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

Amir R. Majithia\(^a,b,c\), Jason Flannick\(^a\), Peter Shahinian\(^d\), Michael Guo\(^e\), Mark-Anthony Bray\(^f\), Pierre Fontanillas\(^g\), Stacey B. Gabriel\(^h\), GoT2D Consortium\(^j\), NHGRI JHS/FHS Allelic Spectrum Project\(^j\), SIGMA T2D Consortium\(^j\), T2D-GENES Consortium\(^k\), Evan D. Rosen\(^c,e\), and David Altshuler\(^a,b,c,d,f,1\)

*Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142; \(^b\)Department of Molecular Biology and Diabetic 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 lipodystrophy, 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 case/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; www.pnas.org/lookup/suppl/doi:10.1073/pnas.1410428111/-/DCSupplemental)

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.


The authors declare no conﬂict of interest.

1This Direct Submission article had a prearranged editor.

2Freely available online through the PNAS open access option.

3To whom correspondence should be addressed. Email: altshuler@mgl.harvard.edu.

4See 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.
53 nonsynonymous \textit{PPARG} variants were observed. Only one of these variants (the well-studied \textit{PPARG} P12A variant, rs1801282) demonstrated a minor allele frequency greater than 1% in any ancestry group we studied (\textit{SI Appendix}, Table 2). As expected, carriers of the common \textit{PPARG} P12A variant showed a reduced risk of T2D, consistent with previous studies.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|}
\hline
Location on chromosome 3 & Base change & Amino acid change & Ancestral geography & Counts in controls & Counts in T2D cases & Bioinformatic prediction$^\dagger$ & OR (95% CI) & \textit{P} \\
\hline
12458632 & G $\rightarrow$ T & A417S & European & 0 & 1 & Deleterious & & \\
12447449 & G $\rightarrow$ T & D230Y & South Asian & 0 & 1 & Deleterious & & \\
12447410 & G $\rightarrow$ A & E217K & Hispanic & 0 & 1 & Deleterious & & \\
12458359 & G $\rightarrow$ A & E326K & Hispanic & 0 & 1 & Deleterious & & \\
12434116 & T $\rightarrow$ G & F162V & European & 0 & 1 & Deleterious & & \\
12434114 & G $\rightarrow$ A & G161D & European & 0 & 2 & Deleterious & & \\
12434179 & C $\rightarrow$ T & H183Y & Hispanic & 1 & 0 & Deleterious & & \\
12434133 & C $\rightarrow$ G & I167M & European, European American & 1 & 1 & Deleterious & & \\
12458374 & A $\rightarrow$ G & L361F & European American & 1 & 1 & Deleterious & 2.11 & 0.12 \\
12475403 & C $\rightarrow$ T & P426L & European & 0 & 1 & Deleterious & & \\
12475486 & C $\rightarrow$ G & P454A & European, European American & 4 & 2 & Deleterious & & \\
12422871 & C $\rightarrow$ T & Q121* & European American & 1 & 1 & Deleterious & & \\
12422929 & G $\rightarrow$ A & R140H & Hispanic, African American & 1 & 1 & Deleterious & & \\
