Meta-analysis of gene expression in human pancreatic islets after in vitro expansion

B. Kutlu,1 A. G. Kayali,2 S. Jung,3 G. Parnaud,4 D. Baxter,1 G. Glusman,1 N. Goodman,1 L. A. Behie,3 A. Hayek,2 and L. Hood1

1Institute for Systems Biology, Seattle, Washington; 2Whittier Institute/Department of Pediatrics, University of California at San Diego, La Jolla, California; 3The University of Calgary, Calgary, Canada; and 4Centre Médical Universitaire, Geneva, Switzerland

Submitted 7 April 2009; accepted in final form 16 July 2009

Meta-analysis of gene expression in human pancreatic islets after in vitro expansion. Physiol Genomics 39: 72–81, 2009. First published July 21, 2009; doi:10.1152/physiolgenomics.00063.2009.—Pancreatic islet transplantation as a potential cure for type 1 diabetes (T1D) cannot be scaled up due to a scarcity of human pancreas donors. In vitro expansion of β-cells from mature human pancreatic islets provides an alternative source of insulin-producing cells. The exact nature of the expanded cells produced by diverse expansion protocols and their potential for differentiation into functional β-cells remain elusive. We performed a large-scale meta-analysis of gene expression in human pancreatic islet cells, which were processed using three different previously described protocols for expansion and for which redifferentiation was attempted. All three expansion protocols induced dramatic changes in the expression profiles of pancreatic islets; many of these changes are shared among the three protocols. Attempts at redifferentiation of expanded cells induce a limited number of gene expression changes. Nevertheless, these fail to restore a pancreatic islet-like gene expression pattern. Comparison with a collection of public microarray datasets confirmed that expanded cells are highly comparable to mesenchymal stem cells. Genes induced in expanded cells are also enriched for targets of transcription factors important for pluripotency induction. The present data increase our understanding of the active pathways in expanded and redifferentiated islets. Knowledge of the mesenchymal stem cell potential may help development of drug therapeutics to restore β-cell mass in T1D patients.

In this study, we have tested three different protocols to expand human pancreatic islets in monolayer, followed by maneuvers to attempt their redifferentiation back to islets. We have characterized the resulting cells in detail by performing microarray analyses with fresh pancreatic islets, expanded islet cells (EXP), and redifferentiated cells. Genes modified by either of three protocols have 70–80% overlap with the genes changed by the other two protocols. Although there are promising changes in the right direction, none of the three protocols could achieve a return to a functional islet state. The expanded cells highly resemble mesenchymal stem cells (MSC), and similar gene regulatory networks seem to be active in both cell types. On the other hand, the EXP are different from MSC in that they seem to retain activity of some islet gene modules. The current results highlight the importance of designing new strategies that take into account the MSC potential of expanded cells.

MATERIALS AND METHODS

Cells

Human pancreatic islets. Human islets were provided through the Juvenile Diabetes Research Foundation Islet Distribution Program by Islet Cell Resource Centers in Geneva (Switzerland), Milan (Italy), and Lille (France). Human islets were cultured in CMRL 1066 supplemented with 10% FCS, 1% glutamine, 5.6 mM glucose, 1 mM HEPES, 110 U/ml penicillin, and 110 μg/ml streptomycin. Islets were obtained from five donors (2 male and 3 female), aged between 37 and 70 yr and body mass index between 22 and 27. We performed microarrays for a subset of islet preparations (2) to measure the effects of culture on normal islet expression profiles. All islet donor information can be found in the Supplementary Document1 (Supplementary Table S1).

Throughout the text, the terms “differentiation” or “redifferentiation” refer to the procedure to obtain functional pancreatic islet cells.

Whittier Protocol

Human pancreatic islets. The Islet Cell Resource Center Basic Science Human Islet Distribution Program and the Islet Transplant Program (University of Illinois at Chicago) provided us with preparations from six donors (3 male and 3 female), aged 25–55 yr, 60–90% pure in mature islets, and with a viability of 80–90%.

Expansion phase. Islets were expanded as previously described (27): 1,000 islets of 50–150 μm in diameter purified by hand-picking after dithizone staining were partially dissociated using Versene (Invitrogen, Carlsbad, CA) (3), to separate “outer” and “inner” populations (5). Removing the outer population minimized the presence of fibroblasts. Cell clusters from the inner population were plated on...
HTB-9 matrix-coated dishes in RPMI-1640 (Mediatech, Herndon, VA) supplemented with 2 mM L-glutamine, 10% FBS, and 25 ng/ml HGF. After confluence, cells were harvested using Versene containing 0.025% trypsin and subcultured (1:2).

Differentiation phase. After four passages (~1 mo expansion), cells were dispersed with Versene and cultured in the same serum-free media as in the National Institutes of Health (NIH) protocol on tissue culture-treated six-well plates, with or without HTB-9 matrix. After 1 wk of culture on HTB-9-coated plates, cells were harvested and forced to reaggregate overnight (40).

