Genetic Insulin Resistance is a Potent Regulator of Gene Expression and Proliferation in Human iPS Cells

Running Title: Insulin Resistance Alters iPS Cell Function

Salvatore Iovino\textsuperscript{1,3}, Alison M. Burkart\textsuperscript{1,3}, Kristina Kriauciunas\textsuperscript{1}, Laura Warren\textsuperscript{1}, Katelyn J. Hughes\textsuperscript{1}, Michael Molla\textsuperscript{1}, Youn-Kyoung Lee\textsuperscript{2}, Mary-Elizabeth Patti\textsuperscript{1,*} and C. Ronald Kahn\textsuperscript{1,*}

1. Integrative Physiology and Metabolism Research Division, Joslin Diabetes Center, Harvard Medical School, Boston, MA, 02215, USA.

2. Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA.

3. Contributed equally to this manuscript.

*Correspondence:

C. Ronald Kahn MD
1 Joslin Place, Boston, MA 02215
Phone: 617-309-2635
FAX: 617-309-2593
C.Ronald.Kahn@joslin.harvard.edu

or

Mary-Elizabeth Patti MD
1 Joslin Place, Boston, MA 02215
Phone: 617-309-1966
FAX: 617-309-2593
Mary.Elizabeth.Patti@joslin.harvard.edu

Keywords: iPSC, insulin resistance, insulin signaling, gene expression

Word Count: 4001, 1 Table, 7 Figures
ABSTRACT

Insulin resistance is central to diabetes and metabolic syndrome. To define the consequences of genetic insulin resistance distinct from those secondary to cellular differentiation or in vivo regulation, we generated induced pluripotent stem cells (iPSC) from individuals with insulin receptor mutations and age-appropriate controls, and studied insulin signaling and gene expression in comparison to fibroblasts from which they were derived. iPSC from patients with genetic insulin resistance exhibited altered insulin signaling, paralleling that seen in the original fibroblasts. Insulin-stimulated expression of immediate early genes and proliferation were also potently reduced in insulin resistant iPSC. Global gene expression analysis revealed marked differences in both insulin-resistant iPSC and corresponding fibroblasts as compared to control iPSC and fibroblasts. Patterns of gene expression in patients with genetic insulin resistance were particularly distinct in the two cell types, indicating dependence on not only receptor activity but also the cellular context of the mutant insulin receptor. Thus, iPSC provide a novel approach to define effects of genetically-determined insulin resistance. Our study demonstrates that effects of insulin resistance on gene expression are modified by cellular context and differentiation state. Moreover, altered insulin receptor signaling and insulin resistance can modify proliferation and function of pluripotent stem cell populations.
INTRODUCTION

Induced pluripotent stem cells (iPSC) offer a unique tool for studying human disease (1-3). iPSC can be derived from multiple cell types and differentiated into all three germ layer-derived tissues, thus providing an opportunity to develop patient- and tissue-specific models for molecular analysis (4;5). iPSC and their differentiated derivatives also provide a tool to dissect gene-environment interactions central to complex human diseases, such as type 2 diabetes (T2D).

Insulin resistance is a key feature of T2D, obesity and metabolic syndrome. Longitudinal studies indicate that insulin resistance is heritable and occurs in individuals at risk for T2D many years prior to glucose intolerance (6;7). Both genetic and environmental factors, including over-nutrition and inactivity, can contribute to insulin resistance in individuals at risk for T2D. However, the precise molecular mechanisms underlying insulin resistance, and the extent to which genes versus environment determine risk, remain unknown.

Important insights into insulin action and insulin resistance have been provided by rare inherited syndromes of severe insulin resistance due to mutations in the insulin receptor (INSR), such as Donohue syndrome and type A insulin resistance (8-10). Studies in fibroblasts and transformed lymphocytes from these individuals have demonstrated altered insulin receptor signaling and provided key information about receptor structure and function (11-15). In general, clinical manifestations and signaling defects are more severe in Donohue syndrome, due to the homozygous or compound heterozygous mutations, as compared with heterozygous mutations in type A patients (16). Because these patients are seriously ill, however, it is difficult to study insulin action in classical target tissues, such as muscle, fat or
liver; thus far, in vitro studies have largely employed skin fibroblasts or lymphocytes, limiting the generalizability of these findings to more relevant metabolic tissues.

To define the impact of genetically-determined insulin resistance in pluripotent cells, we generated iPSC from fibroblasts of three patients with INSR mutations and three healthy controls of similar age. Both receptor and post-receptor signaling are similarly disrupted in both iPSC and fibroblasts. Insulin receptor mutations also alter gene expression, but the nature of these changes is dependent on the cellular context. Thus, iPSC are a powerful new tool for the study of insulin resistance that provides an approach to uncover interactions between genetics and cellular environment in the pathogenesis of T2D.
RESEARCH DESIGN AND METHODS

Animal protocols were approved by the Joslin Diabetes Center IACUC. Generation and usage of iPS cells was approved by the Joslin Committee on Human Studies.

Reagents

Primary antibodies included anti-phospho-IGF1R-β (Tyr1135/1136)/INSR-β (Tyr1150/1151), phospho-AKT (Ser473), AKT, phospho-Erk1/2 (Thr202/Tyr204), Erk1/2, and phospho-GSK3α/β (Ser21/9) (Cell Signaling), INSR-β and IGF1R-β (Santa Cruz), GSK3α/β (Millipore), phospho-IRS1 (Y612) (Life Technologies), IRS1 (BD Biosciences), NANOG and OCT4 (Abcam), and SSEA4 and TRA-1-60 (Millipore).

Fibroblasts and iPSC derivation

C1 fibroblasts (BJ) were from ATCC (Manassas, VA); C2, C3, IR-M2 and IR-M3 were from Coriell (Camden, NJ). IR-M1 and IR-M3 were previously characterized (9;17). Fibroblasts were reprogrammed with individual retroviral constructs (OCT4, KLF4, c-MYC, SOX2) and plated on irradiated mouse embryonic fibroblasts (1;2;5). Individual clones were analyzed for pluripotency and teratoma formation using standard techniques (1;2;5). Chromosomal analysis was performed on colcemid-arrested IPSC at Dana- Farber/Harvard Cytogenetics Core (Boston) and Cell Line Genetics (Madison, WI).

