A Polymorphism within the G6PC2 Gene Is Associated with Fasting Plasma Glucose Levels

Nabila Bouatia-Naji1*, Ghislain Rocheleau2*, Leentje Van Lommel3, Katleen Lemaire3, Frans Schuit3, Christine Cavalcanti-Proença1, Marion Marchand1, Anna-Liisa Hartikainen4, Ulla Sovio5, Franck De Graeve1, Johan Rung2, Martine Vaxillaire1, Jean Tichet6, Michel Marre7, Beverley Balkau8, Jacques Weill9, Paul Elliott5, Marjo-Riitta Jarvelin5,10, David Meyre1, Constantin Polychronakos2,11, Christian Dina1, Robert Sladek2, Philippe Froguel1,12†

1 CNRS UMR 8090 Institute of Biology, Pasteur Institute of Lille and Lille 2 Droit et Santé University, Lille, France.
2 Department of Human Genetics, Faculty of Medicine, McGill University and Génome Québec Innovation Centre, Montreal H3A 1A4, Canada.
3 Gene Expression Unit, Dept. Molecular Cell Biology Katholieke Universiteit Leuven, Belgium
4 Department of Obstetrics and Gynaecology, Oulu, Finland
5 Department of Epidemiology and Public Health, Imperial College, London W12 0NN United Kingdom.
6 IRSA, La Riche, France.
7 INSERM U695, Bichat Hospital, Paris, France
8 INSERM U780-IFR69, « Paris Sud » University, Villejuif, France.
9 Pediatric Endocrine Unit, Jeanne de Flandre Hospital, Lille, France
10 Department of Public Health and General Practice, University of Oulu, Finland.
11 Department of Pediatrics, Faculty of Medicine, McGill University, Montreal H3H 1P3, Canada.
12 Genomic Medicine, Hammersmith Hospital, Imperial College London, United Kingdom
* These authors equally contributed to this work.
† To whom correspondence should be addressed. Email: p.froguel@imperial.ac.uk
Abstract

Several studies have shown that healthy individuals with fasting plasma glucose (FPG) levels at the high end of the normal range have a higher risk of mortality. To identify genetic determinants that contribute to inter-individual variation in FPG, we tested 392,935 Single Nucleotide Polymorphisms (SNPs) in 654 normoglycemic subjects for association with FPG and we replicated the most strongly associated SNP (rs560887, p=4×10^-7) in 9,353 subjects. SNP rs560887 maps to intron 3 of the \textit{G6PC2} gene, which encodes glucose-6-phosphatase catalytic subunit-related protein (also known as IGRP), a protein selectively expressed in pancreatic islets. This SNP was associated with FPG (\(\beta=-0.06\) mmol/l per A-allele, combined p=4×10^{-23}) and with pancreatic beta-cell function (Homa-B model, combined p=3×10^{-13}) in three populations; however it was not associated with type 2 diabetes risk. We speculate that \textit{G6PC2}/IGRP regulated FPG by modulating the set-point for glucose-stimulated insulin secretion in pancreatic beta cells.
Recent innovations in genotyping technology have led to the identification of genetic variants associated with increased risk of type 2 diabetes (T2D) (1-4). However, the genetic factors contributing to inter-individual variation in blood glucose levels in the general population are largely unknown. In addition to affecting metabolic health, these variants may also impact coronary heart disease (CHD) risk, for which a linear correlation between FPG and CHD mortality has been described for both diabetic and non diabetic individuals (5, 6). Here we describe the results of a genome-wide association (GWA) study in which we aimed to identify genetic variants involved in glucose homeostasis in the general population.

We analyzed FPG as a quantitative trait in 654 normoglycemic (NG) nonobese individuals using the Illumina Infinium Human1 and Hap300 BeadArrays (Table S1) (1, 7). The strongest association signal was observed with SNP rs560887 ($p=4\times10^{-7}$ adjusted for age, gender and body mass index (BMI) under the additive model). This SNP is part of a 17 Kb linkage disequilibrium (LD) block on chromosome 2 that encompasses the gene and 3’ flanking region of the islet specific glucose-6-phosphatase related protein (IGRP, also called $G6PC2$ (glucose-6-phosphatase catalytic unit 2)), a glycoprotein embedded in the endoplasmic reticulum (ER) membrane (8). Six SNPs within this LD block were genotyped in our GWA analysis. These included a non synonymous SNP (rs492594-L219V), which showed modest association with FPG ($p=0.04$ for rs492594-L219V, which is in low LD with rs560887 at $r^2=0.24$ in CEU, according to HapMap data) and a SNP located in the 3’ flanking region of $G6PC2$ (rs563694), which showed stronger association ($p=2\times10^{-5}$, $r^2=0.76$ with rs560887) (Table S2). We used linear regression analysis to adjust for the effect of rs560887 on all SNPs located within 100 kbp upstream and 100 kbp downstream of rs560887 that showed various levels of association with FPG ($p < 0.05$) and found no SNP which remained associated at significance level of 5% (Table S2). Imputation of genotypes using the LD
structure from HapMap data (CEU population) did not identify any SNP showing potentially stronger association with FPG (Figure 1). Based on these findings, we tested the association of rs560887 with FPG using a complementary subset (Table S1) of the French DESIR population (DESIR NG stage 2, N=3,419) and replicated our original association with FPG ($\beta=-0.06$ mmol/l per A allele; $p=3 \times 10^{-8}$, Table 1).