Pharmaceutical Production Research Facility Protocol

Human pancreatic islets. A population of pancreatic islets isolated from a cadaveric donor was obtained from the laboratory of Dr. Lawrence Rosenberg (University of McGill, Montreal, Quebec, Canada). The population was 70% pure in mature islets, which was determined by dithizone staining.

Expansion protocol. Approximately 2,500 islet equivalents (IEQs) were seeded onto 75-cm² fibronectin-coated tissue culture-treated flasks (BD Biosciences) in a proprietary serum-free medium [Pharmaceutical Production Research Facility (PPRF) medium] developed in the PPRF laboratory. The serum-free PPRF medium was originally designed for human bone marrow-derived MSC and further supplemented with 30% serum-free mesenchymal-conditioned medium. The addition of mesenchymal-conditioned medium played a significant role on the expansion of sparse pancreatic precursor-like cells (7). The majority of islets attached to the surface of the flasks within 48 h, and then the adherent cells started migrating out from the islets. When the cell migration resulted in the formation of large-size colonies, the cells were harvested by trypsinization and replated into new tissue culture plates. Once the cells reached near confluence (80–90%) in monolayer, they were trypsinized, counted, and subcultured at a density of 5,000 cells per cm². Thereafter, cells were serially passage using the same protocol.

Differentiation protocol. For induction of expanded cells to differentiate into islet-like cell clusters, the cells were seeded into 6.0-mm ultra-low attachment dishes (Corning) in serum-free CMRL-1066 supplemented with 4 mM L-glutamine, 1% BSA, insulin (10 µg/ml), transferrin (5.5 µg/ml), sodium selenite (6.7 ng/ml), and 1% antibiotic antimycotic solution at a density of 6.0 × 10⁶ cells per dish. The induction medium was further added with islet neogenesis-associated protein (INGAP) peptide at a concentration of 1.0 µg/ml. It has been previously shown that exposing cells to INGAP leads to the conversion of duct-like structures into islet-like structures (25). The plated individual cells formed clusters within 24 h and further cultured at 37°C and 5% CO₂ for a total of 8 days with medium replacement every other day. For the time-course experiments, multiple dishes were prepared, and each was harvested and stored at days 2, 4, 6, and 8 for microarray analysis.

NIH Protocol

For NIH expansion and differentiation protocols, we cultured the islets with the method of Gershengorn et al. (19, 27).

All experiments with human materials were inspected and approved by the respective institutions’ review boards.

Microarray Experiments and Analysis

Total RNA was isolated from expanded islets as well as redifferentiated expanded cells. Labeled RNA was processed and hybridized in the same facility to minimize technical variation. Additional microarray analyses were performed on functional pancreatic islets cultured for 7 days or immediately processed after isolation. Data were generated by hybridizing and scanning of labeled RNA to HG.U133 Plus 2.0 Affymetrix arrays. Low-level processing, normalization, and statistical analysis of microarray data were performed using R/Bioconductor packages (18). Alternative probe-set annotations (version 11, released Nov. 12, 2008) were used instead of the manufacturer annotations to remove redundant and ambiguous probes (10). Robust multiarray average algorithm was used to normalize microarray data. Differential expression analysis was assessed by fitting linear models to each gene profile and goodness of fit was used to rank genes.

Pathway and Process Analysis

Putatively active pathway and processes were determined using the public sources GOstats (17) and GSEA (43) and the commercial GeneGO (35) tool. All raw and processed data can be accessed at the National Center for Biotechnology Information-Gene Expression Omnibus (NCBI-GEO) repository (GSE15543).

---

**Fig. 1. Outline of samples analyzed for the microarray experiments. Microarray analyses were performed with functional pancreatic islet cells (kept in culture or uncultured), EXP (WHI-EXP, NIH-EXP, PPRF-EXP), and redifferentiated cells (WHI-DIF, NIH-DIF, PPRF-DIF) obtained with 3 protocols (Whittier, NIH, and PPRF). The redifferentiated cells from PPRF protocol were collected at 2, 4, 6, and 8 days after the differentiation procedure. We compared expanded islets cells vs. normal islet cells and redifferentiated cells vs. normal islets or expanded islets. EXP, expanded islet cells; DIF, redifferentiated cells; protocols: WHI, Whittier; PPRF, Pharmaceutical Production Research Facility; NIH, National Institutes of Health.**
**Data Meta-analysis**

Public microarray datasets were queried using NextBio, a database of microarray results (accessed on December 12, 2008). NextBio is a repository of analyzed microarray datasets that allows the investigator to search results and the expression profiles of publicly available microarray datasets. Microarray results from 98 human samples were obtained from relevant datasets deposited at NCBI-GEO. These data were generated using the same or similar platforms (Affymetrix Hgu133 plus 2.0 or Hgu133 A & B). A subset of the following datasets containing data from MSC or embryonic stem cells (ESC) were used for meta-analysis: GSE2248, GSE6029, GSE6460, GSE7234, GSE7332, GSE7637, GSE7879, GSE7888, GSE8113, GSE9520, GSE9865, GSE9894, GSE9921, GSE9940, GSE9941, GSE10315, GSE10435, GSE10438, GSE11310, and GSE11350. Multi-dimensional scaling was used to visualize the similarities and differences across all the samples from different cell types.