Western blot analysis

Fibroblasts were serum-starved overnight in DMEM (25 mM glucose) and 0.1% bovine serum albumin (BSA) then treated with 0 or 100 nM insulin for 10 min and lysed in RIPA buffer. iPS cells were starved in DMEM (25 mM glucose) and 0.5% BSA for 3 hours prior to stimulation with 0 to 100 nM insulin. Protein concentration was determined by BCA (Thermo Scientific). Equal amounts of protein were separated by SDS-PAGE and transferred to nitrocellulose membranes, blocked with 5% nonfat dry milk in Tris-buffered saline with 0.1% Tween, incubated with primary antibodies overnight, and washed prior to incubation
with HRP-conjugated secondary antibodies. Proteins were detected by enhanced chemiluminescence (PerkinElmer Life Sciences). Band intensities were quantified on scanned images using Adobe Photoshop on inverted images; bands were selected using rectangular marquee tool, and histogram function was used to determine the average pixel intensity.

**DNA sequencing**

Genomic DNA was isolated (DNeasy, Qiagen) and *INSR* exons were PCR-amplified (18) (GoTaq PCR, Promega) and sequenced (Dana Farber Cancer Institute Core).

**Proliferation and EdU incorporation**

$1 \times 10^4$ iPSC were grown on 12-well plates in mTesr1 media. Cells were counted at 2-day intervals (Nexcelom). EdU incorporation into DNA was measured using Alexa 488 Click-It® HCS assay (Life Technologies) after incubation for 12-14 hours with 100 nM insulin or 100nM IGF-1.

**Expression analysis**

RNA was extracted using Ribozol (AMRESCO). cRNA was prepared and hybridized to Affymetrix PrimeView microarrays. Data were normalized using Robust Multichip Average (RMA) (20). Heatmaps were created using GENE-E (GenePattern, Broad Institute). Ontology analysis was performed using DAVID (21). Putative transcription factor binding sites were identified using hypergeometric enrichment (Molecular Signatures Database) (22). For PCR, cDNA was synthesized (High Capacity cDNA Reverse Transcription Kit, Life Technologies) and amplified (iTaq Universal SYBR Green Supermix, Bio-Rad; ABI 7900HT, Life Technologies). *GAPDH* and *36B4* were used for normalization, using the $2\Delta\Delta CT$ method (5).

**Statistical analysis**
Data are presented as mean ± SEM. Between-group differences were examined by Student’s t-test (p values) and false discovery rate (FDR) (q values); p<0.05 and q<0.1 were considered significant.

RESULTS

Insulin-resistant human fibroblasts

Skin fibroblasts were obtained from three healthy controls (BJ, GM05400, GM00409, hereafter designated C1, C2, C3) and three individuals with severe insulin resistance (IR-Mut) due to mutations in the insulin receptor gene (INSR) (Minn1, GM10277, GM20034, hereafter IR-M1, IR-M2, IR-M3). DNA sequencing confirmed INSR mutations in all three patients (Table 1). IR-M1 was a compound heterozygote, with a mutation in one allele generating a STOP codon at amino acid 897 in the extracellular domain of the β-subunit, yielding a non-functional INSR. Previous studies (23) indicated the presence of a second mutation in this patient that results in a 95% reduction in INSR expression. DNA sequencing of IR-M2 identified a missense mutation in exon 1 (5C→G) leading to alanine-for-glycine substitution at the second amino acid in the signal peptide, which would be expected to alter membrane trafficking. In IR-M3, sequencing revealed a homozygous mutation in exon 3 (T→C, codon 233), resulting in a leucine-to-proline mutation and impaired INSR trafficking and autophosphorylation (17).

INSR expression and downstream signaling were reduced in fibroblasts from IR-Mut patients. qPCR with primers specific for the A (-exon 11, short) and B (+exon 11, long) isoforms of the α- and β-subunits of INSR revealed a >85% decrease in expression in IR-M1 cells (Fig. 1A, p<0.05), consistent with prior studies (23;24). INSR expression in IR-M2 fibroblasts did not differ from controls, while IR-M3 fibroblasts had a 2-fold increase in INSR
mRNA (Fig. 1A). Insulin-like growth factor-1 receptor (*IGF1R*) mRNA did not differ between control and any IR-Mut fibroblasts (Fig. 1A, right).

Consistent with previous studies (23;24) and mRNA expression, INSR protein was decreased by 95% in IR-M1 fibroblasts. In IR-M2, expression did not differ from controls (Fig. 1B, Supplementary Fig 1A). Mature INSR protein was also decreased by 88% in IR-M3, despite increased INSR mRNA, consistent with impaired proreceptor processing (17). By contrast, IGF1R protein was unchanged in IR-Mut cells (Fig. 1B, Supplementary Fig 1B).

Western blotting with an antibody that recognizes both the phosphorylated INSR and IGF1R revealed a robust 8-fold increase in insulin-stimulated phosphorylation in control fibroblasts (Fig. 1B, left, 3rd row). In contrast, INSR/IGF1R phosphorylation was undetectable in IR-M1 and IR-M3 fibroblasts, but similar to control in IR-M2 (Fig. 1B, right, Supplementary Fig 1C). Nonetheless, insulin-stimulated phosphorylation of insulin receptor substrate-1 (IRS-1) was reduced by 30-40% in all three IR-Mut fibroblasts. In IR-M2 cells, this was due, at least in part, to a marked decrease in IRS-1 protein. Insulin-stimulated phosphorylation of AKT was also decreased in all IR-Mut fibroblasts by ~70%, with no change in AKT protein or the loading control GRB2. Basal and insulin-stimulated phosphorylation of ERK1/2 was decreased by >50% in all IR-Mut.

