Genetic and Functional Assessment of the Role of the rs13431652-A and rs573225-A Alleles in the *G6PC2* Promoter that Strongly Associate With Elevated Fasting Glucose Levels

Supplementary Methods

Genotype/expression correlation in human pancreatic cDNAs.

Isolated pancreatic islets were prepared by collagenase digestion and density gradient purification (1; 2). After isolation, islets were cultured free-floating in M199 culture medium (Sigma-Aldrich) at 5.5 mmol/l glucose concentration. Within 3 days of isolation, the islets were used for RNA extraction. Total RNA was extracted from the islets using the RNeasy Protect Mini Kit (QIAGEN) and quantified by absorbance at A_{260}/A_{280} in a PerkinElmer spectrophotometer (to ensure integrity of RNA. The ratio between absorbance at 260 nm [A_{260}] and absorbance at 280 nm [A_{280}] was higher than 1.65) and its integrity was assessed after electrophoresis in 1.0% agarose gels by ethidium bromide staining. Human mRNA levels were quantified, following reverse transcription, by real-time quantitative RT-PCR (3). We used gene-specific probes and primer pairs for *G6PC2* (ID: Hs01549773_m1) (Applied Biosystems). Each sample was run in duplicate, and the transcript quantity was normalized to the mRNA level of *POLR2A* (ID: Hs00172187_m1).

Supplementary Results

Fusion genes containing each rs13431652 allele, generated in the context of the -8563 to +11 *G6PC2* promoter region, were analyzed by transient transfection of βTC-3 cells. Figure 3 in the main text shows that the rs13431652-G allele was associated with an approximately 25% decrease in promoter activity in comparison to that observed with the rs13431652-A allele. Supplementary Figure 2 shows that this difference was only observed when the *G6PC2* promoter was analyzed in the context of the pGL4 but not the pGL3 vector; the former lacks multiple cryptic transcription factor binding sites in the vector backbone and luciferase gene.

As previously demonstrated (4), two point mutations in the mouse Foxa2 binding site (Suppl. Table 6), at positions implicated in protein binding in βTC-3 cells by *in situ* footprinting (5), disrupt Foxa2 binding (Suppl. Fig. 3). Thus, a 100-fold molar excess of the unlabeled *G6pc2* Foxa WT oligonucleotide competed effectively for Foxa2 binding whereas an oligonucleotide, designated *G6pc2* Foxa MUT (Suppl. Table 6), failed to compete with the labeled probe for formation of the Foxa2-DNA complex (Suppl. Fig. 3). Supplementary Figure 3 shows that the equivalent mutations in the human Foxa2 binding site (Suppl. Table 6) also disrupt Foxa2 binding. Thus, a 100-fold molar excess of the unlabeled *G6PC2* Foxa WT oligonucleotide competed effectively for Foxa2 binding whereas an oligonucleotide, designated *G6PC2* Foxa MUT (Suppl. Table 6), failed to compete with the labeled probe for formation of the Foxa2-DNA complex (Suppl. Fig. 3).

To investigate the functional significance of Foxa2 binding to the human *G6PC2* promoter this site was mutated using site-directed mutagenesis (SDM) in the context of the human -324 to +3 *G6PC2* promoter region. A fusion gene, designated *G6PC2* Foxa SDM, containing a mutation identical to that in the *G6PC2* Foxa MUT oligonucleotide (Suppl. Table 6; Suppl. Fig. 3), was then analyzed by transient transfection of βTC-3 cells. Supplementary Figure 4 shows that mutation of the Foxa binding site resulted in an approximately 75% reduction in the level of reporter gene expression in βTC-3 cells as compared to that directed by the wild type -

324/+3 promoter. Mutation of the Foxa2 binding site in the mouse *G6pc2* promoter had a similar effect (Suppl. Fig. 4).

Fusion genes containing each rs573225 allele, generated in the context of the -324 to +3 *G6PC2* promoter region, were analyzed by transient transfection of βTC-3 cells. Figure 6 in the main text shows that the rs573225-G allele was associated with an approximately 65% increase in promoter activity in comparison to that observed with the rs573225-A allele. Supplementary Figure 5 shows that a similar effect of the rs573225-G allele was also seen in the HIT and Min6 cell lines. Supplementary Figure 6 shows that a similar effect of the rs573225-G allele was also seen when the *G6PC2* promoter was analyzed in the context of both the pGL4 and pGL3 vectors, respectively.

