al. (8) (SNP rs4646953), we assumed a minor allele frequency of 23%, a type 2 diabetes disease prevalence of 8%, and a genotypic relative risk of 1.2 in a multiplicative model. Under these parameters, we estimated that our combined sample of 1,112 case-control pairs would provide >80% power to reject the null hypothesis of no association at  $P < 0.05$ ; this power is further raised by the inclusion of the family-based samples.

Association results for each of the 21 single-marker and 13 multimarker tests are presented in Table 2. No association was observed to any of the SNPs or haplotypes spanning *IDE* with type 2 diabetes in the overall sample. Heterogeneity among subsamples (at  $P < 0.05$ ) was observed for 3 of the 21 single-marker tests (rs12257053, rs4646953, and rs4646958) and 2 of the 13 multimarker tests (tests 28 and 31). Exclusion of the small Canadian sample did not change the heterogeneity results.

It is possible that overmatching in our case-control panels may have prevented us from detecting a true effect on risk of type 2 diabetes, if this effect was mediated through BMI. We therefore assessed whether BMI differed

across genotypes for all 34 tests in our control sample. We obtained nominal  $P$  values for rs967878, rs10882063, and test 34 of 0.03, 0.02, and 0.049, respectively; these  $P$  values became nonsignificant on permutation testing ( $P = 0.25$ ), suggesting that these common variants are unlikely to confer a significant effect on BMI.

**Genotype-phenotype correlations.** We next examined whether variation in glycemic traits might be associated with any of the single-marker and multimarker tests in *IDE*. To limit the number of comparisons and correlate our findings with those of Karamohamed et al. (9) and Gu et al. (10), we restricted our analysis to fasting plasma glucose, fasting insulin levels, and fasting insulin normalized for ambient glucose as reflected in HOMA-IR. Four single-marker tests (rs4641376 and rs12260688 for fasting plasma glucose and rs12355977 for both fasting insulin and HOMA-IR) and two multimarker tests (tests 22 and 24, both for fasting insulin) showed nominal  $P$  values <0.05 (Table 3). The best nominal  $P$  value was 0.009 for rs12260688 and fasting plasma glucose; permutation testing to correct for the multiple hypotheses examined

TABLE 3  
Quantitative traits according to IDE single-marker and multimarker genotypes