**Creation of iPSC lines and analysis of insulin signaling**

To isolate the impact of genetic insulin resistance from the effects of cellular differentiation, we derived iPSC from control and IR-Mut fibroblasts by introduction of OCT4, KLF4, SOX2 and c-MYC (1) (Fig. 1C). When cultured on low adherence plates, these cells formed embryoid bodies (Fig. 1C, bottom). For each iPS line, we confirmed pluripotency using three independent criteria: (1) expression of pluripotency factors OCT4, TRA1-60, NANOG, and SSEA4 (Supplementary Fig. 2A); (2) alkaline phosphatase staining
(Supplementary Fig. 2B); and (3) in vivo teratoma formation, with histological identification of ectodermal, mesodermal, and endodermal layers (Supplementary Fig. 3A). Additionally, all iPSC had normal karyotype (Supplementary Fig. 3B) and no difference in reprogramming efficiency.

\textit{INSR} mRNA was decreased by 95% in IR-M1 and by 60% in IR-M2 and IR-M3 iPSC (Fig. 2A). Mature INSR protein was also decreased in both IR-M1 and IR-M3 iPSC by 73% and 59%, but unchanged in IR-M2 (Fig. 2B/C), similar to the pattern in fibroblasts. In IR-M3 cells there was also a 6-fold increase in proreceptor expression (Supplementary Fig. 3A/B), consistent with the known impairment in proreceptor processing (17).

There were no differences in IGF1R protein (Fig. 2C, Supplementary Fig. 3C) or mRNA (Supplementary Fig. 3D). Western blotting using an antibody that recognizes phosphorylation of both INSR and IGF1R (P-INSR/IGF1R) demonstrated robust insulin stimulation in controls (Fig. 2C/D, right). By contrast, insulin-stimulated phosphorylation of INSR/IGF1R was decreased by 50% in IR-M1 and IR-M2, but unchanged in IR-M3 iPSC (Fig. 2C, right, and 2D).

Given that the anti-P-INSR/IGF1R antibody represents a composite of phosphorylation and abundance of both INSR and IGF1R, these receptors were individually immunoprecipitated and then blotted with anti-phosphotyrosine antibody. All three control iPS showed increased tyrosine phosphorylation of INSR upon insulin stimulation (Fig. 2E), whereas negligible phosphorylation was observed in the three IR-Mut iPSC lines. Anti-INSR western blots of these anti-INSR immunoprecipitates revealed reduced mature INSR expression in M1 and M3, with normal INSR protein in IR-M2 (Fig. 2E, lower row). Furthermore, anti-INSR immunoprecipitation showed negligible insulin-stimulated phosphorylation of the proreceptor (Supplementary Figure 4E). Immunoprecipitation of IGF1R revealed similar levels of IGF1R protein in all lines (Fig. 2F, lower) and robust
insulin-stimulated phosphorylation of IGF1R in both control and IR-M1 and M3 iPSC (Fig. 2F). Interestingly, basal and insulin-stimulated phosphorylation of IGF1R was reduced ~20% in IR-M2 (Fig. 2F), suggesting an effect of the mutant insulin receptor to inhibit IGF1R.

Western blotting demonstrated that insulin-stimulated phosphorylation of AKT (Ser473) was reduced in all IR-Mut iPSC, reaching significance in IR-M2 (Fig. 3A/B) in parallel with a ~60% reduction in AKT protein (p<0.05, Fig. 3A/C). Likewise, both basal and insulin-stimulated phosphorylation of ERK1/2 was decreased in insulin-resistant cells by 40-60% (Fig. 3A/D/E). Protein expression of ERK1 (but not ERK2) was also significantly decreased in IR-M1 and IR-M3, with a similar trend in IR-M2 (Supplementary Fig. 5A/B). These changes in protein expression were present despite unaltered mRNA expression of these molecules (Supplementary Fig. 4D).

**Proliferation and expression of early response genes are impaired in iPSC from individuals with severe insulin resistance**

Insulin exerts potent effects on cell proliferation and transcriptional regulation (25-27). All IR-Mut iPSC had significant reductions in growth (Figure 4A) and reduced insulin-stimulated incorporation of 5-ethynil-2’-deoxyuridine (EdU) into DNA as compared to controls (Fig. 4B, left). These differences were not due to reduced cell attachment (Supplementary Fig. 6), but were associated with dysregulation of mitogenic gene expression. Indeed, control iPSC responded to insulin with 1.7-1.8-fold increases in expression of *EGR1*, *cFOS* and *JUN* (p<0.05, Fig. 4C, left). By contrast, these insulin-stimulated transcriptional responses were nearly absent in IR-Mut (Fig. 4C). IGF-1-stimulated EdU incorporation and expression of mitogenic genes were robust and similar in both control and IR-Mut (Fig. 4B/C, right).
Global gene expression is altered in insulin-resistant iPSC

To investigate the impact of genetic insulin resistance on transcriptional regulation in iPSC, we analyzed global gene expression using microarrays in both fibroblasts and iPSC. All cells were studied at confluence in serum-free medium without added insulin or growth factors. Volcano plots (Fig. 5A) demonstrated the distribution of gene expression differences between IR-Mut and control (X-axis) for fibroblasts and iPSC plotted against corresponding p-values (Y-axis). In general, gene expression differences were less variable in iPSC, as demonstrated by 290 probes with p<0.01 vs. 87 in fibroblasts.

We next compared gene expression in iPSC and their paired fibroblasts for both control (X-axis) and IR-Mut cells (Y-axis, Fig.5B). Probes for which expression is increased in iPSC by >32-fold (5-fold on log₂ scale) are colored in blue, while downregulated genes are colored in red. The top 25 genes within these two groups are shown in Fig. 5C (iPS-enriched, upper, and fibroblast-enriched, lower). Differential expression of representative genes was confirmed by qPCR (Fig. 5D).

Analysis of transcripts overexpressed in both control and IR-Mut iPSC versus paired fibroblasts (Fig. 5B, upper right) revealed mRNAs that encode proteins involved in pluripotency, self-renewal, and development, including LIN28A, OCT4 and NANOG (Fig. 5C/D). Conversely, transcripts enriched in both control and IR-Mut fibroblasts (Fig. 5B, lower left) represent genes typically highly expressed in fibroblasts (e.g. related to extracellular matrix, fibroblast growth factors and metalloproteases, Fig. 5C/D). These patterns were confirmed using DAVID ontology analysis (21) (Supplementary Table 1).

Transcription factor binding motif analysis (using 2 kb upstream sequences) of genes within both fibroblast and iPSC-enriched groups (22) demonstrated enrichment of binding site motifs for several transcription factors related to insulin action and iPSC physiology.
Interestingly, FOXO4, LEF1, and NFATc were common to the two gene groups, consistent with their important role in both somatic and iPS specification.

