



Pseudotime Ordering of Single Human β -Cells Reveals States of Insulin Production and Unfolded Protein Response

Yurong Xin, Giselle Dominguez Gutierrez, Haruka Okamoto, Jinrang Kim, Ann-Hwee Lee, Christina Adler, Min Ni, George D. Yancopoulos, Andrew J. Murphy, and Jesper Gromada

Diabetes 2018;67:1783–1794 | <https://doi.org/10.2337/db18-0365>

Proinsulin is a misfolding-prone protein, making its biosynthesis in the endoplasmic reticulum (ER) a stressful event. Pancreatic β -cells overcome ER stress by activating the unfolded protein response (UPR) and reducing insulin production. This suggests that β -cells transition between periods of high insulin biosynthesis and UPR-mediated recovery from cellular stress. We now report the pseudotime ordering of single β -cells from humans without diabetes detected by large-scale RNA sequencing. We identified major states with 1) low UPR and low insulin gene expression, 2) low UPR and high insulin gene expression, or 3) high UPR and low insulin gene expression. The latter state was enriched for proliferating cells. Stressed human β -cells do not dedifferentiate and show little propensity for apoptosis. These data suggest that human β -cells transition between states with high rates of biosynthesis to fulfill the body's insulin requirements to maintain normal blood glucose levels and UPR-mediated recovery from ER stress due to high insulin production.

β -Cells of the endocrine pancreas produce large amounts of insulin, which is secreted in a finely regulated manner to maintain blood glucose levels within a narrow range. Insulin biosynthesis accounts for >10% of total protein production under basal conditions and increases up to 50% in the stimulated state (1,2). Proinsulin is a misfolding-prone protein because up to 20% of initially synthesized proinsulin fails to reach its mature conformation (3–6). Misfolded proinsulin is refolded or degraded. Under conditions of high insulin demand, proinsulin misfolding can exceed the capacity of the β -cells to handle the

misfolded protein load. This puts pressure on the endoplasmic reticulum (ER) and results in accumulation of misfolded proinsulin and cellular stress. The large demand for proinsulin biosynthesis and folding makes β -cells highly susceptible to ER stress (7–9).

β -Cells are metabolically active, relying on oxidative phosphorylation for ATP generation (10). This generates reactive oxygen species (ROS) and can result in oxidative stress. β -Cells have low antioxidant defense, further increasing their susceptibility to stress (11,12). ER stress and oxidative stress can enhance each other, because protein misfolding results in the production of ROS, and these species can perturb the ER redox state and cause damage to nascent proteins (13).

To counteract stress conditions, the β -cells activate a network of signaling pathways termed the unfolded protein response (UPR). This is composed of three parallel pathways that are initiated by the ER transmembrane proteins IRE1, PERK, and ATF6 (14). These regulators trigger a signaling cascade that enhances protein folding activity, reduces ER workload, and promotes clearance of misfolded proteins. Once cellular stress is cleared and homeostasis restored, this stress sensor program is deactivated to prevent harmful UPR hyperactivation that could otherwise lead to apoptosis (14).

β -Cell heterogeneity is well established at the functional level. Proteome and transcriptome studies confirmed the heterogeneous nature of mouse and human β -cells and revealed genes and pathways characterizing these subpopulations (15). β -Cell heterogeneity likely represents dynamic states rather than stable and distinct subpopulations because functional studies have shown that β -cells transition

Regeneron Pharmaceuticals, Tarrytown, NY

Corresponding author: Jesper Gromada, jesper.gromada@regeneron.com.

Received 29 March 2018 and accepted 9 June 2018.

This article contains Supplementary Data online at <http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db18-0365/-/DC1>.

Y.X. and G.D.G. are co-first authors.

© 2018 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <http://www.diabetesjournals.org/content/license>.

between periods of activity and rest. The transitions are not synchronized between β -cells within individual islets (15).

In this study, we applied large-scale RNA sequencing (RNAseq) to single human pancreatic islet cells from donors without diabetes. The β -cells were ordered according to their pseudotime. This was obtained by projecting each cell onto a trajectory, and the ordered sequence of the cells was used to study dynamic changes in gene expression. This provides a higher resolution view of the gene expression landscape due to their dynamic and heterogeneous nature and helps understand the complex biological processes governing β -cell function. We identified distinct states with low UPR and low or high insulin gene expression as well as cells with high UPR activation and low insulin expression. The gene signatures and enriched pathways for each state are described.

RESEARCH DESIGN AND METHODS

Human Islets

Islets from 12 donors without diabetes were obtained from Prodo Laboratories (Supplementary Table 1). Donor information and diabetes status were obtained from the patient's medical record and, if available, the hemoglobin A_{1c}. Islets were digested at 37°C (10 min) using TrypLE Express (Life Technologies). Cells were filtered (30- μ m), centrifuged, and resuspended (1 \times PBS containing 0.04% BSA). Trypan blue staining revealed $91.2 \pm 3.3\%$ ($n = 12$) cell viability.

RNA Fluorescence In Situ Hybridization

Dissociated cells were placed on slides using Cytospin and fixed (10% neutral buffered formalin) for 35 min. Islets were fixed in the same way, embedded in paraffin, and cut (6- μ m sections). Cells and islet sections were permeabilized and hybridized with mRNA probes (*INS* and *DDIT3*, *FTL*, *HSPA5*, or *SQSTM1*), according to the manufacturer's instructions (Advanced Cell Diagnostics). A fluorescent kit was used to amplify the mRNA signal, and fluorescein, Cy3, and Cy5 were detected using a microscope slide scanner (Axio-Scan.Z1; Zeiss). Fluorescence intensities were determined using the HALO image software module for RNA fluorescence in situ hybridization (FISH) analysis (Indica Laboratories).

Single-Cell RNAseq and Read Mapping

Cells were loaded on a Chromium Single Cell Instrument (10x Genomics). RNAseq libraries were prepared using Chromium Single Cell 3' Library, Gel Beads & Multiplex Kit (10x Genomics). Sequencing was performed on the Illumina NextSeq500 using Read-1 for transcript read and Read-2 for three indices, I7 index for cell barcode, I5 index for sample index, and unique molecular identifier (UMI). Cell Ranger Single-Cell Software Suite (v1.1.0; 10x Genomics) was used for sample demultiplexing, alignment, filtering, and UMI counting. Human B37.3 Genome assembly and the University of California, Santa Cruz gene model were used for alignment. Single-cell sequencing data described in this study can be found within the Gene

Expression Omnibus database using accession number GSE114297.

Single-Cell Data Analysis

Cells were removed if the number of detected genes was <500, total number of UMI was <3,000, or viability score was >0.2 (16). Viability score was defined by the ratio between the sum of *MT-RNR2*, *MT-ND1*, *MT-CO1*, *MT-CO2*, *MT-ATP8*, *MT-ATP6*, *MT-CO3*, and *MT-CYB* expression (UMI) and total UMI. A high score indicates low viability. Mclust (R package) was used to assess cell-cell contamination and identify islet endocrine cell types. *DensityMclust* function estimated bimodal distribution of *GCG*, *INS*, *SST*, *PPY*, and *GHRL* expression, namely, high-expression and low-expression modes. Cells with more than one hormone in the high-expression mode were excluded (e.g., *GCG*^{high} and *INS*^{high}). Cells with a single-hormone in the high-expression mode were identified as *GCG*⁺ (α), *INS*⁺ (β), *SST*⁺ (δ), *PPY*⁺ (PP), and *GHRL*⁺ (ϵ) cells. With the retained single-hormone endocrine and nonendocrine cells, expression data were normalized by the total UMI and scaled by a factor of 10,000 at cell level.

Seurat package identified cell clusters, cell-type subpopulations, and cluster-enriched genes. A total of 1,166 variable genes were used for the principal component analysis. Cell clusters were identified using *FindClusters* (19 principal components and 0.8 resolution). The first two t-Distributed Stochastic Neighbor Embedding dimensions were used to visualize cell clusters. Enriched genes for each cluster were identified by *FindMarkers* (>25% cell detection; $P < 0.05$ and log-scale fold change >0.25). β -Cell subpopulation markers were defined as enriched genes in one subpopulation over the rest of the subpopulations, as described above. Enriched endocrine genes were obtained by comparing endocrine cells (α -, β -, δ -, PP-, and ϵ -cells) with the nonendocrine cells, as described above.