Test	SNP	Fasting glucose (mmol/l)			log(fasting insulin) (pmol/l)			log(HOMA-IR)		
		M/M	M/m	P	M/M	M/m	P	M/M	M/m	P
1	rs2263638	5.49 ± 0.55	5.47 ± 0.55	NS	0.88 ± 0.24	0.84 ± 0.25	NS	0.26 ± 0.25	0.23 ± 0.27	NS
2	rs2209972	5.49 ± 0.56	5.48 ± 0.53	NS	0.86 ± 0.25	0.86 ± 0.24	NS	0.25 ± 0.26	0.24 ± 0.25	NS
3	rs967878	5.53 ± 0.57	5.46 ± 0.53	NS	0.87 ± 0.24	0.85 ± 0.24	NS	0.26 ± 0.26	0.24 ± 0.25	NS
4	rs884526	5.49 ± 0.56	5.49 ± 0.54	NS	0.86 ± 0.25	0.85 ± 0.24	NS	0.25 ± 0.26	0.24 ± 0.25	NS
5	rs12257053	5.49 ± 0.55	5.44 ± 0.53	NS	0.86 ± 0.23	0.84 ± 0.28	NS	0.25 ± 0.25	0.22 ± 0.29	NS
6	rs11186994	5.49 ± 0.56	5.50 ± 0.54	NS	0.86 ± 0.25	0.85 ± 0.24	NS	0.25 ± 0.27	0.24 ± 0.25	NS
7	rs10882063	5.49 ± 0.57	5.47 ± 0.53	NS	0.87 ± 0.23	0.85 ± 0.24	NS	0.26 ± 0.25	0.24 ± 0.26	NS
8	rs2251101	5.49 ± 0.53	5.50 ± 0.60	NS	0.86 ± 0.25	0.86 ± 0.24	NS	0.24 ± 0.26	0.25 ± 0.25	NS
9	rs4646958	5.49 ± 0.56	5.47 ± 0.52	NS	0.86 ± 0.24	0.85 ± 0.27	NS	0.25 ± 0.25	0.24 ± 0.28	NS
10	rs1887922	5.46 ± 0.55	5.51 ± 0.54	NS	0.86 ± 0.25	0.85 ± 0.24	NS	0.24 ± 0.26	0.24 ± 0.25	NS
11	rs2275218	5.49 ± 0.54	5.49 ± 0.55	NS	0.86 ± 0.25	0.87 ± 0.20	NS	0.24 ± 0.26	0.25 ± 0.21	NS
12	rs2250090	5.49 ± 0.54	5.47 ± 0.75	NS	0.86 ± 0.25	0.87 ± 0.23	NS	0.24 ± 0.26	0.26 ± 0.27	NS
13	rs2249960	5.47 ± 0.54	5.55 ± 0.55	NS	0.86 ± 0.25	0.87 ± 0.23	NS	0.24 ± 0.26	0.26 ± 0.25	NS
14	rs11187031	5.52 ± 0.54	5.44 ± 0.56	NS	0.86 ± 0.25	0.86 ± 0.22	NS	0.25 ± 0.26	0.24 ± 0.23	NS
15	rs4641376	5.48 ± 0.55	5.53 ± 0.55	0.046	0.86 ± 0.24	0.86 ± 0.26	NS	0.24 ± 0.25	0.25 ± 0.27	NS
16	rs10882083	5.50 ± 0.55	5.50 ± 0.55	NS	0.87 ± 0.26	0.85 ± 0.24	NS	0.26 ± 0.27	0.23 ± 0.25	NS
17	rs12355977	5.49 ± 0.54	5.43 ± 0.57	NS	0.86 ± 0.25	0.85 ± 0.22	NS	0.25 ± 0.26	0.23 ± 0.23	0.04
18	rs4646953	5.50 ± 0.55	5.48 ± 0.54	NS	0.86 ± 0.24	0.86 ± 0.23	NS	0.25 ± 0.25	0.24 ± 0.24	NS
19	rs10882086	5.49 ± 0.55	5.46 ± 0.54	NS	0.86 ± 0.24	0.87 ± 0.28	NS	0.24 ± 0.25	0.25 ± 0.29	NS
20	rs12252836	5.49 ± 0.54	5.40 ± 0.65	NS	0.86 ± 0.25	0.86 ± 0.23	NS	0.25 ± 0.26	0.24 ± 0.25	NS
21	rs12260688	5.47 ± 0.54	5.69 ± 0.61	0.009	0.86 ± 0.25	0.87 ± 0.22	NS	0.24 ± 0.26	0.27 ± 0.24	NS
Multimarker		1 1	1 2	2 2	1 1	1 2	2 2	1 1	1 2	2 2
22	21, 20, 11	5.49 ± 0.54	5.48 ± 0.59	NS	0.85 ± 0.25	0.90 ± 0.24	NS	0.24 ± 0.26	0.27 ± 0.25	NS
23	20, 8	4.81 ± 0.40	5.43 ± 0.68	NS	1.01 ± 0.19	0.90 ± 0.21	NS	0.41 ± 0.18	0.31 ± 0.22	NS
24	18, 10	5.50 ± 0.57	5.47 ± 0.55	NS	0.88 ± 0.22	0.84 ± 0.28	NS	0.27 ± 0.24	0.22 ± 0.29	NS
25	18, 15	5.17 ± 0.48	5.51 ± 0.56	NS	1.04 ± 0.15	0.88 ± 0.23	NS	0.42 ± 0.16	0.26 ± 0.24	NS
26	16, 8	5.52 ± 0.51	5.48 ± 0.56	NS	0.88 ± 0.27	0.84 ± 0.23	NS	0.27 ± 0.27	0.22 ± 0.24	NS
27	16, 10	5.42 ± 0.56	5.52 ± 0.54	NS	0.88 ± 0.27	0.85 ± 0.24	NS	0.26 ± 0.28	0.24 ± 0.26	NS
28	15, 9	5.54 ± 0.51	5.46 ± 0.52	NS	0.87 ± 0.12	0.86 ± 0.27	NS	0.26 ± 0.13	0.25 ± 0.28	NS
29	15, 14, 11	5.53 ± 0.55	5.49 ± 0.57	NS	0.86 ± 0.24	0.86 ± 0.24	NS	0.25 ± 0.26	0.25 ± 0.25	NS
30	15, 14	5.53 ± 0.55	5.47 ± 0.57	NS	0.86 ± 0.24	0.86 ± 0.24	NS	0.25 ± 0.25	0.25 ± 0.26	NS
31	11, 9	5.50 ± 0.56	5.47 ± 0.54	NS	0.86 ± 0.25	0.86 ± 0.25	NS	0.24 ± 0.26	0.24 ± 0.26	NS
32	10, 7	5.61 ± 0.49	5.52 ± 0.55	NS	0.88 ± 0.23	0.85 ± 0.23	NS	0.28 ± 0.24	0.24 ± 0.25	NS
33	7, 2	5.48 ± 0.53	5.51 ± 0.53	NS	0.88 ± 0.25	0.85 ± 0.24	NS	0.26 ± 0.26	0.24 ± 0.25	NS
34	6, 2	5.50 ± 0.56	5.49 ± 0.54	NS	0.86 ± 0.25	0.85 ± 0.24	NS	0.25 ± 0.27	0.24 ± 0.25	NS

