npg

ORIGINAL ARTICLE

The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming

Athanasia D Panopoulos[1, *], Oscar Yanes[2, 3, *], Sergio Ruiz[1], Yasuyuki S Kida[1], Dinh Diep[4], Ralf Tautenhahn[2], Aída Herrerías[5], Erika M Batchelder[1], Nongluk Plongthongkum[6], Margaret Lutz[7], W Travis Berggren[7], Kun Zhang[6], Ronald M Evans[1, 8], Gary Siuzdak[2], Juan Carlos Izpisua Belmonte[1, 5]

[1]*Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA;* [2]*Departments of Chemistry and Molecular Biology and Scripps Center for Metabolomics and Mass Spectrometry, The Scripps Research Institute, La Jolla, CA 92037, USA;* [3]*Metabolomics Platform of the Spanish Biomedical Research Center in Diabetes and Associated Metabolic Disorders (CIBERDEM), University Rovira i Virgili, 43007 Tarragona, Spain;* [4]*Bioinformatics and Systems Biology Graduate Program, University of California at San Diego, La Jolla, CA 92093, USA;* [5]*Center of Regenerative Medicine in Barcelona, Dr Aiguader, 88, 08003 Barcelona, Spain;* [6]*Department of Bioengineering, Institute for Genomic Medicine and Institute of Engineering in Medicine, University of California at San Diego, La Jolla, CA, 92093, USA;* [7]*Stem Cell Core, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA;* [8]*Howard Hughes Medical Institute, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA*

Metabolism is vital to every aspect of cell function, yet the metabolome of induced pluripotent stem cells (iPSCs) remains largely unexplored. Here we report, using an untargeted metabolomics approach, that human iPSCs share a pluripotent metabolomic signature with embryonic stem cells (ESCs) that is distinct from their parental cells, and that is characterized by changes in metabolites involved in cellular respiration. Examination of cellular bioenergetics corroborated with our metabolomic analysis, and demonstrated that somatic cells convert from an oxidative state to a glycolytic state in pluripotency. Interestingly, the bioenergetics of various somatic cells correlated with their reprogramming efficiencies. We further identified metabolites that differ between iPSCs and ESCs, which revealed novel metabolic pathways that play a critical role in regulating somatic cell reprogramming. Our findings are the first to globally analyze the metabolome of iPSCs, and provide mechanistic insight into a new layer of regulation involved in inducing pluripotency, and in evaluating iPSC and ESC equivalence.

Introduction

Seminal advances in stem cell biology demonstrated that somatic cells could be reprogrammed back into a pluripotent state (termed induced pluripotent stem cells, or iPSCs) by expression of defined transcription factors [1-3]. Numerous studies have examined the genetic and epigenetic profiles of iPSCs, in relation to both their somatic cells, and to embryonic stem cells (ESCs), to gain insight into somatic cell reprogramming, as well as the quantitative and qualitative differences that may exist between these pluripotent cell types [4, 5]. Yet, these important studies are not sufficient for generating a complete picture of the molecular components regulating cellular function.

Metabolism is either directly or indirectly involved with every aspect of cell function. Metabolomic technologies enable the examination and identification of endogenous biochemical reaction products, revealing information on the metabolic pathways and processes occurring within a living cell [6-8]. Here we examine the metabolomics of human iPSCs relative to ESCs and to their somatic cells of origin, which revealed the bioener-

getic changes that take place in somatic cell reprogramming, and their function in regulating the reprogramming process.

Results

The metabolome profiles of iPSCs versus ESCs

To examine the metabolic profile associated with induced pluripotency, we used an untargeted metabolomics approach with liquid chromatography coupled to electrospray ionization quadrupole time-of-flight MS (LC-ESI-QTOF-MS) [9] to analyze the relative abundance of metabolites in ESCs, iPSCs derived from human keratinocytes or fibroblasts, and their respective somatic cells of origin. All iPSCs used in this study (four fibroblast-derived iPSCs: FiPS4F1, FiPS4F2, FiPS4F4, FiPS4F5, and three keratinocyte-derived iPSCs: KiPS4FA, KiPS4FB and KiPS4F2) demonstrated pluripotent gene and protein expression, the ability to generate all three embryonic germ layers *in vivo* and a normal karyotype (Supplementary information, Data S1 and Figure S1) [10-12]. Our MS-based platform enabled us to observe greater than 5 000 metabolite features, defined as molecular entities with a unique mass/charge and retention time value. The relative abundance of metabolites for each cell population was then quantified by comparing the integrated area of each feature [9].

