Pancreatic Differentiation of Stem Cells Reveals Pathogenesis of a Syndrome of Ketosis-Prone Diabetes

Diane Yang,[1,2,3] Sanjeet Patel,[4] Wojciech J. Szlachcic,[5] Jolanta Chmielowiec,[1,2] Diane Scaduto,[6] Nagireddy Putluri,[1] Arun Sreekumar,[1] James Suliburk,[7] Michael Metzker,[8] Ashok Balasubramanyam,[9] and Malgorzata Borowiak[1,2,3,5]

Genetic analysis of an adult patient with an unusual course of ketosis-prone diabetes (KPD) and lacking islet autoantibodies demonstrated a nucleotide variant in the 5′-untranslated region (UTR) of *PDX1*, a β-cell development gene. When differentiated to the pancreatic lineage, his induced pluripotent stem cells stalled at the definitive endoderm (DE) stage. Metabolomics analysis of the cells revealed that this was associated with leucine hypersensitivity during transition from the DE to the pancreatic progenitor (PP) stage, and RNA sequencing showed that defects in leucine-sensitive mTOR pathways contribute to the differentiation deficiency. CRISPR/Cas9 manipulation of the *PDX1* variant demonstrated that it is necessary and sufficient to confer leucine sensitivity and the differentiation block, likely due to disruption of binding of the transcriptional regulator NFY to the *PDX1* 5′-UTR, leading to decreased PDX1 expression at the early PP stage. Thus, the combination of an underlying defect in leucine catabolism characteristic of KPD with a functionally relevant heterozygous variant in a critical β-cell gene that confers increased leucine sensitivity and inhibits endocrine cell differentiation resulted in the phenotype of late-onset β-cell failure in this patient. We define the molecular pathogenesis of a diabetes syndrome and demonstrate the power of multiomics analysis of patient-specific stem cells for clinical discovery.

Ketosis-prone diabetes (KPD) comprises a group of atypical diabetes syndromes characterized by presentation with diabetic ketoacidosis (DKA) (1), which implies severe β-cell dysfunction. These include a syndrome in which adult patients lacking evidence of islet autoimmunity experience prolonged periods of insulin-independent glycemic control following the index DKA before relapsing to insulin dependence (2,3). Our previous metabolomics and stable isotope kinetic investigations of patients with this form of KPD revealed underlying hypercatabolism of branched-chain amino acids (BCAAs), promoting ketogenesis and slowed tricarboxylic acid cycling (4). Here, we performed genetic analysis coupled with detailed multiomics investigations in induced pluripotent stem cells (iPSCs) derived from a KPD patient with an unusual clinical course and a potentially functional variant in the pancreas and duodenum homeobox-1 (PDX1) gene. In the course of differentiation of this patient's iPSCs through defined cellular stages of the pancreatic lineage, we identified both the leucine hypercatabolism common in KPD patients and a specific molecular mechanism that together explain the pathogenesis.

A Hispanic man was diagnosed with diabetes at 39 years of age and placed on metformin. Three years later

he was admitted for an episode of DKA. The medical history was otherwise significant only for occasional bouts of drinking. He was negative for islet cell autoantibodies to GAD65, islet antigen 2, and zinc transporter 8. The fasting serum C-peptide level was undetectable, with no response to glucagon upon repeated testing over several years. The patient has been on a basal-bolus insulin regimen following the first episode of DKA, with poor glycemic control (HbA$_{1c}$ 10–11%) and multiple recurrences of DKA.

The patient's late onset of diabetes without evidence of islet autoimmunity and initial treatment with metformin alone suggest that he retained significant insulin secretory reserve following diagnosis (2–5). Three years later he developed DKA, then never recovered β-cell function and has remained permanently dependent on insulin therapy. He has been euthyroid, without clinical or biochemical evidence of autoimmune disease or liver, kidney, or cardiac dysfunction. He has moderate diabetic retinopathy. His BMI has ranged from 23.5 to 25 kg/m^2. The current treatment regimen includes insulin, canagliflozin, metformin, lisinopril, and atorvastatin.

RESEARCH DESIGN AND METHODS

Human Dermal Fibroblasts Reprogramming to iPSCs

Using a protocol approved by the Baylor College of Medicine Institutional Review Board and after obtaining informed consent, we collected skin samples from participants with punch biopsy. Dermal fibroblasts were isolated, expanded, and frozen in preparation for induction into iPSCs. Fibroblasts were reprogrammed into iPSCs with use of a nonintegration method based on modified mRNA with a StemRNA Reprogramming Kit (Stemgent; Cambridge, MA). Cell lines were evaluated with G-banded karyotyping analysis (Supplementary Fig. 1E). Cells were characterized based on expression of pluripotency genes (REX1, OCT3/4, SOX2, NANOG, and HTERT) evaluated with quantitative PCR (qPCR) and immunofluorescence (Supplementary Fig. 1C and data not shown). Lastly, iPSC lines were differentiated with a spontaneous differentiation assay, and expression of markers of the three germ layers, GATA4, AFP, CTnT, FLK1, PAX6, TUBB3, BRA, and TUJ1, was assessed with qPCR.

Human iPSCs Maintenance

The iPSCs were maintained on Matrigel (BD Biosciences; San Jose, CA) in Essential 8 (E8) media (STEMCell Technologies; Vancouver, Canada). Cells were passaged every 3–5 days at 80% confluency with TrypLE Express (Invitrogen, Carlsbad, CA). After passage, cells were kept in 10 μmol/L Y-27632 (Stemgent) for 24 h.

Pancreatic Human iPSC Differentiation

For initiation of differentiation, the cells were dissociated into single cells and seeded at 150,000 cell/cm^2 in E8 media supplemented with 10 μmol/L Y27632 (Stemgent).

Differentiation began 2 days following seeding. The leucine concentrations used in the differentiation medium were chosen based on the literature and titration curves. The following media and growth factors/small molecules were used for pancreatic differentiation.

0.4 mmol/L Leucine RPMI

Day 1, RPMI-1640 media + 3 μmol/L CHIR99021 (Stemgent) + 100 ng/mL recombinant human (rh) Activin A (R&D Systems; Minneapolis, MN); days 2–3, +100 ng/mL rh Activin A + 0.2% FBS; days 4–5, + 2% FBS + 50 ng/mL keratinocyte growth factor (KGF) (PeproTech; Rocky Hill, NJ); days 6–9, DMEM/B27 + 50 ng/mL KGF + 2 μmol/L retinoic acid (MilliporeSigma; Burlington, MA) + 0.25 μmol/L SANT-1 (MilliporeSigma) + 100 ng/mL rh Noggin (R&D Systems); and days 10–14, DMEM/B27 + 1 μmol/L PdBU+ 1 μmol/L Alk5i + 100 ng/mL rh Noggin (R&D Systems).

1 mmol/L Leucine Protocol

Day 1, MCDB-131 media with 0.1% BSA + 10 mmol/L glucose + 100 ng/mL rh Activin A and 3 μmol/L CHIR99021; days 2–3, MCDB-131 media + 0.1% BSA + 10 mmol/L glucose + 100 ng/mL rh Activin A; days 4–5, MCDB-131 + 0.1% BSA, + 10 mmol/L glucose, + vitamin C (VitC) + 50 ng/mL KGF; days 6–9, MCDB-131 + 2% BSA + 5.5 mmol/L glucose + 0.25 mmol/L VitC + 1:200 Insulin-Transferrin-Selenium (ITS) (Invitrogen) + 50 ng/mL KGF + 2 μmol/L retinoic acid + 0.25 μmol/L SANT-1 + 100 ng/mL rh Noggin; days 10–12, MCDB-131 + 2% BSA + 5.5 mmol/L glucose + 0.25 mmol/L VitC + 1:200 ITS + 0.25 μmol/L SANT-1 + 100 ng/mL rh Noggin + 1 μmol/L PdBU; days 13–15, MCDB-131 + 2% BSA + 5.5 mmol/L glucose + 0.25 mmol/L VitC + 1:200 ITS + 100 ng/mL rh Noggin + 1 μmol/L AlK5i.