**Significant Overlap Analysis of Genes Expressed In Expanded Cells to Normal Pancreatic Islets**

We identified human islet cell-specific genes based on massively parallel signature sequencing (MPSS) analysis of beta-cell gene expression (28) compared with other human tissues. First, we combined the two human MPSS samples for human islet cells into a composite file and computed transcript per million (tpm) values for the genes observed in it. Second, we obtained MPSS raw data files for 32 human tissues from GEO (28) including adrenal gland, bladder, bone marrow, monocytes, peripheral blood lymphocytes, heart, kidney, liver, lung, mammary gland, pancreas, pituitary gland, placenta, prostate, retina, salivary gland, small intestine, spinal cord, spleen, stomach, testis, thymus, thyroid, trachea, uterus, and seven samples derived from brain tissues (amygdala, hypothalamus, cerebellum, corpus callosum, thalamus, caudate nucleus, and whole fetal brain). We combined the data in the seven brain samples into a single composite file representing human brain gene expression and computed tpm values for all samples. Third, we compared gene expression in the composite islet cell file to the other tissues, using the normalized tpm values. We computed for each gene its total tpm as the sum of tpm values over all samples and then computed the fraction of that expression in islet cells, retaining those genes for which islet cells express at least 50% of the total tpm. In addition, we performed a multiple regression analysis over the full matrix of all samples vs. all observed genes. We computed, for each gene in each sample, the ratio of its observed tpm to the expected tpm based on the multiple regressions, as well as an enrichment Z-score. Using a cutoff Z-score of 2 (in addition to the 50% fractional expression) identified 379 human genes specifically enriched in islet cells, and a more stringent cutoff Z-score of 3 retained 73 islet cell specific genes.

**Protein-Protein Interaction Network Analysis**

We prepared a database of mammalian protein-protein interactions (PPIs) using five sources that contain literature-curated physical PPIs: Human Protein Reference Database (HPRD), Molecular Interaction database (MINT), IntAct, MIPS Mammalian Protein-Protein Interaction Database, and Biological General Repository for Interaction Datasets (BioGRID). The integrated database, excluding data from HPRD, is available for download in flat file format at T1Dbase (24): http://isb.t1dbase.org/downloads/data/interactions/. We are required to exclude HPRD because of licensing restrictions imposed by their funding source. Version information of downloaded data from each source can be in the Supplementary Document (Supplementary Table S2).

**RESULTS**

**Characterization of EXP**

Functional pancreatic islet cells were expanded and redifferentiated with three different protocols (Fig. 1): Whittier (WHI), NIH, and PPRF. The cells obtained by these expansion and redifferentiation protocols will be referred to as WHI-EXP, WHI-DIF, NIH-EXP, NIH-DIF, PPRF-EXP, and PPRF-DIF. We characterized in detail the doubling time of NIH-EXP and PPRF-EXP and confirmed that they produce the expected results. The range of average cell doubling time of NIH-EXP derived from three separate donors was between 38.3 and 52.6 h, whereas that of PPRF-EXP was between 35.3 and 41.1 h. The average doubling times of both expanded cells were maintained at least six passages.

<table>
<thead>
<tr>
<th>Total</th>
<th>Increased vs. Islets</th>
<th>Decreased vs. Islets</th>
</tr>
</thead>
<tbody>
<tr>
<td>WHI-EXP</td>
<td>1,236</td>
<td>582</td>
</tr>
<tr>
<td>NIH-EXP</td>
<td>1,370</td>
<td>699</td>
</tr>
<tr>
<td>PPRF-EXP</td>
<td>1,242</td>
<td>732</td>
</tr>
</tbody>
</table>

Protocol group abbreviations: WHI, Whittier; EXP, expanded islet cell; NIH, National Institutes of Health; PPRF, Pharmaceutical Production Research Facility.
To measure the “apparent average” cell doubling times, specific growth rates were calculated, assuming that the entire growth period per passage was under exponential growth phase (i.e., ignoring both the lag phase and any contact inhibited cell growth). Thus the growth rate was first order. Moreover, the average cell doubling time per passage (excluding the primary culture) was estimated based on cell numbers and the specific rate of replication.