$G6PC2$/IGRP has been proposed to modulate the glycolytic pathway and eventually glucose-stimulated insulin secretion by dephosphorylating glucose-6-phosphate generated by glucokinase, the beta-cell glucose sensor (9). $G6pc2$ knockout mice display decreased FPG and normal insulin sensitivity (10). Interestingly, in the DESIR NG cohort the rs560887 A allele (MAF = 0.30) that associates with decreased FPG, is also associated with a decreased fraction of erythrocyte glycosylated haemoglobin (HbA1c) ($\beta=-0.04$; $p=8 \times 10^{-5}$). This reflects a better long-term glucose control, and with increased basal insulin secretion as assessed by the homeostasis model index of pancreatic beta-cell function Homa-B ($\beta=5.46$; $p=2 \times 10^{-5}$), a measure derived from physiological data of FPG and fasting insulin levels (7).

To validate these findings in different European populations, we genotyped young normoglycemic individuals from the Northern Finland Birth Cohort 86 (7) (NFBC86, N=5,073) and confirmed the association between rs560887 and decreased FPG ($\beta=-0.07$ mmol/l per A allele; $p=6 \times 10^{-16}$) and increased Homa-B index ($\beta=3.79$; $p=6 \times 10^{-10}$). In contrast, no association was found between rs560887 and fasting insulin ($p=0.5$ and $p=0.1$, respectively) in either the French or Finnish cohorts. Similarly, neither cohort showed strong association between rs560887 and BMI (DESIR $p=0.1$, NFBC86 $p=0.04$). Adjustment for BMI did not affect the association with FPG, suggesting that the observed effect on FPG is independent of obesity-induced insulin resistance. To test this hypothesis, we studied 861
severely obese NG French children (mean age 11 ± 3 yrs) (7). In this cohort, we still observed association between rs560887 and decreased FPG (β=-0.07 mmol/l per A allele; p=0.004) and increased Homa-B index (β=9.13; p=0.01). As in the French and Finish cohorts, rs560887 did not associate with fasting insulin (p=0.30) or BMI (p=0.60) in obese children.

rs560887 is located in the third intron of G6PC2/IGRP, 26 base-pairs from its boundary with the fourth exon. Splice site modeling shows that the rs560887 A allele may have a lower splicing acceptor score than the G allele (http://www.cbcb.umd.edu/software/GeneSplicer) (11); however, further investigation is required to define the functional relationship between the observed association signal, which may be caused by rs560887 or an unidentified SNP in high LD with rs560887, on G6PC2/IGRP splicing or gene expression.

Longitudinal analysis of the DESIR cohort showed that carriers of the rs560887 A allele had a lower risk of developing mild hyperglycemia (defined as a FPG≥6.1 mmol/l) over a 9-year follow-up period (Hazard ratio=0.83, 95%CI [0.72-0.96], p=0.009). Surprisingly, no association was detected between T2D risk and rs560887 or any SNP within the 17 Kb LD block in our GWA study (1) or other published T2D case-control studies (www.broad.mit.edu/diabetes/, www.wtccc.org.uk/). To confirm this, we analyzed 2,972 additional T2D cases (Table S1). Using the 4,073 DESIR NG individuals as controls and observed no significant association with T2D risk (p=0.6; OR=0.98 95%CI [0.91-1.05]). We note that our case-control analysis achieved statistical power of at least 99.95% to detect a significant effect (MAF= 0.30; OR≥1.20, N=7,045). Previous reports have shown that the SNP most strongly associated with T2D identified so far (rs7903146, located in TCF7L2) does not associate with FPG in European general populations (12, 13), suggesting that the genetic determinants that regulate FPG in physiological states may be different from those
that increase T2D risk. *G6PC2/IGRP* is homologous to two other ER proteins of the glucose-6-phosphatase family that catalyze the conversion of glucose-6-phosphate into glucose, a key enzymatic step for both gluconeogenesis and glycogenolysis (8). In contrast to the hepatocyte specific member (*G6pc1*) that is highly expressed in gluconeogenic tissues but not in pancreas (Figure S1, upper panel), the expression of *G6pc2* is restricted to islets and pancreatic beta-cell line MIN6 (Figure S1, middle panel), as it was shown in humans (14). Since expression of *G6pc2* mRNA is restricted to beta-cells in the mouse, as indicated in previous studies (14) and as confirmed in our own experiments (Figure S1), rs560887 may play an important role in the early pathogenesis of this disease without contributing to overall T2D risk. For example, the rs560887 common allele (G) might alter the glucostat set-point and up-regulates FPG but the development of overt T2D may require additional compromise of the pancreatic beta-cell function.