12434126 & G $\rightarrow$ A & R165T & European & 0 & 2 & Deleterious & & \\
12447479 & C $\rightarrow$ T & R240W & South Asian & 1 & 0 & Deleterious & & \\
12458306 & G $\rightarrow$ T & R308L & European & 0 & 1 & Deleterious & & \\
12458516 & G $\rightarrow$ A & R378K & European & 0 & 1 & Deleterious & & \\
12475399 & C $\rightarrow$ T & R425C & European & 0 & 1 & Deleterious & & \\
12422908 & C $\rightarrow$ A & S133Y & European & 1 & 1 & Deleterious & & \\
12447486 & C $\rightarrow$ T & S410R & Hispanic & 1 & 0 & Deleterious & & \\
12458335 & G $\rightarrow$ A & S47C & East Asian & 0 & 1 & Deleterious & & \\
12475407 & C $\rightarrow$ T & S410C & European & 0 & 1 & Deleterious & & \\
12458613 & C $\rightarrow$ A & S410R & Hispanic & 1 & 0 & Deleterious & & \\
12421260 & C $\rightarrow$ G & S47C & East Asian & 0 & 1 & Deleterious & & \\
12458335 & G $\rightarrow$ A & V318M & European & 0 & 1 & Deleterious & & \\
12475357 & C $\rightarrow$ T & A259V & European American & 1 & 0 & Benign & & \\
12458594 & C $\rightarrow$ T & A404V & Hispanic & 0 & 1 & Benign & & \\
12475457 & C $\rightarrow$ T & A444V & Hispanic & 0 & 1 & Benign & & \\
12421291 & G $\rightarrow$ A & A91T & African American & 3 & 0 & Benign & & \\
12447572 & G $\rightarrow$ A & D271N & European & 0 & 1 & Benign & & \\
12421266 & A $\rightarrow$ C & D49A & Hispanic & 1 & 0 & Benign & & \\
12421267 & T $\rightarrow$ G & D49E & African American & 2 & 2 & Benign & & \\
12475490 & A $\rightarrow$ G & E455G & European American & 1 & 0 & Benign & & \\
12421355 & G $\rightarrow$ A & E79K & European, East Asian & 1 & 4 & Benign & & \\
12393119 & A $\rightarrow$ G & I10V & South Asian & 1 & 1 & Benign & & \\
12434131 & A $\rightarrow$ G & I167V & European & 0 & 1 & Benign & & \\
12447512 & A $\rightarrow$ G & I251V & Hispanic & 0 & 1 & Benign & & \\
12421253 & A $\rightarrow$ T & I45F & African American & 3 & 0 & Benign & & \\
12458511 & G $\rightarrow$ A & M376I & European & 0 & 2 & Benign & & \\
12421279 & G $\rightarrow$ A & M53I & South Asian & 1 & 0 & Benign & & \\
12422880 & A $\rightarrow$ G & N124D & South Asian & 1 & 0 & Benign & & \\
12475424 & C $\rightarrow$ T & P433L & European, Hispanic, African American, East Asian & 5 & 6 & Benign & & \\
12458386 & G $\rightarrow$ C & V335L & African American, Hispanic & 11 & 9 & Benign & & \\
12421262 & G $\rightarrow$ A & V48M & European American & 1 & 0 & Benign & & \\
12421274 & G $\rightarrow$ A & V52I & African American, East Asian, European & 3 & 2 & Benign & & \\
12422856 & T $\rightarrow$ G & Y116D & South Asian & 1 & 0 & Benign & & \\
12458216 & A $\rightarrow$ G & Y278C & European & 0 & 1 & Benign & & \\
\hline
\end{tabular}
\caption{Rare, nonsynonymous variants in \textit{PPARG} identified from 19,752 T2D case-controls}
\end{table}