RNA Sequencing

Control and patient iPSCs and cells differentiated to the definitive endoderm (DE) and pancreatic progenitor (PP) stages were harvested, and RNA extracted using TRIzol (Thermo Fisher Scientific) according to the manufacturer's protocol. The quality of extracted RNA samples was determined with a 2100 Bioanalyzer and the Agilent RNA 6000 Nano Kit (Agilent Technologies; Santa Clara, CA). Samples with RNA integrity numbers (RIN) >9 were used for mRNA library preparation with the TruSeq stranded mRNA Library Prep Kit LT (Illumina, San Diego, CA). The concentration of the libraries was determined using Quantitative PCR KAPA Library Quantification Kit for Illumina (Kapa Biosystems), and the sequencing runs were performed on a NextSeq 500 (Illumina).

CRISPR/Cas9-Mediated PDX1 Variant Correction in Patient iPSCs

The two guide RNAs targeting the *PDX1* sequence (sequences in Supplementary Table 4) were cloned into px335, which contains the sequence of a human codon-optimized

SpCas9 nickase. The DNA constructs, including the two guide RNAs in px335 and a 91 base pairs (bp) ssODN with a corrected PDX1 variant sequence and mutated PAM sequence, were delivered using 4D-Nucleofector (Lonza; Morristown, NJ) with a P3 primary cell kit and CA137 program. The targeted cells were plated in Geltrex-coated cell culture plates (Thermo Fisher Scientific) until colonies formed. The colonies were individually picked and separated into 96-well plates to expand. For verification of gene-targeting success, DNA was extracted, and the PDX1 variant site was amplified using PCR with the forward and reverse primer (Supplementary Table 4). The 867-bp PCR product was digested with BaeG1 and visualized on an agarose gel. The wild-type sequence fragments were 306, 230, 181, 84, and 67 bp; the mutant fragments were 411, 306, 84, and 67 bp.

Introduction of PDX1 Variant Into iCas9 Human Embryonic Stem Cells

Single guide (sg)RNAs and DNA donors were manufactured as ssDNA oligos (MilliporeSigma), and the sequences are provided in Supplementary Table 4. To get functional sgRNAs, the second DNA strand was synthesized using high-fidelity PCR (Q5 Hot Start polymerase; New England Biolabs, Ipswich, MA) with the universal primers T7_F and Tracr_R, and the resulting product was cleaned with a GeneJET PCR Purification Kit (Thermo Fisher Scientific). In vitro transcription was carried out with a MEGAshortscript T7 Transcription Kit and cleaned using a MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific). Human embryonic stem cell (hESC) Hues8-iCas9 cells (3) were cultured to 90% confluency and treated with 2 μg/mL doxycycline 24 h before transfection to induce Cas9 expression. Cells were dissociated into single cells with TrypLE Express (Thermo Fisher Scientific) and resuspended in 5×10^5 cells per 100 μL E8 medium (Thermo Fisher Scientific). iCas9 cells were reversely transfected using Lipofectamine Stem reagent (Thermo Fisher Scientific) with 1 μg each sgRNA (single or multiplexed) and 200 pmol ssDNA donor. First, the Lipofectamine-oligo complexes were plated on a single 24-well plate well coated with Vitronectin (Thermo Fisher Scientific) and filled with E8 medium, and 100 μL cells added. The next day, cells were washed and medium was changed. Cells were incubated with 10 μmol/L Y-27632 for 48 h and with doxycycline for 24 h posttransfection. After 48–72 h, cells were passaged and collected for genotyping. DNA was isolated with a Genomic Mini kit (A&A Biotechnology; Gdynia, Poland) and the PDX1 region was amplified using PCR with GoTaq G2 Hot Start polymerase (Promega; Madison, WI) with 1% DMSO and F+R (PDX1 specific) or F3+R (SNP specific) primer pairs. Primer sequences are provided in Supplementary Table 4. Reaction products were run on 1.5% agarose gel, and following PCR clean-up, F+R products were sequenced using the Sanger method. The sequencing results were deconvoluted

using Inference of CRISPR Edits (ICE) analysis (Synthego; Silicon Valley, CA) to estimate editing efficiency.

Gas Chromatography–Mass Spectrometry Array and Metabolomics Analysis

Extracted samples underwent high-performance liquid chromatography (HPLC) coupled to Agilent 6490 Triple Quadrupole mass spectrometry in the electrospray ionization positive mode. The HPLC column was an XBridge Amide, 3.5 μm, 4.6 × 100 mm (Waters, Milford, MA). Mobile phases A and B were 0.1% formic acid in water and acetonitrile, respectively, with gradient flow 0–3 min 85% B, 3–12 min 30% B, 12–15 min 2% B, and 16 min 95% B and re-equilibration through 23 min to the initial starting condition of 85% B. The flow rate of the solvents for the analysis was 0.3 mL/min. Samples were isolated and data were processed using the quantitative mode of the MassHunter software (Agilent). Exported data were normalized using internal standards; normalized data were used for further analysis. To ensure quality data, quality controls (liver tissue extract) were used to monitor the extraction efficiency, instrument variation, and reproducibility as previously described (6). Normalized abundances of metabolites underwent principal components analysis (PCA) (RStudio, version 1.2.5033) with use of the prcomp package. The R packages were used to make figures. For heat maps, the package pheatmap was used to create figures demonstrating normalized expression or abundances in RNA sequencing and metabolomics experiments, respectively. For the waterfall plot of the second principal component (PC2) loadings, the package ggplot was used on the values derived from the principle components analysis.

mTOR Inhibition

For mTOR inhibition, cells were treated at days 4–9 of differentiation with varying concentrations of broad PI3K/mTOR inhibitor PI-103 (AdooQ BioScience, Irvine, CA) or specific mTOR inhibitor Torin2 (AdooQ BioScience) or with DMSO as a control. PI-103 was added at 20 nmol/L (half-maximal effective concentration for mTORC1 inhibition), 2 nmol/L, and 0.2 nmol/L concentrations and Torin2 at 0.25, 0.025, and 0.0025 nmol/L.

siRNA Knockdown

Dharmacon ON-TARGETplus SMARTpool anti-NFYB or control siRNAs (Horizon Discovery, Cambridge, U.K.) were delivered to differentiating PP cells at 20 nmol/L concentration using Lipofectamine RNAiMAX Reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. Cells were collected at 24–48 h posttransfection for qPCR analyses and at 72 h for flow cytometry.

Data and Resource Availability

All data generated or analyzed during this study are included in this article (and in Supplementary Material).

The RNA-sequencing data sets generated and analyzed during the current study are available in the Gene Expression Omnibus (GEO) repository under accession no. GSE165385.

RESULTS

The patient's genomic DNA was sequenced against a panel of genes associated with monogenic diabetes, and the results were compared with ethnic-specific control DNA sequences. The patient had a single, rare, heterozygous nucleotide variant (6), C>T located 18 bases upstream of the translational start site in the 5'-untranslated region (UTR) of PDX1 (Fig. 1A). This gene is critical for pancreas and β-cell development and maintenance and for insulin production (7,8).

