Rare variants in *PPARG* with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes

Amit R. Majithia[a,b,c], Jason Flannick[a], Peter Shahinian[a], Michael Guo[d], Mark-Anthony Bray[a], Pierre Fontanillas[a], Stacey B. Gabriel[a], GoT2D Consortium[2], NHGRI JHS/FHS Allelic Spectrum Project[2], SIGMA T2D Consortium[2], T2D-GENES Consortium[2], Evan D. Rosen[c,e], and David Altshuler[a,b,c,d,f,1]

[a]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142; [b]Department of Molecular Biology and Diabetes Unit, Massachusetts General Hospital, Boston, MA 02114; [c]Department of Medicine, Harvard Medical School, Boston, MA 02115; [d]Department of Genetics, Harvard Medical School, Boston, MA 02115; [e]Division of Endocrinology and Metabolism, Beth Israel Deaconess Medical Center, Boston, MA 02115; and [f]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited[†] by Edward Scolnick, Massachusetts Institute of Technology, Cambridge, MA, and approved July 21, 2014 (received for review June 5, 2014)

Peroxisome proliferator-activated receptor gamma (PPARG) is a master transcriptional regulator of adipocyte differentiation and a canonical target of antidiabetic thiazolidinedione medications. In rare families, loss-of-function (LOF) mutations in *PPARG* are known to cosegregate with lipodystrophy and insulin resistance; in the general population, the common P12A variant is associated with a decreased risk of type 2 diabetes (T2D). Whether and how rare variants in *PPARG* and defects in adipocyte differentiation influence risk of T2D in the general population remains undetermined. By sequencing PPARG in 19,752 T2D cases and controls drawn from multiple studies and ethnic groups, we identified 49 previously unidentified, nonsynonymous *PPARG* variants (MAF < 0.5%). Considered in aggregate (with or without computational prediction of functional consequence), these rare variants showed no association with T2D (OR = 1.35; $P = 0.17$). The function of the 49 variants was experimentally tested in a novel high-throughput human adipocyte differentiation assay, and nine were found to have reduced activity in the assay. Carrying any of these nine LOF variants was associated with a substantial increase in risk of T2D (OR = 7.22; $P = 0.005$). The combination of large-scale DNA sequencing and functional testing in the laboratory reveals that approximately 1 in 1,000 individuals carries a variant in *PPARG* that reduces function in a human adipocyte differentiation assay and is associated with a substantial risk of T2D.

Type 2 diabetes (T2D) is a common, complex disease caused by insulin resistance in multiple peripheral tissues combined with inadequate beta-cell response. In the general population, a nonsynonymous P12A variant in peroxisome proliferator-activated receptor gamma (PPARG), a transcriptional regulator of adipocyte differentiation and canonical target of antidiabetic thiazolidinediones, has been associated with decreased risk of T2D (1, 2). It has been challenging to document the impact of this common polymorphism on function in human cell-based assays. For P12A, the variant is very common, but the magnitude of effect on disease risk is modest (20% decreased risk of T2D) (3). In rare families with syndromic lipodystrophy, loss-of-function (LOF) mutations in *PPARG* that prohibit adipocyte differentiation in vitro, have been found that segregate with lipodystophy, insulin resistance, and T2D (4, 5). The magnitude of effect on individual and cellular phenotypes is strong, but the mutations are extremely rare. Whether LOF mutations in PPARG play a broader role in T2D, and whether these mutations implicate a role for adipocyte differentiation in T2D, have not previously been characterized.

More generally, exome sequencing now enables the systematic identification of all nonsynonymous variants, common and rare, in population and clinical cohorts. However, interpretation of rare variants—even those that alter protein sequence—is challenging: The overwhelming majority of nonsynonymous variants in any given sample are extremely rare, and only a minority alters protein function. For example, the NHLBI exome Sequencing Project identified 285,000 nonsynonymous variants in 2,440 individuals (6). Eighty-two percent were previously uncharacterized and over half were observed in a single individual. Bioinformatic analysis predicted that only 2% significantly alter protein function.

We hypothesized that individuals in the general population might harbor rare, nonsynonymous variants in *PPARG*, that only a subset of these variants would alter function in an adipocyte differentiation assay, and that such variants might be associated with a risk of T2D. We further hypothesized that the effect of these variants on type 2 diabetes risk in the general population might in some cases be less severe than that estimated in individuals ascertained based on syndromic lipodystrophy (7). To evaluate this hypothesis we sequenced *PPARG* in 19,752 multiethnic T2D cases/control samples, characterized each nonsynonymous variant through parallel bioinformatic and experimental approaches, and compared the T2D risk of individuals carrying benign and LOF variants.

Results

Identification of Nonsynonymous *PPARG* Variants from the Population. After sequencing and analyzing all exons of *PPARG* in 19,752 multiethnic individuals (9,070 T2D cases and 10,682 controls;

Significance

Genome sequencing of individuals in the population reveals new mutations in almost every protein coding gene; interpreting the consequence of these mutations for human health and disease remains challenging. We sequenced the gene *PPARG*, a target of antidiabetic drugs, in nearly 20,000 individuals with and without type 2 diabetes (T2D). We identified 49 previously unidentified protein-altering mutations, characterized their cellular function in human cells, and discovered that nine of these mutations cause loss-of-function (LOF). The individuals who carry these nine LOF mutations have a sevenfold increased risk of T2D, whereas individuals carrying mutations we classify as benign have no increased risk of T2D.

Author contributions: A.R.M., E.D.R., and D.A. designed research; A.R.M., P.S., and M.G. performed research; A.R.M., M.-A.B., S.B.G., G.C., N.J./F.A.S.P., S.T.C., and T.-G.C. contributed new reagents/analytic tools; A.R.M., J.F., and P.F. analyzed data; and A.R.M. and D.A. wrote the paper.

The authors declare no conflict of interest.

[†]This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: altshuler@molbio.mgh.harvard.edu.

[2]See *SI Appendix* for a complete list of the investigators of the GoT2D Consortium, NHGRI JHS/FHS Allelic Spectrum Project, SIGMA T2D Consortium, and T2D-GENES Consortium.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1410428111/-/DCSupplemental.

SI Appendix, Table 1), 53 nonsynonymous *PPARG* variants were observed. Only one of these variants (the well-studied *PPARG* P12A variant, rs1801282) demonstrated a minor allele frequency greater than 1% in any ancestry group we studied (*SI Appendix*, Table 2). As expected, carriers of the common *PPARG* P12A variant showed a reduced risk of T2D, consistent with previous

Table 1. Rare, nonsynonymous variants in *PPARG* identified from 19,752 T2D case/controls

Location on chromosome 3	Base change	Amino acid change	Ancestral geography	Counts in controls	Counts in T2D cases	Bioinformatic prediction[†]	OR (95% CI)	*P*
12458632	G > T	A417S	European	0	1	Deleterious		
12447449	G > T	D230Y	South Asian	0	1	Deleterious		
12447410	G > A	E217K	Hispanic	0	1	Deleterious		
12458359	G > A	E326K	Hispanic	0	1	Deleterious		
12434116	T > G	F162V	European	0	1	Deleterious		
12434114	G > A	G161D	European	0	2	Deleterious		
12434179	C > T	H183Y	Hispanic	1	0	Deleterious		
12434133	C > G	I167M	European, European American	1	1	Deleterious		
12458374	A > G	I331V	South Asian	1	0	Deleterious		
12475511	A > G	K462R	Hispanic	0	1	Deleterious		
12475583	A > C	K486T	South Asian	1	1	Deleterious		
12434164	C > A	L178I	European	1	5	Deleterious		
12458466	G > C	L361F	European American	1	1	Deleterious	2.11	0.12
12475403	C > T	P426L	European	0	1	Deleterious	(0.82–5.45)	
12475486	C > G	P454A	Hispanic	4	2	Deleterious		
12422871	C > T	Q121*	European American	1	0	Deleterious		
12422929	G > A	R140H	Hispanic, African American	1	1	Deleterious		
12434126	G > C	R165T	European	0	2	Deleterious		
12447479	C > T	R240W	South Asian	1	0	Deleterious		
12458306	G > T	R308L	European	0	1	Deleterious		
12458516	G > A	R378K	European	0	1	Deleterious		
12475399	C > T	R425C	European	0	1	Deleterious		
12422908	C > A	S133Y	European	0	1	Deleterious		
12447507	C > G	S249*	European	0	1	Deleterious		
12458613	C > A	S410R	Hispanic	1	0	Deleterious		
12421260	C > G	S47C	East Asian	0	1	Deleterious		
12458335	G > A	V318M	European	0	1	Deleterious		
12447537	C > T	A259V	European American	1	0	Benign		
12458594	C > T	A404V	Hispanic	0	1	Benign		
12475457	C > T	A444V	European American	1	0	Benign		
12421391	G > A	A91T	African American	3	0	Benign		
12447572	G > A	D271N	European	0	1	Benign		
12421266	A > C	D49A	Hispanic	1	0	Benign		
12421267	T > G	D49E	African American	2	2	Benign		
12475490	A > G	E455G	European American	1	0	Benign		
12421355	G > A	E79K	European, East Asian	1	4	Benign		
12393119	A > G	I10V	South Asian	1	1	Benign		
12434131	A > G	I167V	European	0	1	Benign		
12447512	A > G	I251V	Hispanic	0	1	Benign		
12421253	A > T	I45F	African American	3	0	Benign		
12458511	G > A	M376I	European	0	2	Benign		
12421279	G > A	M53I	South Asian	1	0	Benign		
12422880	A > G	N124D	South Asian	1	0	Benign		
12475424	C > T	P433L	Hispanic, European	0	2	Benign		
12458611	A > T	S410C	European	0	1	Benign		
12421343	A > C	T75P	Hispanic	1	3	Benign		
12458209	G > A	V276I	European, Hispanic, African American, East Asian	5	6	Benign		
12458386	G > C	V335L	African American, Hispanic	11	9	Benign		
12421262	G > A	V48M	European American	1	0	Benign		
12421274	G > A	V52I	African American, East Asian, European	3	2	Benign		
12422856	T > G	Y116D	South Asian	1	0	Benign		
12458216	A > G	Y278C	European	0	1	Benign		

The variant position is based on human genome build NCBI36/hg18, and the amino acid position is based on the NCBI protein reference sequence NP_005028.4. Release notes for this genome build are available at www.ncbi.nlm.nih.gov/genome/guide/human/release_notes.html#b36. CI, confidence interval.

*Stop codon.

[†]Criteria for deleterious: A variant must have a mammalian conservation LOD score >10 and be categorized as damaging by Condel (17) (*Methods*).

reports (odds ratio = 0.85; 95% confidence interval 0.78–0.93; P = 0.0006) (3).