To assess the impact of insulin receptor mutations on expression patterns in both cell types, we plotted the ratio of gene expression in IR-Mut to control in fibroblasts (X-Axis, Fig. 6A) versus the corresponding ratio in iPSC (Y-axis, Fig. 6A, colored points indicate expression differences ≥2-fold). mRNAs along the diagonal represent genes regulated by insulin resistance similarly in fibroblasts and iPSC, i.e., independent of cellular context. Interestingly, only 46 genes met these criteria, with 12 upregulated in both IR-Mut fibroblasts and iPSC (group A), and 34 downregulated in both IR-Mut cell types (group B). Thus, groups A and B represent genes regulated by insulin resistance in fibroblasts and iPSC independent of differentiation state (heatmaps, Figure 6B). We validated selected genes using qRT-PCR (Fig. 6C). For example, insulin signaling-related gene adapter protein 1 sigma 2 (AP1S2) is upregulated in IR-Mut fibroblasts and iPSC, while receptor-type tyrosine-protein phosphatase PCP-2 (PTPRU) is downregulated. Transcription factor binding motif analysis for the top 100 genes (p<0.05) with similar expression patterns to those in groups A and B revealed an enrichment of binding sites for transcription factors involved in glucose homeostasis and/or insulin action, such as E2F1, FOXF2, SP1, JUN, and CREB1 (Supplementary Table 3).

In contrast to the modest number of genes coordinately regulated by insulin resistance in both fibroblasts and iPSC, the majority of insulin resistance-regulated genes were differentially expressed in only fibroblasts (Fig. 6A, groups C/D) or iPSC (groups E/F). Indeed, many mRNAs showed robust differences in expression as a function of insulin resistance in fibroblasts (C/D), but were unaltered in iPSC, or vice versa (E/F).

To clarify potential regulatory and functional differences between subgroups of genes altered in response to insulin resistance in a cell context-dependent manner, we compared
those genes altered in fibroblasts, but unchanged in iPSC (groups C/D, 294/262 mRNAs, respectively) with genes altered in iPSC, but unchanged in fibroblasts (groups E/F, 44/85, respectively). For each group, the top 30 differentially expressed mRNAs (all p<0.05, FDR <0.1) are presented as heatmaps in Figure 7, with qPCR validation of representative genes in adjacent graphs. Group C was upregulated by insulin resistance in fibroblasts but unaltered in iPSC. This pattern was often linked to either constitutively low expression (e.g., SGMS2) or high expression (e.g., EPB41L4B) in iPSC. Likewise, Group D was downregulated by insulin resistance in fibroblasts, but unregulated in iPSC (e.g., IRS1, MYO10). Ontology analysis of genes regulated by insulin resistance specifically in fibroblasts (C/D) reveals genes involved in development (homeobox, cell proliferation, development), signaling (IGF, FGF2R, WNT), and fibroblast function (extracellular matrix) (Supplementary Table 4). Ontology analysis of genes regulated by insulin resistance in iPSC (E/F) also revealed enrichment of genes in IGF, FGF2R and WNT signaling pathways.

Transcription factor binding motifs enriched in promoters of insulin resistance-regulated genes revealed several with regulatory roles in development (e.g. MYC-associated zinc finger protein (MAZ), paired box gene 4 (PAX4) and paired-like homeodomain transcription factor 2 (PITX2)) or insulin action (e.g. forkhead box protein O4 (FOXO4), lymphoid enhancer-binding factor-1 (LEFI) and myocyte enhancer factor 2A (MEF2A), Supplementary Table 5). Interestingly, several motifs were identified in both fibroblast and iPSC-specific groups (FOXO4, NFAT/NFATc, TAF, REPIN1, MYOD1, MEF2A), suggesting the potential for insulin resistance to confer unique transcriptional regulation, either suppression or activation, depending on the cell context or differentiation state.


DISCUSSION

Insulin resistance is central to the pathogenesis of T2D, obesity, and metabolic syndrome. Insulin resistance precedes and predicts the onset of T2D and is even present in offspring of individuals with T2D, indicating an hereditary component (6;7). Determining the contributions of genetic factors to insulin resistance in humans has been challenging given the limited access to relevant tissues, especially during preclinical stages of disease. While there has been some success studying human insulin resistance using cultured myoblasts (28-30), the majority of cellular studies have utilized fibroblasts (24;31) or circulating blood cells (32;33), which show little or no metabolic response to insulin. iPSC provide a new tool for the study of human disease that they can be derived from a variety of cells, passaged indefinitely in culture, and studied either in the pluripotent state or after differentiation into muscle, fat, or other tissues (1-5;34).

To explore the utility of iPSC to study insulin resistance, we have created and characterized iPSC from patients with severe insulin resistance due to mutations in the insulin receptor and compared these to the fibroblasts from which they were derived. These IR-Mut iPSC exhibit major defects in signaling, proliferation, and gene expression, but exhibit important differences from the fibroblasts of the same individuals, indicating that cellular context is a potent modifier, even in genetically-determined insulin resistance.

Each patient had a distinct insulin receptor mutation, resulting in different defects in iPSC. IR-M1 is a compound heterozygote in which a nonsense mutation in one allele produces a truncated receptor lacking the tyrosine kinase, while the other allele reduces $INSR$ transcription. This receptor cannot signal, but can bind insulin and form hybrids with normal insulin/IGF-1 receptors (24). IR-M2 has a mutation in the INSR signal sequence, which impairs membrane trafficking. Interestingly we observe more severe reductions in receptor expression and phosphorylation in IR-M2 iPSC than in fibroblasts. This is similar to the
previously reported Asp15Lys mutation in this region, in which insulin binding defects differed between fibroblasts and lymphoblasts (35). IR-M3 has a homozygous missense mutation in the α-subunit that results in impaired proreceptor processing and membrane transport (17). Consistent with this, we observed accumulation of the proreceptor in IR-M3 iPSC, although it was not phosphorylated upon insulin stimulation, suggesting it may not have been normally transported to the plasma membrane. Interestingly, for this patient, levels of INSR mRNA were reduced in iPSC but increased in fibroblasts, consistent with cell-specific INSR expression patterns previously observed (35) and indicating differentiation dependence.