To determine how the PDX1 variant might contribute to severe β-cell dysfunction in this patient, we derived iPSCs from the patient's skin fibroblasts, differentiated them in the pancreatic lineage, and studied stage-specific functional defects in comparison with iPSCs from an unrelated healthy adult control subject without diabetes. We obtained informed consent for skin biopsies under a Baylor College of Medicine Institutional Review Board–approved protocol to study patient-specific iPSCs. We used a nonintegrating reprogramming approach to generate iPSCs from skin fibroblasts (Supplementary Fig. 1A). Reprogramming efficiency was similar in the cells of both participants (data not

shown). The PDX1 −18 C>T variant was confirmed in the iPSCs of the patient (Supplementary Fig. 1B).

iPSCs from the patient and control subject had normal karyotypes and doubling times and expressed pluripotency markers such as OCT3/4, SOX2, REX1, NANOG, and HTERT (Supplementary Figs. 1C–E and 2A and data not shown). Both iPSC lines were differentiated in standard culture medium with 0.4 mmol/L leucine through defined stages of the pancreatic lineage, each verified using well-characterized protein markers (Fig. 1B) (9,10). Both patient and control iPSCs transformed to DE at similar rates, as shown by the percentage of cells positive for the DE-specific markers SOX17 and FOXA2 using flow cytometry (Fig. 1C and Supplementary Fig. 2B). However, at the PP stage, the percentage of PDX1$^+$ cells diverged, with the patient's cells showing a diminished capacity to differentiate beyond the DE stage (Fig. 1C and D and Supplementary Fig. 2C). We extended this analysis to six additional patient-specific iPSC clonal cell lines and observed the same differentiation defect at the PP stage (Supplementary Fig. 3). Hence, we focused on the DE-to-PP transition for further analyses.

We previously showed that patients with autoantibody-negative KPD and initially preserved insulin secretory reserve (denoted "A−β+" KPD) have a metabolomic signature of altered plasma levels of BCAAs due to kinetic defects in leucine/isoleucine catabolism (4). We hypothesized that
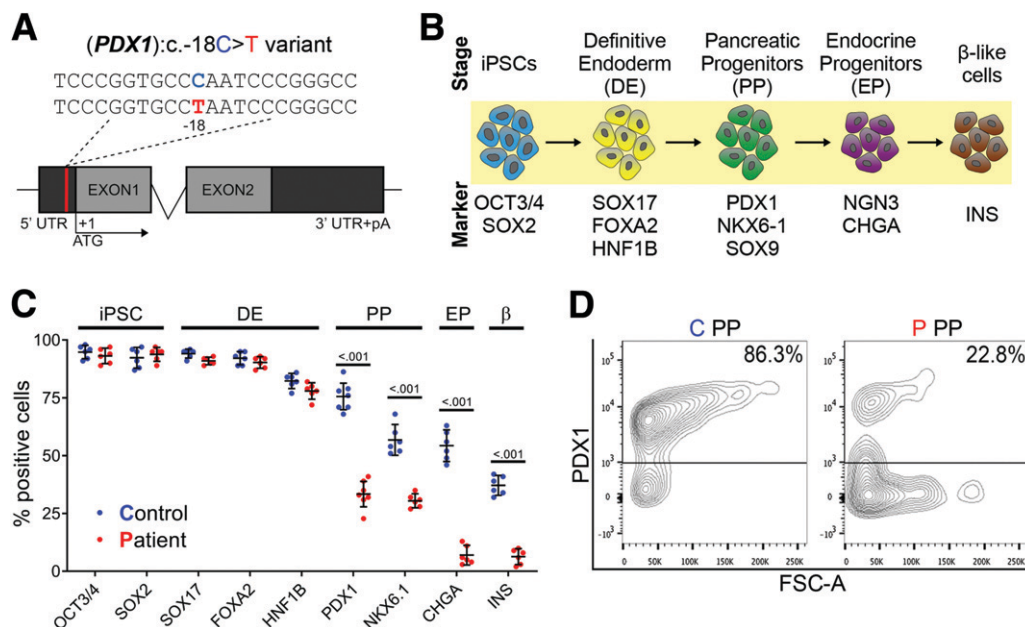

Figure 1—iPSCs derived from a KPD patient with a single nucleotide variant in the PDX1 gene 5'-UTR have diminished potential to differentiate into pancreatic progenitors. A: Schematic depiction of the single nucleotide variant in the PDX1 5'-UTR of the patient with KPD. B: Scheme of in vitro iPSC differentiation along the pancreatic lineage through DE, PP, EP, and β-like cell stages. C: Quantification of stage-specific marker protein expression during pancreatic in vitro differentiation for control and patient iPSCs. Each data point represents flow cytometry analysis of the selected protein. Data are presented as mean ± 95% CI. P values were determined by two-way ANOVA. n = 7. D: Representative flow cytometry plot showing percentage of live cells expressing the PDX1 pancreatic cell marker in control subject (C) and patient (P) cells at the PP stage. n = 8.

this plasma signature reflects a generalized phenomenon in tissues that metabolize BCAAs; hence, there would be a similar signature in the DE and PP stages of the pancreas-differentiated iPSCs. The complete metabolomics data are presented in Supplementary Table 1 and show that, indeed, leucine metabolites were increased in the patient's DE stage cells. Given the high-dimensional nature of metabolomics data, we used PCA to identify metabolic drivers of the in vitro cellular phenotype in the patient. PC1 describes metabolites explaining differences in cell stage rather than differences between patient and control subject cells at the iPSC or DE stage (Fig. 2A). By contrast, PC2 encompasses a signature that differentiates patient and control cells at the DE stage. Because PCA is agnostic with respect to the components that define differences, we analyzed the loadings for PC2, shown as a waterfall plot in Fig. 2B. α-Ketoisocaproic acid (α-KICA), a catabolite of leucine, was a key contributor to the difference between patient and control at the DE stage. Isoleucine/leucine (isobaric amino acids) were also prominent contributors. The scale in Fig. 2C represents z scores for the metabolites of iPSCs and DE cells, and Fig. 2D shows fold changes in isoleucine/leucine and α-KICA, indicating ~30% higher levels of these metabolites in the patient's DE cells compared with control cells.

To investigate whether increased leucine metabolism affects the differentiation process in the patient, we cultured the iPSCs in media containing either the standard (0.4 mmol/L) or increased (1 mmol/L) leucine concentration and induced differentiation. Control cells showed 85–89% efficiency in differentiating to the PP stage (defined by PDX1 expression) when cultured at either concentration of leucine. By contrast, the percentage of PDX1$^+$ cells was as low as 26% at 1 mmol/L leucine in the patient's cells, and this defect was only partially rescued (55% differentiation efficiency) when 0.4 mmol/L leucine concentration was used (Fig. 3A). We also measured expression of other markers of the PP stage. In the control subject, 50–70% of the cells cultured with 1 mmol/L leucine differentiated to the PP stage coexpressing FOXA2, SOX9, or NKX6.1 and PDX1. In the patient, only 6–10% of cells cultured with 1 mmol/L leucine differentiated to the PP stage showing coexpression of these markers, suggesting that induction of PPs is very inefficient for the patient-specific iPSCs due to increased sensitivity to leucine (Supplementary Fig. 4A). Subsequently, the patient PP cells could not progress to the endocrine progenitor (EP) stage, shown by lack of NGN3 expression (Supplementary Fig. 4B). When the leucine concentration was lowered to 0.4 mmol/L, the differentiation block in the patient's cells was overcome sufficiently to generate some NGN3$^+$ EPs (Supplementary Fig. 4B) and INS$^+$ β-like cells but with low efficiency (8%) compared with control cells (34%) (Supplementary Fig. 4C and D).