The alignment analysis of human islet cells between this study and three previous studies (17–19) was performed by Seurat using canonical correlation analysis. The union of the top 2,000 variable genes of this study and one published islet study were used to correlate and integrate data from the two studies. Common cell types and subpopulations were aligned between the two studies and visualized in two t-Distributed Stochastic Neighbor Embedding dimensions.

Pseudotime Trajectory Reconstruction

β -Cell subpopulation markers were used by monocle v2.4 (R package) to construct single-cell pseudotime ordering using the default setting. The β -cell trajectory includes two branch points, and branch-dependent significant genes were identified by the *BEAM* function in monocle. Three sets of genes were derived to capture differential expression between the following branches: *INS*^{lo}UPR^{hi} and *INS*^{hi}UPR^{lo}, *INS*^{lo}UPR^{lo} and *INS*^{hi}UPR^{lo}, and *INS*^{lo}UPR^{hi} and *INS*^{lo}UPR^{lo}. Significant genes were defined ($q < 1^{-10}$). Cells from two donors were excluded in the pseudotime

analysis, because an initial analysis including these cells identified a skewed root state formed almost exclusively by these cells (Supplementary Fig. 1A and B). Cells from 10 other donors exhibited relatively uniform distribution in the trajectory (Supplementary Fig. 1C).

Biological Process Score Calculation

Gene sets of UPR, apoptosis, senescence, and cell cycle were obtained from IPA Ingenuity, SABiosciences (<https://www.qiagen.com/us/search/rt2-profiler-pcr-arrays/>) and Dominguez et al. (20) (Supplementary Table 2). A score for each process was average of scaled UMI of all genes in the gene set. Score distribution was estimated by random selection of the same number of genes for the specific gene set with 1,000 iterations. The empirical *P* value was calculated against the distribution of each score.

Pathway Enrichment

Cell markers and branch-dependent genes were analyzed for pathway enrichment with GO, KEGG, and REACTOME by clusterProfiler (R package). Top enriched gene sets were selected based on the *P* value.

RESULTS

Human β -Cell Subpopulations

Islets isolated from 12 donors without diabetes underwent single-cell RNAseq analysis (Supplementary Table 1). β -Cells were identified based on their *INS* expression, as detailed in the RESEARCH DESIGN AND METHODS. Clustering of the β -cells according to their transcriptome profiles revealed four subpopulations (Fig. 1A). Three of the subpopulations were similar to each other with only a small number of uniquely enriched genes in each subpopulation: 18 in subpopulation 1, 33 in subpopulation 2, and 18 in subpopulation 3 (Supplementary Table 3). The fourth subpopulation (5.1% of the β -cells; $n = 363$) was more distinct, with 431 enriched genes. In total, 488 genes were enriched in the subpopulations (Supplementary Table 3). Pathway analysis revealed that genes involved in protein folding and ER stress response are highly enriched in the fourth subpopulation (Fig. 1B). Three other single-cell RNAseq studies (17–19) have also identified clusters of human β -cells with an enriched UPR signature (Supplementary Fig. 2). To exclude the possibility that subpopulation 4 represents artificially stressed cells that arose during the single-cell isolation process, we performed RNA FISH on intact islets and dissociated single cells for *DDIT3*, *FTL*, *HSPA5*, and *SQSTM1* that were highly enriched in this subpopulation. We identified subsets of *INS*-positive cells with high expression of these genes at similar frequencies in the intact islet and dissociated single cells (Supplementary Fig. 3). This suggests that induction of the stress response and UPR program is unlikely to originate from the single-cell dissociation process.

In addition to the β -cells described above, we identified 310 cells (4.4%) that clustered with them but express low *INS* in all four subpopulations (17% on average of β -cells)

(Fig. 1C). The existence of low *INS*-expressing cells (5.6% [$n = 34,575$]) was confirmed by RNA FISH (Fig. 1D and E). The *INS* mRNA level was lower in the fourth subpopulation ($P = 3.1 \times 10^{-58}$), because 30% of the cells are low *INS*-expressing cells. Conversely, *INS* expression was higher in the third subpopulation ($P = 2.4 \times 10^{-12}$) compared with the rest of the β -cells (Fig. 1C and Supplementary Table 3).

We investigated the differentiation state of the β -cells and found that highly enriched endocrine genes were equally expressed in subpopulation 4 and the other β -cell populations (Fig. 1F). Collectively, these data show that human β -cells segregate into four subpopulations. The cells are fully differentiated, but a subpopulation is characterized by lower *INS* expression and high levels of UPR-related genes.

Pseudotime Ordering of Human β -Cells

Although cell clustering is useful to identify subtypes, reconstructing cell states in continuous processes is difficult. We therefore used a trajectory analysis to derive pseudotime of the β -cells. To guide construction of the trajectory, we used the 488 subpopulation markers obtained from unbiased clustering. The trajectory constitutes two decision points and five states named by their key features: 1) root ($n = 1,079$), 2) average *INS* ($n = 317$), 3) high *INS*-low UPR ($INS^{hi}UPR^{lo}$; $n = 2,399$), 4) low *INS*-low UPR ($INS^{lo}UPR^{lo}$; $n = 1,028$), and 5) low *INS*-high UPR ($INS^{lo}UPR^{hi}$; $n = 1,418$) (Fig. 2A and B). Expression analysis identified genes whose expressions are different between the branches at decision points 1 and 2 (Supplementary Table 4). Figure 2C–E shows heat maps of the scaled expression of all differentially expressed genes for each branch and pathways with enriched gene sets (Supplementary Tables 5–7). Branch-dependent transcription factor expression and the function of key pathways are described below.

Transcription Factor Expression in Human β -Cell Pseudotime States

We detected 1,316 transcription factors (1,637 annotated) in human β -cells. Figure 3 shows transcription factors with differential expression in at least one branch, and their associated gene cluster names are shown in Fig. 2C–E. We detected transcription factors associated with ER stress and UPR regulation (*ATF3*, *ATF4*, *DDIT3*, *XBP1*, and *CREB3*) (14,21,22), important for β -cell maturation and insulin expression (*ISL1*, *PDX1*, *MAFA*, *MAFB*, *NEUROD1*, *NKX2-2*, and *SIX3*) (23–28), ribosomal biogenesis (*GTF3A*), and mitochondrial biogenesis (*TFB2M*) (29,30). We also found that *NFE2L2*, which is better known as NRF2 and a master regulator of the antioxidant response, is differentially expressed between the $INS^{lo}UPR^{lo}$ and $INS^{hi}UPR^{lo}$ branches (31). Interestingly, NRF2 protects β -cells from oxidative stress, lipotoxicity, and DNA damage (31,32).

Insulin Biosynthesis and Stress Recovery Defines Human β -Cell States

INS expression was significantly different between the three pseudotime branches: $q = 6.6 \times 10^{-117}$ between

