Evaluation of the Association of IGF2BP2 Variants with Type 2 Diabetes in French Caucasians

Konsta Duesing¹, Ghazaleh Fatemifar¹, Guillaume Charpentier², Michel Marre³⁴, Jean Tichet⁵, Serge Hercberg⁶, Beverley Balkau⁷⁸, Philippe Froguel¹⁹ and Fernando Gibson¹

¹Genomic Medicine, Imperial College London, Hammersmith Campus, Du Cane Rd, London W12 0NN, UK
²Endocrinology-Diabetology Unit, Corbeil Hospital, Corbeil, France
³Endocrinology-Diabetology, Bichat Hospital, Paris, France
⁴INSERM U695, Paris, France
⁵Institut Régional Pour la Santé, Tours, France
⁶U557 Inserm/U1125 Inra/Cnam/University Paris 13, CRNH IdF, F-93017 Bobigny, France
⁷INSERM U780-IFR69, Villejuif, France
⁸Paris Univ-Sud, Orsay, France
⁹CNRS 8090, Institut de Biologie de Lille, Institut Pasteur, CHU Lille, France

Correspondence should be addressed to:
Fernando Gibson (email: fernando.gibson@imperial.ac.uk)

Received for publication 17 December 2007 and accepted in revised form 17 April 2008.

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

Copyright American Diabetes Association, Inc., 2008
Evaluation of the Association of IGF2BP2 Variants

**OBJECTIVE:** We performed a comprehensive genetic association study of common variation spanning the IGF2BP2 locus, in order to replicate the association of the ‘confirmed’ type 2 diabetes susceptibility variants rs4402960 and rs1470579 in the French Caucasian population, and to further characterise the susceptibility variants at this novel locus.

**RESEARCH DESIGN AND METHODS:** We genotyped a total of 21 tagging SNPs spanning the IGF2BP2 locus in our type 2 diabetes case-control cohort comprising in 3,093 French Caucasian subjects.

**RESULTS:** IGF2BP2 variants rs4402960 and rs1470579 were not associated with type 2 diabetes in the present study (P = 0.632 and P = 0.896, respectively). Meta-analysis of genotype data from over 34,000 subjects demonstrated that our inability to replicate rs4402960/rs1470579 was consistent with the findings from several previous GWAS datasets that were under-powered to detect this modest association signal (OR 1.14). We obtained novel evidence that rs9826022, a borderline rare variant (5% MAF) in the 3’ downstream region, was associated with type 2 diabetes (P = 0.0002; OR 1.53 [95% CI 1.22-1.91]). This result was corroborated by the meta-analysis of 10,542 genotypes from the current study and GWAS datasets using both fixed (P = 9.47 x 10^-6; OR 1.30 [95% CI 1.16-1.46]) and random effects (P = 0.001; 1.30 [95%CI 1.11-1.52]) calculations.

**CONCLUSIONS:** We were unable to replicate the confirmed rs4402960/rs1470579 susceptibility variants, but found novel evidence for a rare variant in the 3’ downstream region of IGF2BP2. Further genetic and functional studies are required to identify the aetiological IGF2BP2 variants.
The insulin-like growth factor 2 mRNA binding protein 2 (IGF2BP2) gene on chromosome 3q27 is a paralog of IGF2BP1, a known regulator of IGF2 gene expression. Genome wide association studies (GWASs) carried out by the Finland–United States Investigation of NIDDM Genetics (FUSION) (1), the Wellcome Trust Case Control Consortium (WTCCC) (2) and the Diabetes Genetics Initiative (DGI) (3) groups each found modest evidence that single nucleotide polymorphisms (SNPs) in the IGF2BP2 region are associated with type 2 diabetes. The subsequent meta-analysis of primary and replication datasets from these GWASs corroborated these findings and identified two strongly correlated IGF2BP2 variants, rs1470579 and rs4402960, as ‘confirmed’ type 2 diabetes susceptibility variants (1-3). By contrast, the French/Canadian GWAS (4) typed 10 SNPs across the IGF2BP2 locus, including rs1470579, in 1363 subjects, but found no nominal (P <0.05) association signals at IGF2BP2. In an attempt to replicate the IGF2BP2 association findings in the French Caucasian population in a larger study and to further characterise the susceptibility variants at this novel locus, we have performed an association study of HapMap Phase II tag SNPs spanning the IGF2BP2 locus in 3,093 French Caucasian subjects.