Data are means ± SD, where SD = 0 indicates a single observation. Fasting plasma glucose (mmol/l), fasting insulin (pmol/l), and insulin resistance by HOMA-IR were determined in 857 nondiabetic Scandinavian subjects (435 male). The latter two were logarithmically transformed. These measures were compared by ANOVA depending on genotype at each of 21 single-marker and 13 multimarker tags. The components of the multimarker tags (*bottom panel*) are numbered as in the *top panel* and correspond to those in Table 2. The best nominal *P* value (*P* < 0.01 for rs12260688 and fasting plasma glucose) was not statistically significant after 1,000 permutations (*P* = 0.44). M/M, homozygotes for the major allele; M/m, heterozygotes; m/m, homozygotes for the minor allele. 1 1, two copies of the multimarker haplotype; 1 2, one copy of the multimarker haplotype; 2 2, zero copies of the multimarker haplotype.

yielded a nonsignificant empirical  $P$  value for this comparison ( $P = 0.44$ ). When insulin AUC was examined in place of fasting insulin as a measure of insulin levels, we only found one nominally significant difference across genotypic groups: homozygotes for test 22 appeared to have slightly higher insulin AUC than heterozygotes, in a direction opposite from that seen for fasting insulin ( $P = 0.03$ ). Given these inconsistent findings, the number of hypotheses tested (34 tests  $\times$  3 phenotypes, albeit not independent), the small number of observations in several of the nominally significant groups, the lack of clear genotypic models to account for genetic risk in our data, and the results of our permutation procedure, we find six  $P$  values  $<0.05$  and one  $P$  value  $<0.01$  consistent with statistical fluctuations for those given levels of statistical significance.

**Comparisons with previous results.** As shown in Table 2 and in agreement with the level of skepticism evidenced by Groves et al. (8), we failed to replicate the association of rs4646953 with type 2 diabetes. When the genotyping results from Groves et al. were combined with ours by meta-analysis, the overall odds ratio in favor of the major allele was reduced to 0.97 (95% CI 0.89–1.07; two-sided  $P = 0.59$ ). This meta-analysis did pass the test for heterogeneity ( $P = 0.07$ ), consistent with the assumption that the various samples were drawn from the same underlying distribution.

We then attempted to replicate the association of the CC haplotype formed by rs2209772 and rs1887922 with type 2 diabetes (9). The frequency of the CC haplotype formed by the above two SNPs was similar in our samples (range 14–17%). When the CC haplotype was compared with all other possibilities at those loci, the odds ratio for association with type 2 diabetes was 1.01 (95% CI 0.87–1.16; one-sided  $P = 0.46$ ). Karamohamed et al. had also found that male carriers of the alternate haplotype TT at those same SNPs had lower fasting plasma glucose; in our sample, we noted that the 96 individuals who were homozygous for the TT haplotype appeared to have a lower mean fasting plasma glucose than the 28 CC homozygotes ( $5.42 \pm 0.53$  [mean  $\pm$  SD] vs.  $5.62 \pm 0.50$  mmol/l, respectively; one-sided  $P = 0.04$ ), although when only men were studied this difference did not reach nominal statistical significance ( $5.44 \pm 0.56$  vs.  $5.55 \pm 0.58$ , respectively; one-sided  $P = 0.24$ ).