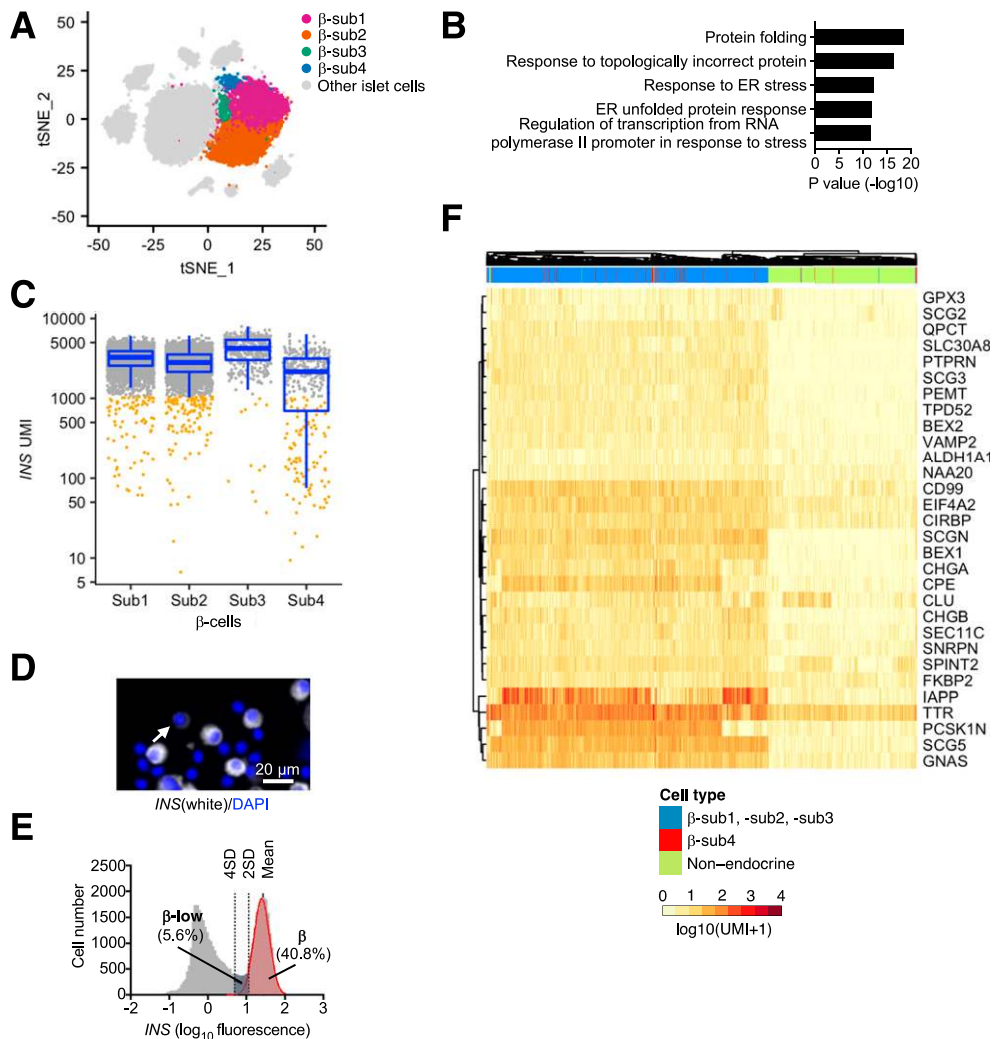


Figure 1—Human β -cell subpopulations. **A**: Human β -cells segregated into four subpopulations (β -sub1, β -sub2, β -sub3, and β -sub4). Other human islet cells are shown in gray. **B**: Top 5 pathways enriched for β -sub4 subpopulation-enriched genes ($P < 1^{-10}$). **C**: Boxplot of *INS* expression in the β -cell subpopulations. Each gray dot represents an individual β -cell. Each orange dot represents an individual low *INS*-expressing cell, with a total of 310 cells (β -sub1 = 81, β -sub2 = 116, β -sub3 = 12, and β -sub4 = 101). The difference of *INS* expression in each β -cell subpopulation was tested by comparing one subpopulation with the other three. *INS* had significant higher expression in β -sub3 ($P = 2.4 \times 10^{-12}$) and significantly lower expression in β -sub4 ($P = 3.1 \times 10^{-58}$). **D**: Representative RNA-FISH image of *INS* (white) and DAPI (blue). Cells with low levels of *INS* are indicated with a white arrow. **E**: Histogram of *INS* \log_{10} fluorescence intensity determined by RNA-FISH analysis. Two thresholds (mean 2^*SD and mean 4^*SD) were used to define β -cells with normal (β) or low *INS* expression (β -low). **F**: Heat map of highly expressed endocrine cell markers showed no differences in expression between the four β -cell subpopulations. Nonendocrine cells are shown for reference.

$INS^{hi}UPR^{lo}$ and $INS^{lo}UPR^{hi}$, 2.6×10^{-60} between $INS^{hi}UPR^{lo}$ and $INS^{lo}UPR^{lo}$, and 1.9×10^{-28} between $INS^{lo}UPR^{lo}$ and $INS^{lo}UPR^{hi}$ (Supplementary Table 4). The $INS^{hi}UPR^{lo}$ state constitutes 38% of all pseudotime-ordered β -cells. Within this state, *INS* expression was higher than in the other states and increased from 16 to 67% along the branch. At the tip of the branch, we detected extreme expressing cells with up to twofold higher *INS* expression than the average of the state (Fig. 4A). An important event to consider in single-cell studies is the possible capture of doublet cells. Such an event would be expected to have a higher distribution of gene number and total UMI compared with singlets. Supplementary Fig. 4A and B shows that β -cells in the $INS^{hi}UPR^{lo}$ do not have

such higher gene number distribution compared with cells in the rest of the states. Therefore, the extreme *INS* expression is unlikely to reflect the capture of doublets. Among the established activators of insulin biosynthesis, we found *MAFA* was significantly expressed between the branches ($q = 5.1 \times 10^{-14}$ between $INS^{lo}UPR^{lo}$ and $INS^{lo}UPR^{hi}$ and 2.3×10^{-13} between $INS^{hi}UPR^{lo}$ and $INS^{lo}UPR^{hi}$) (Fig. 3A and B and Fig. 4B). The *MAFA* expression pattern confirms its importance for regulation of *INS* expression.

Genes used to reflect UPR activity were significantly expressed along pseudotime branches, including *ATF4*, *CALR*, *DDIT3*, *EIF2A*, *HSP90B1*, *HSPA5*, *HSPH1*, *NFE2L2*, *PPP1R15A*, *VCP*, and *XPB1* (Supplementary Table 4). To