Whether *G6PC2/IGRP* has the same catalytic activity as the hepatic glucose-6-phosphatase *G6PC1* is a matter of debate (14). When IGRP dephosphorylates glucose-6-phosphate, it opposes the action of the beta-cell glucose sensor glucokinase and hence moderates the glycolytic pathway and eventually glucose-stimulated insulin secretion (9). We hypothesize that *G6PC2/IGRP* is part of the same glucose phosphorylation pathway that includes the glucokinase gene (*GCK*), responsible for the monogenic form of early onset diabetes MODY2 (15) and familial hypoglycemia-hyperinsulinemia phenotypes (16). This pathway also includes the glucokinase regulatory protein gene (*GCKR*). Thus, *G6PC2/IGRP* may antagonize the glucokinase activity in beta-cell, in a similar manner to the GCKR action as a potent inhibitor of glucokinase activity in hepatocytes (17). Recently, SNPs in both *GCK* (rs1799884/-30G/A) (18) and *GCKR* (rs1260326/P446L) (2, 19, 20) were found to modulate FPG in European population-based cohorts. To study the interaction of rs560887 with
rs1260326-P446L and rs1799884/-30G/A, we analyzed their combined effect in the French DESIR cohort and observed an additive effect (interaction p=0.56) of the three SNPs on FPG (p=5×10^{-21}, Figure 2). Individuals carrying more than four alleles associated with low FPG (10.7% of the DESIR population), show a mean 0.24 mmol/l (4.5%) decrease of FPG compared to subjects carrying only one allele (Figure 2).

G6PC2 rs560887 is associated with significant reductions in both FPG and HbA1c. There is strong evidence suggesting that even small changes in blood glucose have major impact on cardiovascular morbidity and mortality (21). Meta-analyses have showed a linear relationship with no threshold between increasing FPG and CHD risk (6), with individuals whose FPG is lower than 5.5 mM having the lowest mortality rate (5, 22). In contrast, subjects whose FPG lies between 5.6 and 6mM, while still considered normoglycemic, show a 30% increased risk of heart failure and a 60% increased risk of nephropathy (23). In Asians, a slight reduction in FPG from 5.5 mM to 5mM is associated with a 25% reduction in CHD risk (24). Even tight blood glucose control, reflected by HbA1c between 5.1-5.6% (normal value <6.1%), is associated with a 30% higher mortality compared to a lower HbA1c (25). It was concluded that “over 80% of the population excess mortality is explained by HbA1c > 5% and occur in 70 % of the population having HbA1c between 5-6.9%” (23). In this context, Marz et al previously reported that the glucokinase rs1799884/-30 A allele was associated with T2D and increased risk for CHD in both diabetics and non-diabetics (26). Thus, the 3 SNPs we identified in this study, which additively modulate FPG by 0.24 mM in the general population, may have a non negligible impact on human health.

In summary, we found a strong association between SNP rs560887, located in the G6PC2/IGRP gene with FPG and this association was replicated in three independent
populations (N=9,353). The magnitude of the rs560887 effect on FPG is remarkably similar in each cohort (explaining ~1% of FPG variance in each population (heterogeneity test p=0.95). Each copy of the A allele results in a 0.06 mmol/l decrease (overall meta-analysis p=4×10^{-23}). Our findings reconfirm the role of the glucose phosphorylation pathway in glucose homeostasis in humans in the general population.
REFERENCES

7. Materials and methods are available as supporting material on Science Online.
27. We thank O. Lantieri for comments on the manuscript, V. Vatin, M. Deweirder, F. Allegaert, S. Gaget, S. Gallina, J.-C. Chèvre and J. Delplanque for technical assistance. The study was supported ALFEDIAM-Servier, ANR-06 PHYSIO-037-02, EURODIA LSHM-CT-2006-518153, Genome Canada, Genome Quebec, KU-Leuven GOA/2004/11, Juvenile Diabetes Research Foundation 1-2006-182, Imperial College London, and the Academy of Finland.
Figure 1. Genomic context and association with fasting plasma glucose of the G6PC2 gene variant rs560887 on chromosome 2q. The upper panel shows the genomic context of rs560887 that maps to intron 3 of G6PC2. The middle panel shows the association magnitude (-log_{10}(p-value)) with fasting plasma glucose of the SNPs from the Illumina 317 mapping array (Tag-SNPs) (dots) and the HapMap CEU data imputed SNPs (plus signs), not directly genotyped around (±100KB) rs560887. The lower panel shows the LD structure (defined by pairwise $r^2$) between SNPs in the CEU population from HapMap phase II using the Haploview software.