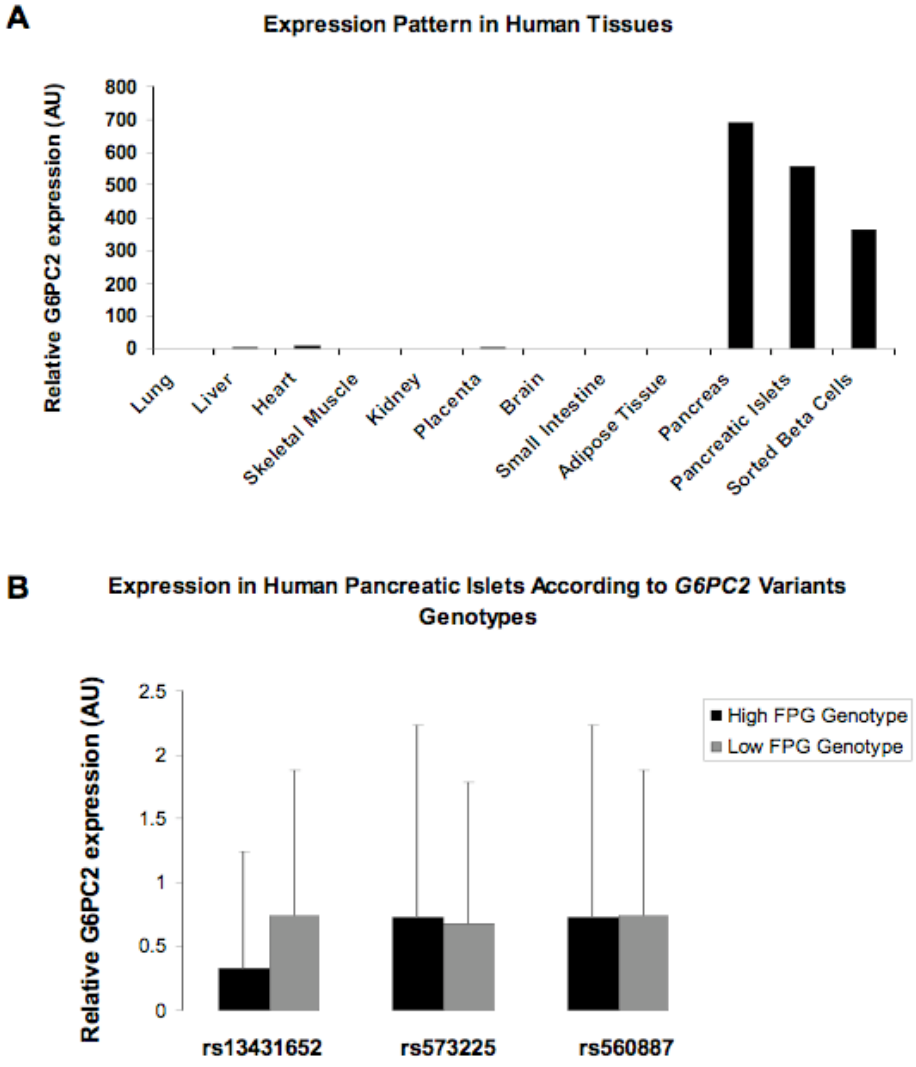
Dos Santos et al. (6) suggested that rs573225 is an epiSNP because methylation of the rs573225-G allele, which is located at a GpC dinucleotide within the Foxa2 binding site, affected Foxa2 binding. However, in contrast to the well-studied methylation of CpG dinucleotides (7), we can find no reports of the existence of methylated GpC dinucleotides in mammals, though such a modification does exist in fish (8). It is therefore highly unlikely that rs573225 is an epiSNP. Nevertheless, the sequence GATC is located 3' of rs573225 and partially overlaps the Foxa2 binding site (Table 1). Since GATC is the target sequence for the bacterial Dam methylase this site was likely to be methylated during the generation of plasmid DNA. This then raised the possibility that bacterial methylation was artificially affecting Foxa2 binding and hence function in fusion gene experiments (9). To address this possibility the fusion gene experiments were repeated using plasmids grown in SCS110 Dam and Dcm methylase deficient bacteria (Stratagene). Supplementary Figure 7 shows that Dam methylation of the Foxa2 binding site did not alter the stimulatory effect of rs573225 on fusion gene expression.