Previous studies have demonstrated that iPSC gene expression profiles and methylation patterns are influenced by prolonged culture *in vitro* [13-15]. For example, iPSCs that have undergone a low number of passages retain aspects of epigenetic memory from their somatic source, which influences their differentiation potential [14, 15]. Conversely, as iPSCs remain in culture and increase in passage number, they more closely genetically and functionally resemble ESCs [13, 15]. Thus, we first examined the global metabolic profile of iPSCs at early and late passages compared to ESC controls. Importantly, since culture conditions can influence the metabolite compositions within cells, and variations can occur between cell lines, we compared identical iPSCs at early (p.16) and late (> p.41) passages that had been grown in chemically-defined medium (i.e. mTeSR1) [16], to enable us to focus on metabolites that are more intrinsically distinct between cell types rather than culture-condition induced. While the global metabolic signature of early passage iPSCs was relatively close to ESC controls (5.3% difference), the metabolome signatures of late passage iPSCs grown in the same defined conditions was significantly closer to ESCs (0.23% difference). These data demonstrate that the metabolite profiles of iPSCs adapt to a more ESC-like state the longer they remain in culture (Figure 1).

Metabolic differences observed between iPSCs and ESCs reveal novel pathways important in somatic cell reprogramming

These findings are the first demonstrating that the global metabolomic profiles of iPSCs and ESCs are very similar, in concurrence with the overall genetic, epigenetic and functional similarities that have been reported comparing these pluripotent cell types [4]. Yet, metabolite differences do exist, and given the fact that recent studies have demonstrated that seemingly minor variations between iPSCs and ESCs may have significant phenotypic consequences [17], we next identified metabolites that differed between these cell types. We identified eight metabolites that showed a greater than two-fold difference with a *P*-value < 0.01 between ESCs and late passage iPSCs. We had previously shown that high levels of ω-6 and ω-3 fatty acids (e.g., arachidonic acid, docosahexanoic acid and linoleic acid) were characteristics of ESCs, and that inhibition of oxidative pathways was important for maintaining pluripotency [9]. Here, we found that the levels of unsaturated fatty acid metabolites (i.e., arachidonic acid, linoleic acid, docosapentaenoic acid and adrenic acid) were lower in iPSCs compared to ESCs, demonstrating differences in metabolites that are important in regulating oxidative pathways between pluripotent cell types (Figure 2A). Interestingly, we also observed differences in metabolites involved in the S-adenosyl methionine (SAM) cycle, such as SAM, 5′methylthioadenosine, hypoxanthine and inosine. More specifically, these metabolites were all significantly elevated in iPSCs compared to ESCs (Figure 2A). The methyl group of SAM is the most common substrate in transmethylation reactions such as DNA methylation. Of note, previous studies have shown that when compared to ESCs, human iPSCs have elevated global levels of methylation [18]. Thus, metabolite levels may also play a role in regulating the epigenetic changes occurring during reprogramming and/or the epigenetic variations observed between iPSCs and ESCs.