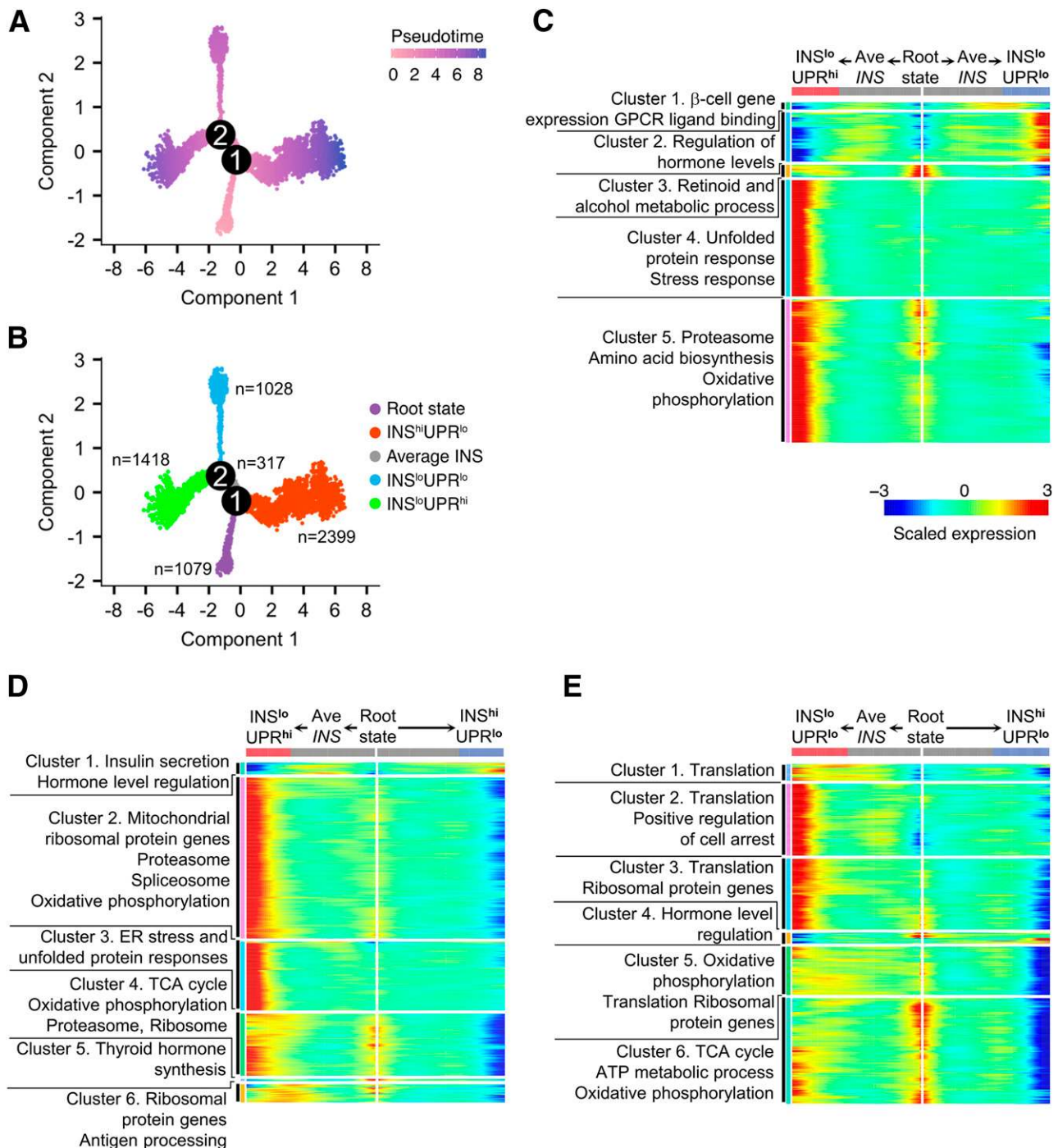
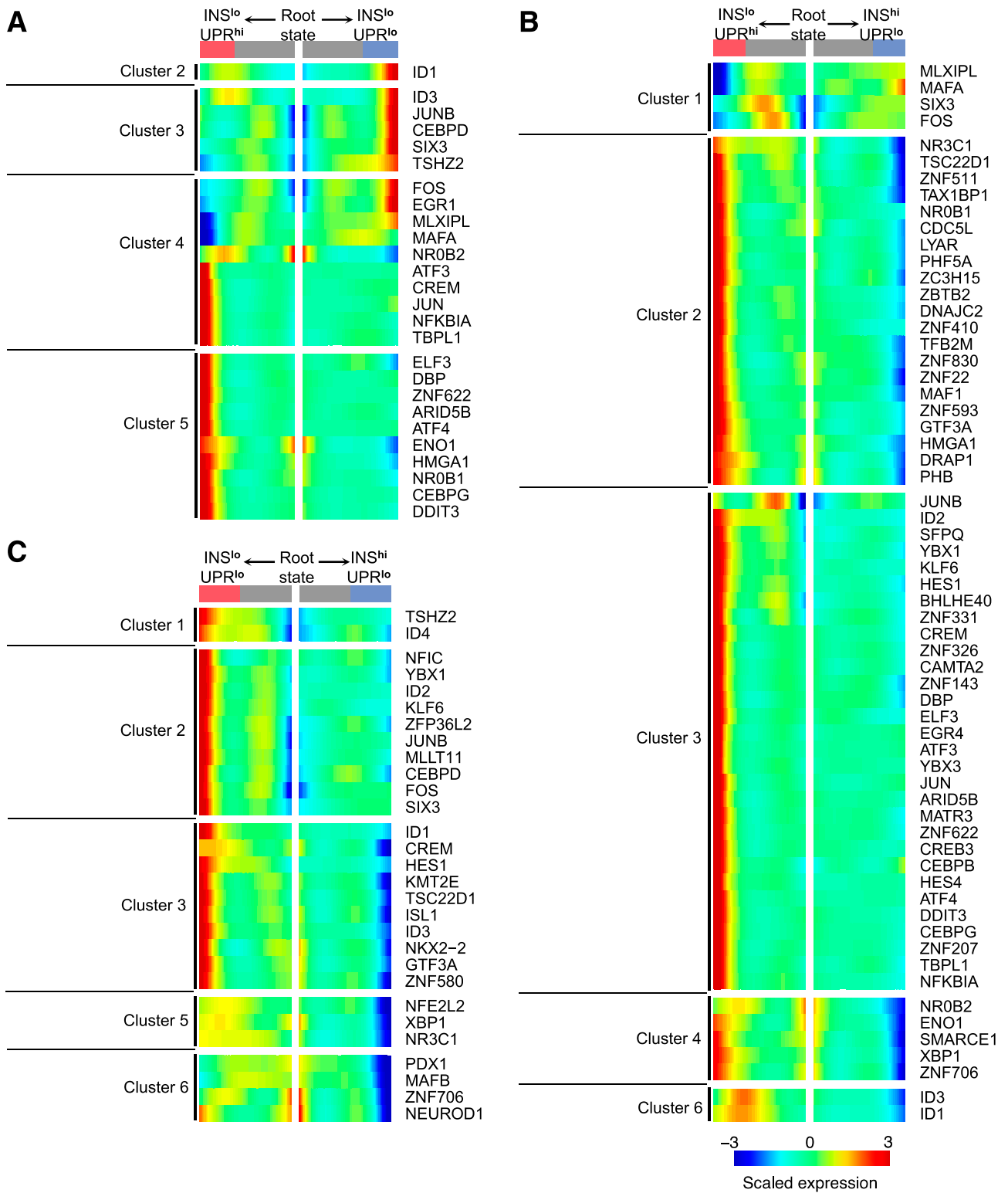


Figure 2—Pseudotime analysis identifies five β -cell states. **A:** Pseudotime trajectory was reconstructed in the 6,241 β -cells, which contains two branch points. Cells are highlighted by pseudotime ranging from 0 to 8.9. **B:** States within the pseudotime trajectory are emphasized by different colors. The number of cells per state are indicated. Analysis of branch-dependent genes is presented for $INS^{lo}UPR^{hi}$ and $INS^{lo}UPR^{lo}$ branches (**C**), $INS^{hi}UPR^{lo}$ and $INS^{lo}UPR^{hi}$ branches (**D**), and $INS^{hi}UPR^{lo}$ and $INS^{lo}UPR^{lo}$ branches (**E**). Each heat map presents genes differentially expressed between two branch comparisons, and each row represents the expression level of a gene along the branch trajectory. Enriched pathways are summarized for each gene cluster. GPCR, G-protein-coupled receptor; TCA, tricarboxylic acid.

investigate how UPR activation was related to *INS* expression, we plotted a UPR score along the pseudotime. Surprisingly, the UPR score revealed an inverse expression pattern to *INS* along the branches with extremes at the tip of the states (Fig. 4C). This was

unexpected, given that insulin is the major secretory load in the β -cells that may cause ER stress and UPR activation. However, UPR is a protective mechanism of the cell to cope with stress, and we hypothesize that the $INS^{lo}UPR^{hi}$ state reflects recovery from stress coupled



Downloaded from <http://diabetesjournals.org/diabetes/article-pdf/67/9/1783/541281/db180365.pdf> by guest on 27 August 2022

Figure 3—Transcription factors significantly expressed in each pseudotime branch. Transcription factors differentially expressed in each gene cluster are presented for INS^{lo}UPR^{hi} and INS^{lo}UPR^{lo} branches (A), INS^{hi}UPR^{lo} and INS^{lo}UPR^{hi} branches (B), and INS^{hi}UPR^{lo} and INS^{lo}UPR^{lo} branches (C). The gene cluster names are consistent with those shown in Fig. 2C–E.

with downregulation of insulin expression. A comparable stress-coping mechanism of UPR activation leading to downregulation of insulin gene expression has been shown in mouse β -cells (33).

Consistent with this notion, we failed to observe upregulation of apoptosis markers in the INS^{lo}UPR^{hi} state, arguing against the presence of chronic stress with negative consequences for the β -cells. Overall, apoptosis was