References

1. Del Guerra S, Lupi R, Marselli L, Masini M, Bugliani M, Sbrana S, Torri S, Pollera M, Boggi U, Mosca F, Del Prato S, Marchetti P: Functional and molecular defects of pancreatic islets in human type 2 diabetes. *Diabetes* 54:727-735, 2005
2. Lupi R, Del Guerra S, Fierabracci V, Marselli L, Novelli M, Patane G, Boggi U, Mosca F, Piro S, Del Prato S, Marchetti P: Lipotoxicity in human pancreatic islets and the protective effect of metformin. *Diabetes* 51 Suppl 1:S134-137, 2002
3. Masini M, Bugliani M, Lupi R, del Guerra S, Boggi U, Filipponi F, Marselli L, Masiello P, Marchetti P: Autophagy in human type 2 diabetes pancreatic beta cells. *Diabetologia* 52:1083-1086, 2009
4. Martin C, Flemming B, Wang Y, Oeser J, O'Brien R: Foxa2 and MafA Regulate Islet-specific Glucose-6-Phosphatase Catalytic Subunit-Related Protein (IGRP/G6PC2) Gene Expression. *J Mol Endocrinol* 41:315-328, 2008
5. Bischof LJ, Martin CC, Svitek CA, Stadelmaier BT, Hornbuckle LA, Goldman JK, Oeser JK, Hutton JC, O'Brien RM: Characterization of the Mouse Islet-Specific Glucose-6-Phosphatase Catalytic Subunit-Related Protein Gene Promoter by In Situ Footprinting. Correlation with Fusion Gene Expression in the Islet Derived bTC-3 and Hamster Insulinoma Tumor Cell Lines. *Diabetes* 50:502-514, 2001
6. Dos Santos C, Bougneres P, Fradin D: An SNP in a methylatable Foxa2 binding site of the G6PC2 promoter is associated with insulin secretion in vivo and increased promoter activity in vitro. *Diabetes* 58:489-492, 2009

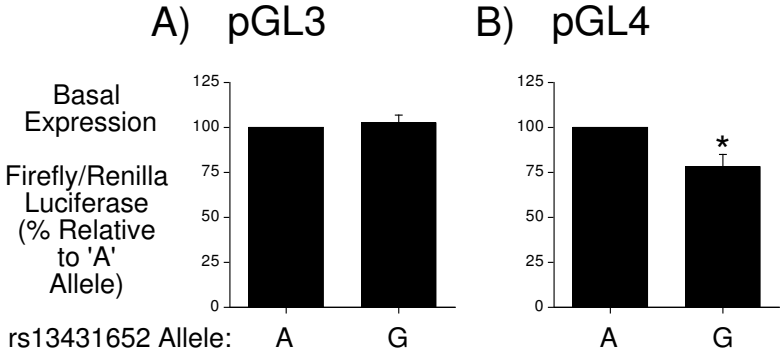
7. Miranda TB, Jones PA: DNA methylation: the nuts and bolts of repression. *J Cell Physiol* 213:384-390, 2007
8. Pontecorvo G, De Felice B, Carfagna M: Novel methylation at GpC dinucleotide in the fish Sparus aurata genome. *Mol Biol Rep* 27:225-230, 2000
9. Tronche F, Rollier A, Bach I, Weiss MC, Yaniv M: The rat albumin promoter: cooperation with upstream elements is required when binding of APF/HNF1 to the proximal element is partially impaired by mutation or bacterial methylation. *Mol Cell Biol* 9:4759-4766, 1989
10. Overdier DG, Porcella A, Costa RH: The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. *Mol Cell Biol* 14:2755-2766, 1994

Supplementary Figure 1. Analysis of G6PC2 gene expression. *G6PC2* expression was analyzed in human tissues (Panel A) and in pancreatic islets, relative to *G6PC2* genotypes (Panel B). *G6PC2* expression was analyzed by real-time PCR and data is normalised according to expression levels of the house keeping gene *POLR2A*.