Despite differing mutations, insulin-stimulated phosphorylation of INSR was significantly reduced in all IR-Mut iPSC. However, insulin was able to mediate some downstream signaling, in part via IGF1R phosphorylation, which was similar in IR-M1 cells and increased in IR-M3 iPSC. In IR-M2 cells, IGF1R phosphorylation was reduced, suggesting that the mutated INSR functions in a dominant-negative manner to modulate IGF1R signaling (36), potentially contributing to more severe downstream signaling defects. The importance of IGF1R signaling to compensate for reduced INSR activity in IR-Mut iPSC also supports a greater reliance of iPSC on IGF1R-mediated signaling.

Different INSR mutations also cause both cell type- and pathway-specific impairments in downstream signaling. Phosphorylation of AKT, the kinase linked to many metabolic actions of insulin, was reduced in all three IR-Mut fibroblast lines. In iPSC, AKT phosphorylation was significantly impaired only in IR-M2, with similar trends in the other two lines. Phosphorylation of GSK3 was unaffected by INSR mutations, suggesting compensation from other signaling pathways. By contrast, ERK expression and basal phosphorylation were altered in both IR-Mut iPSC and fibroblasts, with more dramatic decreases in ERK1 than ERK2. Although ERK1 and ERK2 are often considered functionally
redundant, isoform-specific and differentiation-dependent differences have been observed, including preferential roles of ERK2 in myogenic differentiation (37) and regulation of epithelial-to-mesenchymal transitions (38). Differential regulation of ERK signaling in iPSC is interesting as ERK regulates transcription, stem cell proliferation and differentiation (39;40). Together these data indicate that insulin receptor mutations can selectively perturb specific downstream components of the insulin signaling pathway, and that the cellular context adds a further layer of regulation by modulating expression and function of insulin signaling molecules.

Given that insulin is a potent regulator of transcription (27), we assessed the impact of this genetic insulin resistance by analyzing gene expression in control and IR-Mut fibroblasts and iPSC. Not surprisingly, expression patterns differ markedly in fibroblasts vs. iPSC. Fibroblasts strongly express genes related to extracellular matrix synthesis and wound healing, whereas iPSC express pluripotency-related genes, such as OCT4, NANOG, ZIC, and the micro-RNA-binding protein LIN28 (41). LIN28 itself modulates insulin action by inhibiting microRNA let-7, which represses multiple components of the insulin/PI 3-kinase/mTOR pathway (42).

Insulin resistance imparted robust effects on gene expression. A small subset of genes was concordantly highly regulated by insulin resistance in both fibroblasts and iPSC (Groups A/B). These genes were predominantly related to cell growth and signaling, including IGFI-2, IGFBPs, EGFR, bFGFR, BMP2 and SERPINE1. However, most effects on gene expression were cell-type specific, with major dysregulation of expression present either only in IR-Mut fibroblasts (Groups C/D) or only in iPSC (Groups E/F). In many cases, loss of regulation in either fibroblasts or iPSC appears to be linked to constitutively high or low expression of the regulated gene, which in turn maybe linked to cellular differentiation state or epigenetic regulation of transcription. Analysis suggested that insulin resistance may
impact regulation of developmental pathways (HOX, MEIS), potentially affecting the capacity of iPSC to differentiate along specific developmental lineages (43-45).

At a functional level, insulin resistance resulted in a significant reduction in iPSC proliferation and insulin-stimulated DNA synthesis. Reduced proliferation is particularly important in iPSC, as self-renewal is a critical and defining characteristic of stem cells, and altered growth rates could have a major impact on differentiation (46). Impairment in insulin signaling through the ERK pathway and reductions in insulin-stimulated expression of the early-growth response genes are likely contributors to reduced proliferation. Importantly, these defects were specifically due to defective insulin signaling, as mutant lines remained responsive to IGF-1 stimulation. Thus, these data underscore the importance of insulin signaling for proliferation of stem cells. Although PI 3-kinase has been shown to be important for pluripotency in iPSC (46), we did not observe significant differences in markers of pluripotency, despite reduced insulin action within this pathway.

Collectively, our data support the hypothesis that insulin resistance in stem cells may be a novel mechanism contributing to diabetes pathogenesis. Reductions in insulin signaling could decrease the size of stem cell populations, disrupt developmental trajectories, alter differentiation, and potentially reduce tissue-resident stem cell numbers, thus reducing regenerative responses to injury during adult life (47). Consistent with this hypothesis, we have previously demonstrated that nutritional signals can reduce the size and functional capacity of stem cell pools in mice at risk for diabetes (48). Going forward, iPSC will be an important tool to dissect pathways contributing to diabetes by leveraging their potential capacity to differentiate into tissue-specific insulin resistant cell lines.

In summary, iPSC generated from fibroblasts from patients with severe insulin resistance have defects in signaling, gene expression, and proliferation. Thus, genetically-determined insulin resistance is an important modifier of stem cell function and could
contribute to disease pathogenesis. iPSC are also a potent tool for studying the impact of cellular context on insulin resistance and dissecting the contributions of genetics and differentiation-dependent effects of insulin resistance to disease risk.
ACKNOWLEDGMENTS

The authors have no conflicts of interest. The authors acknowledge the valuable support provided by C. Cowan and L. Daherent of HSCI iPSC Core Facility for assistance with generation of human iPSC. This work was supported by Joslin DRC grant (P30DK036836), NIH R01DK31036 (CRK), funding from the Novo-Nordisk Foundation (CRK/MEP), and pilot funding from the Harvard Stem Cell Institute. AB was supported by T32DK007260 and the American Diabetes Association mentor-based fellowship (MEP).

SI and AB designed, researched, and analyzed data and wrote the manuscript. KK, LW, KJH, and MM researched data. YKL provided materials. MEP and CRK oversaw the project, contributed to discussion and helped write the manuscript. CRK is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
REFERENCES


45. Orvis, GD, Hartzell, AL, Smith, JB, Barraza, LH, Wilson, SL, Szulc, KU, Turnbull, DH, Joyner, AL: The engrailed homeobox genes are required in multiple cell lineages to


TABLES

Table 1. Summary of fibroblasts, including donor age and gender, mutation analysis, and functional domain of insulin receptor mutations.