The remaining 52 variants were observed in 120 individuals (Table 1), yielding an aggregate frequency of 0.6% in the population. The most frequently occurring variant in any ethnic group, p.V335L, was observed at a frequency of 0.7% (20 individuals of African-American ancestry). The majority of the variants (33 of 52 or 63%) were observed in only a single individual. Nonsynonymous variants were identified in every ancestry group sampled: European, East Asian, South Asian, European American, African American, and Hispanic. Some variants were specific to individuals from one ethnic background, whereas others were observed in individuals across multiple ethnic backgrounds. Every individual with a rare, nonsynonymous *PPARG* was heterozygous for that variant and carried only one rare, nonsynonymous *PPARG* variant. The *PPARG* variants identified were distributed across the entire protein-coding region and included variants in all previously identified functional domains. Two variants (p.Q121* and p.S249*) were predicted to result in premature termination of the protein. Of the variants we identified, 49 are previously unidentified and 3 (p.R165T, p.V318M, and p.R425C) have been previously reported to segregate with disease in families with familial partial lipodystrophy 3 (FPLD3).

Association analysis for T2D was performed comparing individuals carrying any rare missense *PPARG* variant (with frequency <1% in the study sample) to those who carried no such variant; no significant association was observed (odds ratio of 1.36; 95% confidence interval 0.87–2.11; P > 0.17). Next, variants were classified as benign or deleterious (Table 1) based on bioinformatic annotation combining computational prediction, evolutionary conservation, and variant frequency (restricted to variants observed in a single individual or the less stringent minor allele frequency <0.1%). The strongest association was for variants classified as deleterious (odds ratio of 2.11; 95% confidence interval 0.82–5.45); again, the result was not nominally significant (P > 0.12) despite nearly 20,000 samples.

Functional Assessment of Nonsynonymous *PPARG* Variants. Recognizing that the majority of rare protein-coding variants are benign or very mildly deleterious, and that computational prediction remains imperfect (8), we set out to experimentally characterize the function of each nonsynonymous *PPARG* variant by genetic complementation in an assay measuring differentiation of human preadipocytes. Specifically, we developed a quantitative adipocyte differentiation assay in human Simpson–Golabi–Behmel syndrome (SBGS) preadipocytes by combining high-content microscopy with a custom automated image analysis pipeline (Fig. 1*A*). This assay compared favorably with standard triglyceride quantification methods using Oil Red O staining and extraction (Fig. 1*B*) with the advantages of accelerated throughput and an explicit measurement of total cell number. To isolate the effect of exogenous *PPARG* variants on adipocyte differentiation, preadipocytes were exposed to a submaximal differentiation mixture that only permitted differentiation in the

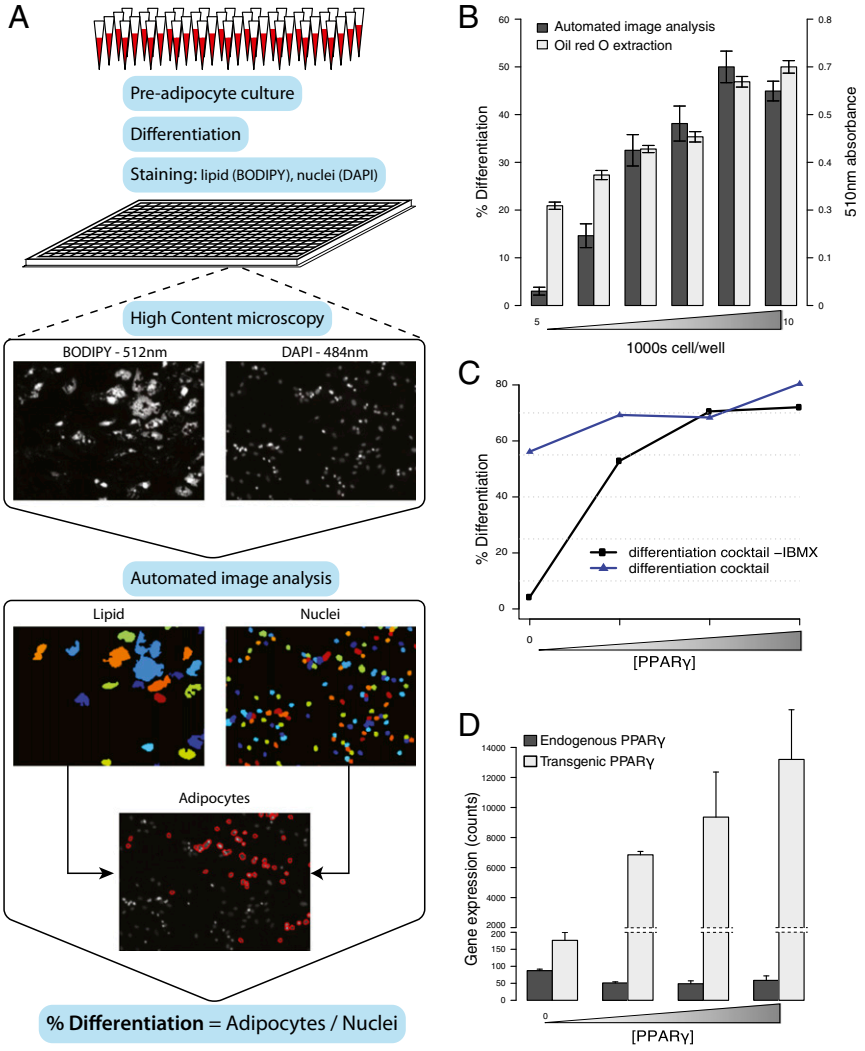


Fig. 1. High-throughput quantification of adipocyte differentiation in response to exogenous PPARγ. (*A*) Preadipocytes are cultured in 96-well plates, differentiated for 8 d, and stained for lipid (BODIPY) and nuclei (DAPI). Each well is imaged in a high-content microscope for lipid and nuclei. Adipocytes and undifferentiated cells are identified by the overlay of lipid and nuclei from automated image analysis. (*B*) Preadipocytes were plated at increasing density and differentiated. Parallel samples were subjected to image-based differentiation measurement or Oil Red O staining followed by lipid extraction and spectrophotometric quantification. (*C*) Preadipocyte differentiation in response to increasing doses of exogenous PPARγ with and without 3-isobutyl-1-methylxanthine (IBMX). (*D*) Gene expression levels in preadipocytes of endogenous and exogenous PPARγ in response to increasing doses of exogenous PPARγ. Error bars indicate ±1 SEM.

presence of functioning, exogenous PPARG (Fig. 1*C*) and maintained endogenous PPARG at background levels (Fig. 1*D*).

Each nonsynonymous *PPARG* variant identified from population-based sequencing was engineered into a construct in vitro, and tested for its ability to rescue adipocyte differentiation in SGBS preadipocytes (Fig. 2*A*). The empirical distribution of WT PPARG function in this assay was defined using multiple independent replicates of WT PPARG clones, with reduced function in the assay defined as adipocyte differentiation index falling below this null distribution in a one-tailed *t* test with a $P < 0.05$ threshold. Variants previously reported to be benign (Fig. 2*A*, blue bars) and to cause LOF (Fig. 2*A*, red bars) were generated and tested in parallel as positive and negative controls, respectively. Among these previously characterized variants, those characterized as benign (Fig. 2*A*, blue bars) stimulate adipocyte differentiation with similar efficacy as WT PPARG whereas those known to cause reduced protein activity (Fig. 2*A*, red bars) show decreased ability to stimulate adipocyte differentiation to varying degrees. Consistent with prior work, variants reported to segregate with disease in FPLD3 families show the most severe LOF with those that reside in the DNA binding domain (p.R165T, p.C159Y, and p.Y151C), almost completely inactivating PPARG (9–11).

Using this assay we classified the 53 missense variants observed in population screening. Forty-one of the rare missense variants were scored as benign when tested in the assay: they stimulated adipocyte differentiation in a manner that fell within the 95% confidence interval based on replicates of WT PPARG. (The common P12A variant was at the lower limit of the normal range.) However, 12 variants fell below the 95% confidence limit for WT PPARG constructs. Of the 12 with reduced activity, 3 were previously reported as LOF mutations observed in patients with lipodystrophy, and 9 were previously unidentified. Novel variants with reduced function were identified in the DNA binding, the hinge, and the ligand-binding domains of PPARG (Fig. 2*B*). Notably, whereas all previously identified mutations in the DNA-binding domain (from families segregating FPLD3) completely inactivate PPARG, in study samples ascertained for common disease, two partial LOF variants were observed in the DNA-binding domain (p.R140H and p.E217K).

Each variant that displayed reduced activity in the assay was retested for the ability to stimulate adipocyte differentiation in the presence of varying doses of the PPARG agonist rosiglitazone. Consistent with previous reports (13), and the lack of clinical efficacy of thiazolidinediones in FPLD3, complete LOF variants are unresponsive even to 100-fold increased doses. In contrast, some of the variants observed as having reduced activity in the cellular assay (e.g., p.R140H, p.E217K, p.Y278C, and p.M376I) were rescued to WT levels using a higher dose (two- to fivefold) of rosiglitazone (Fig. 2*B*).

LOF Nonsynonymous *PPARG* Variants and T2D Risk in the Population. Based on the experimental classification of variants in the adipocyte differentiation assay, we repeated the analysis of association to T2D in individuals carrying benign and functional *PPARG* variants (Fig. 3). Of the 102 individuals harboring variants classified as benign, half occurred in cases and half in controls (52 T2D cases and 50 controls). In contrast, of the 16 individuals harboring variants that cause reduced function in the assay, 14 occurred in cases of T2D and only 2 in controls. The estimated risk of T2D was 1.17-fold (95% confidence interval 0.68–2.02) in carriers of a benign *PPARG* variant and 7.22-fold (95% confidence interval 1.79–29.02; $P = 0.005$) in carriers of a *PPARG* variant with reduced function in the assay. We examined the phenotypic characteristics of these 16 carriers (where phenotypic data were available; Table 2), but did not observe compelling evidence that these individuals were extreme outliers in the measured parameters.