In studies performed at the Hayek lab the doubling time of the NIH-EXP increased throughout the expansion phase. At passages 0, 1, 2, 3, and 4, the doubling times were 52.2, 95.86, 198.45, 208.13, and 290 h, respectively, as reported

### Table 2. Average gene expression signals in PPRF-EXP and PPRF-DIF in log2 scale

<table>
<thead>
<tr>
<th>Entrez Gene</th>
<th>Symbol</th>
<th>Description</th>
<th>PPRF-EXP</th>
<th>PPRF-DIF-2d</th>
<th>PPRF-DIF-4d</th>
<th>PPRF-DIF-6d</th>
<th>PPRF-DIF-8d</th>
</tr>
</thead>
<tbody>
<tr>
<td>3630</td>
<td>INS</td>
<td>insulin</td>
<td>6.69</td>
<td>7.54</td>
<td>8.07</td>
<td>8.35</td>
<td>8.86</td>
</tr>
<tr>
<td>2641</td>
<td>GCG</td>
<td>glucagon</td>
<td>4.84</td>
<td>5.64</td>
<td>5.21</td>
<td>6.05</td>
<td>6.54</td>
</tr>
<tr>
<td>26750</td>
<td>SST</td>
<td>somatostatin</td>
<td>6.52</td>
<td>7.24</td>
<td>7.34</td>
<td>7.63</td>
<td>7.66</td>
</tr>
</tbody>
</table>

The table contains average gene expression signals in PPRF-EXP and the time series of attempted redifferentiation using the PPRF protocol after 2, 4, 6, and 8 days (d). Although there is almost a 4-fold increase in insulin gene expression at the end of 8d of differentiation, the final value is ~70-fold less than signal intensities observed in wild-type islets. DIF, redifferentiated cell.
expressed at high levels in functional islets (28). E-cadherin
levels of cell cycle and mitotic genes that are normally not
NKX2–2, and PAX6. On the other hand, we observed an increase
creases in islet-enriched transcription factors ISL1, NEUROD1,
Supplementary Document). This is accompanied by strong de-
are downregulated (Fig. 4). In the expanded cells, insulin,
transcription factors. Gene expression changes induced in ex-
tary Document).
Fig. 5 and Table 2, also see Additional Files 4 – 6, Supplemen-
the direction of change in some key genes, such as insulin and
the same criteria used for expanded cells. On the other hand,
between the lists of genes deemed significantly altered using
profiles. However, the redifferentiation protocols induced mod-
entiated was attempted clearly showed that none of the proce-
found in the supplementary materials (Additional File 9).
Around 14% of genes differentially expressed in NIH-EXP
Table 1). The proportion of increased or decreased genes was
similar for the three protocols. The direction of change (in-
crease or decrease) was always consistent among the protocols.
14% of genes differentially expressed in NIH-EXP
were unique to the NIH protocol. Similarly, 15 and 30% of
modified genes were unique to the WHI-EXP and PPRF-EXP,
respectively. (Additional Files 1–3). Approximately 50% of the
genes detected as changed in each protocol were also modified
by the other two expansion methods. The full list of genes that
are inversely correlated with functional islet profiles can be
found in the supplementary materials (Additional File 9).
The gene expression profiles of the cells for which rediffer-
ated was attempted clearly showed that none of the proce-
dures succeeded in rescuing normal pancreatic islet expression
profiles. However, the redifferentiation protocols induced modest
changes in gene expression profiles. There was no overlap
between the lists of genes deemed significantly altered using
the same criteria used for expanded cells. On the other hand,
the direction of change in some key genes, such as insulin and
other islet hormones, is encouraging in some protocols (see
Fig. 5 and Table 2, also see Additional Files 4–6, Supplementary
Document).

Changes Induced by the Islet Expansion Process

Pancreatic β-cells are highly differentiated cells responsible
for regulating blood glucose concentrations. This is achieved
by the activation of β-cell-specific transporters, receptors, and
transcription factors. Gene expression changes induced in ex-
panding islet cells suggest that genes involved in islet functions
are downregulated (Fig. 4). In the expanded cells, insulin,
along with the other islet-specific hormones glucagon and
somatostatin, is decreased significantly (Supplementary Table S3,
Supplementary Document). This is accompanied by strong de-
creases in islet-enriched transcription factors ISL1, NEUROD1,
NKK2–2, and PAX6. On the other hand, we observed an increase
in levels of cell cycle and mitotic genes that are normally not
expressed at high levels in functional islets (28). E-cadherin
gene, an epithelial cell marker is decreased, while levels of
N-cadherin are increased in all three samples; this switch
suggests that cells undergo an epithelial to mesenchymal cell
transition (EMT). This switch was accompanied by a decrease
of epithelial cell markers (claudin-5, -7, -10, mucin-1, and
PECAM-1) and an increase in mesenchymal markers (fi-
bronectin, matrix metallopeptidase 2, disabled homolog 2,
Dickkopf homolog 1, transforming growth factor-β2) (14, 16,
20, 38, 46). Transcription factors that are active in gene
networks regulating the EMT process (TWIST2 and SNAI2)
were also increased in all EXP cells (Supplementary Table S4)
(45). Activin is a molecule involved in the proliferation or
differentiation of endocrine cell progenitors in the pancreas,
depending on the concentration that is present. Its elevated
expression in all the EXP may suggest that these cells indeed
have progenitor potential (4). Bone morphogenic proteins
(BMPs) are pleiotropic proteins that regulate proliferation,
differentiation, and migration of several cell types. They have
been shown to regulate neuronal stem cell niches (22) and limit
precursor number. Noggin, an inhibitor of the BMPs, is elev-
ated in all three groups of expanded cells, again suggesting
that progenitor precursor regulatory pathways are active in
these cells (8).