We extended our analysis to a heterologous stem cell line that has high propensity to differentiate into pancreatic cells. H1 cells (an hESC line) showed progressive

reduction in PDX1$^+$ PP cells as the leucine concentration in the medium increased from 4 to 10 mmol/L (Fig. 3B and Supplementary Fig. 2D). Differentiation of control iPSCs in the presence of 10 mmol/L leucine also resulted in ~50% reduction in the number of PDX1$^+$ cells (Supplementary Fig. 5). Thus, high levels of ambient leucine inhibit differentiation of embryonic stem cells and iPSCs, and the patient's iPSCs have increased sensitivity to this inhibition, even at low concentrations of leucine. We assessed the PDX1 protein level per cell in control and patient cells differentiated in media with varying leucine concentrations. Using antibody staining and flow cytometry, we observed a positive correlation between leucine concentration and average PDX1 intensity per cell in the control PDX1$^+$ cells (Supplementary Fig. 6A). We did not see a similar dose response in the patient PDX1$^+$ cells, which expressed less PDX1 per cell when differentiated in higher leucine concentration, suggesting aberrant in vitro differentiation of the patient's cells. Indeed, at the end of the PP stage, the patient cells underwent cell death (Supplementary Fig. 6B).

We performed RNA sequencing of the cells to identify molecular pathways that might inhibit the patient's iPSCs from transitioning from DE to PP. We sequenced cells from three independent pancreas differentiation and iPSC clones (Supplementary Fig. 7A). To avoid imputing genes differentially expressed at a single stage with no effect on a subsequent stage, we used a likelihood ratio test to identify genes different between patient and control subject at each stage of the differentiation process. We performed pathway analysis of functionally annotated genes among the >700 such genes that were identified (Supplementary Table 2). Differences in mRNA levels of this gene subset at each stage are shown in Fig. 3C. Analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and WikiPathways platforms revealed the patient had a gene signature enriched in the mTOR, AKT, and PI3K pathways (Fig. 3D and Supplementary Fig. 7B and C), linking high leucine to abnormal mTOR signaling (11). Levels of activated kinases in this pathway (phosphorylated [phospho]-PDK1, phospho-AKT, and phospho-mTOR) were increased in the patient's late DE cells cultured at 1 mmol/L leucine (Fig. 3E and F). For phospho-mTOR, the increase was twofold compared with control cells (Fig. 3F). This observation is generalizable, since leucine caused this outcome in hESCs, albeit at higher leucine concentrations, as evidenced by increased phospho-AKT and phospho-mTOR in H1 cells cultured in medium containing 10 mmol/L leucine (Fig. 3G). Next, we tested whether reduction in mTOR activity could reveal differential changes in PDX1 expression in control cells compared with patient cells. Rapamycin, a well-known mTORC1 inhibitor, caused massive cell death at any concentration tested in both control and patient cells (data not shown). With use of PI-103, a weak inhibitor of PI3K, or Torin2, a potent and selective mTOR inhibitor, the
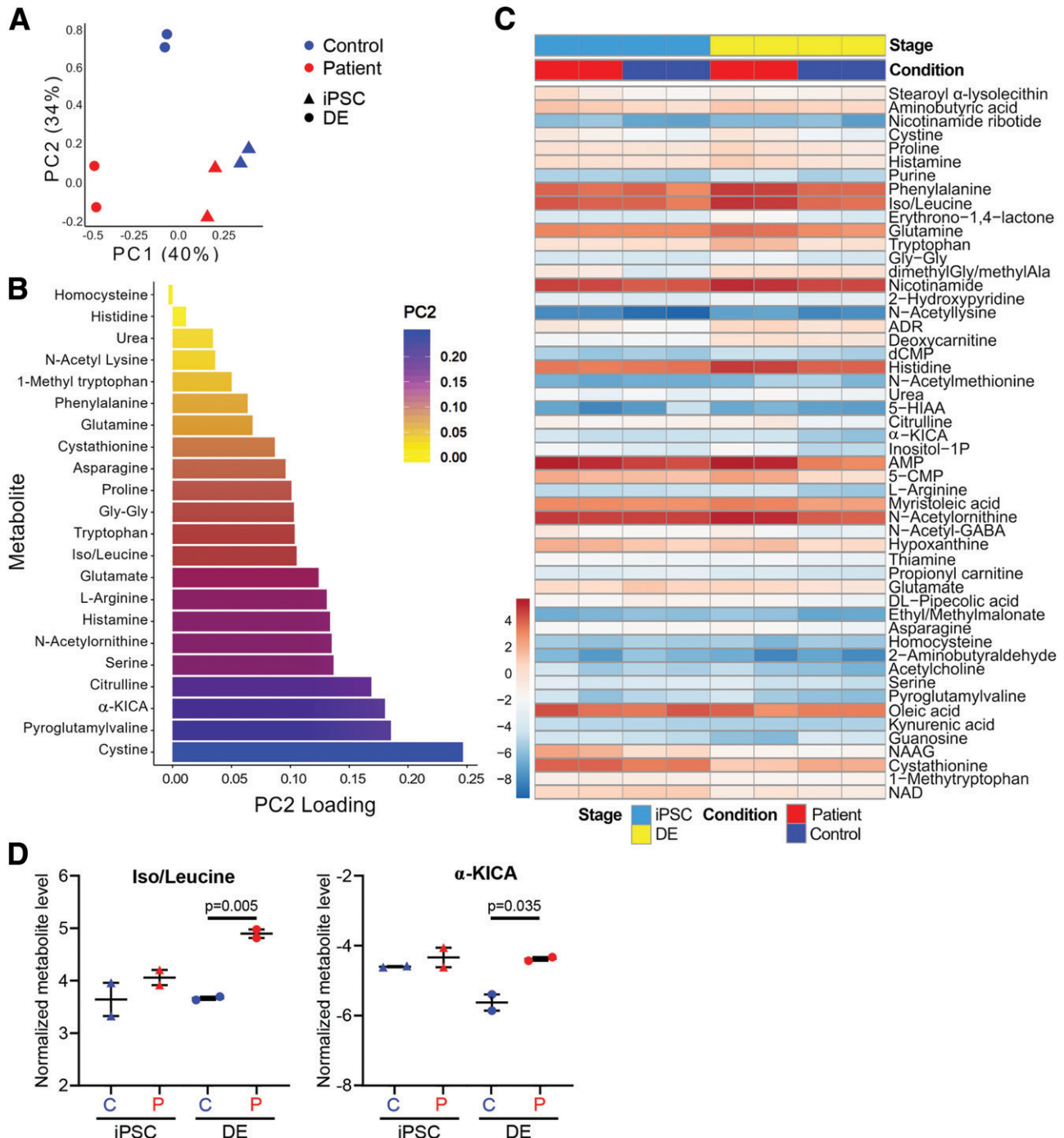
Figure 2—Leucine metabolism is altered in DE cells derived from patient iPSCs. *A*: PCA of control and patient cells at the iPSC and DE stages; dimension reduction using PCA of the metabolomics data shows that PC2 separates patient and control cells at the DE stage. *n* = 2. *B*: Rank ordering of the PC2 loadings shows prominent signals of α-KICA and iso/leucine. *C*: Row-normalized heat map of all metabolites confirms higher level of leucine and isoleucine in patient DE cells compared with control DE cells. *D*: Normalized iso/leucine and α-KICA levels in control and patient iPSCs and DE indicate higher content of these metabolites in patient DE. *n* = 2.

control cells demonstrated a dose-dependent increase in PDX1 expression level per cell, whereas the patient cells demonstrated a small but significant decrease in Pdx1 expression level per cell (Supplementary Fig. 8). Again, these data suggest that the patient's cells at the PP stage are functionally aberrant and do not represent true pancreatic progenitor cells.