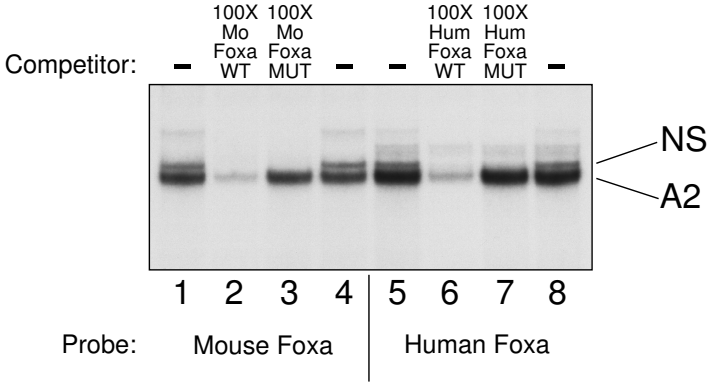
Supplementary Figure 2. The human *G6PC2* rs13431652-G allele is associated with decreased *G6PC2* promoter activity relative to the rs13431652-A allele in the context of the pGL4 but not the pGL3 vector.

βTC-3 cells were transiently co-transfected, as described in Research Design and Methods, using a lipofectamine solution containing human *G6PC2*-luciferase fusion genes in the pGL3 MOD **(Panel A)** or pGL4 MOD **(Panel B)** vectors (2 μg) and an expression vector encoding *Renilla* luciferase (0.5 μg). The *G6PC2*-luciferase fusion genes represented the rs13431652-A or -G alleles present in the context of the human *G6PC2* promoter sequence located between -8563 and +11. Following transfection, cells were incubated for 18-20 hr in serum-containing medium. The cells were then harvested and firefly and *Renilla* luciferase activity was assayed as described in Research Design and Methods. Results are presented as the ratio of firefly:*Renilla* luciferase activity, expressed as a percentage relative to the value obtained with the rs13431652-A allele, and represent the mean of 9-12 experiments ± S.E.M., each using an independent preparation of each fusion gene plasmid, assayed in triplicate. *, $p < 0.05$ *versus* rs13431652-A allele.



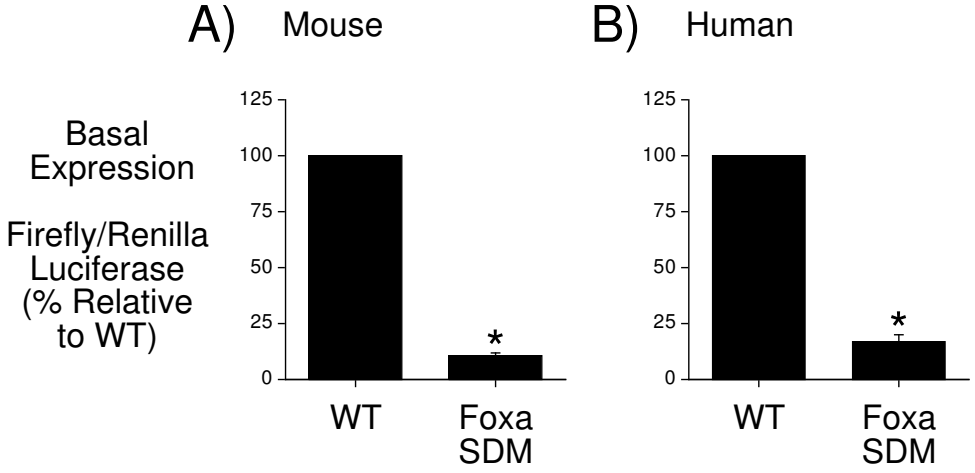
Supplementary Figure 3. Comparison of Foxa2 binding to the labeled human *G6PC2* -265/-246 and mouse *G6pc2* -247/-228 promoter regions *in vitro*.

Labeled oligonucleotides representing the wild-type (WT) human *G6PC2* -265/-246 or mouse *G6pc2* -247/-228 promoter regions were incubated in the absence or presence of a 100-fold molar excess of the indicated unlabeled WT or mutant (MUT) competitors (Suppl. Table 6). βTC-3 nuclear extract was then added and protein binding was analyzed using the gel retardation assay as described in Research Design and Methods. In the representative audioradiograph shown, only the retarded complexes are visible and not the free probe, which was present in excess. The two major complexes detected represent Foxa2 (A2) and non-specific (NS) binding (4).