<table>
<thead>
<tr>
<th>Cell Lines</th>
<th>Age</th>
<th>Gender</th>
<th>Mutation</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 (BJ)</td>
<td>neonatal</td>
<td>M</td>
<td>No Mutation</td>
<td>-</td>
</tr>
<tr>
<td>C2 (GM05400)</td>
<td>6 year old</td>
<td>M</td>
<td>No Mutation</td>
<td>-</td>
</tr>
<tr>
<td>C3 (GM00409)</td>
<td>7 year old</td>
<td>M</td>
<td>No Mutation</td>
<td>-</td>
</tr>
<tr>
<td>IR-M1 (Minn1)</td>
<td>1 month old</td>
<td>F</td>
<td>Arg897 $\rightarrow$ Stop (one allele)</td>
<td>$\beta$-subunit (one allele)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>unknown mutation (second allele)</td>
<td>Decreased expression (second allele)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exon 14</td>
<td></td>
</tr>
<tr>
<td>IR-M2 (GM10277)</td>
<td>15 year old</td>
<td>F</td>
<td>Ala2 $\rightarrow$ Gly (both alleles)</td>
<td>$\alpha$-subunit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exon 1</td>
<td>Signal sequence</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L1 domain</td>
</tr>
<tr>
<td>IR-M3 (GM20034)</td>
<td>3 month old</td>
<td>M</td>
<td>Leu233 $\rightarrow$ Pro (both alleles)</td>
<td>$\alpha$-subunit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exon 3</td>
<td>L2 domain</td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Figure 1. Insulin signaling in fibroblast cell lines and generation of iPSC

(A) Gene expression analysis of INSR alpha and beta subunits and IGF1R, expressed relative to the average of controls (n=3).

(B) Fibroblasts were serum starved overnight before 10-minute stimulation with 100 nm insulin. Western blot analysis of insulin signaling in control and IR-Mut fibroblasts are shown. Specific antibodies are indicated adjacent to the respective image. Images are representative of 2 independent experiments.

(C) Bright-field images of control and patient iPSC colonies and embryoid bodies, with bar = 100 µm. All values represent mean ± SEM. *p<0.05 vs. controls.

Figure 2. Expression and phosphorylation of INSR in iPSC

(A) mRNA expression of INSR by qRT-PCR, expressed as percent of maximum value (n=3 experiments).

(B) Quantification of INSR protein level from western blot analysis, expressed as percentage of maximum value (n=3).

(C) iPSC were serum starved for 3 hours before 10-minute stimulation with 0, 10, or 100 nm insulin. Representative Western blot analysis of INSR and IGF1R expression and phosphorylation are shown (n=3).

(D) Quantification of western blot analysis of INSR/IGFIR phosphorylation, represented as percent of maximum value (n=3).

(E) iPSC were serum starved for 3 hours before 5-minute stimulation with 0 or 100 nm insulin. Immunoprecipitation of INSR using a β-subunit-specific anti-insulin receptor antibody followed by immunoblotting with an anti-phosphotyrosine antibody (upper row) and anti-INSR antibody (lower row) was performed. Representative blots are shown (n=2).
(F) iPSC were serum starved for 3 hours before 5-minute stimulation with 0 or 100 nm insulin. Immunoprecipitation of IGF1R using a β-subunit-specific anti-IGF1R antibody and immunoblotting with an anti-phosphotyrosine antibody (upper row), anti-phospho INSR/IGF1R (middle row), and anti-IGF1R antibody (lower row) was performed. Representative blots are shown (n=2). For panels A/B/D, all values indicate mean ± SEM. *p<0.05 vs. controls.

Figure 3. Regulation of downstream insulin signaling in iPSC

(A) iPSC were serum starved for 3 hours before 10-minute stimulation with 0 or 100 nm insulin. Western blot analysis was performed for key insulin signaling pathway proteins, as indicated. Representative blots (n=3).

(B) Quantitation of western blot analysis for AKT Ser473 phosphorylation and (C) expression, shown as percent of maximum value (n=3).

(D, E) Quantification of ERK1 and ERK2 phosphorylation, expressed relative to the average of controls (n=3). All values represent mean ± SEM. *p<0.05 vs. controls.

Figure 4. Mitogenesis of iPSC

(A) Proliferation in iPSC lines was assessed every 2 days over 10 days by cell counting. Data represent the mean of triplicates for each of the 3 control and 3 IR-Mut lines; all values represent mean ± SEM. *p<0.05, **p<0.01 vs. controls (n=2).

(B) EdU incorporation in iPSC lines after insulin and IGF-1 stimulation (100nM). Bars represent percentage (%) of incorporated EdU compared to unstimulated iPSC (dash line). All values represent mean ± SD. *p<0.05, **p<0.01, ***p<0.001 vs. unstimulated, #p<0.05 vs. controls.

(C) qRT-PCR analysis for insulin-stimulated EGR1, cFOS, and JUN mRNA expression, relative to unstimulated cells (dash line). All values represent mean ± SD. *p<0.05, **p<0.01 vs. unstimulated, #p<0.05, ##p<0.01 vs. controls.
Figure 5. Microarray analysis of fibroblast and iPSC lines

(A) Volcano plots showing distribution of differential expression in fibroblasts (top) and iPSC (bottom), with ratio of IR-Mut/Control on x-axis and p value on Y-axis.

(B) Comparison of gene expression (ratio, iPSC/fibroblasts) for the IR-Mut (Y-axis) and control lines (X-axis). Colored dots at each end of the distribution mark genes with expression ratio > or < 5-fold (log₂ scale).

(C) Heatmaps indicating expression differences for the top 25 genes in either iPS or fibroblast-enriched groups (log₂ FC > or < 5), normalized by row.

(D) Expression of LIN28A, USP44, PUO5F1, MMP1, FN1 and COL15A1 was analyzed by qRT-PCR, using GAPDH as housekeeping gene, *p<0.05, **p<0.01, ***p<0.001 average of iPSC vs. fibroblasts.