Discussion

Based on a multiethnic sample of nearly 20,000 individuals, we estimate that (*i*) approximately 6 in 1,000 individuals carry an inherited rare coding variant in *PPARG*, (*ii*) 20% of these variants demonstrate reduced function in an adipocyte differentiation assay, and (*iii*) individuals who are heterozygous for the latter class of variants have an estimated sevenfold increased risk

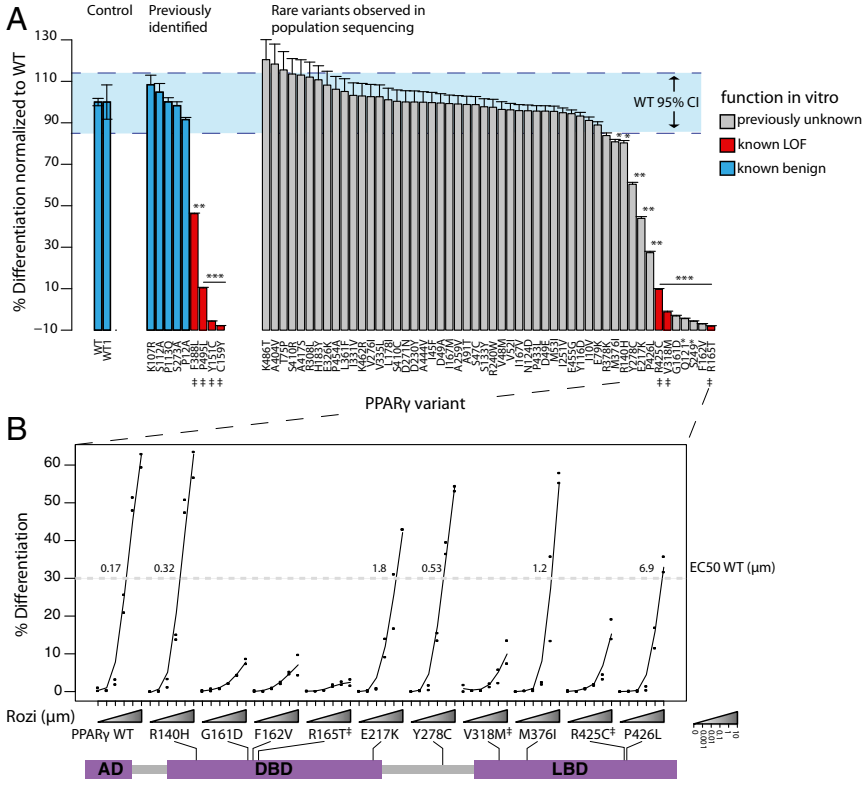


Fig. 2. Experimental characterization of rare PPARγ variants identified from population sequencing. (*A*) Each PPARγ variant was generated and tested for its ability to rescue adipocyte differentiation in vitro. From left to right PPARγ variants are sorted by in vitro function in three groups: (*i*) WT from independent experiments, (*ii*) previously identified synthetic and human mutations, and (*iii*) variants identified in population based exon resequencing. Blue dashed lines denote the 95% confidence interval of WT function. (*B*) Roziglitazone (rozi) dose–response of PPARγ variants identified as LOF. The amino acid position along the PPARγ protein is shown. EC_{50} WT denotes the rozi dose required to achieve 50% of maximal WT response. AD, activation domain; DBD, DNA-binding domain; LBD, ligand-binding domain. Error bars indicate ±1 SEM. Significant differences compared with WT are noted: $^{*}P < 0.05$; $^{**}P < 0.005$; $^{***}P < 0.0001$. ‡Variants identified in families with partial lipodystrophy.

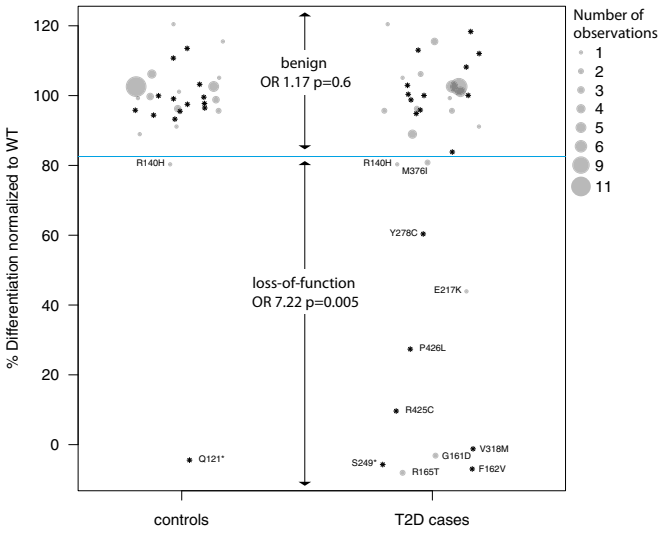


Fig. 3. T2D case/control status in multiethnic individuals harboring nonsynonymous *PPARG* variants, according to PPARG function in vitro. Each point represents an individual variant; point size denotes the number of individuals carrying that variant. Function in vitro was determined by the ability of each variant to rescue adipocyte differentiation in comparison with WT PPARG. The blue dashed line indicates the threshold for a one-tailed *t* test below which variants are classified as LOF compared with WT PPARG ($P < 0.05$). Odds ratios and *P* values for T2D case status among individuals carrying benign and LOF variants were calculated as described in *Methods*. *Variants observed only in a single case or control individual.

of T2D. Based on available clinical data, T2D patients who carry such mutations have increased risk of T2D but may lack the highly penetrant, extreme syndromic features observed in mutation carriers who were ascertained based on lipodystrophy that segregates in families.

Compared with the P12A variant, which has a smaller effect size but a 150-fold higher frequency, these rare variants contribute very modestly to the overall population burden of disease. However, given their larger individual effect sizes, such variants may prove useful for clinical risk prediction. An aggregated score of common genetic variants at 18 loci (including PPARG P12A) provided a 2.6-fold increased risk in individuals in high-score versus low-score groups (12); our data suggests that functional variants in *PPARG* may have effects larger than fivefold. However, only 0.1% of individuals with T2D are estimated to carry such rare variants in *PPARG*, and it is expected that the few individuals in the 0.1% tail of the distribution of risk based on common variants might similarly have larger magnitude of risk. A score that combines common and rare variants will be more predictive than an approach that considers only rare variants, or only common variants, alone.

The data presented here are consistent with the hypothesis that some patients with the common form of T2D have partial defects in adipocyte function attributable to mutations in PPARG. Some of the variants we observed in PPARG cause reduced function in the adipocyte differentiation assay that is as severe as the PPARG mutations associated with FPLD3. Other protein variants in PPARG cause a milder degree of dysfunction and can be rescued to WT levels by elevated doses of PPARG agonists (Fig. 2*B*). Based on the response to rosiglitazone in the adipocyte differentiation assay, we hypothesize that individuals with mild LOF variants in *PPARG* might respond positively to PPARG agonists, because their individual risk of disease was substantially increased by a genetic variant that could be rescued in vitro by PPARG agonists. Administration of rosiglitazone to individuals with severe LOF PPARG mutations who manifest lipodystrophy, insulin resistance, and T2D showed unclear therapeutic benefit for glycemia or insulin resistance, but this might be because mutations conferring complete LOF are not responsive to PPARG agonists (13).

This study has multiple limitations, including a cross-sectional case-control design and the extent of phenotypic characterization of mutation carriers. We are unable to detect any physiologic correlate in LOF *PPARG* variant carriers, which could indicate that the phenotype is not severe, or reflect the lack of more detailed characterization to date such as by dual-energy X-ray absorptiometry-based (DEXA) body composition. The individuals in this study were not ascertained based on extreme phenotypes such as lipodystrophy, nor demonstrate unusual features in the available

Table 2. Clinical and biochemical characteristics of individuals carrying LOF variants in *PPARG*

PPARG variant	Effect on PPARG function	T2D status	Ethnicity	Age	Sex	BMI	Waist-to-hip ratio	SystolicBP	DiastolicBP	Total cholesterol	LDL	HDL	Triglycerides
R165T	Severe	Case	European	40	F	23.6		125	82.5	184			201
R165T	Severe	Case	European	74	M	33.6		150	115	189	100	35	280
F162V	Severe	Case	European	65	M	25.3	0.92	160	85	268	188	53	135
S249*	Severe	Case	European	55	F	21.4		177.5	102.5	228	145	41	211
Q121*	Severe	Control	Caucasian American	36–62†	F	20.0†		96†	67†	184†	114†	65†	12†
G161D	Severe	Case	European	54	M	25.2	0.94	149	84	256	173	46	183
G161D	Severe	Case	European	82	M	23.7	0.98	150	90	203	125	30	236
V318M	Severe	Case	European	55	F	29.3	0.88						
R425C	Severe	Case	European	50	M	26.1		110	65	180		28	395
P426L	Mild	Case	European	49	F	24.5		134	77	217	137	36	223
E217K	Mild	Case	Hispanic	61	F	21.5	0.90			243	158	39	230
Y278C	Mild	Case	European	69	F	25.8	0.96	146	82	215	129	51	181
R140H	Mild	Case	Hispanic	55	F	31.0	0.96	128	81	202	139	37	126
R140H	Mild	Control	African American	67	F	33.2		130	77	249	186	41	119
M376I	Mild	Case	European	39	M	24.3	0.92	125	89	216	136	44	178
M376I	Mild	Case	European	44	M	26.5	1.03	135	86	193	104	57	164

Units of measurement are as follow: age is in years; systolic and diastolic blood pressure are in millimeters of mercury; total cholesterol, LDL, HDL, and triglycerides are in milligrams per deciliter. BMI, body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein.
†This individual had longitudinal measurements obtained over 30 y of follow up. The average values over this period are reported.

data (*SI Appendix*, Fig. 1), but we cannot rule out partial lipodystrophy, which can manifest subtly and easily escape clinical detection. Finally, this study assesses one cellular function of PPARG—adipocyte differentiation. It is possible that some missense variants may alter other cellular functions of PPARG and influence glycemic physiology.

The requirement for experimental characterization before association analysis is consistent with other studies in which functional characterization of rare mutations was needed to discover the relationship to disease (14, 15). This is in contrast to genome-wide association studies of common variants, where the combination of frequency and effect size is sufficient to discover associations without assumptions as to the in vitro assay that will predict clinical impact. Generalization of a genotype-function-phenotype approach to rare variants presents several challenges, in particular the definition of in vitro functional assays that are relevant to the clinical phenotype of interest. With genome sequencing becoming readily available, the key to clinically interpreting rare variants may turn out to be the laboratory assays and computational methods to discriminate benign from functional variants.

Methods

Sample Ascertainment. We studied 19,752 individuals (9,070 cases and 10,682 controls) from multiple ancestries as part of five candidate gene or whole-exome sequencing studies: the Genetics of Type 2 Diabetes (GoT2D) study, the Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples (T2D-GENES) study, the SIGMA (Slim Initiative in Genomic Medicine for the Americas) T2D Consortium, and the Framingham and Jackson Heart Study Allelic Spectrum project (FHS/JHS). For each study, individuals were drawn from previously described cohorts shown in *SI Appendix*, Table 1. Details on sample sequencing and *PPARG* variant identification are provided in *SI Appendix, Supplementary Methods, Sequencing, Variant Calling, Data QC, and Variant Annotation*. These sequencing studies were approved by the Massachusetts Institute of Technology committee on the use of humans as experimental subjects. Informed consent was obtained from the subjects.