**RESEARCH DESIGN AND METHODS**

**Case-control subjects**

All subjects were of French Caucasian ancestry. Individuals identified by Sladek et al. (4) to lie outside the HapMap CEU ancestry cluster were excluded from the study. Type 2 diabetic case subjects were known diabetic patients. Normoglycemic control subjects were selected to have a fasting blood glucose concentration <7.0 mM (5). Case subjects were composed of: (i) 372 probands from diabetic families (6), recruited in Lille; and (ii) 1083 patients with a family history of T2D recruited at the Corbeil-Essonne Hospital. Control subjects were composed of: (i) 353 normoglycemic parents from T2D families; (ii) 543 subjects from the SUVIMAX (Supplementation en Vitamines et Minéraux Antioxydant) prospective population-based cohort study (7); and (iii) 742 subjects selected from the DESIR (Data from an Epidemiologic Study on the Insulin Resistance Syndrome) cohort, a large prospective study of insulin resistance in French subjects (8). Informed consent was obtained from all subjects and the study was approved by the local ethics committees.

**Statistical Power**

The case-control cohort comprised 1,455 type 2 diabetic subjects (age, 60 ± 12 years; BMI, 29.0 ± 6.0 kg/m²; male/female, 56:44%) and 1,638 normoglycemic subjects (age, 54 ± 13 years; BMI, 24.1 ± 3.3 kg/m²; male/female, 43:57%). At α = 0.05, this sample size provided 76% power (9) to detect the type 2 diabetes susceptibility variants rs1470579 and rs4402960, assuming an allele frequency of 0.30, a disease prevalence of 0.1, a heterozygote relative risk of 1.14 (1-3), a multiplicative model and a 100% genotype call rate.

**IGF2BP2 tag SNP selection**

The genomic target region for tag SNP selection was extended 10 kb upstream and downstream of the NCBI36 IGF2BP2 locus (chr3:186,844,221..187,025,521). A total of 19 HapMap Phase II multimarker tagging SNPs (HapMap Data Release 21a/ Jan07) with r² and minor allele frequency (MAF) thresholds of 0.8 and 0.05, respectively, were identified for genotyping. In addition, the two GWAS-identified susceptibility variants rs4402960 and rs1470579 (1-3) were added to the genotyped SNP set, making a total of 21 genotyped SNPs.
SNP genotyping

Genotyping was performed with the Sequenom MassARRAY iPLEX system (10). SNP genotype frequencies were tested for accordance with Hardy-Weinberg equilibrium using chi-squared analysis. Quality control: all 21 genotyped SNPs exhibited a call rate >90% and a Hardy-Weinberg P >0.05, with well defined genotype clusters. There was no evidence (at $\alpha = 0.01$) of differential call rates across cases and controls for any SNP (online supplementary Table 2).

Statistical analyses

To test for association of IGF2BP2 SNPs with type 2 diabetes, chi-squared analysis of allele and genotype counts was performed. Pairwise SNP linkage disequilibrium (LD) values were calculated from the genotype data of the control cohort with Haploview (11). Quantitative metabolic phenotypes, body mass index (BMI), waist-hip ratio (WHR), fasting serum levels of triacylglycerol, total- and HDL-cholesterol, glucose, insulin, apolipoprotein A-I (APOA1) and apolipoprotein B (APOB), measured in 1,539 normoglycemic subjects from the control cohort, were log transformed and adjusted for age, sex and BMI, as appropriate. SNPs were tested for association with adjusted quantitative traits using SPSS 14.0 with the ANOVA test under a codominant model. Quantitative trait association p-values are presented uncorrected for multiple testing. Combined analysis of association datasets was carried out with the Mantel-Haenszel (fixed effects) meta-analysis method. Inter-study heterogeneity was assessed with Cochrans’s Q statistic and the $I^2$ metric (12; 13). All calculations were performed using R (v2.5.1) statistical software and the meta (v0.8-2) package (14). Association analysis of SNPs captured by multimarker tags was carried out with the PLINK software package (15). Haplotype association with was performed with the WHAP (v2.09) software package (16).

RESULTS & DISCUSSION

A total of 21 HapMap Phase II multimarker tag SNPs ($r^2 \geq 0.8$; MAF ≥ 0.05) spanning the IGF2BP2 locus, including the susceptibility variants rs4402960 and rs1470579, were tested for association with type 2 diabetes in 3,093 French subjects. The allele and genotype counts for all SNPs are presented in online supplementary Tables 1 and 2, respectively. Figure 1 shows that SNPs rs4402960 and rs1470579 exhibited very strong LD ($r^2 = 0.95$) in agreement with the GWAS (1-3) and HapMap data (17). However, the allele frequencies of SNPs rs4402960 and rs1470579 were not significantly different in the case and control groups (rs4402960, P = 0.632; rs1470579, P = 0.896) indicating that these variants were not associated with type 2 diabetes in the present study (Table 1). None of the three SNPs captured by multimarker tags (rs4575929, P = 0.159; rs4686692, P = 0.566; and rs16860216, P = 0.972;) were associated with type 2 diabetes (online supplementary Table 3).