Pathways and Processes Active in Expanded Cells

To identify the active processes in expanded cells, we
performed GO term enrichment analysis of genes significantly
altered in response to expansion of islets. Development and
regeneration processes are significantly active in expanded
cells (Table 3). We also identified active pathways involved in
development as well as cell adhesion and cell cycle (Table 4).
Of note, two pathways related to EMT were also found to be
significantly activated in EXP.

Overlap of Genes Active in EXP With Publicly Available
Microarray Studies

Querying published microarray datasets showed that genes
active in EXP appear also to be active in MSC or ESC. Given
Table 3. GO terms enriched for the list of active genes
in expanded islets compared with functional islets

<table>
<thead>
<tr>
<th>Process</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anatomical structure morphogenesis</td>
<td>1.32E-12</td>
</tr>
<tr>
<td>Organ development</td>
<td>2.10E-12</td>
</tr>
<tr>
<td>Anatomical structure development</td>
<td>1.50E-10</td>
</tr>
<tr>
<td>System development</td>
<td>1.80E-10</td>
</tr>
<tr>
<td>Organ morphogenesis</td>
<td>4.82E-10</td>
</tr>
<tr>
<td>Multicellular organisinal development</td>
<td>2.31E-09</td>
</tr>
<tr>
<td>Developmental process</td>
<td>9.82E-09</td>
</tr>
<tr>
<td>Wound healing</td>
<td>4.23E-08</td>
</tr>
<tr>
<td>Cell motion</td>
<td>7.31E-08</td>
</tr>
<tr>
<td>Localization of cell</td>
<td>7.31E-08</td>
</tr>
<tr>
<td>Cellular developmental process</td>
<td>1.03E-07</td>
</tr>
<tr>
<td>Regulation of biological quality</td>
<td>1.11E-07</td>
</tr>
<tr>
<td>Actin filament-based process</td>
<td>1.41E-07</td>
</tr>
<tr>
<td>Regeneration</td>
<td>2.06E-07</td>
</tr>
<tr>
<td>Embryonic skeletal morphogenesis</td>
<td>2.96E-07</td>
</tr>
<tr>
<td>Organ regeneration</td>
<td>4.09E-07</td>
</tr>
<tr>
<td>Cell cycle phase</td>
<td>9.73E-07</td>
</tr>
<tr>
<td>Regulation of phosphate metabolic process</td>
<td>1.14E-06</td>
</tr>
<tr>
<td>Regulation of phosphorus metabolic process</td>
<td>1.14E-06</td>
</tr>
<tr>
<td>Actin cytoskeleton organization</td>
<td>1.25E-06</td>
</tr>
</tbody>
</table>

P values were calculated by hypergeometric tests and indicate the probability of finding the term enriched by chance. GO, Gene Ontology.
the poor redifferentiation potential of the EXP, we wanted to explore further the nature of these cells. For this purpose, we performed NextBio (www.nextbio.com) queries that allowed us to identify microarray study results that have significant similarities with the list of activated genes common to all EXP. The list of genes that are increased in all expanded cell types vs. islets has highly significant overlap with studies that were mostly performed with either MSC or ESC (Table 5). In all cases, the correlation with various types of stem cells suggests that expanded cells have common gene modules active in these progenitor cells.

**Meta-analysis of EXP with MSC and ESC**

Based on the results of microarray database query, we performed an unbiased meta-analysis of the present set of data with microarray data from several studies from human MSC and ESC (see MATERIALS AND METHODS). We included in our analysis ~100 arrays that were performed with the same or very similar Affymetrix human chips. The scatter plot of mean whole microarray gene expression values in expanded cells vs. either MSC or ESC indicate that the former cells are closer to EXP. Indeed, the Pearson correlation coefficient of mean expanded gene expression profiles vs. MSC is 0.967, while it is lower 0.872 for expanded vs. ESC (Supplementary Fig. S1). We used multidimensional scaling to visualize this complex data set (Fig. 5). This visualization suggests that expanded cells are more similar to MSC than to ESC, and that pancreatic islet profiles are different from all other cell types studied. When included in the visualization microarray data from liver (representing an unrelated cell type), the expanded cells clustered together with other stem cells and away from islets or liver cells (data not shown). When we limited the analysis to the