To determine how the observed unsaturated fatty acid and methylation metabolite differences between iPSCs and ESCs related to the levels found in somatic cells, we evaluated the metabolite levels of key components from each pathway in fibroblasts, FiPSCs and ESCs (Figure 2B), and tested the role of these metabolites in reprogramming experiments. Using a human fibroblast-like reporter cell line in which green fluorescent protein (GFP) expression is driven by the endogenous *OCT4* promoter (dFibOCT4^GFP cells) [19], we found that treatment with metabolites from each pathway inhibited somatic cell
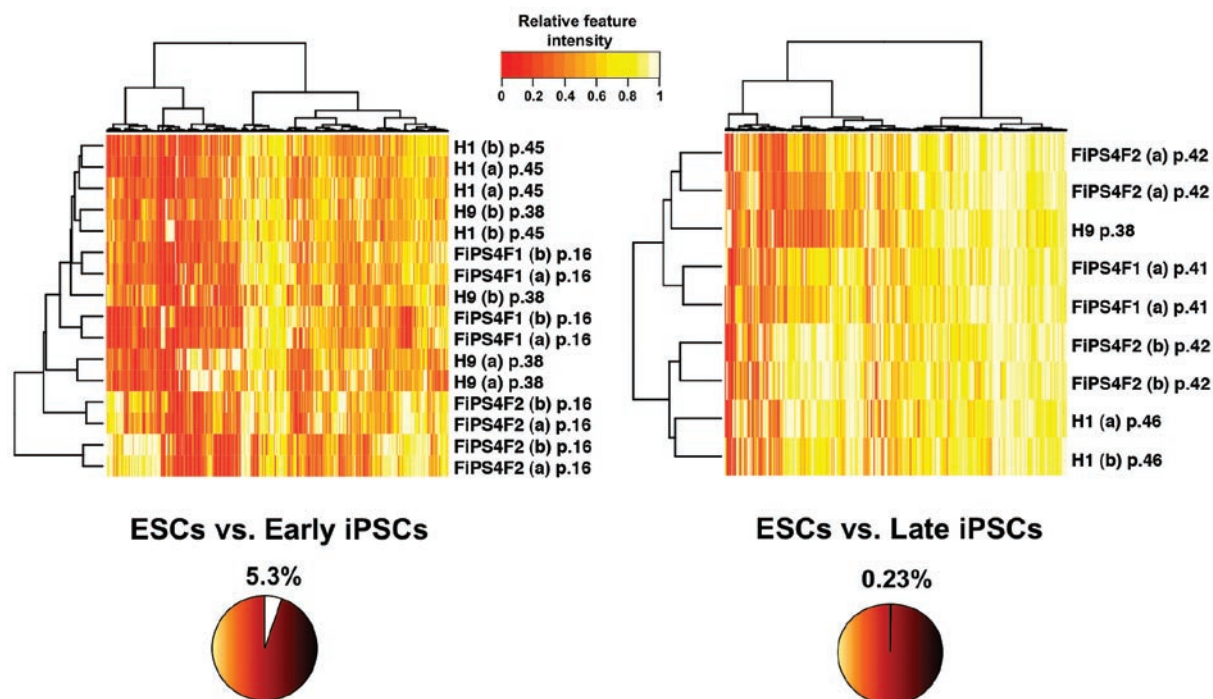
Figure 1 The metabolome profiles of iPSCs versus ESCs. Heat maps of metabolite features (> 5 000) in the indicated iPSC and ESC lines grown in chemically defined conditions at early (left panel) and late passage (right panel). The percentage of metabolite feature differences in iPSCs compared to ESCs is indicated below each respective heat map. Biological duplicates (e.g., (a) vs (b)) and experimental duplicates (e.g., (a) vs (a)) were performed.

reprogramming (Figure 2C). Parallel experiments performed with keratinocytes produced similar effects (Supplementary information, Figure S2A), demonstrating that this was not unique to a specific somatic cell type. Furthermore, the observed effects on reprogramming efficiencies were not due to metabolite-induced apoptosis of the starting somatic population (Supplementary information, Data S1 and Figure S2B). Lastly, treatment of the cells with metabolites at later stages of reprogramming after colony appearance (day 10) did not affect colony number (Supplementary information, Figure S2A). These results indicated that the observed effects on reprogramming efficiencies were due to inhibition of the process of reprogramming itself, and not to an impairment of iPSC colony survival or growth. These combined results demonstrated that the metabolome of iPSCs and ESCs is very similar, but that metabolite differences detected between these cell types reveal important metabolic pathways involved in somatic cell reprogramming.

Induced pluripotency is characterized by changes in metabolites involved in cellular respiration

To gain insight into the metabolism changes that occur in somatic cell reprogramming, we examined the metab-olome of iPSCs relative to their parental cells of origin. Each data set of feature intensities was visualized using a multidimensional scaling (MDS) plot (Figure 3A), such that the proximity of data points is indicative of overall similarities between cell populations. The MDS plot showed a clustering of seven iPSCs (FiPSCs and KiPSCs) and ESCs (H1, H9 and Hues6 cell lines), highlighting that the metabolomes of these cell populations are similar. Fibroblast or keratinocyte parental cells from which the iPSCs originated, however, were strikingly different, and were also distinct from one another (Figure 3A). Thus, various cell types display unique metabolite profiles, and iPSCs and ESCs share a characteristic pluripotent metabolomic signature.