Our inability to replicate the confirmed rs4402960/rs1470579 association result can be attributed to a lack of power to detect this modest signal. An examination of the published association evidence for these variants (Fig. 2A and 2B; online supplementary Tables 4 and 5) illustrates this point and demonstrates that our results are not inconsistent with those of previous studies. Of the nine published rs4402960 datasets, the three statistically well-powered studies (those with $\geq 90\%$ power) all obtained an association for this variant, while the six under-powered studies showed either no association or a weak association with type 2 diabetes. Overall, the combined data shows a 3% difference in allele frequency between the case and control groups in over 34,000
Evaluation of the Association of IGF2BP2 Variants

subjects, which equates to very strong evidence of association \( (P = 1.9 \times 10^{-14}; \text{OR} 1.13 [95\% \text{ CI} 1.10-1.17]) \). Similarly for the rs1470579 variant, the under-powered datasets were either non-significant or weakly associated with type 2 diabetes. The combined data shows a 2% allele frequency difference in over 22,000 case-control subjects, and a clearly significant association with type 2 diabetes \( (P = 2.6 \times 10^{-9}; \text{OR} 1.13 [95\% \text{ CI} 1.09-1.18]) \). All of this serves as a reminder that the meta-analysis of individually under-powered studies has an invaluable role to play in the identification and confirmation of susceptibility variants of small effect.

The between-study heterogeneity metric \( I^2 \) \((12; 13)\) was calculated for these two variants (Fig. 2D). Heterogeneity was moderate for rs4402960 \( (I^2 = 21\%) \) in agreement with a recent study (18). For rs1470579, the meta-analysed signal is clearly driven by the DGI Replication set ‘S’ result \( (P = 3.73 \times 10^{-8}) \). In accordance with this standout result and the smaller number of studies available for this SNP, heterogeneity was higher \( (I^2 = 58\%) \); the random effects odds ratio gave a mere \( P = 0.001 \) compared to the Mantel-Haenszel \( P = 2.56 \times 10^{-9} \); and Cochran’s Q statistic was also statistically significant \( (P = 0.037) \).

We obtained novel evidence that rs9826022 in the 3′ downstream region \( (P = 0.0002; \text{OR} 1.53 [95\% \text{ CI} 1.22-1.91]) \) was associated with type 2 diabetes (Table 1). This result survived Bonferroni correction for the number of SNPs tested \( (P = 0.003) \) and we sought confirmation in the publicly available GWAS data. The WTCCC (http://www.wtccc.org.uk/) and DGI (http://www.broad.mit.edu/diabetes/) GWASs did not directly type the rs9826022 variant, but instead typed rs9878208, a rs9826022 proxy (HapMap \( r^2 = 1 \)). Meta-analysis of this data (Fig. 2C; online supplementary Table 6) provided support for the association although the random effects evidence \( (P = 0.001) \) was weaker than that produced by the Mantel-Haenzel analysis \( (P = 9.47 \times 10^{-6}) \). The heterogeneity between these three studies was moderate \( (I^2 = 44\%) \). The disparity between the fixed and random effects may indicate that rs9826022 is not the ‘causative’ variant, but is merely in partial LD with the true susceptibility variant; or it may simply reflect the ‘winner’s curse’ result of the present study and the inherently larger variance of the genetic effect of rare variants in moderately sized studies. The rs9826022 result will clearly require confirmation in further large, independent studies before a definitive assessment of the contribution of this rare variant to type 2 diabetes susceptibility can be made.

The only other nominal association signal, rs9864104 \( (P = 0.012; \text{OR} 1.19 [95\% \text{ CI} 1.04-1.37]) \), was modest and disappeared upon multiple test correction. Since rs9826022 and rs9864104 were in low-moderate LD \( (r^2 = 0.26) \) we carried out haplotype analysis of these variants. Two-SNP haplotypes containing the rare allele of rs9826022 showed a virtually identical frequency and p-value as the single point rs9864104 analysis (online supplementary Table 7) indicating that the haplotype analysis did not add anything to the single point analysis, and that the weak rs9864104 signal was caused by this variant being in partial LD with rs9826022. There were no significant differences in SNP allele frequencies between men and women, and no association with type 2 diabetes was uncovered by stratifying for sex (data not shown).