### Table 4. Active pathways in expanded cells

<table>
<thead>
<tr>
<th>GeneGO Maps</th>
<th>Cell Process</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell adhesion: role of tetraspanins in the integrin-mediated cell adhesion</td>
<td>cell adhesion</td>
<td>1.47E-05</td>
</tr>
<tr>
<td>Cytoskeleton remodeling: TGF, WNT, and cytoskeletal remodeling</td>
<td>cell differentiation</td>
<td>0.000161</td>
</tr>
<tr>
<td>Cell cycle: chromosome condensation in prometaphase</td>
<td>cell cycle</td>
<td>0.000182</td>
</tr>
<tr>
<td>Development: TGF-β-induced EMT via SMADs</td>
<td>intracellular receptor-mediated signaling pathway</td>
<td>0.000310</td>
</tr>
<tr>
<td>Cytoskeleton remodeling: cytoskeleton remodeling</td>
<td>cell adhesion</td>
<td>0.00518</td>
</tr>
<tr>
<td>Cell adhesion: integrin-mediated cell adhesion and migration</td>
<td>cell adhesion</td>
<td>0.000625</td>
</tr>
<tr>
<td>Cytoskeleton remodeling: integrin outside-in signaling</td>
<td>cell adhesion</td>
<td>0.000645</td>
</tr>
<tr>
<td>Cell adhesion: endothelial cell contacts by nonjunctional mechanisms</td>
<td>cell adhesion</td>
<td>0.000999</td>
</tr>
<tr>
<td>Cell adhesion: chemokines and adhesion</td>
<td>cell adhesion</td>
<td>0.001456</td>
</tr>
<tr>
<td>Cell cycle: nucleocytoplasmic transport of CDK/cyclins</td>
<td>cell cycle</td>
<td>0.001456</td>
</tr>
</tbody>
</table>

**Broad Institute Gene Set Pathway**

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Cell Process</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSA04110_CELL_CYCLE</td>
<td>cell cycle</td>
<td>0</td>
</tr>
<tr>
<td>GIPATHWAY</td>
<td>cell cycle</td>
<td>0</td>
</tr>
<tr>
<td>HSA04115_P53_SIGNALING_PATHWAY</td>
<td>cell proliferation</td>
<td>0</td>
</tr>
<tr>
<td>ST_WNT_BETA_CATENIN_PATHWAY</td>
<td>cell differentiation</td>
<td>0</td>
</tr>
<tr>
<td>ARFPATHWAY</td>
<td>cell cycle arrest</td>
<td>0</td>
</tr>
<tr>
<td>FASPATHWAY</td>
<td>cell cycle arrest</td>
<td>0</td>
</tr>
<tr>
<td>HSA04510_FOCAL_ADHESION</td>
<td>cell adhesion</td>
<td>0</td>
</tr>
<tr>
<td>ATMPATHWAY</td>
<td>cell cycle</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 5. List of study results that have significant overlap with studies performed with mesenchymal and embryonic stem cells

<table>
<thead>
<tr>
<th>Study Name (Ref. no.)</th>
<th>P Value</th>
<th>Cell Type</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Differences between differentiated and undifferentiated ESC (39)</td>
<td>3.6E-44</td>
<td>MSC</td>
<td>MSC-derived hESC (VUB01/SA01)</td>
</tr>
<tr>
<td>Human bone marrow MSC (13)</td>
<td>4.7E-42</td>
<td>MSC</td>
<td>Bone marrow MSC in culture</td>
</tr>
<tr>
<td>Gene expression profile during human CD4+ T cell differentiation (30)</td>
<td>8.9E-37</td>
<td>Thymic stromal</td>
<td>Thymic stromal cells</td>
</tr>
<tr>
<td>Reprogramming of human somatic cells to pluripotency with defined factors (37)</td>
<td>1.5E-29</td>
<td>Induced pluripotent</td>
<td>Fetal lung fibroblasts</td>
</tr>
<tr>
<td>Generation of pluripotent stem cells from adult human testis (9)</td>
<td>3.2E-29</td>
<td>Induced pluripotent</td>
<td>Human adult germ stem cells</td>
</tr>
<tr>
<td>Human MSC (2)</td>
<td>3.9E-22</td>
<td>MSC</td>
<td>Human ES cell-derived mesenchymal precursor</td>
</tr>
<tr>
<td>Murine ESC, neural precursor cells, and embryonic fibroblasts (32)</td>
<td>1.4E-21</td>
<td>ESC</td>
<td>Mouse embryonic fibroblasts</td>
</tr>
<tr>
<td>Expression data from cultured human ESC (44)</td>
<td>1.1E-14</td>
<td>MSC</td>
<td>Undifferentiated human ESC</td>
</tr>
<tr>
<td>Unique molecular signature of multipotent adult progenitor cells (47)</td>
<td>1.7E-14</td>
<td>ESC</td>
<td>Mouse marrow stromal cells</td>
</tr>
<tr>
<td>Cord-blood derived MSC populations (31)</td>
<td>5.5E-14</td>
<td>MSC</td>
<td>UCB-derived MSC</td>
</tr>
<tr>
<td>Expression studies on MSC (29)</td>
<td>1.4E-10</td>
<td>MSC</td>
<td>MSC grown for 7 days</td>
</tr>
<tr>
<td>Carcinoma-associated fibroblast-like differentiation of human MSC (33)</td>
<td>1.6E-8</td>
<td>MSC</td>
<td>Bone marrow-derived human MSC</td>
</tr>
<tr>
<td>Transformed glial progenitor cells (12)</td>
<td>2.6E-8</td>
<td>ESC</td>
<td>Embryonic progenitor SC</td>
</tr>
<tr>
<td>Gene regulatory networks along the small intestinal crypt-villus axis (42)</td>
<td>41.0E-7</td>
<td>MSC</td>
<td>Mouse E12 mesenchyme</td>
</tr>
</tbody>
</table>