Figure 2. Combined effects of G6PC2 rs560887, GCKR rs1260326-P446L and GCK rs1799884-30G variants on fasting plasma glucose levels. Data are presented as mean [CI 95%] and the $p$-value is for the $\beta$ coefficient in the linear regression model (including age, sex and BMI as covariates) of the number of predisposing alleles on FPG levels (rs560887G allele, risk allele frequency = 0.30; GCKR rs1260326-P446 allele, risk allele frequency = 0.44; GCK -30G rs1799884 allele, risk allele frequency = 0.82)
Table 1. Association of rs560887 with fasting plasma glucose levels in normoglycemic individuals from a genome-wide association study and replication analyses. Association was tested in the additive model, adjusted for age, sex and BMI. All individuals were normoglycemic (FPG<6.1 mmol/l). * Per-A allele effect size was estimated from the regression coefficient β. †Overall meta-analysis includes replication populations (DESIR NG, NFBC86 and obese children). MAF: minor allele frequency. DESIR Controls: non-obese normoglycemic individuals. FPG: fasting plasma glucose. NFBC86: North Finland Birth Cohort 1986.

<table>
<thead>
<tr>
<th>FPG</th>
<th>N</th>
<th>MAF</th>
<th>GG</th>
<th>GA</th>
<th>AA</th>
<th>Variance (%)</th>
<th>Per-A allele effect size* (mmol/l) (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>DESIR controls (stage 1)</td>
<td>654</td>
<td>0.33</td>
<td>5.15 ± 0.33</td>
<td>5.06 ± 0.37</td>
<td>4.93 ± 0.34</td>
<td>3.90</td>
<td>-0.10 (-0.14; -0.06)</td>
<td>4×10⁻⁷</td>
</tr>
<tr>
<td>DESIR NG (stage 2)</td>
<td>3,419</td>
<td>0.30</td>
<td>5.26 ± 0.45</td>
<td>5.19 ± 0.46</td>
<td>5.14 ± 0.47</td>
<td>0.74</td>
<td>-0.06 (-0.08; -0.04)</td>
<td>3×10⁻⁸</td>
</tr>
<tr>
<td>DESIR cohort (stage 1+stage 2)</td>
<td>4,073</td>
<td>0.30</td>
<td>5.24 ± 0.44</td>
<td>5.17 ± 0.45</td>
<td>5.10 ± 0.46</td>
<td>1.05</td>
<td>-0.07 (-0.09; -0.05)</td>
<td>1×10⁻¹²</td>
</tr>
<tr>
<td>NFBC86</td>
<td>5,073</td>
<td>0.30</td>
<td>5.18 ± 0.40</td>
<td>5.10 ± 0.42</td>
<td>5.05 ± 0.40</td>
<td>1.10</td>
<td>-0.07 (-0.08; -0.05)</td>
<td>6×10⁻¹⁶</td>
</tr>
<tr>
<td>Obese children</td>
<td>861</td>
<td>0.28</td>
<td>4.94 ± 0.43</td>
<td>4.88 ± 0.43</td>
<td>4.80 ± 0.47</td>
<td>1.09</td>
<td>-0.07 (-0.11; -0.02)</td>
<td>4×10⁻³</td>
</tr>
<tr>
<td>Overall meta-analysis†</td>
<td>9,353</td>
<td>0.30</td>
<td>5.18 ± 0.43</td>
<td>5.11 ± 0.44</td>
<td>5.06 ± 0.44</td>
<td>0.91</td>
<td>-0.06 [-0.08; -0.05]</td>
<td>4×10⁻³²</td>
</tr>
</tbody>
</table>
rs560887
P-value = 5e-21

Coeff = -0.06 (-0.07 - -0.04)
mean mRNA expression signal (normalised for beta actin)

- islets (males)
- islets (females)
- MIN6 cells
- acinar tissue
- skeletal muscle
- diaphragm
- heart
- adipose tissue
- liver
- kidney
- small intestine
- lung
- brain cortex
- brain stem
- eye
- adrenal gland
- pituitary
- salivary gland
- seminal vesicles
- testis
- bone marrow
- thymus
- spleen
- foetus ED16
- placenta
- ES-cells