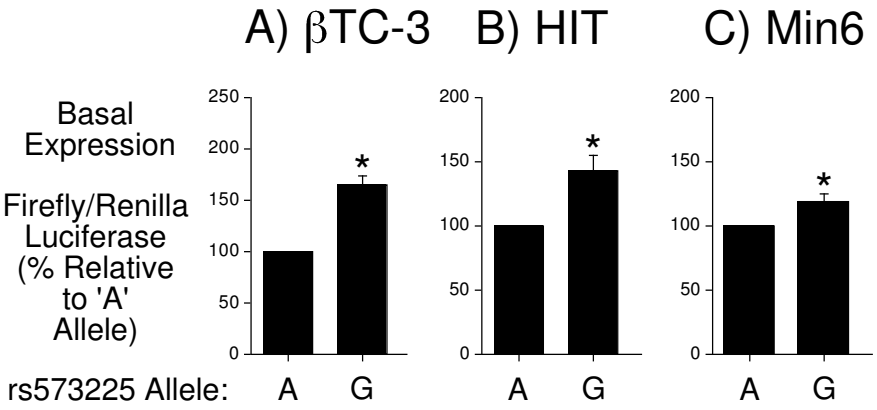
Supplementary Figure 4. Marked disruption of Foxa2 binding reduces human *G6PC2* and mouse *G6pc2* promoter activity.

βTC-3 cells were transiently co-transfected, as described in Research Design and Methods, using a lipofectamine solution containing various *G6PC2*-luciferase fusion genes in the pGL3 MOD vector (2 μg) and an expression vector encoding *Renilla* luciferase (0.5 μg). The *G6PC2*-luciferase fusion genes represented either the wild-type (WT) mouse promoter sequence, located between -306 and +3 (**Panel A)**, the WT human promoter sequence (rs573225-A allele), located between -324 and +3 (**Panel B)**, or the same sequences with a site-directed mutations (SDMs) in the Foxa2 binding sites. The mutations were identical to those used in the gel retardation analysis (Suppl. Fig. 3). Following transfection, cells were incubated for 18-20 hr in serum-containing medium. The cells were then harvested and firefly and *Renilla* luciferase activity was assayed as described in Research Design and Methods. Results are presented as the ratio of firefly:*Renilla* luciferase activity, expressed as a percentage relative to the value obtained with the WT fusion genes, and represent the mean of 3 experiments ± S.E.M., each using an independent preparation of each fusion gene plasmid, assayed in triplicate. *, $p < 0.05$ *versus* WT.



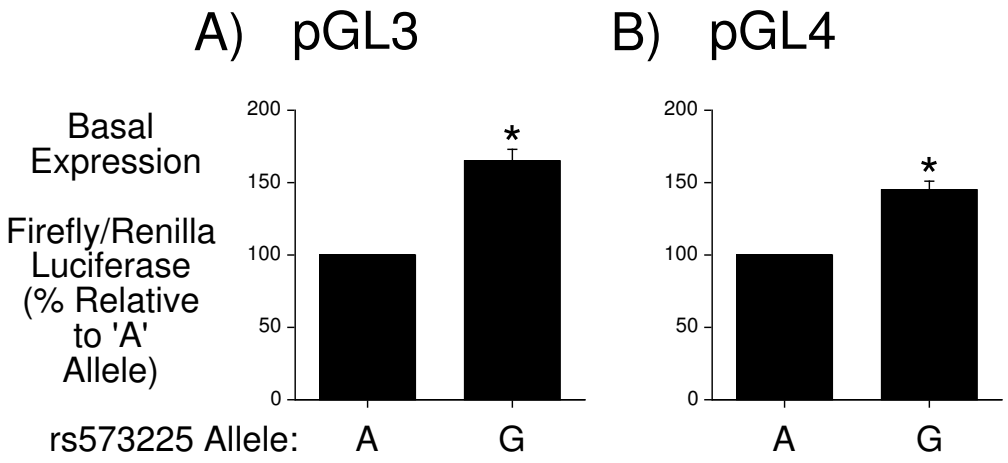
Supplementary Figure 5. The human *G6PC2* rs573225-G allele is associated with increased *G6PC2* promoter activity relative to the rs573225-A allele.