Top 14 microarray study results that have significant overlap with the list of expanded genes are listed. P value is a hypergeometric test based score of significant overlap with the comparison list. Tissue type refers to the type of cell the study was performed with. Correlation: cell type that was correlated with the expanded cells. MSC, mesenchymal stem cell; ESC, embryonic stem cell.
subset of genes whose expression changed significantly in the three expansion protocols, we observed similar results as when studying the full gene expression data set.

We next examined the overlap of gene signatures in EXP with MSC or ESC. This analysis showed that EXP and MSC share 902 genes (up- or downregulated vs. ISL) that are not regulated (or changed in the opposite direction) in ESC. Similarly, 917 genes are shared by DIF and MSC but remain unchanged in ESC. There is a large overlap (>81%) between the two lists (Additional File 10). There was little (<10% of MSC) overlap between EXP/DIF and ESC. The overlap analysis strongly suggests the similarity between expression profiles EXP and MSC (Table 6).

Interestingly, insulin, other islet hormone genes (glucagon, somatostatin and pancreatic polypeptide), and other genes abundantly or specifically expressed in pancreatic islets (such as islet amyloid polypeptide, ISL LIM homeobox 1, neurogenic differentiation 1, insulinoma-associated 1) are among the top genes that are preferentially expressed by more than fourfold in expanded vs. mesenchymal cells (see Additional File 7). This list has a significant overlap either with genes expressed robustly in human pancreatic islets (top 5% by abundance, including islet-specific and housekeeping genes) identified by Beta Cell Gene Atlas (28) or genes that are expressed preferentially in pancreatic islets (Z = 2, see MATERIALS AND METHODS) compared with other human tissues (hypergeometric test of overlap, P values <9.9E-04 and 9.6E-11, respectively). The overlap is more significant when using a more stringent criterion (Z = 3) for pancreatic islet specificity (P value 1.4E-12).

Analysis of the PPI Networks Induced by Expansion Protocols

Finally, we investigated the biological networks active in expanded cells by integrating the physical PPI datasets with the transcriptome data of EXP. We identified a core network of 121 genes that are overlapping in EXP and MSC (see the previous section) (Fig. 6). The same network analysis with genes shared by EXP and ESC did not return any interactions. The GO terms associated with the network nodes are strongly enriched (hypergeometric P value between 1.5E-06 and 5.1E-14 for the top 10 GO terms) for terms related to organ development, morphogenesis, and cell adhesion (Supplementary Table S5).

DISCUSSION

Redifferentiation of pancreatic β-cells after expansion in vitro represents an alternative source of pancreatic islet cells for transplantation. In this study, we have shown that EXP produced by different protocols are genetically more similar to MSC than to cells from unperturbed islets. Thus, the present data represent an important step in characterization of the expanded cells in terms of exploring their differentiation potential to functional pancreatic islets.

Phenotypic transition of nondividing functional islet cells to expanding cells leads to large changes involving increase or decrease in levels of expression of ∼1,250 genes, most of which were shared by two or three of the protocols tested. Our data analyses indicate that there is considerable overlap among the changes induced by each of the three protocols we tested. The direction of changes in gene expression was always