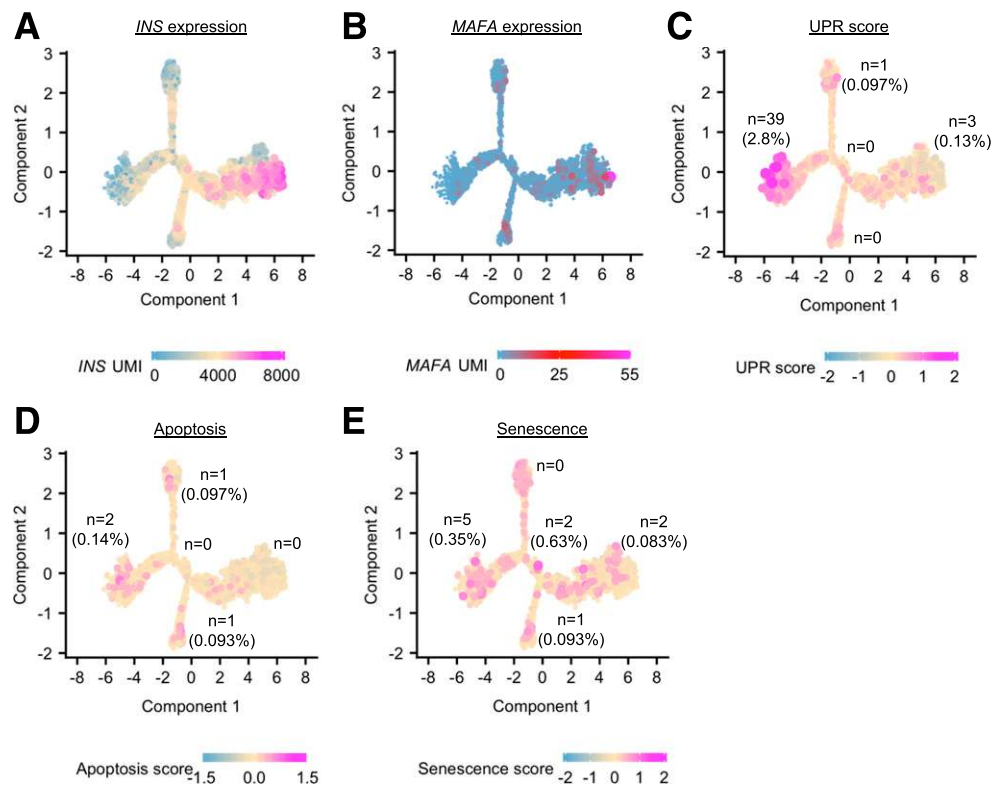


Figure 4—Pseudotime of human β -cells reveals dynamic states of *INS* expression and stress recovery. Expression pattern in pseudotime ordering of *INS* (A) and *MAFA* (B). Composite of factors important for each of the following programs were calculated into a score and its value plotted in pseudotime ordering: UPR (C), apoptosis (D), and senescence (E). Each dot represents a cell colored by the level of composite score.

negligible, with only two β -cells in the $INS^{lo}UPR^{hi}$ state and one cell in the $UPR^{lo}INS^{lo}$ and root states exhibiting high scores (Fig. 4D). Viability of the cells was also assessed using a score as previously described (16) and further supported the previous findings (Supplementary Fig. 4C). Lastly, we explored the possibility that the pseudotime states reflect the β -cell's age. No evidence was found to support this because cellular senescence was not increased overall (Fig. 4E). Collectively, the data suggest that β -cells undergo cycles of insulin biosynthesis and stress recovery.

Stress Response in Human β -Cells

β -Cells in the $INS^{lo}UPR^{hi}$ state showed activation of the antioxidant defense programs. We detected increased expression of superoxide dismutase isoenzymes, whose function is to catalyze the conversion of superoxide anions into hydrogen peroxide (34). In particular, *SOD1* and *SOD2* were enriched in the $INS^{lo}UPR^{hi}$ state. Antioxidant system genes involved in glutathione metabolism (*CD44*, *GCLM*, *GSTP1*, *GLRX2*, *GLRX3*, *GPX1*, *GPX2*, *GPX3*, *GPX4*, and *SLC3A2*), thioredoxin metabolism (*PRDX1*, *PRDX2*, *PRDX6*, *TXN*, *TXNL1*, and *TXNRD1*), quinone detoxification (*NQO1*), and iron storage (*FTH1*, *FTL*, and *PCBP1*) were also enriched in the $INS^{lo}UPR^{hi}$ state and protect against ROS (Fig. 5) (34–37).

As mentioned above, NRF2 (*NFE2L2*) is a master regulator of the antioxidant response and controls the

expression of the superoxide dismutase isoenzymes and genes involved in glutathione and thioredoxin production and utilization as well as iron detoxification and storage (31). Factors promoting the activation or stabilization of NRF2 (*SQSTM1*, *PARK7*, *CDKN1A*, and *MANF*) were enriched in the $INS^{lo}UPR^{hi}$ state (38–41). In addition, expression of genes encoding the inhibitor of differentiation proteins (*ID1-3*) showed progressive upregulation in the $INS^{lo}UPR^{hi}$ state (Fig. 5B). *ID1-3* are important antioxidant response factors crucial for β -cell survival under stress and are associated with positive regulation of the NRF2-interacting proteins, called small MAF proteins (42). Thus, β -cells cope with stress by reducing insulin expression and activating NRF2-dependent antioxidant defense mechanisms.

$INS^{lo}UPR^{hi}$ State Promotes Human β -Cell Proliferation

Mouse β -cells proliferate under conditions of UPR activation and low insulin production (43,44). Because these conditions are hallmarks of β -cells in the $INS^{lo}UPR^{hi}$ state, we investigated the existence and location of proliferating β -cells using a proliferation score consisting of 67 cell-cycle genes (Supplementary Table 2). Interestingly, most proliferating human β -cells were found at the tip of the $INS^{lo}UPR^{hi}$ state, with only few cells in the $INS^{hi}UPR^{lo}$ and root states. More proliferating cells were

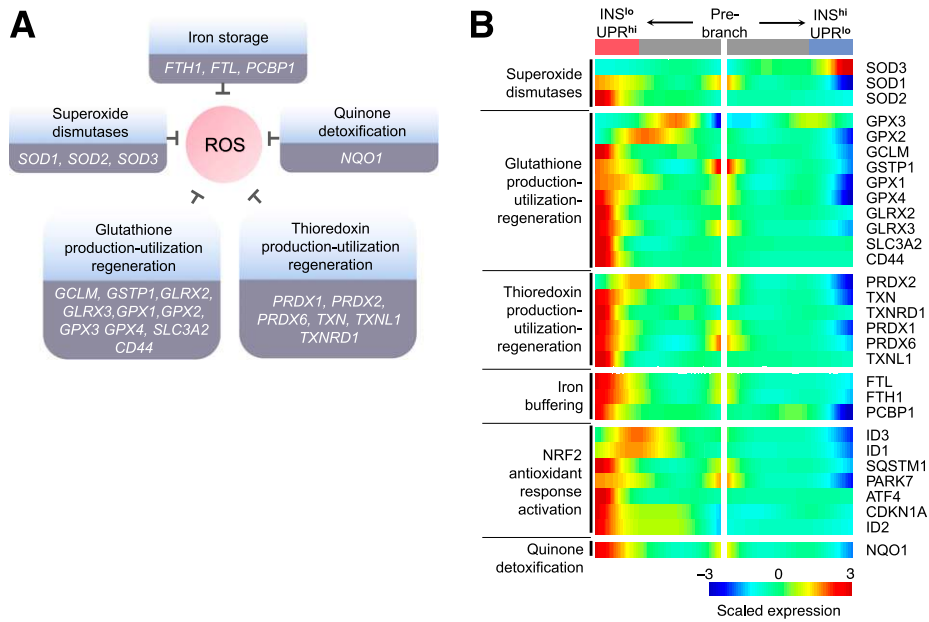


Figure 5—During stress recovery, human β -cells activate an antioxidant program. *A*: Schematic summarizing the different pathways contributing to cell protection from ROS. The first is iron sequestration by ferritin light and heavy chain (*FTL*, *FTH*) and poly(rC)-binding protein 1 (*PCBP1*). The second is superoxide dismutases (*SOD1*, *SOD2*, *SOD3*). The third is glutathione production, utilization, and regeneration, which includes the glutamate–cysteine ligase complex modifier subunit (*GCLM*), glutathione *S*-transferase (*GSTP1*), glutaredoxins (*GLRX2*, *GLRX3*), glutathione peroxidases (*GPX1*, *GPX2*, *GPX3*, *GPX4*), and the system Xct structural component and stabilizer unit (*SLC3A2*, *CD44*). The fourth is composed of thioredoxin production, utilization, and regeneration, which is regulated by thioredoxin (*TXN*), thioredoxin reductase 1 (*TXNRD1*), thioredoxin like-1 (*TXNL1*), and peroxiredoxins (*PRDX1*, *PRDX2*, *PRDX6*). The fifth is quinone detoxification by NAD(P)H quinone dehydrogenase 1 (*NQO1*). *B*: Expression heat map showing significantly expressed genes between $INS^{lo}UPR^{hi}$ and $INS^{hi}UPR^{lo}$ branches, which are presented in *A*.

in the G1S than in the G2M cell cycle phase (Fig. 6A and B). Among UPR activation and *INS* expression, we found low *INS* expression was associated with proliferation (Fig. 6C).