Bioinformatic Assessment of Nonsynonymous *PPARG* Variants. Variants were bioinformatically classified as pathogenic if they met the following three criteria: (*i*) occurred at an evolutionarily conserved site [logarithm of the odds (LOD) >10 based on an alignment of 29 mammalian genomes] (16), (*ii*) computationally predicted as protein damaging by the consensus mutation analysis tool Consensus Deleteriousness Score (Condel) (17), and (*iii*) private to one study individual and not observed in the 1000 Genomes project (18). If they did not meet all of these criteria, they were classified as computationally benign. A second, less stringent bioinformatics classification scheme, where rare variants (i.e., minor allele frequency <0.1%) were classified as pathogenic if they fulfilled criteria *i* and *ii* here above, was also tested.

Rescue of Adipocyte Differentiation by in Vitro *PPARG* Variant Constructs. Each *PPARG* variant was recreated in vitro by PCR mutagenesis and packaged into lentiviruses. These lentiviruses were used to transduce SGBS pre-adipocytes exposed to a submaximal stimulation for adipocyte differentiation. In this assay, preadipocytes differentiate only when provided with functional, exogenous *PPARG* (Fig. 1*C*). Details are provided in *SI Appendix, Supplementary Methods, Rescue of Adipocyte Differentiation by in Vitro* PPARG *Variant Constructs*.

High-Throughput Adipocyte Differentiation Assay. To measure adipocyte differentiation at the end of 8 d of exposure to differentiation mixture and *PPARG* variants, cells were fixed in 4% (wt/vol) PFA, washed in PBS, and stained with boron-dipyrromethene (BODIPY; Sigma) (1 μg/mL) to stain lipids and DAPI (1 μg/mL) to stain nuclei. Stained cells were imaged with a high-content fluorescence microscope (Molecular Devices IXM) at 4× at 512 and 484 nm, corresponding respectively to the peak emission spectra of BODIPY and DAPI. The obtained images were analyzed using a custom analysis pipeline developed in CellProfiler (19) to identify total numbers of adipocytes and undifferentiated cells. The ratio of adipocytes to total cells is the percentage of differentiation (Fig. 1*A*).

Statistical Analysis. In the experimental classification of *PPARG* variants, differentiation scores for variants were compared with differentiation scores for unmutated *PPARG*. Variants were classified experimentally as LOF if they demonstrated decreased ability to stimulate adipocyte differentiation compared with a series of WT controls as assessed by a one-tailed Student *t* test with equal variances and a *P* value threshold of 0.05. Association tests were performed to assess the diabetes risk of variant carriers relative to noncarriers. An identical aggregate gene-based analysis was repeated for each variant annotation: experimental LOF, experimental benign, bioinformatically deleterious, and bioinformatically benign. Details are provided in *SI Appendix, Supplementary Methods, Association Tests*.

ACKNOWLEDGMENTS. We thank John Doench, Joseph Avruch, Suzanne Jacobs, Noel Burtt, and Victor Rusu for helpful discussions, laboratory assistance, and manuscript review. We also thank Anne Carpenter, who is supported by National Institutes of Health (NIH) Grant GM089652, for assistance with image analysis. Funding for the GoT2D and T2D-GENES studies was provided by Grants 5U01DK085526 (Multiethnic Study of Type 2 Diabetes Genes), DK088389 (Low-Pass Sequencing and High-Density SNP Genotyping for Type 2 Diabetes), and U54HG003067 (Large-Scale Sequencing and Analysis of Genomes), as well as NIH U01 Grants DK085501, DK085524, DK085545, and DK085584. The Jackson Heart Study (JHS) is supported by Contracts HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, and HHSN268201300050C from the National Heart, Lung, and Blood Institute and National Institute on Minority Health and Health Disparities. The NHGRI JHS/FHS Allelic Spectrum Project was supported by grants from the National Human Genome Research Institute of the US NIH [Medical Sequencing Program Grant U54 HG003067 (to E. Lander, Broad Institute Principal Investigator)]. The SIGMA T2D Consortium, a joint US–Mexico project, was funded by the Carlos Slim Health Institute.

1. Deeb SS, et al. (1998) A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. *Nat Genet* 20(3):284–287.
2. Stumvoll M, Häring H (2002) The peroxisome proliferator-activated receptor-γ2 Pro12Ala polymorphism. *Diabetes* 51(8):2341–7.
3. Altshuler D, et al. (2000) The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26(1):76–80.
4. Barroso I, et al. (1999) Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* 402(6764):880–883.
5. Agostini M, et al. (2006) Non-DNA binding, dominant-negative, human PPARgamma mutations cause lipodystrophic insulin resistance. *Cell Metab* 4(4):303–311.
6. Tennessen JA, et al.; Broad GO; Seattle GO; NHLBI Exome Sequencing Project (2012) Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 337(6090):64–69.
7. Flannick J, et al. (2013) Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes. *Nat Genet* 45(11):1380–1385.
8. Sunyaev SR (2012) Inferring causality and functional significance of human coding DNA variants. *Hum Mol Genet* 21(R1):R10–R17.
9. Auclair M, et al. (2013) Peroxisome proliferator-activated receptor-γ mutations responsible for lipodystrophy with severe hypertension activate the cellular renin-angiotensin system. *Arterioscler Thromb Vasc Biol* 33(4):829–838.
10. Jeninga E, Gurnell M (2009) Functional implications of genetic variation in human PPAR [gamma]. *Trends Endocrinol Metab* 20(8):380–387.
11. Visser ME, et al. (2011) Characterisation of non-obese diabetic patients with marked insulin resistance identifies a novel familial partial lipodystrophy-associated PPARγ mutation (Y151C). *Diabetologia* 54(7):1639–1644.
12. Meigs JB, et al. (2008) Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 359(21):2208–2219.
13. Savage DB, et al. (2003) Human metabolic syndrome resulting from dominant-negative mutations in the nuclear receptor peroxisome proliferator-activated receptor-gamma. *Diabetes* 52(4):910–917.
14. Ji W, et al. (2008) Rare independent mutations in renal salt handling genes contribute to blood pressure variation. *Nat Genet* 40(5):592–599.
15. Bonnefond A, et al.; Meta-Analysis of Glucose and Insulin-Related Traits Consortium (MAGIC) (2012) Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nat Genet* 44(3):297–301.
16. Lindblad-Toh K, et al.; Broad Institute Sequencing Platform and Whole Genome Assembly Team; Baylor College of Medicine Human Genome Sequencing Center Sequencing Team; Genome Institute at Washington University (2011) A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 478(7370):476–482.
17. González-Pérez A, López-Bigas N (2011) Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. *Am J Hum Genet* 88(4):440–449.
18. Abecasis GR, et al.; 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. *Nature* 467(7319):1061–1073.
19. Carpenter AE, et al. (2006) CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 7(10):R100.

Supplemental Appendix

Author list:

The GoT2D Consortium: Jason Flannick[1,9], Alisa K. Manning[1], Christopher Hartl[1], Vineeta Agarwala[1], Pierre Fontanillas[1], Todd Green[1], Eric Banks[1], Mark DePristo[1], Ryan Poplin[1], Khalid Shakir[1], Timothy Fennell[1], Pål R. Njølstad[6], David Altshuler[1,2,9,10,12,13,14], Noël Burtt[1], Stacey Gabriel[1], Christian Fuchsberger[25], Hun Min Kang[25], Xueling Sim[25], Clement Ma[25], Adam Locke[25], Thomas Blackwell[25], Anne Jackson[83], Tanya M. Teslovich[25], Heather Stringham[25], Peter Chines[25], Phoenix Kwan[25], Jeroen Huyghe[25], Adrian Tan, Goo Jun[25], Michael Stitzel[84], Richard N. Bergman[11], Lori Bonnycastle[54], Jaakko Tuomilehto[106], Francis S. Collins[54], Laura Scott[25], Karen Mohlke[71], Gonçalo Abecasis[25], Michael Boehnke[25], Tim Strom[85], Christian Gieger[86], Martina Müller-Nurasyid[86], Harald Grallert[86], Jennifer Kriebel[86], Janina Ried[86], Martin Hrabé de Angelis[86], Cornelia Huth[86], Christa Meisinger[86], Annette Peters[86], Wolfgang Rathmann[86], Konstantin Strauch[86], Thomas Meitinger[85], Jasmina Kravic[99], Peter Algren[99], Claes Ladenvall[99], Tiinamaija Toumi[99], Bo Isomaa[99], Leif Groop[99], Kyle Gaulton[24], Loukas Moutsianas[24], Manny Rivas[24], Richard Pearson[24], Anubha Mahajan[24], Inga Prokopenko[24,76], Ashish Kumar[24], John Perry[24], Bryan Howie[50], Martijn van de Bunt[51], Kerrin Small[87], Cecilia Lindgren[24], Gerton Lunter[24], Neil Robertson[24], William Rayner[24], Andrew Morris[72], David Buck[24], Andrew Hattersley[59], Tim Spector[87], Gil McVean[24], Tim Frayling[59], Peter Donnelly[24], Mark McCarthy[24,51,69]

National Human Genome Research Institute Jackson Heart Study Framingham Heart Study (NHGRI JHS/FHS) Allelic Spectrum Project: Namrata Gupta[1], Herman Taylor[96,97,98], Ervin Fox[96], Christopher Newton-Cheh[1,93,103,104], James G Wilson[44], Christopher J O'Donnell[103,104,111], Sekar Kathiresan[1,93,103,104], Joel Hirschhorn[1,91,112], J.G. Seidman[91], Stacey Gabriel[1], Christine Seidman[91,114,115] David Altshuler[1,2,9,10,12,13,14]

The SIGMA Type 2 Diabetes Genetics Consortium: Amy L. Williams[1,2], Suzanne B. R. Jacobs[1], Hortensia Moreno-Macías[3], Alicia Huerta-Chagoya[4,5], Claire Churchhouse[1], Carla Márquez-Luna[6], Humberto García-Ortíz[6], María José Gómez-Vázquez[4,7], Noël P. Burtt[1], Stephan Ripke[1,15], Alisa K. Manning[1], Benjamin Neale[1,15], Daniel O. Stram[11], Juan Carlos Fernández-López[6], Sandra Romero-Hidalgo[6], Irma Aguilar-Delfín[6], Angélica Martínez-Hernández[6], Federico Centeno-Cruz[6], Elvia Mendoza-Caamal[6], Cristina Revilla-Monsalve[16], Sergio Islas-Andrade[16], Emilio Córdova[6], Eunice Rodríguez-Arellano[17], Xavier Soberón[6], María Elena González-Villalpando[8], Kristine Monroe[11], Lynne Wilkens[18], Laurence N. Kolonel[18], Loic Le Marchand[18], Laura Riba[5], María Luisa Ordóñez-Sánchez[4], Rosario Rodríguez-Guillén[4], Ivette Cruz-Bautista[4], Maribel Rodríguez-Torres[4], Linda Liliana Muñoz-Hernández[4], Tamara Sáenz[4], Donají Gómez[4], Ulices Alvirde[4], Robert C. Onofrio[19], Wendy M. Brodeur[19], Diane Gage[19], Jacquelyn Murphy[1], Jennifer Franklin[19], Scott Mahan[19], Kristin Ardlie[19], Andrew T. Crenshaw[19], Wendy Winckler[19], David Altshuler[1,2,9,10,12,13,14], Jose C. Florez[1,9,10], Christopher A. Haiman[11], Brian E. Henderson[11], Carlos A. Aguilar-Salinas[4], Clicerio González-Villalpando[8], Lorena Orozco[6] Teresa Tusié-Luna[4,5]