Finally, we explored whether diplotype combinations based on the SNPs studied by Gu et al. (10) were associated with different levels of fasting plasma insulin, HOMA-IR, or BMI. We reconstructed the haplotypes examined by Gu et al. with our own genotypic data, filling in the SNPs we had not genotyped in our disease samples with the single-marker and multimarker predictors derived from Tagger. Not surprisingly, the haplotype frequencies of our Scandinavian samples essentially agreed with the Swedish samples of Gu et al. (10) (Fig. 2 of the online appendix). We found no statistically significant differences in fasting plasma insulin, HOMA-IR, or BMI based on the diplotype combinations of the Scandinavian subjects in our sample, whether all individuals or only males were considered. As an additional test of the best single-marker results of Gu et al. (10), we evaluated 2-h insulin levels according to genotype at rs2251101: we found no statistically significant difference among the three genotypic groups [ $\log(2\text{-h insulin}) = 1.56 \pm 0.24$ ,  $1.55 \pm 0.29$ , and  $1.54 \pm 0.33$  for CC, CT, and TT, respectively;  $P > 0.5$ ]; these values remained essentially identical when only males were considered (data not shown).

## DISCUSSION

We set out to test the previous hypotheses of nominal association of SNPs and haplotypes in *IDE* with type 2 diabetes and related traits (8–10). We confirmed that linkage disequilibrium among common variants in the *IDE* gene is extensive (8–10,27–29) but failed to detect a statistically significant association of any tagging SNP or multimarker haplotype with type 2 diabetes in our samples. We could not replicate the association of rs4646953 with type 2 diabetes detected by Groves et al. in one U.K. sample (but not in another sample in the same study) (8), and when all genotypic data were combined by meta-analysis the association did not reach nominal statistical significance.

We also failed to replicate the association of the CC haplotype formed by SNPs rs2209772 and rs1887922 with type 2 diabetes (9) and did not find any significant variation in fasting plasma insulin, HOMA-IR, or BMI based on diplotype load as described by Gu et al. (10). In contrast to the most significant single-marker finding of Gu et al. (10), 2-h insulin levels did not vary according to genotype at rs2251101. We did detect a slight increase in fasting plasma glucose in carriers of the CC haplotype formed by SNPs rs2209772 and rs1887922 when compared with carriers of the TT haplotype; however, given our marginal  $P$  value, the small numbers of homozygous individuals and the fact that our analysis was performed in the entire population rather than in males alone, we do not consider this finding a convincing replication of the result obtained by Karamohamed et al. (9).

Lack of reproducibility usually implies that either the first report was a false-positive result due to statistical fluctuation or that the second study was a false-negative attempt at replication. The former is quite common in genetic association studies, especially when multiple SNPs, genetic models, and phenotypes are examined; this is often due to the adoption of inappropriately relaxed thresholds when declaring statistical significance (30). We note that the three studies that have examined *IDE* for association with type 2 diabetes thus far (8–10) do not report findings that are consistent with each other and therefore cannot be considered true replications. The present study constitutes the first deliberate attempt at testing the various hypotheses posed by the original reports.

We believe our study represents a reasonable attempt at replication. When the initial association is real, replication can fail because of 1) insufficient power due to a small sample size (12), 2) heterogeneity among the combined subsamples, and/or 3) heterogeneity within subsamples, possibly due to population stratification. To address the first concern, even though our study was well-powered to replicate the model of Groves et al. (8), it is possible that the initial association may have been an overestimate; in either case, the best estimate of the odds ratio is bounded by the 95% CIs reported in our meta-analysis (0.89–1.07). In terms of sample heterogeneity, although 5 of 34 tests examined in this study did indicate some sample heterogeneity, the great majority of tests employed in this and in previous studies that have analyzed these samples have passed formal tests of heterogeneity (14,15). Finally, we note that these samples have shown consistent results when genotyped for the two most widely reproduced associations with type 2 diabetes, P12A in *PPARG* (13) and

E23K in *KCNJ11* (15), and our results also overlap those of other groups for the promoter SNPs in *HNF4A* (14).

Although overmatching in our samples may have prevented us from detecting a genetic effect on type 2 diabetes if this was mediated through BMI, we did not observe a sizeable impact of genotype on BMI. Finally, we have not evaluated the role of rare variants in *IDE*, and it is therefore possible that polymorphisms with minor allele frequency <5% may confer risk of type 2 diabetes in selected populations.