NROB1 and *ZNF143* were among the enriched transcription factors in this state and play important roles for embryonic stem cell self-renewal and proliferation (Fig. 3A and B). *ZNF143* is a positive regulator of multiple cell-cycle genes and is a component of a transcriptional network that regulates cell proliferation (45,46). Furthermore, *HES1*

expression was increased in the stressed β -cells and regulates proliferation in progenitor cells through inhibition of *CDKN1B* (Fig. 3B) (47). *HES4* was also enriched in this state and is associated with maintenance of stem cell features (48). Lastly, *LYAR* expression was higher in the $INS^{lo}UPR^{hi}$ state and is involved in the regulation of mouse β -cell proliferation (Fig. 3B) (49). Together, these results extend the observations in mice that low insulin expression but not UPR activation signals for human β -cells to proliferate.

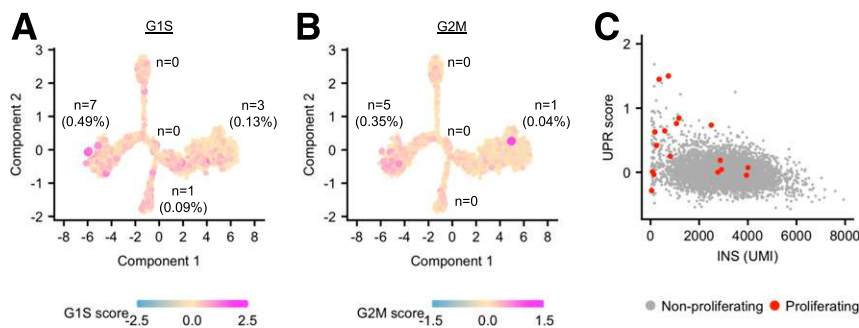


Figure 6—Human β -cells proliferate more in a state of low *INS* expression. Scores of composites of cell cycle regulatory genes were calculated and used to identify cells in G1S phase (*A*) and G2M phase (*B*). Each dot represents a cell colored by the score of the appropriate gene set. Percentage of cells for each cell cycle phase relative to total cells for the state is shown. *C*: β -Cells from all pseudotime states were plotted against their UPR score and *INS* expression (UMI). Proliferating β -cells, which are highlighted in red, were identified by empirical $P < 0.001$ for the G1S or G2M cell cycle scores.

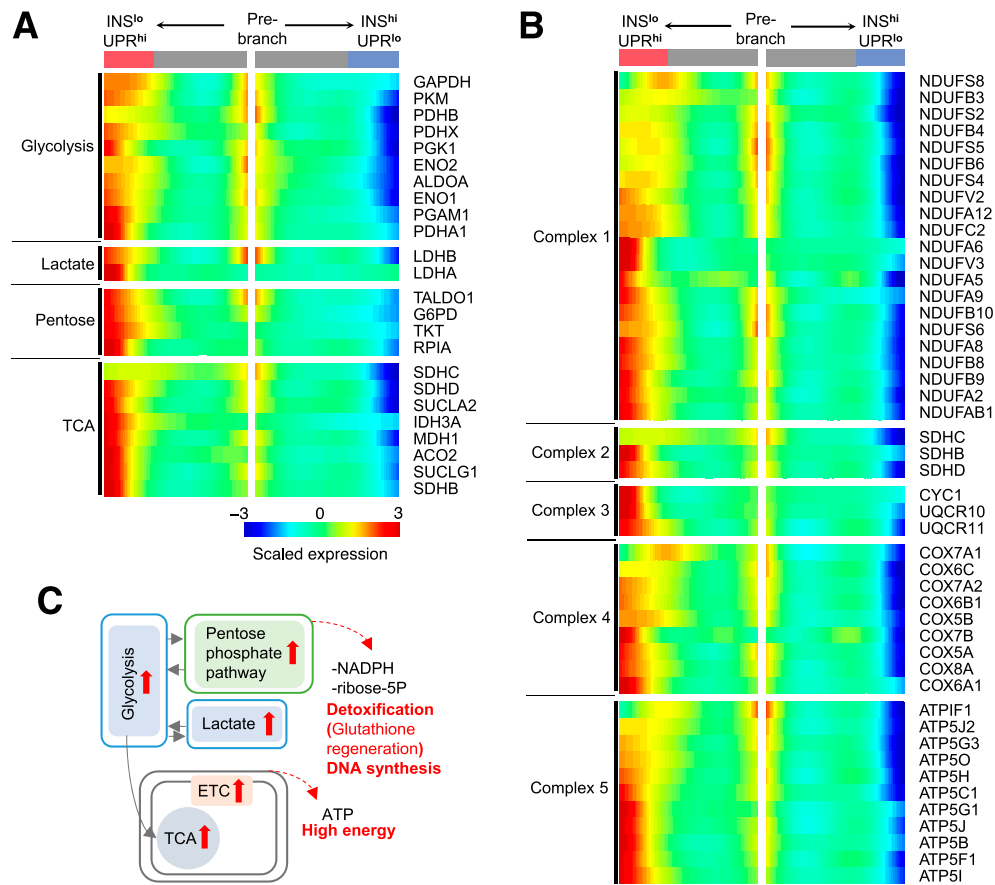


Figure 7—Stress recovery increases metabolic demand in human β -cells. **A**: Heat map of significantly expressed genes between the INS^{hi}UPR^{lo} and INS^{lo}UPR^{hi} branches. The genes are involved in glycolysis, lactate production, pentose phosphate pathway, or tricarboxylic acid (TCA) cycle. **B**: Complexes of electron transport chain (ETC) are presented as individual clusters. **C**: Working model summarizing the findings presented in **A** and **B**.

Human β -Cell Stress Recovery Is Metabolically Demanding

A surprising observation is that β -cells in the INS^{lo}UPR^{hi} state had high expression of genes involved in energy production. In particular, we found upregulation of genes in the glycolytic pathway, tricarboxylic acid cycle, and the electron transport chain (Fig. 7). In addition, β -cells in the INS^{lo}UPR^{hi} state showed upregulation of genes (*G6PD*, *RPIA*, *TALDO*, and *TKT*) encoding enzymes in the pentose phosphate pathway (Fig. 7A). This pathway provides ribose-5-phosphate and NADPH for nucleic acid synthesis. NADPH is used as a reducing agent in the synthetic steps of fatty acids and steroid hormones along with several detoxification systems; for example, NADPH is the limiting substrate for glutathione reductase (50). More importantly, *G6PD* is the rate-limiting enzyme in the pentose phosphate pathway and increases nucleotide production favoring cell survival and proliferation (51,52). Mice deficient in *G6PD* are glucose intolerant and have smaller islets (51). These data suggest that *G6PD* might play a role in human β -cell growth and survival. Collectively, the data suggest that β -cells in the INS^{lo}UPR^{hi} state use more energy than cells undertaking higher levels of *INS* expression. The

shift toward a more active pentose phosphate pathway is likely advantageous for stress recovery and in providing building blocks for cell proliferation.

DISCUSSION

In this study, we detected heterogeneity among human β -cells. Pseudotime analysis ordered the β -cells into five states with varying degrees of insulin gene expression and UPR activation. Unexpectedly, we found that UPR was activated in a subset of β -cells expressing low levels of *INS* (INS^{lo}UPR^{hi}) and appear to represent a state of recovery from stress. The high levels of *INS* expression observed in the INS^{hi}UPR^{lo} state are likely to represent a state of active production. Although we did not observe many changes in the expression of factors known to participate in its biosynthesis, it is possible that these are regulated at the posttranslational level or are not rate limiting in this setting. We found little evidence for apoptosis and dedifferentiation among the β -cells in the different states. Markers of β -cell dedifferentiation, including *NEUROG3*, *OCT4*, and *NANOG*, were barely detected or absent under these nondiabetic conditions (53). Interestingly, most of the β -cells with a high proliferation score were found in