We next sought to identify metabolites that distinguished iPSCs from their somatic cells of origin. Metabolites were structurally identified based on accurate mass, retention time and MS/MS data. Previous studies have demonstrated that somatic cells primarily utilize mitochondrial oxidation for proliferation, while pluripotent cells favor glycolysis [20-22]. Many of the metabolites we identified in somatic cells, which showed decreased levels upon reprogramming, were pyrimidine and purine nucleotides, and tricarboxylic acids involved
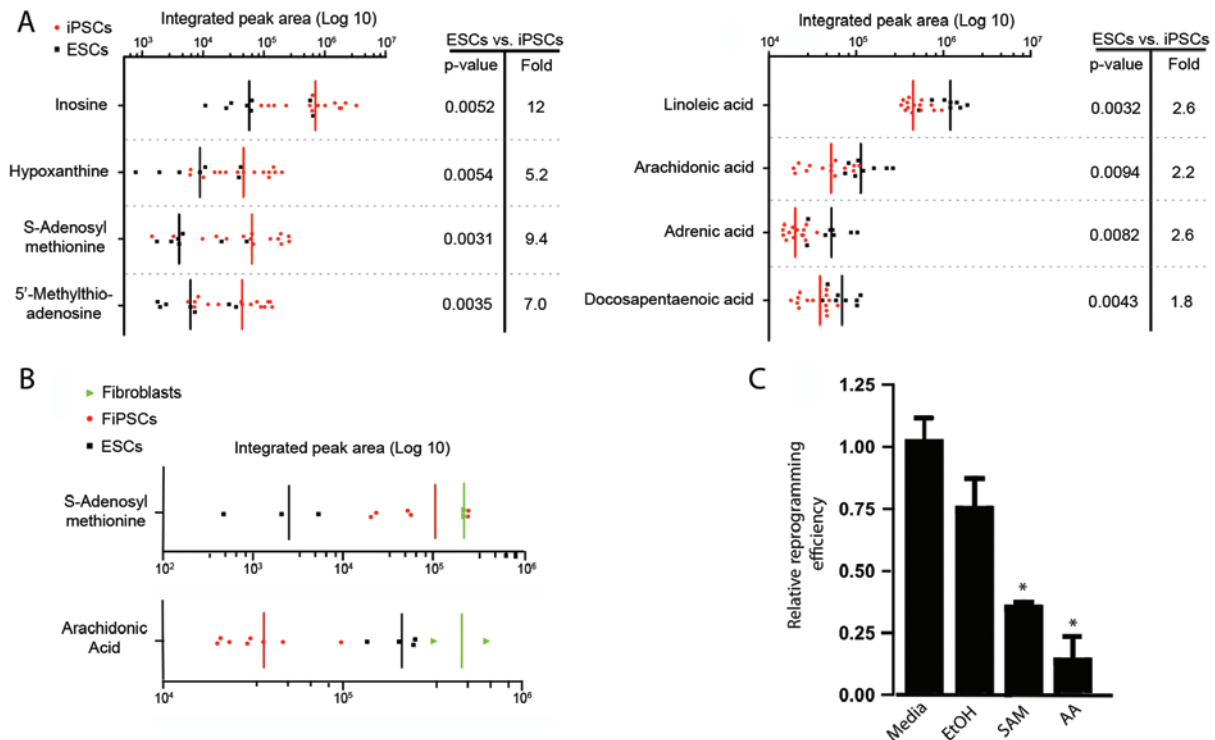
Figure 2 Metabolic differences between iPSCs and ESCs reveal additional pathways important for reprogramming. **(A)** Metabolites that differ between ESCs and iPSCs (grown in chemically defined conditions and isolated at late passage) as identified by accurate mass and MS/MS data. Fold values comparing the median integrated peak intensities are shown. **(B)** Integrated peak intensities for a sample of metabolites identified in ESCs, FiPSCs and parental fibroblast somatic populations. **(C)** dFib-OCT4GFP cells [19] infected with retroviruses encoding KLF-4, OCT4, SOX2 and c-MYC (KOSM) were grown in ESC medium in the presence or absence of arachidonic acid (AA, 0.1 μM), S-adenosyl methionine (SAM, 0.5 mM), or ethanol (EtOH) alone as a control. The number of GFP-positive colonies were counted ~16 days after the initial infection, and are plotted for each condition relative to controls. Error bars depict the SEM. *$P < 0.05$.