βTC-3 (**Panel A**), HIT (**Panel B**) and Min6 (**Panel C**) cells were transiently co-transfected, as described in Research Design and Methods, using a lipofectamine solution containing various *G6PC2*-luciferase fusion genes in the pGL3 MOD vector (2 μg) and an expression vector encoding *Renilla* luciferase (0.5 μg). The *G6PC2*-luciferase fusion genes represented the rs573225-A or G alleles present in the context of the human *G6PC2* promoter sequence located between -324 and +3. Following transfection, cells were incubated for 18-20 hr in serum-containing medium. The cells were then harvested and firefly and *Renilla* luciferase activity was assayed as described in Research Design and Methods. Results are presented as the ratio of firefly:*Renilla* luciferase activity, expressed as a percentage relative to the value obtained with the rs573225-A allele, and represent the mean of 3 experiments ± S.E.M., each using an independent preparation of each fusion gene plasmid, assayed in triplicate. *, $p < 0.05$ *versus* rs573225-A allele.



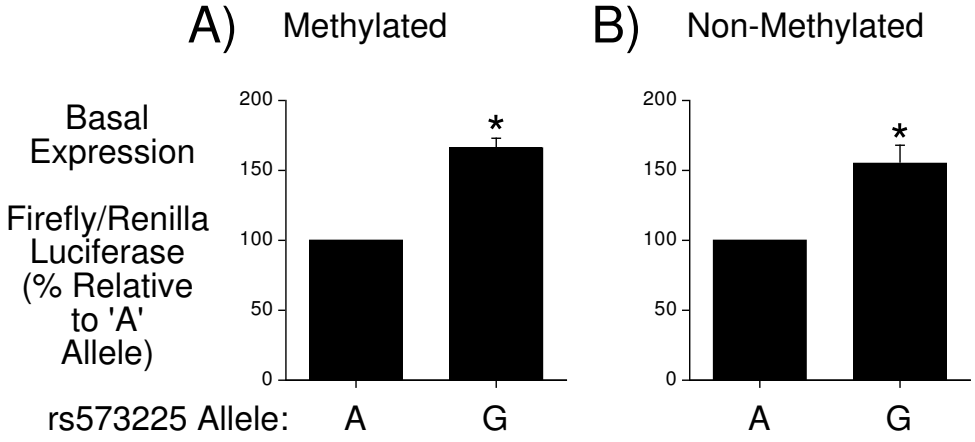
Supplementary Figure 6. The human *G6PC2* rs573225-G allele is associated with increased *G6PC2* promoter activity relative to the rs573225-A allele in the context of the pGL3 and pGL4 vectors.

βTC-3 cells were transiently co-transfected, as described in Research Design and Methods, using a lipofectamine solution containing human *G6PC2*-luciferase fusion genes in the pGL3 MOD **(Panel A)** or pGL4 MOD **(Panel B)** vectors (2 μg) and an expression vector encoding *Renilla* luciferase (0.5 μg). The *G6PC2*-luciferase fusion genes represented the rs573225-A or -G alleles present in the context of the human *G6PC2* promoter sequence located between -324 and +3. Following transfection, cells were incubated for 18-20 hr in serum-containing medium. The cells were then harvested and firefly and *Renilla* luciferase activity was assayed as described in Research Design and Methods. Results are presented as the ratio of firefly:*Renilla* luciferase activity, expressed as a percentage relative to the value obtained with the rs573225-A allele, and represent the mean of 3 experiments ± S.E.M., each using an independent preparation of each fusion gene plasmid, assayed in triplicate. *, $p < 0.05$ *versus* rs573225-A allele.



Supplementary Figure 7. The human *G6PC2* rs573225-G allele is associated with increased *G6PC2* promoter activity relative to the rs573225-A allele independent of DNA methylation status.