We next tested whether the 5′-UTR variant in *PDX1* is necessary and sufficient to confer the unique sensitivity of the patient's iPSC cells to leucine-induced differentiation
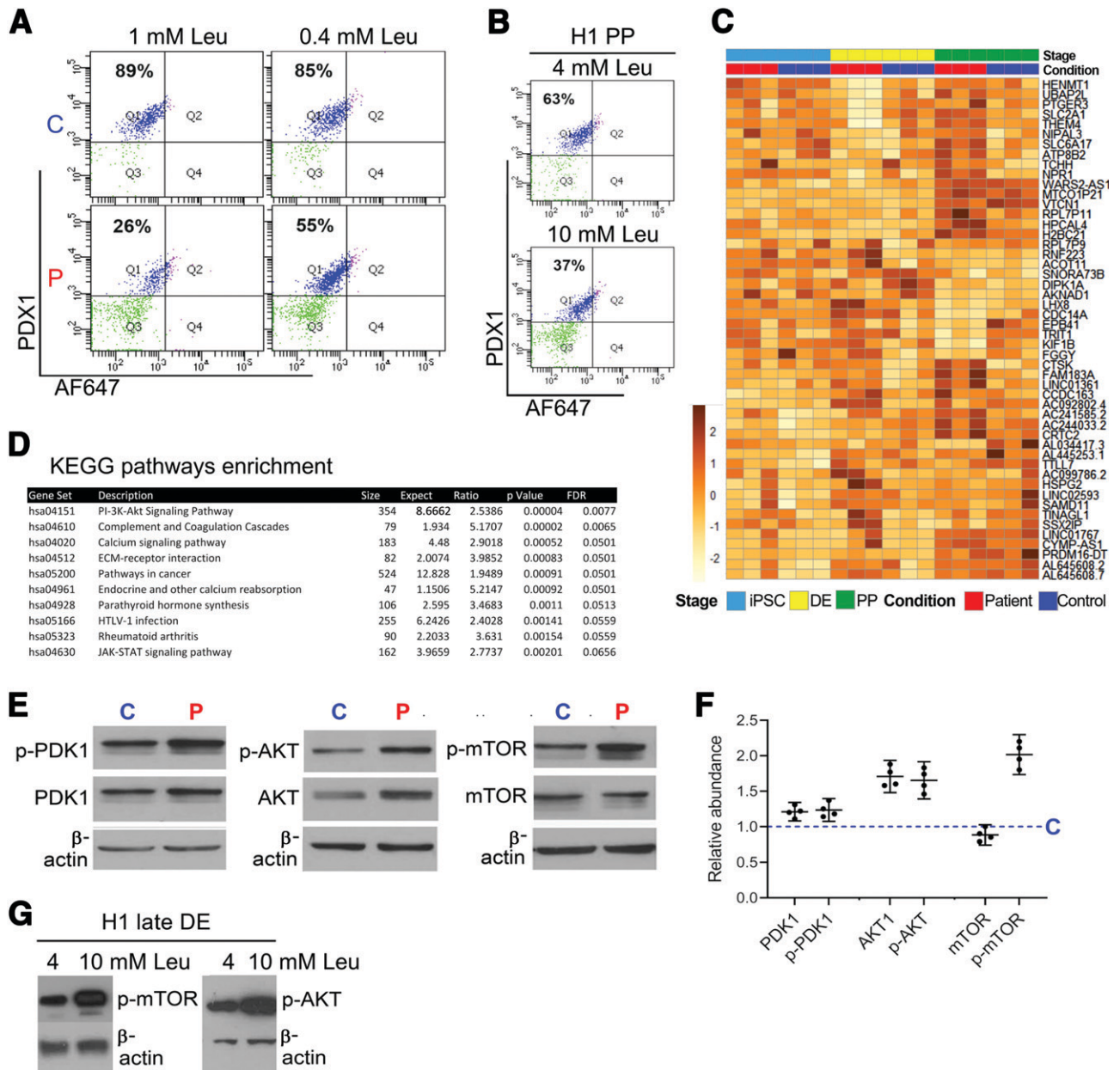
Figure 3—(PDX1)c.−18 C>T confers increased susceptibility to leucine-mediated inhibition of differentiation to pancreatic progenitors. *A*: PDX1 expression induction in patient iPSCs is lower than that of control iPSCs when differentiated in media containing 1 mmol/L leucine and still lower when differentiated in media containing 0.4 mmol/L leucine. PDX1 expression is higher and unchanged at either leucine concentration in control iPSCs. *B*: High leucine levels (4–10 mmol/L) impair PDX1 expression in H1 hESC-derived PP cells. *C*: Heat map of mRNA levels of the top 100 of ∼700 genes statistically different between patient and control cells across three stages (iPSC, DE, and PP), with use of a likelihood ratio test. *D*: KEGG pathway analysis confirms the enrichment of the PI3K-AKT pathway in the patient's cells. *E*: Phospho-PDK1, -AKT, and -mTOR levels are increased in the patient's late DE cells compared with the control subject's late DE cells when cultured with 1 mmol/L leucine in the medium. β-Actin was used as loading control. *F*: Quantification of immunoblots from *E* confirms increased levels of phospho-PDK1, phospho-AKT, and phospho-mTOR in the patient's late DE cells. $n = 4$. *G*: High leucine levels cause increased phosphorylation of mTOR and AKT in H1 hESC-derived late DE cells. C, control subject; P, patient; p-, phosphorylated.