in the Krebs cycle (Figure 3B). The overall pattern of changes from somatic cells to iPSCs was indicative of a reversion of energy production metabolism. In accordance with these previous studies, we speculated that the main source of acetyl-CoA feeding into the Krebs cycle, contributing to the energy supply of somatic cells, is mitochondrial oxidation. Conversely, since pluripotent cells rely on glycolysis for their energy production, we next hypothesized that a conversion to glycolytic metabolism would occur in reprogramming.

Genes involved in glycolysis and oxidative phosphorylation undergo epigenetic and gene expression changes following somatic cell reprogramming
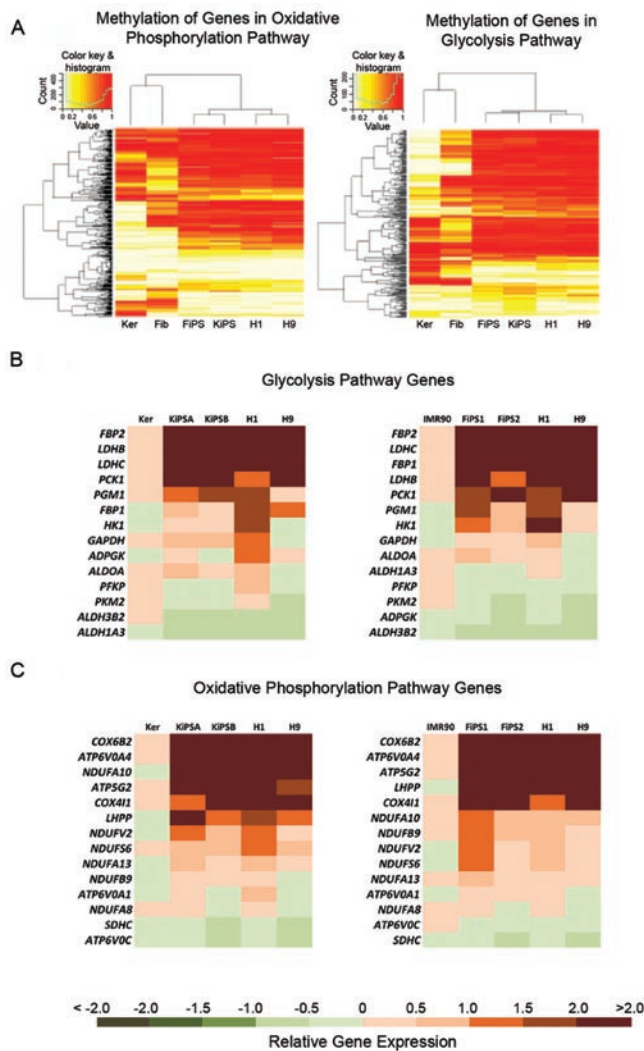
Our analysis demonstrated that cellular respiration changes were occuring at the metabolomic level. To next examine additional levels of regulation that may either influence, or be influenced by, these metabolite changes, we analyzed the epigenetic status of genes involved in the glycolysis and mitochondrial oxidative