The T2D-GENES Consortium: Gonçalo Abecasis[25], Marcio Almeida[27], David Altshuler[1,2,9,10,12,13,14], Jennifer L. Asimit[49], Gil Atzmon[26], Mathew Barber[50], Nicola L. Beer[51], Graeme I. Bell[50,52], Jennifer Below[35], Tom Blackwell[25], John Blangero[27], Michael Boehnke[25], Donald W. Bowden[28], Noël

Burtt[1], John Chambers[29,30,31], Han Chen[53], Peng Chen[22], Peter S.Chines[54], Sungkyoung Choi[55], Claire Churchhouse[1], Pablo Cingolani[56], Belinda K. Cornes[57], Nancy Cox[50,52], Aaron G. Day-Williams[49], Ravindranath Duggirala[27], Josée Dupuis[53], Thomas Dyer[27], Shuang Feng[25], Juan Fernandez-Tajes[24], Teresa Ferreira[24], Tasha E. Fingerlin[58], Jason Flannick[1,9], Jose Florez[1,9,10], Pierre Fontanillas[1], Timothy M. Frayling[59], Christian Fuchsberger[25], Eric R. Gamazon[52], Kyle Gaulton[24], Saurabh Ghosh, Anna Gloyn[51], Robert L. Grossman[52,60], Jason Grundstad[60], Craig Hanis[35], Allison Heath[60], Heather Highland[35], Momoko Hirokoshi[24], Ik-Soo Huh[55], Jeroen R. Huyghe[25], Kamran Ikram[21,57,61,62], Kathleen A. Jablonski[63], Young Jin Kim[64], Goo Jun[25], Norihiro Kato[65], Jayoun Kim[55], C. Ryan King[66], Jaspal Kooner[30,31,36], Min-Seok Kwon[55], Hae Kyung Im[66], Markku Laakso[37], Kevin Koi-Yau Lam[22], Jaehoon Lee[55], Selyeong Lee[55], Sungyoung Lee[60], Donna M. Lehman[47], Heng Li[1], Cecilia M. Lindgren[24], Xuanyao Liu[22,42], Oren E. Livne[50], Adam E. Locke[25], Anubha Mahajan[24], Julian B. Maller[24,67], Alisa K. Manning[1]. Taylor J. Maxwell[35], Alexander Mazoure[68], Mark I. McCarthy[24,51,69], James B. Meigs[10,70], Byungju Min[55], Karen L. Mohlke[71], Andrew Morris[72], Solomon Musani[73], Yoshihiko Nagai[68], Maggie C.Y. Ng[28], Dan Nicolae[50,52,54], Sohee Oh[55], Nicholette Palmer[28], Taesung Park[55], Toni I. Pollin[75], Inga Prokopenko[24,76], David Reich[1,2], Manuel A. Rivas[24], Laura J. Scott[25], Mark Seielstad[20], Yoon Shin Cho[32], E-Shyong Tai[21,22,23], Xueling Sim[25], Robert Sladek[68,77], Philip Smith[78], Ioanna Tachmazidou[49], Tanya M. Teslovich[25], Jason Torres[50,52], Vasily Trubetskoy[50,52], Sara M. Willems[79], Amy L. Williams[1,2], James G. Wilson[44], Steven Wiltshire[80], Sungho Won[81], Andrew R. Wood[59], Wang Xu[23], Yik Ying Teo[39,40,41,42,43], Joon Yoon[55], Jong-Young Lee[38], Matthew Zawistowski[25], Eleftheria Zeggini[49], Weihua Zhang[29], Sebastian Zöllner[25,82]

Author Affiliations:

1. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA.
2. Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.
3. Universidad Autonoma Metropolitana, 14387 Tlalpan, Mexico City, Mexico.
4. Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Sección XVI, Tlalpan, 14000 Mexico City, Mexico.
5. Instituto de Investigaciones Biomédicas, UNAM. Unidad de Biología Molecular y Medicina Genómica, UNAM/INCMNSZ, Coyoacán, 04510 Mexico City, Mexico.
6. Instituto Nacional de Medicina Genómica, Tlalpan, 14610 Mexico City, Mexico.
7. Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Nuevo León 66451, México.
8. Centro de Estudios en Diabetes, Unidad de Investigacion en Diabetes y Riesgo Cardiovascular, Centro de Investigacion en Salud Poblacional, Instituto Nacional de Salud Publica, 01120 Mexico City, Mexico.
9. Center for Human Genetic Research and Diabetes Research Center (Diabetes Unit), Massachusetts General Hospital, Boston 02114, Massachusetts, USA.
10. Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA.
11. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California 90089, USA.

12. Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.
13. Department of Molecular Biology, Harvard Medical School, Boston, Massachusetts 02114, USA.
14. Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
15. Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.
16. Instituto Mexicano del Seguro Social SXXI, Cuauhtémoc, 06720 Mexico City, Mexico.
17. Instituto de Seguridad y Servicios Sociales para los Trabajadores del Estado, Cuauhtémoc, 06030 Mexico City, Mexico.
18. Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii 96813, USA.
19. The Genomics Platform, The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA.
20. University of California San Francisco, San Francisco, California 94143, USA.
21. Duke National University of Singapore Graduate Medical School, Singapore 169857, Singapore.
22. Saw Swee Hock School of Public Health, National University of Singapore, Singapore 117597, Singapore.
23. Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore.
24. Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK.
25. Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA.
26. Department of Medicine, Department of Genetics, Albert Einstein College of Medicine, Bronx, New York 10461, USA.
27. Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas 78227, USA.
28. Center for Genomics and Personalized Medicine Research, Center for Diabetes Research, Department of Biochemistry, Department of Internal Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina 27157, USA.
29. Department of Epidemiology and Biostatistics, Imperial College London, London SW7 2AZ, UK.
30. Imperial College Healthcare NHS Trust, London W2 1NY, UK.
31. Ealing Hospital National Health Service (NHS) Trust, Middlesex UB1 3HW, UK.
32. Department of Biomedical Science, Hallym University, Chuncheon, Gangwon-do, 200-702 South Korea.

33. [38]Endocrinology and Metabolism Service, Hadassah-Hebrew University Medical School, Jerusalem, 91120 , Israel.

34. [39]Israel Diabetes Research Group (IDRG), Diabetes Unit, The E. Wolfson Medical Center, Holon, 58100 Israel.

35. Human Genetics Center, University of Texas Health Science Center at Houston, Houston, Texas 77030, USA.

36. [41]National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Hospital, London W12 0HS, UK.

37. [42]Department of Medicine, University of Eastern Finland, Kuopio Campus and Kuopio University Hospital, FI-70211 Kuopio, Finland.

38. Center for Genome Science, Korea National Institute of Health, Osong Health Technology Administration Complex, Chungcheongbuk-do, 363-951, South Korea.

39. Department of Epidemiology and Public Health, National University of Singapore, Singapore 117597, Singapore.

40. Centre for Molecular Epidemiology, National University of Singapore, Singapore 117456, Singapore.

41. Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore 138672, Singapore.

42. Graduate School for Integrative Science and Engineering, National University of Singapore, Singapore 117456, Singapore.

43. Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546, Singapore.

44. Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, Mississippi 39216, USA.

45. [50]Division of Nephrology, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, Texas 78229, USA.

46. Division of Diabetes, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, Texas 78229, USA.

47. Division of Clinical Epidemiology, Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, Texas 78229, USA.

48. Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA.

49. Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1HH, UK.

50. Department of Human Genetics, University of Chicago, Chicago, Illinois 60637, USA.

51. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, OX3 7LJ, UK.

52. Department of Medicine, University of Chicago, Chicago, Illinois 60637, USA.

53. Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts 02115, USA.

54. National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA.

55. Seoul National University, Seoul 110-799, South Korea.

56. McGill Centre for Bioinformatics, McGill University, Montréal, Quebec, H3G 0B1, Canada.

57. Singapore Eye Research Institute, Singapore National Eye Centre, Singapore 168751, Singapore.

58. Department of Epidemiology, Colorado School of Public Health, Aurora, Colorado 80045, USA.

59. Genetics of Complex Traits, University of Exeter Medical School, Exeter, EX4 4SB UK.

60. Institute for Genomics and Systems Biology, University of Chicago, Chicago, Illinois 60637, USA.

61. Department of Ophthalmology, National University of Singapore and National University Health System, Singapore 119228, Singapore.

62. Department of Ophthalmology, Erasmus Medical Center, Rotterdam 3000 CA, the Netherlands.

63. The Biostatistics Center, George Washington University, Rockville, Maryland 20852, USA.

64. Department of Neurology, Konkuk University School of Medicine, Seoul 143-701, South Korea.

65. Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo 162-8655, Japan.

66. Department of Health Studies, University of Chicago, Chicago, Illinois 60637, USA.

67. Department of Statistics, University of Oxford, Oxford, OX1 3TG UK.

68. McGill University, Montréal, Québec H3A 0G4, Canada.

69. Oxford NIHR Biomedical Research Centre, Churchill Hospital, Headington OX3 7LE, UK.

70. [75]General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.