Figure 6. Gene expression analysis of gene groups A-B

(A) Dot plot shows distribution of the ratio of gene expression (IR-Mut/control) in iPSC (Y-axis) vs. that in fibroblasts (X-axis). Letters indicate gene groups with expression ratio >2-fold (log₂ FC > or < 1).

(B) Heatmaps of top 12 significant genes within gene groups A and B, normalized by row (p<0.05).

(C) Expression of B3GALTL, AP1S2, HSPA2 and PTPRU was measured by qRT-PCR, with GAPDH as housekeeping gene, *p<0.05, **p<0.01 average of controls vs. mutants.

Figure 7. Gene expression analysis of gene groups C-F

Heatmaps and mRNA expression of gene groups C-F. Heatmaps represent the top 30 significant genes of gene groups C-F in a row-normalized fashion (p<0.05, q<0.1), with qRT-PCR analysis in adjacent graphs, *p<0.05, **p<0.01 average of controls vs. mutants.
Figure 1

A

mRNA expression (Fold Change)

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>IR-M1</th>
<th>IR-M2</th>
<th>IR-M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>INS R alpha (long)</td>
<td>□</td>
<td>■</td>
<td>▪</td>
<td>▣</td>
</tr>
<tr>
<td>INS R alpha (short)</td>
<td>□</td>
<td>■</td>
<td>▪</td>
<td>▣</td>
</tr>
<tr>
<td>INS R beta</td>
<td>□</td>
<td>■</td>
<td>▪</td>
<td>▣</td>
</tr>
<tr>
<td>IGF1R</td>
<td>□</td>
<td>■</td>
<td>▪</td>
<td>▣</td>
</tr>
</tbody>
</table>

B

Fibroblast

<table>
<thead>
<tr>
<th></th>
<th>Controls</th>
<th>IR-M1</th>
<th>IR-M2</th>
<th>IR-M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulin (nM):</td>
<td>0 100 0 100</td>
<td>0 100 0 100</td>
<td>0 100 0 100</td>
<td></td>
</tr>
<tr>
<td>INS R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGF1R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-INSR/IGF1R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-IRS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-AKT (S473)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P-ERK1/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERK1/2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRB2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C

iPS colonies

Embryoid Bodies

Control

IR-Mut

C1       C2       C3

IR-M1     IR-M2     IR-M3

Scale bars: 100 μm
Figure 4

A

Controls
IR-Mut

B

% of EdU incorporated

Insulin
IGF-1

C

EGR1 mRNA

cFOS mRNA

Jun mRNA
Figure 5

(iPS Fibroblasts) Control IR-Mut Control IR-Mut

(iPS-enriched) Fibroblast-enriched

<table>
<thead>
<tr>
<th>Gene</th>
<th>mRNA</th>
<th>mRNA</th>
<th>mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIN28A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPCAM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POU5F1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USP44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDGF1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TDGF3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POU5F1P3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LITD1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZIC2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZIC5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLC7A3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIN28B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIF26A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBPMS2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KLHL23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZFP42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTX2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPRZ1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTSL2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZIC3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTMR1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESRP1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEMA6A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCGB3A2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPM6B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL3A1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LUM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GREM1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DKK1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNM1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL6A3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GREM2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSG5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMP3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL1R1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FN1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZEB1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANPEP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDKN2C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTBP2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COL15A1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NUPR1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSTN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFB1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FBN1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMP1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGF5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KRTAP1-5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>mRNA</th>
<th>mRNA</th>
<th>mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>COL15A1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Diabetes
**Figure 6**

**Diabetes**

- **A**
  - Scatter plot showing the fold change in mRNA expression of IR-Mut compared to Control for Fibroblasts and iPS cells.
  - The x-axis represents the Fold Change Log2 (IR-Mut/Con), while the y-axis represents the fold change.

- **B**
  - Heatmap for gene expression in different cell types (Fibro, Control, IR-Mut, iPS).

- **C**
  - Bar charts showing the fold change in mRNA expression for specific genes:
    - **B3GALTL mRNA**
    - **AP1S2 mRNA**
    - **HSPA2 mRNA**
    - **PTPRU mRNA**

- **Gene group A**
  - Includes genes like AP1S2, ZNF565, TCTE3, ADH1A, ADAL, NAF1, FLJ16124, APPY1, GPR112, etc.

- **Gene group B**
  - Includes genes like P2RX7, HFCFC1R1, HSPA2, DNAE1L2, CACNA2D4, PTPRU, HFCFC1R1, PPF18P2, RBM44, HLA-C, LOH3CR2A, OCEL1, etc.
Figure 7

Gene group C

Gene group D

Gene group E

Gene group F

A

B

C

D

iPS Fibroblasts

Control IR-Mut Control IR-Mut

IRS1 mRNA

IRS1 mRNA

Gene group C

Gene group D

Gene group E

Gene group F
SUPPLEMENTARY INFORMATION

SUPPLEMENTARY FIGURES

A

B

C

D

E

F
Supplementary Figure 1. Quantification of western blot analysis of fibroblasts (related to Figure 1)

(A-F) Quantification of western blot analysis for control and IR-Mut fibroblasts. Data are expressed as mean ± SEM, relative to average of controls (n=2).
Supplementary Figure 2. Pluripotency markers in control and IR-Mut iPSC (related to Figure 1)

(A) Expression of nuclear NANOG, OCT4 and cytoplasmic SSEA4, TRA1-60 in iPSC. Images show merged colors between blue (Hoechst) and red (indicated protein). Scale bar represents 100 µM.

(B). Alkaline phosphatase activity was measured using a colorimetric kit. Images are representative of 3 wells.
Supplementary Figure 3. Teratoma formation and karyotyping of iPSC (related to Figure 1)

(A) iPSC were injected into SCID mice for in vivo teratoma formation. All three germ layers were observed: ectoderm (pigmented cells and primitive neural tissue), mesoderm (cartilage and smooth muscle), and endoderm (primitive gut and respiratory tissue). Representative images are shown.

(B) Karyotyping analysis revealed normal karyotype for all six iPS lines.
Supplementary Figure 4. mRNA and protein expression key insulin signaling components in iPSC (related to Figure 2)

(A) iPSC were serum starved for 3 hours before 10-minute stimulation with 0, 10, or 100 nm insulin. Western blot for mature INSR and proreceptor expression, representative of 3 independent experiments.