phosphorylation pathways in iPSCs and ESCs relative to their somatic cells of origin. As shown in Figure 4A, bisulfite sequencing revealed a large number of epigenetic changes occurring in genes regulating both pathways. Hierarchical clustering of the methylation levels of each gene group demonstrated that iPSCs clustered with one another and with ESCs, and were distinct from their parental sources. We next evaluated if some of the genes involved in either the glycolytic or oxidative pathways, that had undergone epigenetic changes in both somatic source comparisons (i.e., keratinocytes vs KiPS, and fibroblasts vs FiPS), also changed in gene expression (see Supplementary information, Tables S1 and S2). Similar to what we had observed in our methylation analysis, iPSCs acquired a gene expression signature that was similar to ESCs and distinct from their parental source. Furthermore, following reprogramming, both somatic cell types similarly upregulated or downregulated the expression of genes involved in regulating glycolytic or oxidative pathways (Figure 4B and 4C). These combined
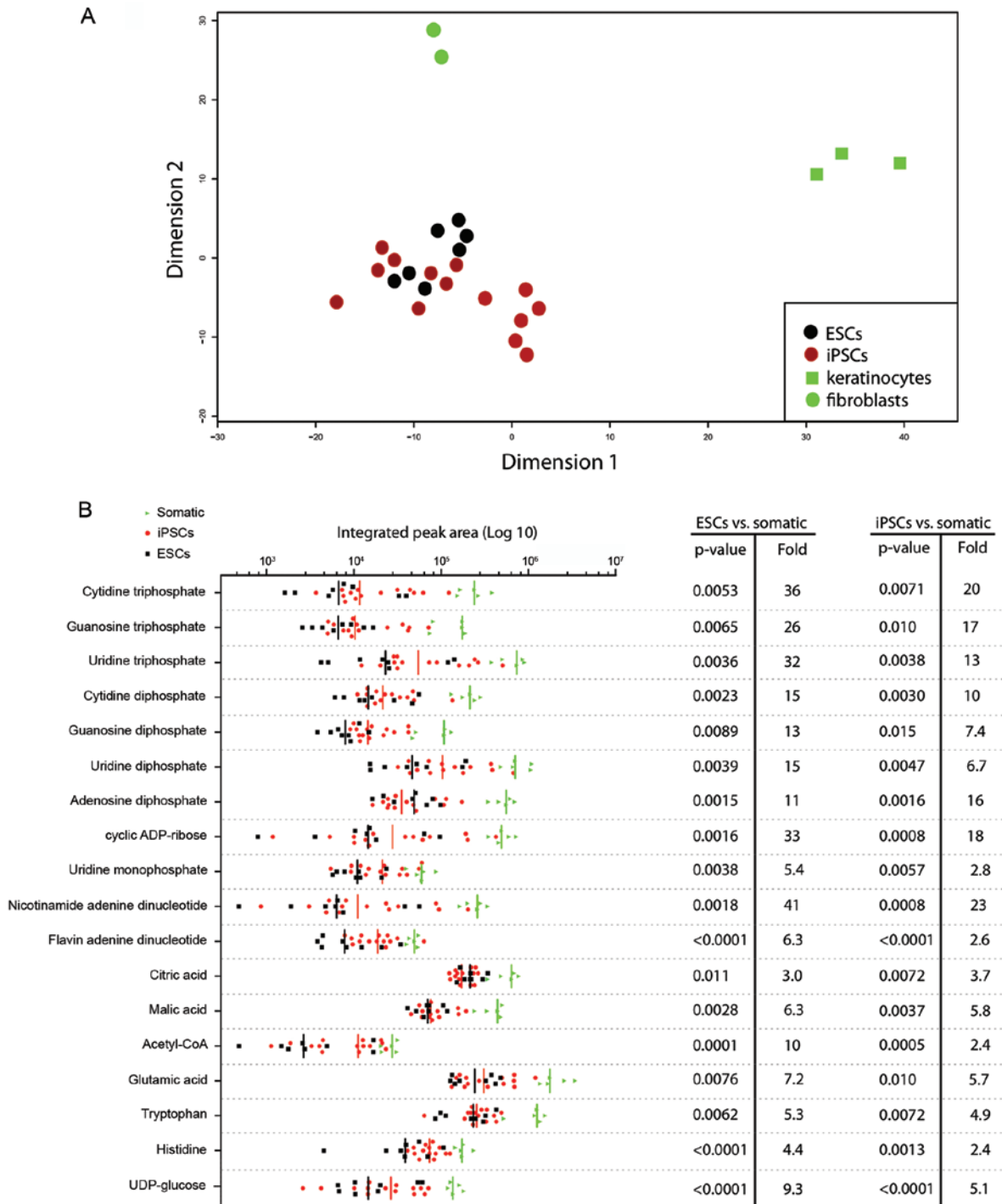
Figure 3 Pluripotent cells have a distinct metabolic signature characterized by changes in metabolites involved in cellular respiration. **(A)** Two-dimensional representation of the XCMS matrix of retention time, *m/z*, and feature intensity values using a multidimensional scaling (MDS) plot for ESCs, iPSCs (derived from human keratinocytes and fibroblasts) and their somatic cells of origin. Data points for cell populations producing similar features are closer to one another than data points for cell populations producing more dissimilar features. Each cell line was analyzed by at least two biological and experimental replicates. **(B)** Quantification of metabolites from ESCs, iPSCs and somatic cell populations of origin (keratinocytes and fibroblasts). Data points and bars represent the integrated peak intensity for each sample and the median intensity values. Fold values indicate the difference in integrated peak intensity (median value) for selected metabolites in ESCs and iPSCs relative to somatic populations. Identification is based on accurate mass and MS/MS data. Each cell line was analyzed by at least two biological and experimental replicates.