We therefore must conclude, in agreement with Groves et al. (8), that analysis of >6,600 samples provides no compelling evidence that common variation in the *IDE* gene contributes to type 2 diabetes susceptibility in humans. Nevertheless, we hope that the comprehensive haplotype structure and set of tag SNPs provided in this report will be useful for further tests in type 2 diabetes or in other common diseases where IDE is thought to play a role, such as Alzheimer's disease (28,29).

#### ACKNOWLEDGMENTS

J.C.F. is supported by National Institutes of Health Research Career Award 1 K23 DK65978-02. S.W. is supported by a Wellcome Trust Career Development Fellowship. T.T. is a Research Fellow at the Academy of Finland. D.A. is a Charles E. Culpeper Scholar of the Rockefeller Brothers Fund and a Burroughs Wellcome Fund Clinical Scholar in Translational Research, the latter of which supported this work. D.A., M.J.D., and J.N.H. are recipients of The Richard and Susan Smith Family Foundation/American Diabetes Association Pinnacle Program Project Award. L.G., T.T., and the Botnia Study are principally supported by the Sigrid Juselius Foundation, the Academy of Finland, the Finnish Diabetes Research Foundation, The Folkhalsan Research Foundation, the European Community (BM4-CT95-0662, GIFT), the Swedish Medical Research Council, the Juvenile Diabetes Foundation Wallenberg Foundation, and the Novo Nordisk Foundation.

We thank R. Tanzi for his encouragement in initiating this project, the Botnia research team for clinical contributions, and the members of the Altshuler, Hirschhorn, Daly, and Groop labs for helpful discussions.