(B) Quantification of western blot analysis for proreceptor expression. Data are expressed as the ratio of the expression of the proreceptor to the mature receptor (n=3).

(C) Quantification of western blot analysis for IGF1R expression. Data are expressed as percentage of the maximum value (n=3).
(D) mRNA expression for key insulin signaling molecules was analyzed by qRT-PCR using specific primers, as indicated. Data are expressed as fold change relative to the average of the control iPSC (n=3).

(E) iPSC were serum starved for 3 hours before 5-minute stimulation with 0 or 100 nm insulin. Immunoprecipitation of INSR using a β-subunit-specific anti-INSR antibody and immunoblotting with an anti-phospho INSR/IGF1R (left panel) and anti-INSR antibody (right panel) was performed. Representative blots are shown (n=2).

All values represent mean ± SEM. * p<0.05.
Supplementary Figure 5. Quantification of ERK protein expression and relative phosphorylation (related to Figure 3)

Quantification of western blots for expression and phosphorylation of (A, C) ERK1, and (B, D) ERK2 protein. Data are expressed as mean ± SEM, relative to average of controls (n=3). * p<0.05.
Supplementary Figure 6. Attachment assay of iPSC (related to Figure 4)

Attachment assay of iPSC; cells were plated onto matrigel-coated dishes and counted 1 hour post seeding by Nexcelom automatic cellometer.
**Supplementary Table 1. DAVID – pathway analysis of iPS- and fibroblasts-enriched gene groups (related to Figure 5)**

DAVID pathway analysis of the top 100 mRNAs of iPS- and fibroblast-enriched groups (log$_2$ FC $>0$ or $<5$). The table shows pathways significantly overrepresented (p<0.001, q<0.001), and indicates example of genes in the pathways and the percentage of differentially expressed genes found in each enriched pathway. Pathways common to the two groups are highlighted in bold.
Supplementary Table 2. Transcription factor motif analysis of iPS- and fibroblast-enriched gene groups (related to Figure 5)

We utilized the molecular signatures database (C3, GenePattern) to identify transcription factor binding site motifs over represented in putative 2 kb upstream promoters (enrichment threshold p<0.001) for the top 100 iPS- and fibroblast-enriched probes (log2 FC > or < 5). Examples of genes responsible for the enrichment and the percentage of promoters with these motifs are provided. Putative regulatory transcription factors common to both fibroblast- and iPS–enriched gene groups are highlighted in bold.

<table>
<thead>
<tr>
<th>TF pattern in gene group</th>
<th>TF motif p&lt;0.001</th>
<th>Example of regulated genes</th>
<th>% of genes regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fibroblasts-enriched</td>
<td>MLL17 (FOXO4)**</td>
<td>TFAP2A, TNFRSF19</td>
<td>26%</td>
</tr>
<tr>
<td></td>
<td>MAZ</td>
<td>MEIS1, SNAP25, PPARG, COL18A1</td>
<td>24%</td>
</tr>
<tr>
<td></td>
<td>LEF1*</td>
<td>SALL1, COL18A1, PPARG</td>
<td>21%</td>
</tr>
<tr>
<td></td>
<td>TAF+</td>
<td>ABCA1, MEIS1, PAX3, SNAP25, PPARG</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>NFAT/NFATC**</td>
<td>SNAP25, ALDH1A1, MEIS1</td>
<td>14%</td>
</tr>
<tr>
<td></td>
<td>FOXA1**</td>
<td>MEIS1, SNAP25, MAB21L1</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td>GATA1*</td>
<td>ALDH1A1, ZFPM2</td>
<td>10%</td>
</tr>
<tr>
<td>iPS-enriched</td>
<td>LEF1*</td>
<td>SALL2, ZIC3, SEMA6A</td>
<td>23%</td>
</tr>
<tr>
<td></td>
<td>TCF3**</td>
<td>SEMA6A, ANK3, ZIC2-3, GPM6B, LRRN1</td>
<td>22%</td>
</tr>
<tr>
<td></td>
<td>MLL17 (FOXO4)**</td>
<td>NTS, ZIC2-5, FOXD3</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>NFAT/NFATC**</td>
<td>GPR98, ADCY2, USP44</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>PAX4**</td>
<td>ZIC2-5, SEMA6A, POUF5F1, GPM6B</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>PUOF1F1'</td>
<td>ZIC2-5, SEMA6A, OTX2, FOXD3, OTX2</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td>MEIS1**</td>
<td>OTX2, ZIC5, GCNT2, SCNN1A</td>
<td>5%</td>
</tr>
</tbody>
</table>

*TF involved in glucose homeostasis/insulin action
+TF Involved in ES/IPSC physiology
Bold: shared transcription factor
Supplementary Table 3. Transcription factor motif analysis of gene groups A and B (related to Figure 6)

Transcription factor motif analysis (C3, mSig database) analyzed putative promoter (2000 base pairs upstream the transcription start site) of top 100 statistically significant (nominal p<0.05) mRNAs of groups A and B. The table shows significant transcription factor motifs (p<0.001), example of regulated genes in the groups and the percentage of genes regulated by the transcription factor. Transcription factors found common between the two gene groups are highlighted in bold.
Supplementary Table 4. DAVID pathway analysis of groups C, D, E, and F (related to Figure 7)

Results of DAVID pathway analysis of probes from groups C, D, E, and F (all with log₂ FC > or < 1), including significantly enriched pathways (p<0.001, q<0.001), example genes, and the percentage of genes in the pathway. The pathways common to fibroblasts and iPS comparisons are highlighted in bold.
Supplementary Table 5. Transcription factor binding motif analysis of gene groups C, D and E, F
(related to Figure 7)

Transcription factor motif analysis (mSig database) analyzed putative promoter (2000 base pairs upstream the transcription start site) of mRNAs of groups C, D and E, F (all with log2 FC > or < 1).

The table shows significant motifs for transcription factors (p<0.001), example of regulated genes and the percentage of putative regulated genes that are recognized by the mSig database algorithm. Transcription factors found common between the two gene groups are highlighted in bold.