71. Department of Genetics, University of North Carolina-Chapel Hill, Chapel Hill, North Carolina 27599, USA.

72. Department of Genetic Medicine, Manchester Academic Health Sciences Centre, Manchester M13 9NT, UK.

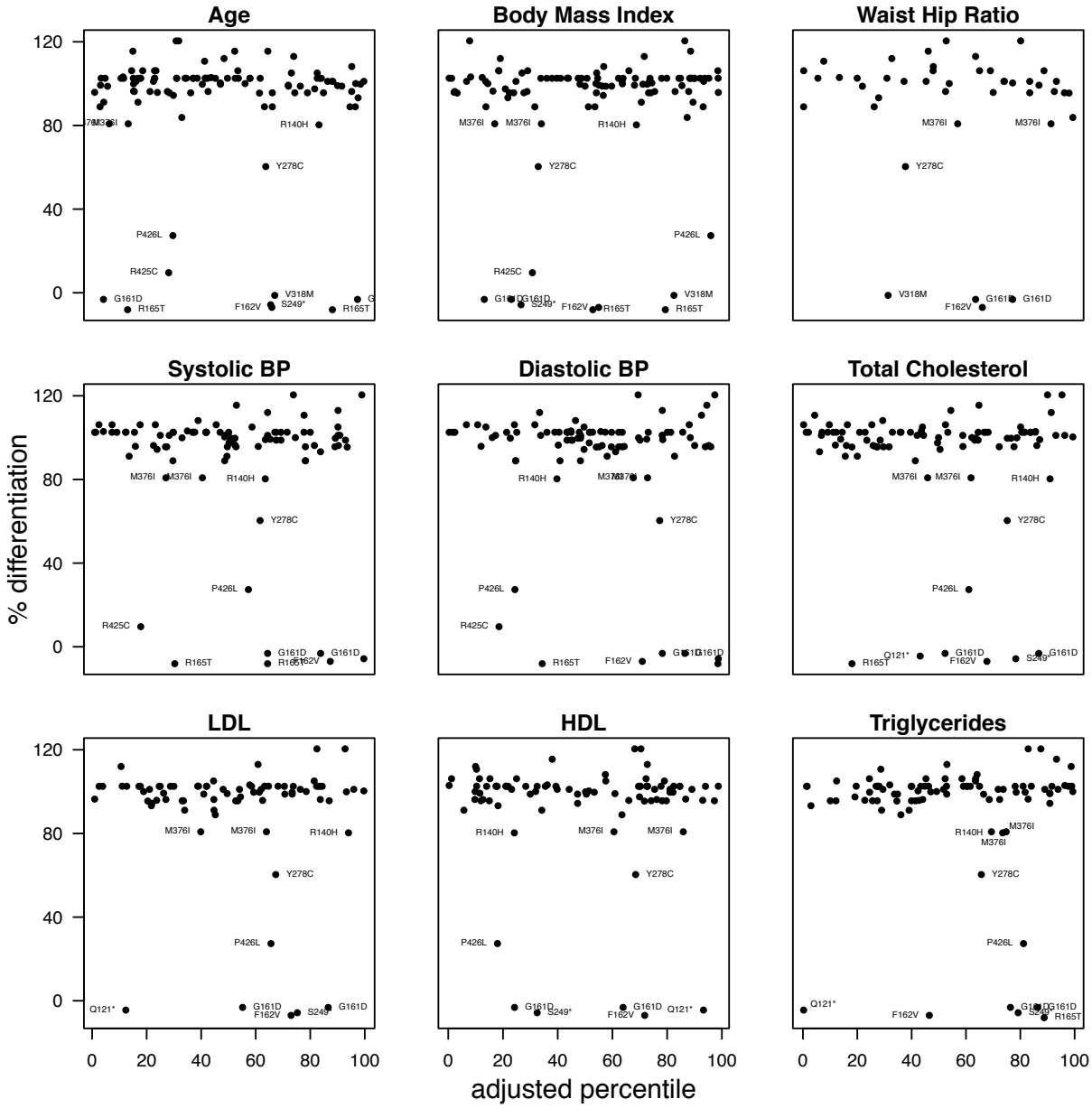
73. Department of Medicine, University of Mississippi Medical Center, Jackson, Mississippi 39126, USA.

74. [79]Department of Statistics, University of Chicago, Chicago, Illinois 60637, USA.

75. Department of Medicine, Program for Personalized and Genomic Medicine, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA.

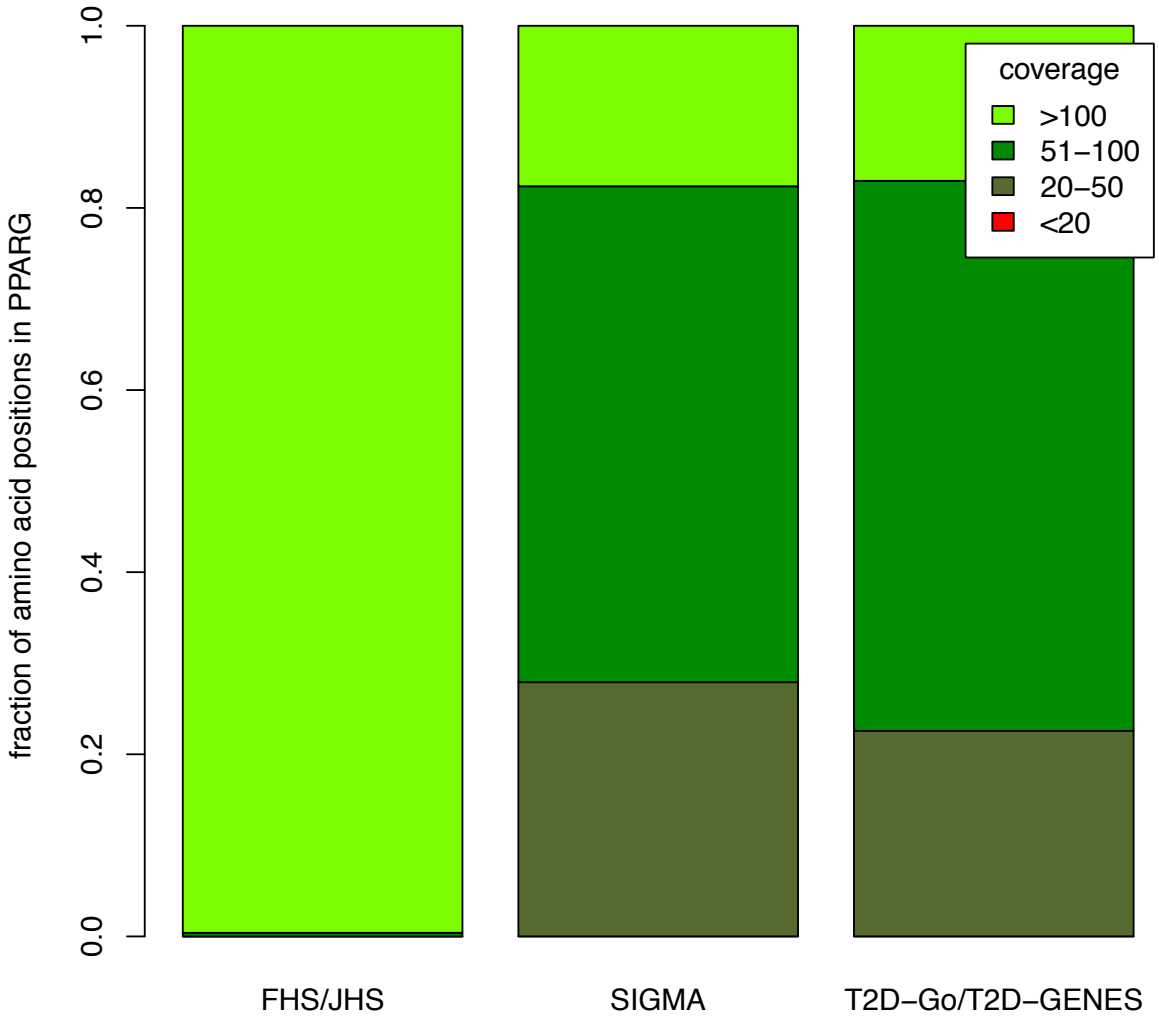
76. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, 751 05 Uppsala, Sweden.
77. Department of Medicine, Royal Victoria Hospital, Montréal, Québec H3A 1A1, Canada.
78. National Institute of Diabetes and Digestive and Kidney Disease, National Institutes of Health, Bethesda, MD 20817, USA.
79. Department of Genetic Epidemiology, Erasmus Medical Center, Rotterdam 3000 CA, the Netherlands.
80. Centre for Medical Research, Western Australian Institute for Medical Research, The University of Western Australia, Nedlands WA 6008, Australia.
81. Chung-Ang University, Seoul 156-756, South Korea.
82. Department of Psychiatry, University of Michigan, Ann Arbor, Michigan 48109, USA.
83. Department of Pathology, University of Michigan, Ann Arbor, Michigan 48109, USA.
84. Department of Biochemistry & Molecular Biology, Pennsylvania State University, Pennsylvania, PA
85. Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany
86. Institute of Genetic Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany
87. Kings College of London, Department of Twin Research and Genetic Epidemiology, London England
88. Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA
89. Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA
90. Diabetes Unit, Massachusetts General Hospital, Boston, MA, USA
91. Department of Genetics, Harvard Medical School, Boston, MA, USA
92. KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway
93. Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA
94. Department of Medicine, Harvard Medical School, Boston, MA, USA
95. General Medicine Division, Massachusetts General Hospital, Boston, MA, USA
96. Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA
97. Jackson State University, Jackson, MS, USA
98. Tougaloo College, Tougaloo MS, USA
99. Department of Clinical Sciences, Diabetes and Endocrinology, Clinical Research Centre, Lund University, Malmö, Sweden

100. Cardiovascular and Metabolic Diseases Practice, Prescient Life Sciences, London, UK
101. Center for Medical Genetics and Molecular Biology, Haukeland University Hospital, Bergen, Norway
102. Cardiovascular and Metabolic Diseases Research Unit, Pfizer Inc., Cambridge, MA, USA
103. National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham,
MA, USA
104. Cardiology Division, Massachusetts General Hospital, Boston, MA, USA
105. Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland
106. Department of Public Health, University of Helsinki, Helsinki, Finland
107. Department of Medicine, Helsinki University Central Hospital and Research Program for Molecular Medicine
108. The Gade Institute, University of Bergen, Bergen, Norway
109. Department of Pathology, Haukeland University Hospital, Bergen, Norway
110. Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA
111. Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD, USA
112. Divisions of Genetics and Endocrinology and Program in Genomics, Children's Hospital, Boston, MA, USA
113. Applied Quantitative Genotherapeutics, Pfizer Inc., South San Francisco, CA, USA
114. Division of Cardiovascular Medicine, Brigham and Women’s Hospital, Boston, MA, USA
115. Howard Hughes Medical Institute, Chevy Chase, MD, USA
116. Finnish Institute for Molecular Medicine (FIMM), Helsinki University, Helsinki , Finland



Supplementary Figure 1. **Anthropometric and metabolic traits of individuals harboring rare PPARγ variants identified from population sequencing.**

For each trait, the value for a given individual, if known, is represented on the abscissa as a percentile calculated within that individual's ascertainment cohort and with respect to T2D case/control status. On the ordinate, the ability of that individual's PPARγ variant to rescue adipocyte differentiation in vitro is shown as % differentiation. PPARγ variants shown to cause loss-of-function (see Figure 2) are explicitly labeled.