inhibition. CRISPR/Cas9-mediated correction of the variant in the patient's iPSCs (Fig. 4*A* and *B*) increased the expression of PDX1, SOX9, and NKX6.1, key markers of the PP stage, when cultured in a medium containing 1 mmol/L leucine (Fig. 4*C*). After correction, quantification of the PDX1$^+$ cells showed a level of PDX1 expression in the patient's cells similar to that of the control cells (Fig. 4*D*). Furthermore, when corrected patient cells were differentiated through the endocrine lineage to β-like cells, we saw an efficiency of induction to CHGA$^+$ and INS$^+$ cells similar to that for control iPSCs (Fig. 4*E*). The reverse experiment was performed to determine whether the variant-induced sensitivity to
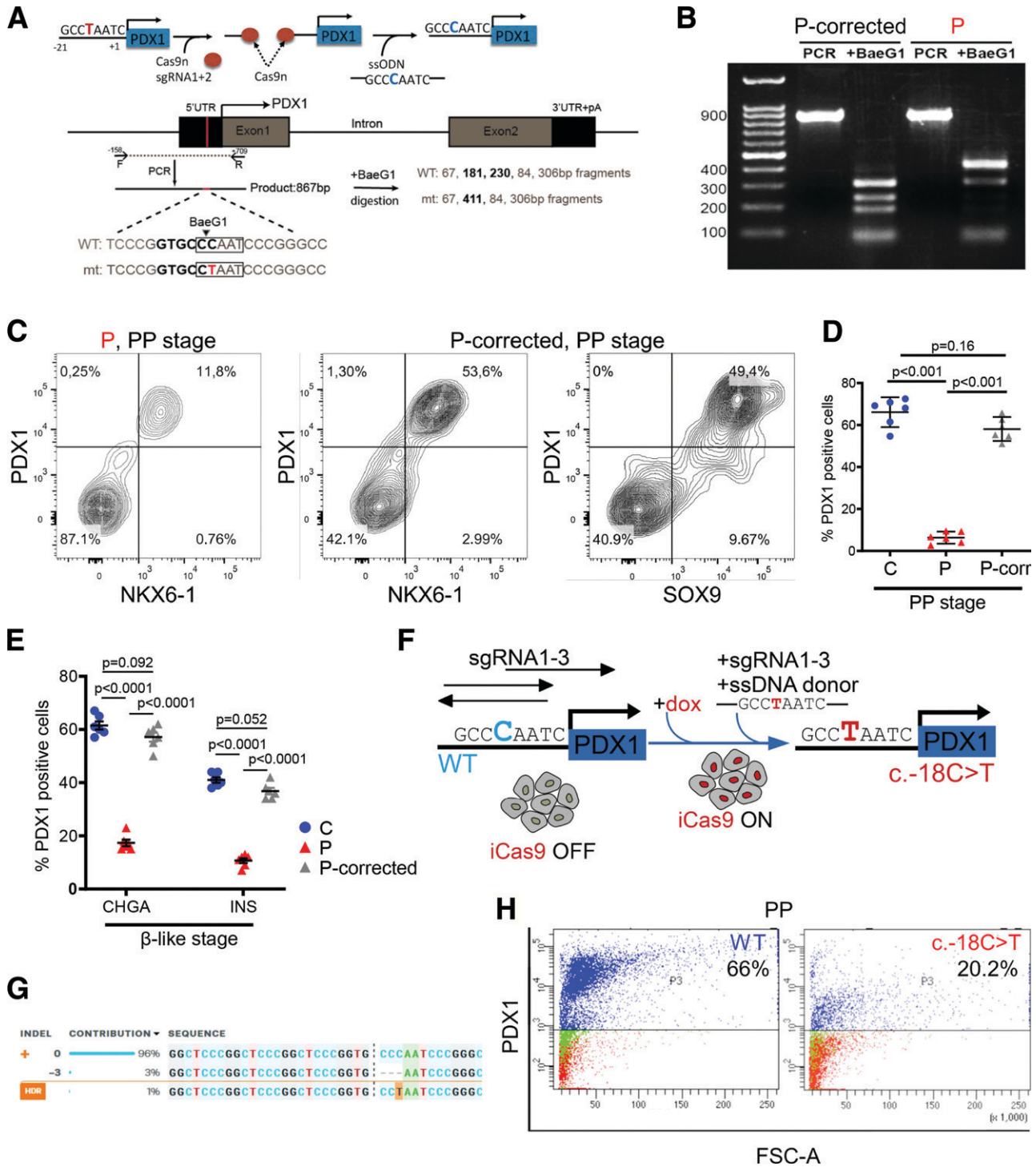
Figure 4—Correction of the 5′-UTR *PDX1* variant in patient iPSCs rescues differentiation potential at the PP stage, and de novo knock-in of the 5′-UTR variant in control iPSCs recapitulates the abnormal phenotype. *A*: Scheme of CRISPR/Cas9-mediated gene editing to correct the nucleotide variant in the −18 position of the *PDX1* 5′-UTR in the patient's cells. *B*: Restriction enzyme digestion of PCR products shows that the pattern of fragments cut by BaeG1 in the corrected *PDX1* 5′-UTR sequence differs from that in the patient's original sequence. *C*: Correction of the *PDX1* 5′-UTR variant restores efficient induction of PDX1$^+$, NKX6.1$^+$, and SOX9$^+$ PP cells from iPSCs. Representative flow cytometry plots show the percentage of PDX1$^+$/NKX6.1$^+$ and PDX1$^+$/SOX9$^+$ cells at the PP stage in the original (P) compared with the corrected (P-corrected) patient iPSCs. *D*: Quantification of PP cells confirms restoration of percent PDX1$^+$ cells in patient-corrected (P-corr) cells to the level of control subject (C) cells. $n = 6$. *E*: Quantification of CHGA$^+$ and INS$^+$ cells shows similar efficiency of endocrine cell induction from patient-corrected cells compared with control cells. $n = 6$. *F*: Scheme of strategy to introduce the patient-specific variant (c.−18C.T) into the *PDX1* 5′-UTR of HUES8 cells expressing inducible Cas9. *G*: ICE analysis of CRISPR editing events with Sanger sequencing of genomic DNA shows 1% of targeted cells have homology-directed repair and the c.−18C>T substitution. *H*: Introduction of the c.−18C>T PDX1 variant in wild-type HUES8 hESC leads to defective induction of PDX1 expression during in vitro pancreas differentiation. Representative plot, $n = 4$. mt, mutant; P, patient; WT, wild type.

leucine is specific to the patient's cellular background or a generalizable phenomenon. We used the iCas9 wild-type hESC line, HUES8, with an inducible Cas9 system under doxycycline control (12) to introduce the −18 C>T variant in *PDX1*. We then performed single cell colony expansion with genomic sequencing of PCR products to identify cells containing the *PDX1* variant (Fig. 4*F* and *G*) in only one of the two alleles. Introducing this variant into a different cellular background produced the same differentiation defect; the number of PDX1$^+$ cells decreased to less than one-third of the number generated from PP-stage cells with wild-type PDX1 cultured with 1 mmol/L leucine (Fig. 4*H*). Thus, the *PDX1* 5′-UTR variant that inhibits differentiation also confers leucine sensitivity, leading to a severe differentiation defect in the transition from DE to PP.

Next, we investigated the mechanism whereby the −18 bp 5′-UTR variant impacts PDX1 expression during iPSC-directed differentiation. Using immunofluorescence, we noted fewer PDX1$^+$ cells in differentiating patient cells than in differentiating control cells. We confirmed lower total PDX1 protein level in patient cells compared with control subject cells by immunoblot analysis (Fig. 5*A*). We then investigated Pdx1 protein level per cell using antibody staining and flow cytometry and determined that the patient's PDX1+ cells express less protein; thus, there is less robust PDX1+ induction in the patient's cells not only at the population level but also at the individual cell level (Fig. 5*B*). We then analyzed the wild-type DNA sequence around the 5′-UTR region where the −18 bp variant resides and found a binding motif for the transcriptional regulator protein NF-YA (which absolutely requires the 5-nt sequence CCAAT) and is part of the NF-Y trimeric protein complex (13,14) (Fig. 5*C*). The NF-Y complex behaves as a pioneer factor that regulates cell state transitions during differentiation by creating accessible chromatin when a cell type–specific master regulator like PDX-1 is required for expression (15). Specifically, the transactivation domain-containing NF-YA subunit binds the CCAAT sequence, whereas NF-YB and NF-YC bind the histone fold, likely contributing the pioneer factor function of the trimeric complex (16,17). Electromobility shift assay revealed that NF-YA does indeed bind to the wild-type DNA sequence in vitro but does not bind to the mutant DNA sequence (CTAAT) (Fig. 5*D*). We hypothesized that the NF-Y complex behaves as an activator of the Pdx1 promoter; hence, loss of NF-YB would decrease NF-YA binding to the wild-type sequence beyond the transcription start site and decrease wild-type *PDX1* transcription. We therefore knocked down NF-YB in the control cells and found that, indeed, PDX1 expression decreased, even with a limited (∼20%) knockdown of NF-YB (Fig. 5*E* and *F*). In patient cells, the effect of the knockdown was allele dependent in the heterozygous state; we saw no decrease in the expression of the mutant *PDX1* allele but decreased expression of the wild-type *PDX1* allele. We then performed flow cytometry and saw

the same effect of partial NF-YB knockdown on PDX1 protein expression (Fig. 5*G*). Collectively, these experiments suggest that transcriptional regulation of *PDX1* is affected in the patient's cells due to the 5′-UTR mutant. The patient-specific C>T alteration within the binding site for NF-Y disrupts NF-YA binding, displacing a transcriptional regulator of *PDX1* and decreasing PDX1 expression.