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 \url{www.ncbi.nlm.nih.gov/genome/guide/human/release_notes.html#b36}. CI, confidence interval.

$^\ast$Stop codon.

$^\dagger$Criteria for deleterious: A variant must have a mammalian conservation LOD score >10 and be categorized as damaging by Condel (17) (Methods).
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 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. 1A). This assay compared favorably with standard triglyceride quantification methods using Oil Red O staining and extraction (Fig. 1B) 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 PPARG.](image-url)
presence of functioning, exogenous PPARG (Fig. 1C) and maintained endogenous PPARG at background levels (Fig. 1D).

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. 2A). 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. 2A, blue bars) and to cause LOF (Fig. 2A, red bars) were generated and tested in parallel as positive and negative controls, respectively. Among these previously characterized variants, those characterized as benign (Fig. 2A, blue bars) stimulate adipocyte differentiation with similar efficacy as WT PPARG whereas those known to cause reduced protein activity (Fig. 2A, 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 (12). 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. 2B). 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. 2B).

**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 of T2D.
Fig. 3. T2D case/control status in multiethnic individuals harboring non-synonymous 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.

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

<table>
<thead>
<tr>
<th>PPARG variant</th>
<th>Effect on PPARG function</th>
<th>T2D status</th>
<th>Ethnicity</th>
<th>Age</th>
<th>Sex</th>
<th>BMI</th>
<th>SystolicBP</th>
<th>DiastolicBP</th>
<th>Total cholesterol</th>
<th>LDL</th>
<th>HDL</th>
<th>Triglycerides</th>
</tr>
</thead>
<tbody>
<tr>
<td>R165T</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>40</td>
<td>M</td>
<td>23.6</td>
<td>125</td>
<td>82.5</td>
<td>184</td>
<td>201</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R165T</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>74</td>
<td>M</td>
<td>33.6</td>
<td>150</td>
<td>115</td>
<td>189</td>
<td>100</td>
<td>35</td>
<td>280</td>
</tr>
<tr>
<td>F162V</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>65</td>
<td>M</td>
<td>25.3</td>
<td>160</td>
<td>85</td>
<td>268</td>
<td>188</td>
<td>53</td>
<td>135</td>
</tr>
<tr>
<td>S249*</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>55</td>
<td>F</td>
<td>21.4</td>
<td>177.5</td>
<td>102.5</td>
<td>228</td>
<td>145</td>
<td>41</td>
<td>211</td>
</tr>
<tr>
<td>Q121*</td>
<td>Control</td>
<td>Caucasian</td>
<td>36–62</td>
<td>F</td>
<td>20.0</td>
<td>96*</td>
<td>67*</td>
<td>184*</td>
<td>114</td>
<td>65*</td>
<td>12*</td>
<td></td>
</tr>
<tr>
<td>G161D</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>54</td>
<td>M</td>
<td>25.2</td>
<td>149</td>
<td>84</td>
<td>256</td>
<td>173</td>
<td>46</td>
<td>183</td>
</tr>
<tr>
<td>G161D</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>82</td>
<td>M</td>
<td>23.7</td>
<td>150</td>
<td>90</td>
<td>203</td>
<td>125</td>
<td>30</td>
<td>236</td>
</tr>
<tr>
<td>V318M</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>55</td>
<td>F</td>
<td>29.3</td>
<td>110</td>
<td>65</td>
<td>180</td>
<td>28</td>
<td>395</td>
<td></td>
</tr>
<tr>
<td>R245C</td>
<td>Severe</td>
<td>Case</td>
<td>European</td>
<td>50</td>
<td>M</td>
<td>26.1</td>
<td>134</td>
<td>77</td>
<td>217</td>
<td>137</td>
<td>36</td>
<td>223</td>
</tr>
<tr>
<td>P426L</td>
<td>Mild</td>
<td>Case</td>
<td>European</td>
<td>49</td>
<td>F</td>
<td>24.5</td>
<td>243</td>
<td>158</td>
<td>230</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E217K</td>
<td>Mild</td>
<td>Case</td>
<td>Hispanic</td>
<td>61</td>
<td>F</td>
<td>21.5</td>
<td>146</td>
<td>82</td>
<td>215</td>
<td>129</td>
<td>51</td>
<td>181</td>
</tr>
<tr>
<td>Y278C</td>
<td>Mild</td>
<td>Case</td>
<td>European</td>
<td>69</td>
<td>F</td>
<td>25.8</td>
<td>128</td>
<td>81</td>
<td>202</td>
<td>139</td>
<td>37</td>
<td>126</td>
</tr>
<tr>
<td>R140H</td>
<td>Mild</td>
<td>Case</td>
<td>Hispanic</td>
<td>55</td>
<td>M</td>
<td>31.0</td>
<td>130</td>
<td>77</td>
<td>249</td>
<td>186</td>
<td>41</td>
<td>119</td>
</tr>
<tr>
<td>R140H</td>
<td>Mild</td>
<td>Control</td>
<td>African</td>
<td>67</td>
<td>F</td>
<td>33.2</td>
<td>125</td>
<td>89</td>
<td>216</td>
<td>136</td>
<td>44</td>
<td>178</td>
</tr>
<tr>
<td>M376I</td>
<td>Mild</td>
<td>Case</td>
<td>European</td>
<td>39</td>
<td>M</td>
<td>24.3</td>
<td>135</td>
<td>86</td>
<td>193</td>
<td>104</td>
<td>57</td>
<td>164</td>
</tr>
<tr>
<td>M376I</td>
<td>Mild</td>
<td>Case</td>
<td>European</td>
<td>44</td>
<td>M</td>
<td>26.5</td>
<td>135</td>
<td>86</td>
<td>193</td>
<td>104</td>
<td>57</td>
<td>164</td>
</tr>
</tbody>
</table>

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, in which 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 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. 1C). 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-dipyromethene (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 4x 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. 1A).

**Statistical Analysis.** In the experimental classification of PPARG variants, differentiation scores for variants were compared with differentiation scores for functional PPARG. Variants 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 GM089852, for assistance with image analysis. Funding for the GoT2D and T2D-GENES studies was provided by Grants SU01DK085526 (Multiethnic Study of Type 2 Diabetes Genes), DK088389 (Low-Pass Sequencing and High-Density SNP Genotyping for Type 2 Diabetes), and US4HG003067 (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/JHS Alleric Spectrum Project was supported by grants from the National Human Genome Research Institute of the 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.