Supplementary Figure 2: **Sequencing read coverage of PPARG2 from five gene sequencing studies.** Coverage is indicated by amino acid for the PPARG2 protein (NP_005028). Samples from GoT2D and T2D-GENES are shown together as they were jointly analyzed. Samples from FHS and JHS are shown together as they were jointly analyzed.

Supplementary Table 1: T2D case/control samples sequenced for *PPARG*						
Consortium name	Ethnicity	Geography	Study	Reference	Cases	Controls
GoT2D	European	Sweden	Malmo Preventive Project	[1-6]	670	236
GoT2D	European	Finland	The Botnia Study	[7-9]	500	268
GoT2D	European	Finland	FUSION	[10]	470	474
GoT2D	European	UK	UKT2D	[11-13]	329	332
GoT2D	European	Germany	KORA	[14]	97	91
T2D-GENES	European	Finland	METSIM	[15]	487	501
T2D-GENES	European	USA	Ashkenazim Study	[16]	506	385
T2D-GENES	South Asian	UK	LOLIPOP	[17, 18]	531	539
T2D-GENES	South Asian	Singapore	Singapore Indian Chinese Cohort Study	[19]	564	585
T2D-GENES	East Asian	Korea	KARE	[20]	522	554
T2D-GENES	East Asian	Singapore	The Singapore National Health Survey	[21-24]	476	592
T2D-GENES	Hispanic	USA	The San Antonio Family Diabetes/Gallbladder Study	[25-28]	245	182
T2D-GENES	Hispanic	USA	Starr County Health Study	[29]	754	705
T2D-GENES	African-American	USA	Wake Forest Study	[30]	540	571
FHS/JHS/T2D-GENES	African-American	USA	Jackson Heart Study	[31]	341	1357
FHS/JHS	European	USA	Framingham Heart Study	[32]	225	1338

SIGMA-T2D	Hispanic	USA	Multiethnic Cohort Study	[33]	483	441
SIGMA-T2D	Hispanic	Mexico	UNAM/INCMNSZ Diabetes Study	[33]	551	546
SIGMA-T2D	Hispanic	Mexico	Diabetes in Mexico Study	[33]	509	459
SIGMA-T2D	Hispanic	Mexico	Mexico City Diabetes Study	[33]	270	526

References corresponding to Supplementary Table 1:

1. Lindholm, E., et al., *Classifying diabetes according to the new WHO clinical stages.* European journal of epidemiology, 2001. **17**(11): p. 983-989.
2. Parker, A., et al., *A gene conferring susceptibility to type 2 diabetes in conjunction with obesity is located on chromosome 18p11.* Diabetes, 2001. **50**(3): p. 675-680.
3. Berglund, G., et al., *The Malmo Diet and Cancer Study. Design and feasibility.* Journal of internal medicine, 1993. **233**(1): p. 45-51.
4. Berglund, G., et al., *Long-term outcome of the Malmö preventive project: mortality and cardiovascular morbidity.* Journal of internal medicine, 2000. **247**(1): p. 19-29.
5. Lyssenko, V., et al., *Clinical risk factors, DNA variants, and the development of type 2 diabetes.* New England Journal of Medicine, 2008. **359**(21): p. 2220-2232.
6. Bøg-Hansen, E., et al., *Risk factor clustering in patients with hypertension and non-insulin-dependent diabetes mellitus. The Skaraborg Hypertension Project.* Journal of internal medicine, 1998. **243**(3): p. 223-232.
7. Groop, L., et al., *Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects.* Diabetes, 1996. **45**(11): p. 1585-1593.
8. Forsen, T.e.a., *The diabetes registry in vaasa hospital district - direva.* The Journal of General Practitioner, 2011. **6**: p. 17-21.
9. Isomaa, B., et al., *A family history of diabetes is associated with reduced physical fitness in the Prevalence, Prediction and Prevention of Diabetes (PPP)-Botnia study.* Diabetologia, 2010. **53**(8): p. 1709-1713.
10. Scott, L.J., et al., *A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants.* Science, 2007. **316**(5829): p. 1341-1345.
11. Burton, P.R., et al., *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.* Nature, 2007.
12. Spector, T.D. and F. Williams, *The UK adult twin registry (TwinsUK).* Twin Research and Human …, 2006.

13. Voight, B., et al., *Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis.* Nature genetics, 2010.
14. Wichmann, H., C. Gieger, and T. Illig, *KORA-gen-resource for population genetics, controls and a broad spectrum of disease phenotypes.* Gesundheitswesen, 2005.
15. Stancakova, A., et al., *Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men.* Diabetes, 2009. **58**(5): p. 1212-21.
16. Atzmon, G., et al., *Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry.* … of Human Genetics, 2010.
17. Chahal, N.S., et al., *Ethnicity-related differences in left ventricular function, structure and geometry: a population study of UK Indian Asian and European white subjects.* Heart (British Cardiac Society), 2010. **96**(6): p. 466-471.
18. Chahal, N.S., et al., *Does subclinical atherosclerosis burden identify the increased risk of cardiovascular disease mortality among United Kingdom Indian Asians? A population study.* American heart journal, 2011. **162**(3): p. 460-466.
19. Lavanya, R., et al., *Methodology of the Singapore Indian Chinese Cohort (SICC) eye study: quantifying ethnic variations in the epidemiology of eye diseases in Asians.* Ophthalmic epidemiology, 2009. **16**(6): p. 325-336.
20. Cho, Y.S., et al., *A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits.* Nature genetics, 2009. **41**(5): p. 527-534.
21. Cutter, J., B.Y. Tan, and S.K. Chew, *Levels of cardiovascular disease risk factors in Singapore following a national intervention programme.* Bulletin of the World Health Organization, 2001. **79**(10): p. 908-915.
22. Hughes, K., et al., *Central obesity, insulin resistance, syndrome X, lipoprotein (a), and cardiovascular risk in Indians, Malays, and Chinese in Singapore.* Journal of epidemiology and …, 1997.
23. Hughes, K., et al., *Cardiovascular diseases in Chinese, Malays, and Indians in Singapore. II. Differences in risk factor levels.* … of Epidemiology and …, 1990.
24. Tan, C.E., S.C. Emmanuel, and B.Y. Tan, *Prevalence of diabetes and ethnic differences in cardiovascular risk factors. The 1992 Singapore National Health Survey.* Diabetes, 1999.
25. Coletta, D.K., et al., *Genome-wide linkage scan for genes influencing plasma triglyceride levels in the Veterans Administration Genetic Epidemiology Study.* Diabetes, 2009. **58**(1): p. 279-284.
26. Hunt, K.J., et al., *Genome-Wide Linkage Analyses of Type 2 Diabetes in Mexican Americans The San Antonio Family Diabetes/Gallbladder Study.* Diabetes, 2005.
27. Knowler, W.C., et al., *The Family Investigation of Nephropathy and Diabetes (FIND): design and methods.* Journal of diabetes and its complications, 2005. **19**(1): p. 1-9.
28. Mitchell, B.D., et al., *Genetic and environmental contributions to cardiovascular risk factors in Mexican Americans. The San Antonio Family Heart Study.* Circulation, 1996. **94**(9): p. 2159-2170.
29. HANIS, C.L., R.E. FERRELL, and S.A. BARTON, *Diabetes among Mexican Americans in Starr County, Texas.* American Journal of …, 1983.
30. Yu, H., et al., *Linkage analysis between loci in the renin-angiotensin axis and end-stage renal disease in African Americans.* Journal of the American Society of Nephrology : JASN, 1996. **7**(12): p. 2559-2564.

31. Taylor, H.A., et al., *Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study.* Ethnicity & disease, 2005. **15**(4 Suppl 6): p. S6-4-17.
32. Kannel, W.B., et al., *An investigation of coronary heart disease in families. The Framingham offspring study.* Am J Epidemiol, 1979. **110**(3): p. 281-90.
33. Consortium, T.S.T.D., *Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico.* Nature, 2013.

Supplementary table 2: **Allele frequency (%) of PPARG variants by ethnicity**

PPARG variant	European	African.American	Hispanic	South.Asian	East.Asian
p.A259V	0.014	0	0	0	0
p.A404V	0	0	0.018	0	0
p.A417S	0.014	0	0	0	0
p.A444V	0.014	0	0	0	0
p.A91T	0	0.11	0	0	0
p.D230Y	0	0	0	0.045	0
p.D271N	0.014	0	0	0	0
p.D49A	0	0	0.018	0	0
p.D49E	0	0.14	0	0	0
p.E217K	0	0	0.035	0	0
p.E326K	0	0	0.018	0	0
p.E455G	0.014	0	0	0	0
p.E79K	0.058	0	0	0	0.047
p.F162V	0	0	0	0	0
p.G161D	0.029	0	0	0	0
p.H183Y	0	0	0.018	0	0
p.I10V	0	0	0	0.09	0
p.I167M	0.029	0	0	0	0
p.I167V	0.014	0	0	0	0
p.I251V	0	0	0.018	0	0
p.I331V	0	0	0	0.045	0
p.I45F	0	0.11	0	0	0
p.K462R	0	0	0.018	0	0
p.K486T	0	0	0	0.09	0
p.L178I	0.087	0	0	0	0
p.L361F	0.029	0	0	0	0
p.M376I	0.029	0	0	0	0
p.M53I	0	0	0	0.045	0
p.N124D	0	0	0	0.045	0
p.P426L	0.014	0	0	0	0
p.P433L	0.014	0	0.018	0	0
p.P454A	0	0	0.11	0	0
p.Q121*	0.014	0	0	0	0
p.R140H	0	0.036	0.071	0	0

p.R165T	0.029	0	0	0	0
p.R240W	0	0	0	0.045	0
p.R308L	0.014	0	0	0	0
p.R378K	0.014	0	0	0	0
p.R425C	0.014	0	0	0	0
p.S133Y	0.014	0	0	0	0
p.S249*	0.014	0	0	0	0
p.S410C	0.014	0	0	0	0
p.S410R	0	0	0.018	0	0
p.S47C	0	0	0	0	0.047
p.T75P	0	0	0.071	0	0
p.V276I	0.029	0.25	0.018	0	0.047
p.V318M	0.014	0	0	0	0
p.V335L	0	0.71	0.11	0	0
p.V48M	0.014	0	0	0	0
p.V52I	0.014	0.11	0	0	0.047
p.Y116D	0	0	0	0.045	0
p.Y278C	0.014	0	0	0	0