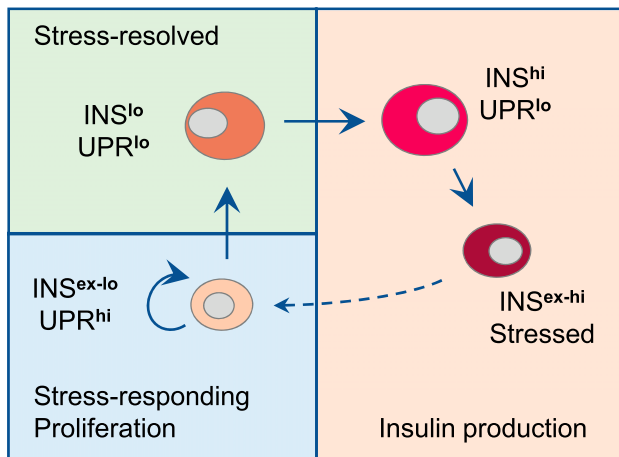


Figure 8—Human β -cells traverse through dynamic states to meet insulin demands. Model showing that β -cells undergoing extremely high insulin biosynthesis (INS^{ex-hi}) are likely to become stressed and retire for a period of recovery ($INS^{ex-lo}UPR^{hi}$) that entails UPR activation and low INS expression. This process is followed by a transitioning state where INS is still low and UPR activity is decreased ($INS^{lo}UPR^{lo}$); β -cells are then ready to resume high insulin expression ($INS^{hi}UPR^{lo}$). Under these states, proliferation was preferentially found under a state of low INS expression and high UPR activation.

the $INS^{lo}UPR^{hi}$ state. This is similar to the situation in mouse β -cells (43,44). We also found that β -cells in the $INS^{lo}UPR^{hi}$ state have higher expression of genes involved in energy metabolism than cells in the other states. Based on these data, we postulate that human β -cells transition between states of high insulin production, which is likely to cause cellular stress due to the high propensity of pro-insulin to misfold, and states of low insulin biosynthesis and increased UPR-mediated stress recovery (Fig. 8). Oscillation between periods of activity and recovery is supported by studies demonstrating functional heterogeneity among β -cells (15).

Humans are endowed with a large number of pancreatic islets. Each islet contains $\sim 50\%$ β -cells, and each cell expresses thousands of insulin-containing granules (54,55). Only a small fraction of the insulin secretory granules is typically released to maintain normal blood glucose levels (55). This suggests spare β -cell capacity and is supported by partial pancreatectomy studies demonstrating that humans only start to develop impaired glucose tolerance after removal of approximately half of their pancreas and, consequently, half of the β -cell mass (56,57). Assuming that the remaining β -cells oscillate between periods of high insulin biosynthesis (activity) and UPR-mediated stress recovery (rest), the partial pancreatectomy studies suggest that less than half of the β -cells are active at any given time. Furthermore, β -cells in the active state have different glucose thresholds for stimulation of insulin secretion (15). Collectively, these data show that β -cells undergo periods of activity and rest, which requires spare capacity. A snapshot of the dynamic processes that β -cells undergo could be interpreted as

static conditions rather than transition states. This is an important consideration when single-cell RNAseq data are analyzed. Therefore, applying pseudotime analysis allowed us to follow the progression of gene expression changes leading to each of the five interconnected β -cell states. We postulate that β -cells transition between states of activity with high insulin expression and recovery. We provide unprecedented insights into the human β -cell transcriptome and the gene changes associated with these functional states.

MAFA is a transcription factor regulating insulin gene transcription (25) and was one of few genes with increased expression in the $INS^{hi}UPR^{lo}$ state. Assuming active insulin secretion in this state, the lack of expression changes in genes encoding the β -cell stimulus-secretion coupling and secretory machinery indicate that they are not rate limiting at the transcriptional level for the secretion of the newly formed insulin. On the contrary, we observed many genes with increased expression in the $INS^{lo}UPR^{hi}$ state. In particular, we found genes involved in reducing cellular stress as well as in providing energy and substrates for the stress recovery processes. We also detected increased rates of β -cell proliferation in the $INS^{lo}UPR^{hi}$ state. Interestingly, our data do not support a correlation between proliferation and UPR activity but rather with low insulin gene expression. We did not detect dedifferentiated human β -cells, even among the lowest insulin expressing cells. This excludes the possibility that dedifferentiation is a trigger of β -cell proliferation.

Pseudotime analysis of single human β -cell RNAseq data from patients with type 2 diabetes could provide valuable information on shifts in the time that β -cells spend in the different states and the associated changes in gene expression. It is tempting to speculate that during chronic high blood glucose and insulin demand, β -cells spend a larger proportion of their time in the high insulin gene expression state and less time in the recovery state. This could lead to chronic stress and eventually cell death. Indeed, patients with type 2 diabetes have 20–65% reduced β -cell mass (58). It would also be interesting to investigate the effects of diabetes medications on the distribution of human β -cells between the states to predict the potential for maintenance of good glycemic control for extended periods of time or risk for β -cell exhaustion and failure.

The identification of functional states is not unique to the β -cells among the pancreatic islet cells. In this sequencing effort, we found human α -cells segregated into two closely related clusters with similar *GCG* expression and UPR scores and a subpopulation with strong expression of cell cycle genes and low UPR score. The reason why human β -cells but not human α -cells become stressed and need to transition between states of activity and rest is an important subject for future investigations.

Acknowledgments. The authors thank Samantha Intriligator (Regeneron Pharmaceuticals) for her help with preparing the manuscript.

Funding. The studies were funded by Regeneron Pharmaceuticals.

Duality of Interest. All authors are employees and shareholders of Regeneron Pharmaceuticals. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. Y.X., G.D.G., H.O., J.K., A.-H.L., and J.G. analyzed the data. Y.X., G.D.G., H.O., J.K., and J.G. designed the studies. Y.X., G.D.G., A.-H.L., G.D.Y., A.J.M., and J.G. wrote the manuscript. G.D.G., J.K., C.A., and M.N. conducted the studies. Y.X. and J.G. are the guarantors of this work and, as such, had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