DISCUSSION

Altered BCAA metabolism plays an important role in the pathogenesis of T2D (11,18), and we have shown that accelerated leucine catabolism, fueling ketone production, is a biochemical hallmark of KPD patients (4). The present data show that exposure to high levels of leucine causes a PP differentiation defect in both this KPD patient's iPSCs and a heterologous ESC line. Leucine hypercatabolism associated with mTOR hyperactivation likely sets the stage for a PP differentiation defect, since multiple translational programs are overactivated by hyperphosphorylation of mTOR (19), including some that inhibit pancreatic endocrine cell development (20). The Bernal-Mizrachi group demonstrated a clear link between leucine concentration and the development of pancreatic progenitors, likely mediated by the mTOR pathway (21,22). In KPD patients, this defective leucine metabolism and mTOR signaling appear to render PP cells sensitive to leucine-mediated inhibition and induce β-cell dysfunction that is usually not severe enough to cause early-onset, complete, irreversible β-cell failure. In this patient, an additional "hit"—a variant in *PDX1* that diminishes PDX1 expression at a critical stage of endocrine cell differentiation and augments sensitivity to the inimical effects of leucine catabolism on PP differentiation—caused him to lapse over time to complete β-cell failure and presentation with DKA.

Diabetic phenotypes associated with other known variants in *PDX1* manifest as either late-onset mild monogenic diabetes if heterozygous (the rare maturity-onset diabetes of the young 4 syndrome [7]) or neonatal diabetes if homozygous (23–25). The heterozygous 5′-UTR variant in this patient had slower and less dramatic effects on β-cell differentiation than those that underlie *PDX1*-associated neonatal diabetes. Variants in β-cell function genes that are not highly deleterious (early stops, splicing defects), or inherited in a heterozygous fashion with some rescue by the normal allele, may permit relatively normal function for many years, manifesting clinical disease only later in life. In contrast, highly deleterious or homozygous variants give rise to disease early in life. Differences in phenotypic severity observed in iPSC disease models in vitro and disease progression in vivo have been noted previously (26). The patient's late onset of non-insulin-requiring diabetes with subsequent progression to β-cell failure resembles the chronology of heterozygous *Pdx1*$^{+/-}$ mice that manifest progressive β-cell loss with
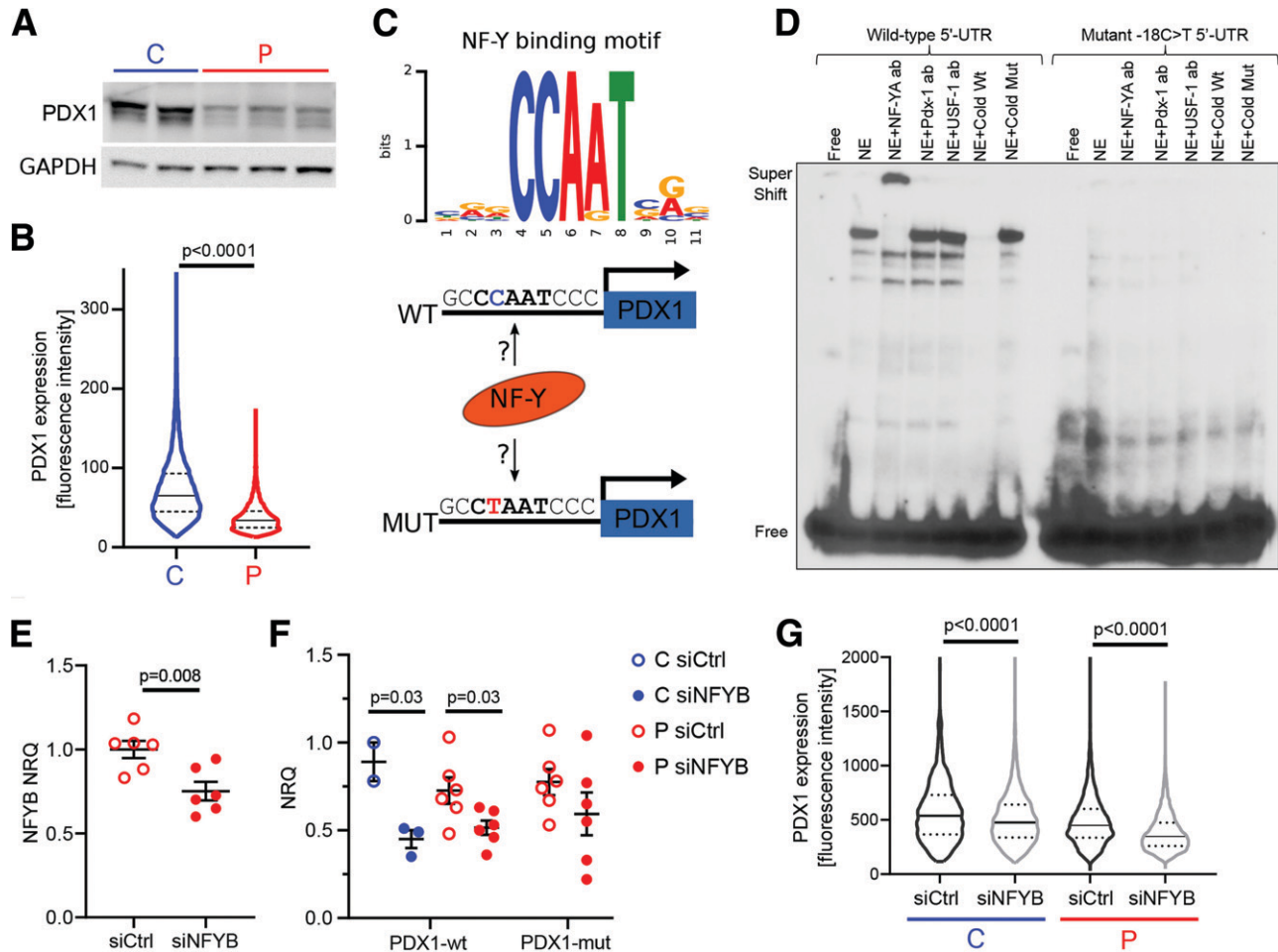
Figure 5—(PDX1)c.−18 C>T mutation alters NF-Y transcription factor binding to the PDX1 promoter. A: Global expression of PDX1 is lower in patient PP cells than in control subject PP cells as assessed by Western blot. $n = 2$ for control subject (C), $n = 3$ for patient (P) biological repeats; blots were performed in two technical repeats. B: Violin plots of PDX1 signal intensities in individual PDX1$^+$ PP cells analyzed by flow cytometry. PDX1 expression per cell is significantly decreased in patient compared with control subject cells. Solid line, median; dotted lines, 1st and 3rd quartiles. $n = 10,041$ for C and 5,679 for P, each with three biological repeats. C: Scheme of NF-Y transcription factor family members (NF-YA and NF-YB) binding motif (factorbook.org) discovered as part of the Encyclopedia of DNA Elements (ENCODE) project, which suggests potential regulation of *PDX1* transcription by NF-Y family members and altered regulation by the c.−18 C>T mutant allele. D: Electrophoretic mobility shift assay shows binding of nuclear extract (NE) to the 5'-UTR sequence of the *PDX1* gene (lane 2). The specific binding protein in the NE is NF-YA, as shown by appearance of a supershift band upon treatment with NF-YA–specific antibody (ab) (lane 3). Neither DNA-NE binding nor the supershift occurs with the mutant 5'-UTR DNA (lanes 9 and 10). The supershift band is not present with negative control anti-PDX1 (lanes 4 and 11) and anti–USF-1 (lanes 5 and 12) antibodies. "Cold" (lanes 6–7 and 13–14) indicates competition with unlabeled oligonucleotides. E: qPCR results showing silencing of the *NFYB* transcript in patient PP cells ($n = 6$) 48 h after anti-NFYB siRNA or control siRNA delivery. NRQ, normalized relative quantity. F: qPCR results showing decrease in *PDX1* wild-type (Wt) allele transcripts in control ($n = 3$) and patient ($n = 6$) PP cells 48 h after anti-NFYB siRNA or control siRNA delivery. The mutant allele (*PDX1*-mut) appears unaffected. G: Violin plots of PDX1 signal intensities in individual PDX1$^+$ PP cells analyzed by flow cytometry. PDX1 expression per cell is significantly decreased in siNFYB compared with siCTRL-treated cells, in both control subject and patient cells, 72 h post–siRNA delivery. $n = 3,875$ and 2,502 for C siNFYB and siCTRL, respectively, and $n = 6,294$ and 2,837 for P siNFYB and siCTRL. Analyzed cells are from three biological repeats.