Supplementary table 3: **Sequencing quality metrics for PPARG variants**

Location on Ch3	base change	amino acid change	DP	GQ
12393119	A>G	p.I10V	63	99
12393119	A>G	p.I10V	20	99
12421253	A>T	p.I45F	108	99
12421253	A>T	p.I45F	394	99
12421253	A>T	p.I45F	350	99
12421260	C>G	p.S47C	97	99
12421262	G>A	p.V48M	92	99
12421266	A>C	p.D49A	53	99
12421267	T>G	p.D49E	406	99
12421267	T>G	p.D49E	106	99
12421267	T>G	p.D49E	98	99
12421267	T>G	p.D49E	283	99
12421274	G>A	p.V52I	101	99
12421274	G>A	p.V52I	102	99
12421274	G>A	p.V52I	135	99
12421274	G>A	p.V52I	178	99
12421274	G>A	p.V52I	521	99
12421279	G>A	p.M53I	103	99
12421343	A>C	p.T75P	70	99
12421343	A>C	p.T75P	149	99
12421343	A>C	p.T75P	155	99
12421343	A>C	p.T75P	191	99
12421355	G>A	p.E79K	73	99

12421355	G>A	p.E79K	133	99
12421355	G>A	p.E79K	82	99
12421355	G>A	p.E79K	91	46
12421355	G>A	p.E79K	112	99
12421391	G>A	p.A91T	95	99
12421391	G>A	p.A91T	256	99
12421391	G>A	p.A91T	413	99
12422856	T>G	p.Y116D	58	99
12422871	C>T	p.Q121*	65	99
12422880	A>G	p.N124D	21	99
12422908	C>A	p.S133Y	170	99
12422929	G>A	p.R140H	6	20
12422929	G>A	p.R140H	262	99
12434114	G>A	p.G161D	21	40
12434114	G>A	p.G161D	55	99
12434116	T>G	p.F162V	65	99
12434126	G>C	p.R165T	33	23
12434126	G>C	p.R165T	52	99
12434131	A>G	p.I167V	91	99
12434133	C>G	p.I167M	69	99
12434133	C>G	p.I167M	31	99
12434164	C>A	p.L178I	93	99
12434164	C>A	p.L178I	58	99
12434164	C>A	p.L178I	91	99
12434164	C>A	p.L178I	58	99
12434164	C>A	p.L178I	86	99
12434164	C>A	p.L178I	76	99
12434179	C>T	p.H183Y	87	99
12447410	G>A	p.E217K	58	99
12447410	G>A	p.E217K	70	99
12447449	G>T	p.D230Y	67	99
12447449	G>T	p.D230Y	49	99
12447479	C>T	p.R240W	81	99
12447507	C>G	p.S249*	35	99
12447512	A>G	p.I251V	50	99
12447537	C>T	p.A259V	46	99
12447572	G>A	p.D271N	56	99
12458209	G>A	p.V276I	32	99
12458209	G>A	p.V276I	28	99
12458209	G>A	p.V276I	83	99
12458209	G>A	p.V276I	88	99
12458209	G>A	p.V276I	29	99
12458209	G>A	p.V276I	26	99
12458209	G>A	p.V276I	36	99
12458209	G>A	p.V276I	28	99

12458209	G>A	p.V276I	39	99
12458209	G>A	p.V276I	101	99
12458209	G>A	p.V276I	15	99
12458216	A>G	p.Y278C	23	99
12458306	G>T	p.R308L	233	99
12458335	G>A	p.V318M	74	99
12458359	G>A	p.E326K	100	99
12458374	A>G	p.I331V	93	99
12458386	G>C	p.V335L	101	99
12458386	G>C	p.V335L	72	99
12458386	G>C	p.V335L	80	99
12458386	G>C	p.V335L	52	99
12458386	G>C	p.V335L	199	99
12458386	G>C	p.V335L	220	99
12458386	G>C	p.V335L	81	99
12458386	G>C	p.V335L	190	99
12458386	G>C	p.V335L	162	99
12458386	G>C	p.V335L	87	99
12458386	G>C	p.V335L	111	99
12458386	G>C	p.V335L	92	99
12458386	G>C	p.V335L	78	99
12458386	G>C	p.V335L	243	99
12458386	G>C	p.V335L	255	99
12458386	G>C	p.V335L	232	99
12458386	G>C	p.V335L	82	99
12458386	G>C	p.V335L	270	99
12458386	G>C	p.V335L	76	99
12458386	G>C	p.V335L	66	99
12458466	G>C	p.L361F	56	99
12458466	G>C	p.L361F	88	99
12458511	G>A	p.M376I	34	99
12458511	G>A	p.M376I	56	99
12458516	G>A	p.R378K	38	99
12458594	C>T	p.A404V	66	99
12458611	A>T	p.S410C	13	84
12458613	C>A	p.S410R	67	99
12458632	G>T	p.A417S	40	99
12475399	C>T	p.R425C	124	99
12475403	C>T	p.P426L	125	99
12475424	C>T	p.P433L	61	99
12475424	C>T	p.P433L	136	99
12475457	C>T	p.A444V	84	99
12475486	C>G	p.P454A	63	99
12475486	C>G	p.P454A	62	99
12475486	C>G	p.P454A	87	99

12475486	C>G	p.P454A	83	99
12475486	C>G	p.P454A	63	99
12475486	C>G	p.P454A	70	99
12475490	A>G	p.E455G	129	99
12475511	A>G	p.K462R	98	99
12475583	A>C	p.K486T	41	99
12475583	A>C	p.K486T	42	21

Variant position is based on human genome build NCBI36/hg18, and amino acid position is based on the protein reference sequence NP_005950.
DP = read depth at position, GQ = phred quality -10log_10p(genotype call is wrong)

Supplementary methods:

Sequencing, variant calling, data QC, and variant annotation

DNA libraries were barcoded using the Illumina index read strategy and sequenced with an Illumina HiSeq2000 following target capture with the Agilent SureSelect Human All Exon platform. Reads were mapped to the human genome hg19 with the BWA algorithm [1] and processed with the Genome Analysis Toolkit (GATK) [2] to recalibrate base quality-scores and perform local realignment around known indels. Target coverage for each sample was also computed with the GATK (Supplementary figure 2). Single nucleotide variants (SNVs) and small insertions and deletions (indels) were called with the Unified Genotyper module of the GATK and filtered to remove SNVs with annotations indicative of technical artefacts (such as strand-bias, low variant call quality, or homopolymer runs). In concordance with a previously published analysis framework[2], samples with fewer than 76% of targeted bases covered to 20x, with an abnormally high number of non-reference alleles or heterozygosity, or with an abnormally low concordance with prior SNP array genotypes (based on the distribution across all samples) were excluded from analysis. Any sample genotype at a site with fewer than 10x coverage in the sample was not included in analysis (i.e. set as missing) (Supplementary table 3). Variants were annotated with the Variant Effect Predictor [3]. Non-synonymous variants identified in *PPARG*

(transcript isoform 2, ENST00000287820) were advanced for downstream experimental characterization and statistical analysis. Raw sequence read data for each variant carrier was examined as a quality control step; variants or genotypes that had visual signatures of sequencing artifacts – such as reads of poor mapping quality, evidence for variation supported by only reads on one strand of the genome, or additional called variants nearby – were excluded from further analysis.

Rescue of adipocyte differentiation by in vitro PPARG variant constructs

To recreate variants *in vitro*, cloned human *PPARG2* (the adipocyte predominant isoform) sequences were engineered in parallel by PCR mutagenesis (Stratagene Quikchange II) to create a series of constructs, one per variant. To enable controlled titration of *PPARG* dosage, these constructs, were cloned into a custom designed lenti-viral plasmid backbone containing a doxycycline inducible promoter (Tet-ON). Each variant plasmid was packaged into lentivirus using standard protocols (http://www.broadinstitute.org/rnai/public/resources/protocols), matched by titre, and used to transduce Simpson-Golabi Behmel Syndrome (SGBS) pre-adipocytes (a gift from M. Wabitsch). Subsequently, transduced SGBS pre-adipocytes were stimulated to differentiate for 8 days in the presence of 0.1ug/mL doxycyline to activate gene expression of exogenous *PPARG,* 2uM rosiglitazone, and a standard hormonal cocktail described previously lacking the cAMP agonist IBMX[4]. In this assay, pre-adipocytes differentiate only when provided with functional, exogenous *PPARG* (Figure 1c). Levels of endogenous *PPARG* and exogenous *PPARG* were monitored with monitored with probes (Nanostring nCounter) directed against the native 3'UTR and the lenti-viral expression 3'UTR

respectively with no increase in endogenous *PPARG* expression over background levels in pre-adipocytes (Figure 1d).

Association tests

Analysis was performed separately for three groups of individuals – those sequenced for the T2D-GENES and GoT2D projects, those sequenced for the SIGMA-T2D project, and those sequenced for the FHS/JHS projects – and then combined via a meta-analysis to produce the final estimated significance and effect sizes. As a control, diabetes risk of variant carriers relative to non-carriers of the common *PPARG* rs1801282 variant was assessed in parallel.

For each group of individuals, each individual was scored according to the presence of a variant with the relevant annotation (e.g., carriers assigned a 1 and non-carriers a 0). Analysis was then performed using a linear mixed-model, regressing T2D status on the genotype score, within the EPACTS software package (http://www.sph.umich.edu/csg/kang/epacts/). A kinship matrix for the analysis was computed using independent SNPs (MAF >1%). For the T2D-GENES/GoT2D and SIGMA-T2D analyses, SNPs for the kinship matrix were obtained from the exome-wide sequencing data produced. For the PMBL and FHS/JHS projects, SNPs were obtained from available genome-wide SNP array data on the same subjects.

Point estimates for odds ratios were computed via the logistic regression score test (http://stattech.wordpress.fos.auckland.ac.nz/files/2012/11/skat-meta-paper.pdf) as implemented in EPACTS, with 10 principal components as covariates (computed via the EIGENSTRAT [5]software package from the same SNPs as for the kinship matrix). These were then transformed into 95% confidence intervals using standard error estimates back calculated

from the p-values produced by the linear mixed model.

The resulting three estimated odds ratios, standard errors, and association p-values were combined via an inverse variance based fixed-effects meta-analysis (as implemented in the METAL software package [6]) to obtain an estimated odds-ratio and p-value for association across the full set of studied individuals.

References for Supplementary Methods:

1. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
2. DePristo, M.A., et al., *A framework for variation discovery and genotyping using next-generation DNA sequencing data.* Nat Genet, 2011. **43**(5): p. 491-8.
3. McLaren, W., et al., *Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor.* Bioinformatics, 2010. **26**(16): p. 2069-70.
4. Fischer-Posovszky, P., et al., *Human SGBS cells - a unique tool for studies of human fat cell biology.* Obes Facts, 2008. **1**(4): p. 184-9.
5. Price, A.L., et al., *Principal components analysis corrects for stratification in genome-wide association studies.* Nat Genet, 2006. **38**(8): p. 904-9.
6. Willer, C.J., Y. Li, and G.R. Abecasis, *METAL: fast and efficient meta-analysis of genomewide association scans.* Bioinformatics, 2010. **26**(17): p. 2190-1.