#### REFERENCES

- Duckworth WC, Bennett RG, Hamel FG: Insulin degradation: progress and potential. *Endocr Rev* 19:608–624, 1998
- Fakhrai-Rad H, Nikoshkov A, Kamel A, Fernstrom M, Zierath JR, Norgren S, Luthman H, Galli J: Insulin-degrading enzyme identified as a candidate diabetes susceptibility gene in GK rats. *Hum Mol Genet* 9:2149–2158, 2000
- Farris W, Mansourian S, Chang Y, Lindsley L, Eckman EA, Frosch MP, Eckman CB, Tanzi RE, Selkoe DJ, Guenette S: Insulin-degrading enzyme regulates the levels of insulin, amyloid  $\beta$ -protein, and the  $\beta$ -amyloid precursor protein intracellular domain in vivo. *Proc Natl Acad Sci U S A* 100:4162–4167, 2003
- Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL, Leach RJ, O'Connell P, Stern MP: Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican Americans. *Am J Hum Genet* 64:1127–1140, 1999
- Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD, Ally DS, Mohlke KL, Silander K, Kohtamaki K, Chines P, Balow Jr J, Birznieks G, Chang J, Eldridge W, Erdos MR, Karanjawala ZE, Knapp JI, Kudelko K, Martin C, Morales-Mena A, Musick A, Musick T, Pfahl C, Porter R, Rayman JB: The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. *Am J Hum Genet* 67:1174–1185, 2000
- Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M, O'Rahilly S, Frayling TM, Bell JI, Lathrop GM, Bennett A, Dhillion R, Fletcher C, Groves CJ, Jones E, Prestwich P, Simecek N, 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 UK 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
- Meigs JB, Panhuysen CIM, Myers RH, Wilson PWF, Cupples LA: A genome-wide scan for loci linked to plasma levels of glucose and HbA<sub>1c</sub> in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. *Diabetes* 51:833–840, 2002
- Groves CJ, Wiltshire S, Smedley D, Owen KR, Frayling TM, Walker M, Hitman GA, Levy JC, O'Rahilly S, Menzel S, Hattersley AT, McCarthy MI: Association and haplotype analysis of the insulin-degrading enzyme (*IDE*) gene, a strong positional and biological candidate for type 2 diabetes susceptibility. *Diabetes* 52:1300–1305, 2003
- Karamohamed S, Demissie S, Volcjak J, Liu C, Heard-Costa N, Liu J, Shoemaker CM, Panhuysen CI, Meigs JB, Wilson P, Atwood LD, Cupples LA, Herbert A: Polymorphisms in the insulin-degrading enzyme gene are associated with type 2 diabetes in men from the NHLBI Framingham Heart Study. *Diabetes* 52:1562–1567, 2003
- Gu HF, Efendic S, Nordman S, Ostenson C-G, Brismar K, Brookes AJ, Prince JA: Quantitative trait loci near the insulin-degrading enzyme (*IDE*) gene contribute to variation in plasma insulin levels. *Diabetes* 53:2137–2142, 2004
- McCarthy MI: Progress in defining the molecular basis of type 2 diabetes mellitus through susceptibility-gene identification. *Hum Mol Genet* 13: R33–R41, 2004
- Lohmueller K, Pearce CL, Pike M, Lander ES, Hirschhorn JN: Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 33:177–182, 2003
- Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, Schaffner SF, Bolk S, Brewer C, Tuomi T, Gaudet D, Hudson TJ, Daly M, Groop L, Lander ES: The common PPAR $\gamma$  Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26:76–80, 2000
- Winckler W, Graham RR, de Bakker PIW, Sun M, Almgren P, Tuomi T, Gaudet D, Hudson TJ, Ardlie KG, Daly MJ, Hirschhorn JN, Groop L, Altshuler D: Association testing of variants in the hepatocyte nuclear factor 4 $\alpha$  gene with risk of type 2 diabetes in 7,883 people. *Diabetes* 54:886–892, 2005
- Florez JC, Burt N, de Bakker PIW, Almgren P, Tuomi T, Holmkvist J, Gaudet D, Hudson TJ, Schaffner SF, Daly MJ, Hirschhorn JN, Groop L, Altshuler D: Haplotype structure and genotype-phenotype correlations of the sulfonylurea receptor and the islet ATP-sensitive potassium channel gene region. *Diabetes* 53:1360–1368, 2004
- Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC: Homeostasis model assessment: insulin resistance and  $\beta$ -cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28:412–419, 1985
- Hosono S, Faruqi AF, Dean FB, Du Y, Sun Z, Wu X, Du J, Kingsmore SF, Egholm M, Lasken RS: Unbiased whole-genome amplification directly from clinical samples. *Genome Res* 13:954–964, 2003
- Paez JG, Lin M, Beroukhi R, Lee JC, Zhao X, Richter DJ, Gabriel S, Herman P, Sasaki H, Altshuler D, Li C, Meyerson M, Sellers WR: Genome coverage and sequence fidelity of  $\phi$ 29 polymerase-based multiple strand displacement whole genome amplification. *Nucleic Acid Res* 32:e71, 2004
- Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P, the International HapMap Consortium: A haplotype map of the human genome. *Nature* 437:1299–1320, 2005
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D: The structure of haplotype blocks in the human genome. *Science* 296:2225–2229, 2002
- Excoffier L, Slatkin M: Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 12:921–927, 1995
- de Bakker PIW, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D: Efficiency and power in genetic association studies. *Nat Genet* 11:1217–1223, 2005
- Purcell S, Chery SS, Sham PC: Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19:149–150, 2003
- Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506–516, 1993
- Boehnke M, Langefeld CD: Genetic association mapping based on discordant sib pairs: the discordant-alleles test. *Am J Hum Genet* 62:950–961, 1998

26. Barrett JC, Fry B, Maller J, Daly MJ: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265, 2005
27. Abraham R, Myers A, Wavrant-DeVrieze F, Hamshere ML, Thomas HV, Marshall H, Compton D, Spurlock G, Turic D, Hoogendoorn B, Kwon JM, Petersen RC, Tangalos E, Norton J, Morris JC, Bullock R, Liolitsa D, Lovestone S, Hardy J, Goate A, O'Donovan M, Williams J, Owen MJ, Jones L: Substantial linkage disequilibrium across the insulin-degrading enzyme locus but no association with late-onset Alzheimer's disease. *Hum Genet* 109:646–652, 2001
28. Prince JA, Feuk L, Gu HF, Johannsson B, Gatz M, Blennow K, Brookes AJ: Genetic variation in a haplotype block spanning *IDE* influences Alzheimer disease. *Hum Mutat* 22:363–371, 2003
29. Blomqvist ME-L, Chalmers K, Andreassen N, Bogdanovic N, Wilcock GK, Cairns NJ, Feuk L, Brookes AJ, Love S, Blennow K: Sequence variants of *IDE* are associated with the extent of  $\beta$ -amyloid deposition in the Alzheimer's disease brain. *Neurobiol Aging* 26:795–802, 2005
30. Wacholder S, Chanock S, Garcia-Closas M, El ghomli L, Rothman N: Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst* 96:434–442, 2004