Figure 4 Genes involved in glycolysis and oxidative phosphorylation undergo changes in methylation and gene expression following the transition from a somatic to a pluripotent state. **(A)** Levels of epigenetic changes occurring in genes regulating glycolysis and oxidative phosphorylation (see Materials and Methods). Heatmap and hierarchical clustering results of keratinocytes, fibroblasts, their respective iPSC lines (KiPS and FiPS) and ESC lines (H1 and H9) using methylation patterns at CpG sites containing a change in methylation in at least one iPSC line. Clustering was performed on a dissimilarity matrix with values equal to the complement of the Pearson's correlation for each pair. **(B-C)** A selected number of genes involved in glycolysis **(B)** or oxidative phosphorylation **(C)**, that demonstrated differentially methylated sites between somatic cells and their respective iPSC counterparts, were measured for expression changes by real-time PCR (see Supplementary information, Tables S1 and S2). A heatmap of gene expression in iPSC lines and ESC controls (H1 and H9) relative to their somatic cell of origin is shown.

results demonstrate that somatic cell reprogramming is accompanied by changes in cellular respiration at the metabolomic, epigenetic and gene expression levels.

The bioenergetic changes associated with inducing pluripotency influence reprogramming efficiency

To evaluate the cellular respiration changes that emerged from our combined analysis, we next tested the cellular bioenergetics of somatic and induced pluripotent cells. Using an extracellular flux analyzer, we measured the basal oxygen consumption rate (OCR; as a measure of mitochondrial respiration) and the extracellular acidification rate (ECAR; as a measure of glycolysis) of keratinocytes, IMR90 fibroblasts and their respective iPSC lines. These results demonstrated that the bioenergetics of iPSCs cluster together as highly glycolytic, and that somatic cells vary in their oxidative:glycolytic ratio (Figure 5A). Since keratinocytes reprogram with

~100-fold greater efficiency and ~2 times faster than fibroblasts, these results also suggested a possible correlation between somatic bioenergetics and reprogramming efficiency, in that it appeared the closer the energy state to iPSCs, the higher the reprogramming efficiency. To test this hypothesis, we examined the OCR/ECAR values (representing the overall bioenergetic state) of somatic cells of varying efficiencies, and compared them to their respective iPSCs and ESCs. Whereas somatic cells demonstrate an overall higher oxidative:glycolytic OCR/ECAR ratio compared to pluripotent cells, indicative of energy production that is dependent on oxidative phosphorylation, there was also a difference in the energy state between somatic cell groups (Figure 5B). Keratinocytes and human umbilical vein endothelial cells (HUVECs), which reprogram with ~100-300-fold greater efficiency and 2-4 times faster than fibroblasts [11, 23], had a statistically significant lower OCR:ECAR ratio than fibroblasts (Figure 5B). These findings suggest that the closer a somatic cell's overall bioenergetics are to pluripotent cells, the higher its reprogramming efficiency.