βTC-3 cells were transiently co-transfected, as described in Research Design and Methods, using a lipofectamine solution containing various *G6PC2*-luciferase fusion genes in the pGL3 MOD vector (2 μg) and an expression vector encoding *Renilla* luciferase (0.5 μg). The *G6PC2*-luciferase fusion genes represented the rs573225-A or -G alleles present in the context of the human *G6PC2* promoter sequence located between -324 and +3. The fusion gene plasmids were grown in methylase containing DH5α (**Panel A)** or methylase deficient SCS110 (**Panel B)** bacterial cells. Following transfection, cells were incubated for 18-20 hr in serum-containing medium. The cells were then harvested and firefly and *Renilla* luciferase activity was assayed as described in Research Design and Methods. Results are presented as the ratio of firefly:*Renilla* luciferase activity, expressed as a percentage relative to the value obtained with the WT fusion genes, and represent the mean of 3 experiments ± S.E.M., each using an independent preparation of each fusion gene plasmid, assayed in triplicate. *, $p < 0.05$ *versus* rs573225-A allele.



Supplementary Table 1. Clinical characteristics of the study populations with successful genotyping for the four SNPs studied.

	DESIR	HAGUENAU	NFBC 1986	Obese children
N (Men / Women)	3483 (1620/1863)	1201 (578/623)	4372 (2137/2235)	476 (228/248)
Age (years)	46.87 +/- 9.99	22.27 +/- 3.95	16 +/- 0	10.89 +/- 3.14
BMI (kg/m2)	24.29 +/- 3.57	22.63 +/- 4.15	21.27 +/- 3.56	28.42 +/- 6.13
FPG (mmol/l)	5.19 +/- 0.44	4.79 +/- 0.36	5.13 +/- 0.41	4.9 +/- 0.46
Fasting Insulin (pmol/l)	43.91 +/- 26.36	37.33 +/- 24.92	76.66 +/- 42.49	14.52 +/- 10.98
Homa%B	90.66 +/- 61.11	87.5 +/- 74.08	137.2 +/- 86.72	220.95 +/- 179.78
Homa-IR	1.67 +/- 0.98	1.11 +/- 0.76	2.46 +/- 1.45	3.22 +/- 2.53
Triglycerides (mmol/l)	1.1 +/- 0.72	1.06 +/- 0.53	0.83 +/- 0.41	1.06 +/- 0.56
HDL-cholesterol (mmol/l)	1.67 +/- 0.43	1.41 +/- 0.34	1.41 +/- 0.29	2.72 +/- 0.73
LDL-cholesterol (mmol/l)	3.53 +/- 0.91	2.82 +/- 0.86	2.25 +/- 0.58	2.73 +/- 0.71
Total-cholesterol (mmol/l)	5.69 +/- 0.98	4.71 +/- 0.94	4.26 +/- 0.79	4.46 +/- 0.84
Waist (cm)	81.8 +/- 11.26	77.99 +/- 11.32	73.85 +/- 8.59	90.7 +/- 16.91
Waist to hip ratio	0.85 +/- 0.09	0.8 +/- 0.08	0.8 +/- 0.06	0.92 +/- 0.08
SBP (mm Hg)	128.37 +/- 15.21	120.99 +/- 10.79	115.84 +/- 12.69	113.01 +/- 14.1
DBP (mm Hg)	78.72 +/- 9.54	62.95 +/- 9.83	67.99 +/- 7.99	66.92 +/- 11.26

Supplementary Table 2. Linkage desequillibrium (r^2) per population between the SNPs studied

DESIR	**rs560887**	**rs13431652**	**rs573225**	**rs853789**
rs560887	1			
rs13431652	0.62	1		
rs573225	0.69	0.91	1	
rs853789	0.55	0.69	0.71	1

NFBC 1986	**rs560887**	**rs13431652**	**rs573225**	**rs853789**
rs560887	1			
rs13431652	0.94	1		
rs573225	0.92	0.95	1	
rs853789	0.78	0.77	0.74	1

Haguenau	**rs560887**	**rs13431652**	**rs573225**	**rs853789**
rs560887	1			
rs13431652	0.75	1		
rs573225	0.90	0.82	1	
rs853789	0.65	0.74	0.65	1

Obese Children	**rs560887**	**rs13431652**	**rs573225**	**rs853789**
rs560887	1			
rs13431652	0.51	1		
rs573225	0.55	0.91	1	
rs853789	0.42	0.66	0.68	1

.

Supplementary Table 3. Haplotype analyses to assess contribution to association with FPG between promoter variants (rs13431652 and rs573225) and intronic variant rs560887 and between *G6PC2* variants (rs13431652, rs573225 and rs560887) and an *ABCB11 variant* (rs853789).