Table 6. Top overlapping genes in EXP and MSC, but not in ESC

<table>
<thead>
<tr>
<th>Entrez Gene Id</th>
<th>Symbol</th>
<th>Description</th>
<th>MSC vs. ISL</th>
<th>ESC vs. ISL</th>
<th>EXP vs. ISL</th>
<th>DIF vs. ISL</th>
</tr>
</thead>
<tbody>
<tr>
<td>3885 KRT34</td>
<td>keratin 34</td>
<td></td>
<td>6.62</td>
<td>0.33</td>
<td>7.01</td>
<td>2.82</td>
</tr>
<tr>
<td>54845 RBM35A</td>
<td>RNA binding motif protein 35A</td>
<td></td>
<td>-8.08</td>
<td>-1.31</td>
<td>-6.09</td>
<td>-6.57</td>
</tr>
<tr>
<td>5806 PTX3</td>
<td>pentrax-related gene, 3</td>
<td></td>
<td>7.40</td>
<td>0.17</td>
<td>5.93</td>
<td>4.37</td>
</tr>
<tr>
<td>4015 LOX</td>
<td>lysyl oxidase</td>
<td></td>
<td>5.64</td>
<td>-0.64</td>
<td>5.80</td>
<td>4.72</td>
</tr>
<tr>
<td>151887 CCDC80</td>
<td>coiled-coil domain containing 80</td>
<td></td>
<td>6.49</td>
<td>0.06</td>
<td>5.76</td>
<td>5.18</td>
</tr>
<tr>
<td>117247 SLC16A10</td>
<td>solute carrier family 16, member 10</td>
<td></td>
<td>-5.25</td>
<td>1.90</td>
<td>-5.41</td>
<td>-5.23</td>
</tr>
<tr>
<td>9358 ITGBL1</td>
<td>integrin, β-like 1 (with EGF-like repeat domains)</td>
<td></td>
<td>7.01</td>
<td>-2.46</td>
<td>5.29</td>
<td>5.24</td>
</tr>
<tr>
<td>6423 SFRP2</td>
<td>secreted frizzled-related protein 2</td>
<td></td>
<td>-5.07</td>
<td>1.79</td>
<td>-5.27</td>
<td>-4.06</td>
</tr>
<tr>
<td>389336 C5orf46</td>
<td>chromosome 5 open reading frame 46</td>
<td></td>
<td>2.19</td>
<td>0.06</td>
<td>5.19</td>
<td>2.69</td>
</tr>
</tbody>
</table>

ISL, islet.
consistent between the three protocols. This finding is in line with previous observations with HL60 progenitor cell differentiation, where cells can follow divergent trajectories to reach a final attractor state (23). This raises the intriguing possibility that pancreatic islets eventually reach a stable expansion attractor state, although individual protocols cause different initial changes. It is also plausible that this might be true for at least a subset of cells that we analyzed. Measurements of expression at the single cell level are likely to provide the ultimate explanation to these observations (21).

Functional analysis of expanded cell data indicates that biological processes involved in “development” and “regeneration” are active in these cells. There is considerable debate in the literature on whether the β-cell expansion process is an EMT (19, 36) as opposed to other accompanying cells taking over the cell culture. In our analysis, EMT genes and related pathways were detected at significant levels in expanded cells (Tables 2 and 3). However, EMT processes have not been observed in mouse islet expansion studies (1, 11, 48). It appears that some of the controversy regarding the feasibility of expansion and redifferentiation of islet cells has been generated by the discrepancies between the mouse and human systems.

To further characterize the expanded cells, we performed a data-driven analysis that involved exploiting the wealth of microarray results in the public domain. Meta-data analysis of
the EXP with a comprehensive set of microarray data from ESC and MSC has shown that expanded cells are closer to MSC. Of note, Davani et al. (11) have previously shown that cells obtained with NIH protocol do have MSC characteristics based on cell surface markers. Our data confirm that this is also the case in cells obtained with two other expansion protocols. In addition, we have identified the pathways and biological processes active in the expanded cells that will be useful for discovering methods that succeed in redifferentiating expanded cells.

Pancreatic islet-specific genes seem to remain expressed in expanded cells albeit at lower levels. These results suggest that either some pancreatic β-cell gene modules are still active in expanded cells or a subpopulation of the original β-cells gain the ability to divide and proliferate. In agreement with the second possibility, lineage trace experiments have demonstrated that cultured human β-cells can dedifferentiate and proliferate in vitro (41). Moreover, the epigenetic structure of insulin gene promoter was demonstrated to keep its active state in expanded cells (34). Attempted redifferentiation of expanded cells is so far not able to bring the gene expression back to the levels present in functional pancreatic islets. Thus, EXP provided the correct stimulus yet to be determined, might still have the potential to differentiate into functional β-cells.

The PPI and gene regulatory network analyses provide insightful summaries of the underlying mechanisms of islet cell expansion. We focused specifically on the gene modules active in EXP and MSC. We have identified an interacting core of 121 genes that are strongly enriched in organ development and morphogenesis (Supplementary Table S5). Moreover, our analysis of the transcription factor networks involved in islet cell expansion process shows that targets of transcription factors and signaling proteins important in pancreas differentiation (SMAD 2/3/4), development (SOX4), and β-cell proliferation (FOXM1) are enriched significantly (Supplementary Table S6) (49, 50). Of note, transcription factors KLF4/5, SOX4, and c-Myc, which are necessary to reprogram skin cells to pluripotent cells, were also detected in the target analysis. However, we observed increased transcriptional activity for only a subset of these transcription factors (E2F3, FOXM1, and RBMS1) in all protocols. The latter group of transcription factors may have common targets with KLF4/5, SOX4, and c-Myc. It also remains to be seen whether the transcriptionally inactive regulators function at the protein level in the expanded cells.

We anticipate that our results will help devise strategies to differentiate the expanded cells into functional islets, whether by genetic modification or by a combination of culture media, matrices, and growth factors.

ACKNOWLEDGMENTS

We are grateful to Jeff Ranish and Marta Janer for critical reading of the manuscript and Buz Marzolf for help with microarrays and submission of data to NCBIG20.

REFERENCES

10. Duong TD, Erickson CA.