- Scheuner D, Kaufman RJ. The unfolded protein response: a pathway that links insulin demand with beta-cell failure and diabetes. *Endocr Rev* 2008;29:317–333
- Schuit FC, In't Veld PA, Pipeleers DG. Glucose stimulates proinsulin biosynthesis by a dose-dependent recruitment of pancreatic beta cells. *Proc Natl Acad Sci U S A* 1988;85:3865–3869
- Liu M, Haataja L, Wright J, et al. Mutant INS-gene induced diabetes of youth: proinsulin cysteine residues impose dominant-negative inhibition on wild-type proinsulin transport. *PLoS One* 2010;5:e13333
- Liu M, Li Y, Cavener D, Arvan P. Proinsulin disulfide maturation and misfolding in the endoplasmic reticulum. *J Biol Chem* 2005;280:13209–13212
- Wang J, Chen Y, Yuan Q, Tang W, Zhang X, Osei K. Control of precursor maturation and disposal is an early regulative mechanism in the normal insulin production of pancreatic β -cells. *PLoS One* 2011;6:e19446
- Wang J, Osei K. Proinsulin maturation disorder is a contributor to the defect of subsequent conversion to insulin in β -cells. *Biochem Biophys Res Commun* 2011;411:150–155
- Eizirik DL, Cardozo AK, Cnop M. The role for endoplasmic reticulum stress in diabetes mellitus. *Endocr Rev* 2008;29:42–61
- Papa FR. Endoplasmic reticulum stress, pancreatic β -cell degeneration, and diabetes. *Cold Spring Harb Perspect Med* 2012;2:a007666
- Vetere A, Choudhary A, Burns SM, Wagner BK. Targeting the pancreatic β -cell to treat diabetes. *Nat Rev Drug Discov* 2014;13:278–289
- Schuit F, De Vos A, Farfari S, et al. Metabolic fate of glucose in purified islet cells. Glucose-regulated anaplerosis in beta cells. *J Biol Chem* 1997;272:18572–18579
- Lenzen S, Drinkgern J, Tiedge M. Low antioxidant enzyme gene expression in pancreatic islets compared with various other mouse tissues. *Free Radic Biol Med* 1996;20:463–466
- Tiedge M, Lortz S, Drinkgern J, Lenzen S. Relation between antioxidant enzyme gene expression and antioxidative defense status of insulin-producing cells. *Diabetes* 1997;46:1733–1742
- Hasnain SZ, Prins JB, McGuckin MA. Oxidative and endoplasmic reticulum stress in β -cell dysfunction in diabetes. *J Mol Endocrinol* 2016;56:R33–R54
- Fonseca SG, Gromada J, Urano F. Endoplasmic reticulum stress and pancreatic β -cell death. *Trends Endocrinol Metab* 2011;22:266–274
- Gutiérrez GD, Gromada J, Sussel L. Heterogeneity of the pancreatic beta cell. *Front Genet* 2017;8:22
- Xin Y, Kim J, Okamoto H, et al. RNA sequencing of single human islet cells reveals type 2 diabetes genes. *Cell Metab* 2016;24:608–615
- Baron M, Veres A, Wolock SL, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. *Cell Syst* 2016;3:346–360.e4
- Segerstolpe Å, Palasantza A, Eliasson P, et al. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab* 2016;24:593–607
- Muraro MJ, Dharmadhikari G, Grun D, et al. A single-cell transcriptome atlas of the human pancreas. *Cell Syst* 2016;3:385–394.e3
- Dominguez D, Tsai YH, Gomez N, Jha DK, Davis I, Wang Z. A high-resolution transcriptome map of cell cycle reveals novel connections between periodic genes and cancer. *Cell Res* 2016;26:946–962
- Liang F, Audas TE, Li Y, et al. Luman/CREB3 induces transcription of the endoplasmic reticulum (ER) stress response protein Herp through an ER stress response element. *Mol Cell Biol* 2006;26:7999–8010
- Jiang HY, Wek SA, McGrath BC, et al. Activating transcription factor 3 is integral to the eukaryotic initiation factor 2 kinase stress response. *Mol Cell Biol* 2004;24:1365–1377
- Ediger BN, Du A, Liu J, et al. Islet-1 is essential for pancreatic β -cell function. *Diabetes* 2014;63:4206–4217
- Gao T, McKenna B, Li C, et al. Pdx1 maintains β cell identity and function by repressing an α cell program. *Cell Metab* 2014;19:259–271
- Hang Y, Stein R. MafA and MafB activity in pancreatic β cells. *Trends Endocrinol Metab* 2011;22:364–373
- Gu C, Stein GH, Pan N, et al. Pancreatic beta cells require NeuroD to achieve and maintain functional maturity. *Cell Metab* 2010;11:298–310
- Gutiérrez GD, Bender AS, Cirulli V, et al. Pancreatic β cell identity requires continual repression of non- β cell programs. *J Clin Invest* 2017;127:244–259
- Arda HE, Li L, Tsai J, et al. Age-dependent pancreatic gene regulation reveals mechanisms governing human β cell function. *Cell Metab* 2016;23:909–920
- Sloan KE, Bohnsack MT, Watkins NJ. The 5S RNP couples p53 homeostasis to ribosome biogenesis and nucleolar stress. *Cell Rep* 2013;5:237–247
- Nicholas LM, Valtat B, Medina A, et al. Mitochondrial transcription factor B2 is essential for mitochondrial and cellular function in pancreatic β -cells. *Mol Metab* 2017;6:651–663
- Ma Q. Role of nrf2 in oxidative stress and toxicity. *Annu Rev Pharmacol Toxicol* 2013;53:401–426
- Abebe T, Mahadevan J, Bogachus L, et al. Nrf2/antioxidant pathway mediates β cell self-repair after damage by high-fat diet-induced oxidative stress. *JCI Insight* 2017;2
- Lipson KL, Fonseca SG, Ishigaki S, et al. Regulation of insulin biosynthesis in pancreatic beta cells by an endoplasmic reticulum-resident protein kinase IRE1. *Cell Metab* 2006;4:245–254
- Espinosa-Diez C, Miguel V, Mennerich D, et al. Antioxidant responses and cellular adjustments to oxidative stress. *Redox Biol* 2015;6:183–197
- Gray JP, Karandrea S, Burgos DZ, Jaiswal AA, Heart EA. NAD(P)H-dependent quinone oxidoreductase 1 (NQO1) and cytochrome P450 oxidoreductase (CYP450OR) differentially regulate menadione-mediated alterations in redox status, survival and metabolism in pancreatic β -cells. *Toxicol Lett* 2016;262:1–11
- Orino K, Lehman L, Tsuji Y, Ayaki H, Torti SV, Torti FM. Ferritin and the response to oxidative stress. *Biochem J* 2001;357:241–247
- Shi H, Bencze KZ, Stemmler TL, Philpott CC. A cytosolic iron chaperone that delivers iron to ferritin. *Science* 2008;320:1207–1210
- Clements CM, McNally RS, Conti BJ, Mak TW, Ting JP. DJ-1, a cancer- and Parkinson's disease-associated protein, stabilizes the antioxidant transcriptional master regulator Nrf2. *Proc Natl Acad Sci U S A* 2006;103:15091–15096
- Jain A, Lamark T, Sjøttem E, et al. p62/SQSTM1 is a target gene for transcription factor NRF2 and creates a positive feedback loop by inducing antioxidant response element-driven gene transcription. *J Biol Chem* 2010;285:22576–22591
- Chen W, Sun Z, Wang XJ, et al. Direct interaction between Nrf2 and p21(Cip1/WAF1) upregulates the Nrf2-mediated antioxidant response. *Mol Cell* 2009;34:663–673
- Zhang J, Tong W, Sun H, et al. Nrf2-mediated neuroprotection by MANF against 6-OHDA-induced cell damage via PI3K/AKT/GSK3 β pathway. *Exp Gerontol* 2017;100:77–86
- Bensellam M, Montgomery MK, Luzuriaga J, Chan JY, Laybutt DR. Inhibitor of differentiation proteins protect against oxidative stress by regulating the antioxidant-mitochondrial response in mouse beta cells. *Diabetologia* 2015;58:758–770
- Sharma RB, O'Donnell AC, Stamateris RE, et al. Insulin demand regulates β cell number via the unfolded protein response. *J Clin Invest* 2015;125:3831–3846
- Szabat M, Page MM, Panzhinskiy E, et al. Reduced insulin production relieves endoplasmic reticulum stress and induces β cell proliferation. *Cell Metab* 2016;23:179–193
- Khalfallah O, Rouleau M, Barbry P, Bardoni B, Lalli E. Dax-1 knockdown in mouse embryonic stem cells induces loss of pluripotency and multilineage differentiation. *Stem Cells* 2009;27:1529–1537

46. Parker JB, Yin H, Vinckevicius A, Chakravarti D. Host cell factor-1 recruitment to E2F-bound and cell-cycle-control genes is mediated by THAP11 and ZNF143. *Cell Rep* 2014;9:967–982
47. Murata K, Hattori M, Hirai N, et al. Hes1 directly controls cell proliferation through the transcriptional repression of p27Kip1. *Mol Cell Biol* 2005;25:4262–4271
48. El Yakoubi W, Borday C, Hamdache J, et al. Hes4 controls proliferative properties of neural stem cells during retinal ontogenesis. *Stem Cells* 2012;30:2784–2795
49. Klochendler A, Caspi I, Corem N, et al. The genetic program of pancreatic β -cell replication in vivo. *Diabetes* 2016;65:2081–2093
50. Riganti C, Gazzano E, Polimeni M, Aldieri E, Ghigo D. The pentose phosphate pathway: an antioxidant defense and a crossroad in tumor cell fate. *Free Radic Biol Med* 2012;53:421–436
51. Zhang Z, Liew CW, Handy DE, et al. High glucose inhibits glucose-6-phosphate dehydrogenase, leading to increased oxidative stress and beta-cell apoptosis. *FASEB J* 2010;24:1497–1505
52. Stanton RC. Glucose-6-phosphate dehydrogenase, NADPH, and cell survival. *IUBMB Life* 2012;64:362–369
53. Talchai C, Xuan S, Lin HV, Sussel L, Accili D. Pancreatic β cell dedifferentiation as a mechanism of diabetic β cell failure. *Cell* 2012;150:1223–1234
54. Ichii H, Inverardi L, Pileggi A, et al. A novel method for the assessment of cellular composition and beta-cell viability in human islet preparations. *Am J Transplant* 2005;5:1635–1645
55. Rorsman P, Renström E. Insulin granule dynamics in pancreatic beta cells. *Diabetologia* 2003;46:1029–1045
56. Kendall DM, Sutherland DE, Najarian JS, Goetz FC, Robertson RP. Effects of hemipancreatectomy on insulin secretion and glucose tolerance in healthy humans. *N Engl J Med* 1990;322:898–903
57. Menge BA, Schrader H, Breuer TG, et al. Metabolic consequences of a 50% partial pancreatectomy in humans. *Diabetologia* 2009;52:306–317
58. Meier JJ, Bonadonna RC. Role of reduced β -cell mass versus impaired β -cell function in the pathogenesis of type 2 diabetes. *Diabetes Care* 2013;36(Suppl. 2):S113–S119