M0 = model that considers no haplotype effect
M1 = model that considers all haplotypes individual effect
Mno1 = model that considers that SNP1 has no effect on haplotype association
Mno2 = model that considers that SNP2 has no effect on haplotype association

We have tested three hypotheses and reported the p-values in the following tests:
A : P-value of the likelihood-ratio test for the comparison of M0 and M1 (general association)
B : P-value of the likelihood-ratio test for the comparison of M1 and Mno1 (effect of SNP1 in the final association)
C : P-value of the likelihood-ratio test for the comparison of M1 and Mno2 (effect of SNP2 in the final association)

	DESIR (N=3,483)			**NFBC 1986 (N=4,372)**			**Haguenau (N=1,201)**			**Obese Ch (N=476)**		
SNP1=rs560887	A	B	C	A	B	C	A	B	C	A	B	C
SNP2=rs13431652	6.39×10^{-17}	7.58×10^{-5}	1.21×10^{-8}	2.43×10^{-14}	0.006	0.472	2.61×10^{-6}	0.219	0.704	0.024	0.477	0.054
SNP2=rs573225	2.06×10^{-15}	0.015	4.09×10^{-7}	4.75×10^{-14}	0.075	0.933	5.57×10^{-6}	0.108	0.460	0.029	0.382	0.067
SNP1=rs853789	A	B	C	A	B	C	A	B	C	A	B	C
SNP2=rs13431652	4.99×10^{-13}	0.0001	0.0009	5.92×10^{-15}	0.11	2.29×10^{-7}	5.88×10^{-6}	0.49	0.008	0.02	0.04	0.39
SNP2=rs560887	1.54×10^{-13}	0.20	0.0003	7.23×10^{-13}	0.20	3.01×10^{-5}	2.01×10^{-5}	0.88	0.03	0.03	0.53	0.53
SNP2=rs573225	6.18×10^{-14}	0.49	0.0001	5.27×10^{-15}	0.008	2.04×10^{-7}	4.58×10^{-5}	0.96	0.06	0.04	0.53	0.82

Supplementary Table 4. Assessment of the effect of rs853789 on FPG and its independency from *G6PC2* variants. [£] Regression model included age, gender, cohort, BMI and * SNP. In these models, rs13431652 rs573225 and rs560887 significance were, $P=1.2\times10^{-10}$, $P=1.8\times10^{-9}$ and $P=1.3\times10^{-7}$, respectively.

	N	Beta (se)	P
DESIR	3,483	-0.070 (0.010)	1.1×10^{-11}
NFBC 1986	4,372	-0.054 (0.009)	2.3×10^{-10}
Haguenau	1,201	-0.061 (0.014)	3.4×10^{-5}
Obese Children	476	-0.091 (0.03)	0.002
Meta-analysis£	9,532	-0.062 (0.006)	1.1×10^{-25}
Meta-analysis conditioned by rs13431652*	9,532	0.002 (0.01)	0.88
Meta-analysis conditioned by rs573225*	9,532	-0.006 (0.01)	0.61
Meta-analysis conditioned by rs560887*	9,532	-0.02 (0.01)	0.038

Supplementary Table 5. The effect size and significance of sex, age and BMI on FPG in the DESIR cohort.

	Beta estimate	SE	P value
Sex	-0.198	0.014	6.3×10^{-43}
Age	0.008	0.001	4.1×10^{-28}
BMI	0.022	0.002	2.9×10^{-27}

Supplementary Table 6. Oligonucleotides used in mouse *G6pc2* and human *G6PC2* gel retardation studies.

The sense strand sequence of the wild type (WT) and mutant (MUT) oligonucleotides used in these studies are shown. Mutated base pairs are shown in bold lower case letters. The consensus Foxa binding motif is taken from Ref. (10).

```
Mo G6pc2 Foxa WT      -247 TTGCAAGCAAACACGATCCA -228

Mo G6pc2 Foxa MUT     -247 TTGCAAcgAAACACGATCCA -228

Hum G6PC2 Foxa WT     -265 TTTCAAACAAACATGATCCA -246

Hum G6PC2 Foxa MUT    -265 TTTCAAcgAAACATGATCCA -246

Foxa Consensus                AAGCCAACATTT
                              G ATA  T A G
                                     C
```