SNPs rs4402960, rs1470579, rs9826022 and rs9864104 were also tested for association with a number of metabolic quantitative phenotypes (online supplementary Table 8). SNPs rs4402960 and rs1470579 presented weak associations with APOA1 \( (P = 0.019 \text{ and } 0.028) \) and rs9864104 was associated with APOA1 \( (P = 0.008) \) and APOB levels \( (P \)
Evaluation of the Association of IGF2BP2 Variants

= 0.002) although there was no linear trend between the three genotype groups. The association of IGF2BP2 variation with apolipoprotein levels may be consistent with the role of the insulin-like growth factor system in regulating lipid metabolism. However, in the absence of replication we emphasise that these quantitative trait associations are of nominal significance and require confirmation in further large studies.

In conclusion, we have carried out a comprehensive association study of common variation spanning the IGF2BP2 locus and type 2 diabetes in French Caucasians. We were unable to replicate the confirmed susceptibility variants rs4402960 and rs1470579 but found novel evidence for a rare variant in the 3′ downstream region of IGF2BP2. Further genetic and functional studies are required to identify the aetiological variants at the IGF2BP2 locus and determine the cellular and physiological mechanisms by which they act to modulate type 2 diabetes susceptibility.

ACKNOWLEDGEMENTS

This work was supported by a Wellcome Grant to F.G. (WT081510MA). We thank the anonymous referees and editorial staff for critically appraising and improving draft versions of the manuscript.
### Evaluation of the Association of IGF2BP2 Variants

Table 1. Association of *IGF2BP2* SNPs with type 2 diabetes: ‘confirmed’ susceptibility SNPs rs4402960 and rs1470579; and SNPs showing nominal association (P <0.05) in French Caucasians.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr Pos (bp) NCBI36</th>
<th>Gene Regiona</th>
<th>Allele</th>
<th>Allele count</th>
<th>P</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>T2D (%)</td>
<td>NG (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs9826022</td>
<td>186,839,954</td>
<td>3′ downstream (+4267 bp)</td>
<td>A</td>
<td>2532 (93.0)</td>
<td>2949 (95.3)</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td>190 (7.0)</td>
<td>145 (4.7)</td>
<td></td>
</tr>
<tr>
<td>rs9864104</td>
<td>186,840,225</td>
<td>3′ downstream (+3996 bp)</td>
<td>C</td>
<td>2260 (81.7)</td>
<td>2628 (84.2)</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T</td>
<td>506 (18.3)</td>
<td>494 (15.8)</td>
<td></td>
</tr>
<tr>
<td>rs4402960</td>
<td>186,994,381</td>
<td>intron 6</td>
<td>G</td>
<td>1786 (67.5)</td>
<td>2111 (68.1)</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T</td>
<td>858 (32.5)</td>
<td>987 (31.9)</td>
<td></td>
</tr>
<tr>
<td>rs1470579</td>
<td>187,011,773</td>
<td>intron 5</td>
<td>A</td>
<td>1753 (67.3)</td>
<td>2120 (67.4)</td>
<td>0.896</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C</td>
<td>853 (32.7)</td>
<td>1024 (32.6)</td>
<td></td>
</tr>
</tbody>
</table>

arelative to the NCBI36 coordinates of the *IGF2BP2* genomic locus (chr3:186,844,221..187,025,521). T2D: type 2 diabetic cases; NG: normoglycemic controls. Chi-squared p-values are shown.
Figure 1. Pattern of linkage disequilibrium between *IGF2BP2* SNPs

Plot of pairwise SNP $r^2$ values calculated from control genotype data. The plot was drawn to scale using LocusView (T. Petryshen, A. Kirby, M. Ainscow, unpublished software).
Evaluation of the Association of IGF2BP2 Variants

Figure 2. Meta-analysis of the association of IGF2BP2 SNPs with type 2 diabetes.

For each study, the point estimate of the odds ratio with 95% CI is shown. In addition, the summary fixed and random effects are shown for rs4402960 (A), rs1470579 (B) and rs9826022 (C). (D) shows Cochran’s Q and $I^2$ (12; 13) statistics for these three variants.
REFERENCES


Evaluation of the Association of IGF2BP2 Variants


9. Purcell S, Cherny SS, Sham PC: Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19:149-150, 2003


Evaluation of the Association of IGF2BP2 Variants


15. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559-575, 2007


18. Ioannidis JP, Patsopoulos NA, Evangelou E: Heterogeneity in meta-analyses of genome-wide