increased stress during adulthood, resulting in ∼50% reduced β-cell mass at 1 year of age (27–29).

The clinical significance of this variant in the proximal promoter region of *PDX1* is underscored by our finding that its minor allele frequency was 3.8% among African Americans and 1.4% among Hispanics in a cohort of 81 patients with KPD and absent β-cell function (frequencies ≥3× higher than in race/ethnicity-matched control subjects) (30). Identification of a potentially pathogenic var-

iant in a β-cell–regulated gene with detailed multiomics investigation of patient-specific iPSCs differentiated into the pancreatic endocrine lineage is a promising approach to identify the mechanism of β-cell failure in patients with complex, atypical forms of diabetes (31).

References

1. Balasubramanyam A, Nalini R, Hampe CS, Maldonado M. Syndromes of ketosis-prone diabetes mellitus. Endocr Rev 2008;29:292–302

2. Maldonado M, Hampe CS, Gaur LK, et al. Ketosis-prone diabetes: dissection of a heterogeneous syndrome using an immunogenetic and beta-cell functional classification, prospective analysis, and clinical outcomes. J Clin Endocrinol Metab 2003;88:5090–5098

3. Mauvais-Jarvis F, Sobngwi E, Porcher R, et al. Ketosis-prone type 2 diabetes in patients of sub-Saharan African origin: clinical pathophysiology and natural history of beta-cell dysfunction and insulin resistance. Diabetes 2004;53:645–653

4. Patel SG, Hsu JW, Jahoor F, et al. Pathogenesis of $A^-\beta^+$ ketosis-prone diabetes. Diabetes 2013;62:912–922

5. Nalini R, Ozer K, Maldonado M, et al. Presence or absence of a known diabetic ketoacidosis precipitant defines distinct syndromes of "A-β+" ketosis-prone diabetes based on long-term β-cell function, human leukocyte antigen class II alleles, and sex predilection. Metabolism 2010;59:1448–1455

6. Haaland WC, Scaduto DI, Maldonado MR, et al. A−β− subtype of ketosis-prone diabetes is not predominantly a monogenic diabetic syndrome. Diabetes Care 2009;32:873–877

7. Holland AM, Hale MA, Kagami H, Hammer RE, MacDonald RJ. Experimental control of pancreatic development and maintenance. Proc Natl Acad Sci U S A 2002;99:12236–12241

8. Stoffers DA, Ferrer J, Clarke WL, Habener JF. Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. Nat Genet 1997;17:138–139

9. Borowiak M, Melton DA. How to make beta cells? Curr Opin Cell Biol 2009;21:727–732

10. Sneddon JB, Borowiak M, Melton DA. Self-renewal of embryonic-stem-cell-derived progenitors by organ-matched mesenchyme. Nature 2012;491: 765–768

11. Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab 2009;9:311–326

12. González F, Zhu Z, Shi ZD, et al. An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell 2014;15:215–226

13. Dorn A, Bollekens J, Staub A, Benoist C, Mathis D. A multiplicity of CCAAT box-binding proteins. Cell 1987;50:863–872

14. Mantovani R. A survey of 178 NF-Y binding CCAAT boxes. Nucleic Acids Res 1998;26:1135–1143

15. Oldfield AJ, Yang P, Conway AE, et al. Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. Mol Cell 2014;55:708–722

16. Nardini M, Gnesutta N, Donati G, et al. Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. Cell 2013;152:132–143

17. Huber EM, Scharf DH, Hortschansky P, Groll M, Brakhage AA. DNA minor groove sensing and widening by the CCAAT-binding complex. Structure 2012;20:1757–1768

18. Hernández-Alvarez MI, Díaz-Ramos A, Berdasco M, et al. Early-onset and classical forms of type 2 diabetes show impaired expression of genes involved in muscle branched-chain amino acids metabolism. Sci Rep 2017;7:13850

19. Son SM, Park SJ, Lee H, et al. Leucine signals to mTORC1 via its metabolite acetyl-coenzyme A. Cell Metab 2019;29:192–201.e7

20. Rachdi L, Aïello V, Duvillié B, Scharfmann R. L-leucine alters pancreatic β-cell differentiation and function via the mTor signaling pathway. Diabetes 2012;61:409–417

21. Alejandro EU, Bozadjieva N, Blandino-Rosano M, et al. Overexpression of kinase-dead mTOR impairs glucose homeostasis by regulating insulin secretion and not β-cell mass. Diabetes 2017;66:2150–2162

22. Elghazi L, Blandino-Rosano M, Alejandro E, Cras-Méneur C, Bernal-Mizrachi E. Role of nutrients and mTOR signaling in the regulation of pancreatic progenitors development. Mol Metab 2017;6:560–573

23. Nicolino M, Claiborn KC, Senée V, Boland A, Stoffers DA, Julier C. A novel hypomorphic PDX1 mutation responsible for permanent neonatal diabetes with subclinical exocrine deficiency. Diabetes 2010;59:733–740

24. De Franco E, Shaw-Smith C, Flanagan SE, et al. Biallelic PDX1 (insulin promoter factor 1) mutations causing neonatal diabetes without exocrine pancreatic insufficiency. Diabet Med 2013;30:e197–e200

25. Caetano LA, Santana LS, Costa-Riquetto AD, et al. PDX1 -MODY and dorsal pancreatic agenesis: new phenotype of a rare disease. Clin Genet 2018;93:382–386

26. Shi ZD, Lee K, Yang D, et al. Genome editing in hPSCs reveals GATA6 haploinsufficiency and a genetic interaction with GATA4 in human pancreatic development. Cell Stem Cell 2017;20:675–688.e6

27. Ahlgren U, Jonsson J, Edlund H. The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. Development 1996;122:1409–1416

28. Ahlgren U, Jonsson J, Jonsson L, Simu K, Edlund H. beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes. Genes Dev 1998;12:1763–1768

29. Zhu Z, Li QV, Lee K, et al. Genome editing of lineage determinants in human pluripotent stem cells reveals mechanisms of pancreatic development and diabetes. Cell Stem Cell 2016;18:755–768

30. Scaduto D, Ozer K, Nalini R, et al. Characterization of a novel 5′-UTR putative CCAAT-box variant in the human PDX-1 gene promoter within a ketosis-prone diabetes (KPD) cohort (Abstract). Diabetes 2011;60 (Suppl. 1)

31. Cardenas-Diaz FL, Osorio-Quintero C, Diaz-Miranda MA, et al. Modeling monogenic diabetes using human ESCs reveals developmental and metabolic deficiencies caused by mutations in HNF1A. Cell Stem Cell 2019;25:273–289.e5