We next determined if manipulating bioenergetics could affect reprogramming. Using human fibroblast-like dFibOCT4[GFP] reporter cells [19], we tested the effect of the glycolysis inhibitor 2-deoxy-D-glucose (2-DG) or the glycolysis stimulator D-fructose-6-phosphate (F6P) on reprogramming efficiency. As shown in Figure 5C, 2-DG treatment inhibited GFP+ colony formation, whereas F6P treatment increased reprogramming efficiency, in agreement with recent reports [24, 25]. Importantly, examination of the bioenergetic changes following treatment with each metabolic modulator demonstrated that 2-DG inhibited ECAR values, while F6P induced them (Figure
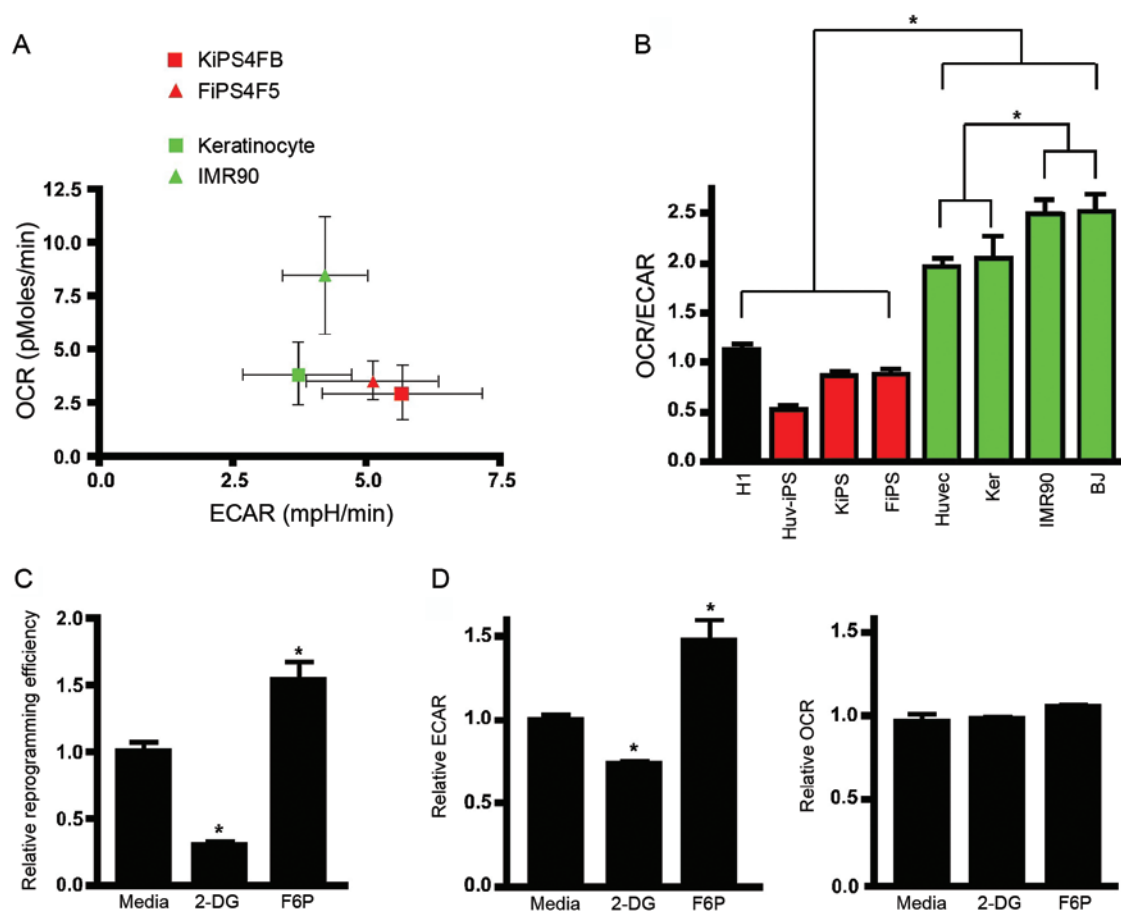
Figure 5 Reprogramming induces a bioenergetic conversion from an oxidative to a glycolytic state. **(A)** Basal oxygen consumption rate (OCR, indicative of mitochondrial oxidative phosphorylation) plotted versus basal extracellular acidification rate (ECAR, representing glycolysis) for IMR90 fibroblasts, keratinocytes and their respective iPSCs (FiPS4F5, KiPS4FB). Results represent the average of four independent experiments performed in triplicate. **(B)** OCR/ECAR ratios of H1 ESCs, HUVECs, IMR90 and BJ fibroblasts, and iPSCs from HUVECs (Huv-iPS), keratinocytes (KiPS) or IMR90 fibroblasts (FiPS) are shown. Results were determined from the average of five independent experiments performed in triplicate. **(C)** dFib-OCT4^{GFP} cells [19] were infected with retroviruses encoding KLF-4, OCT4, SOX2 and c-MYC (KOSM). Equivalent numbers of KOSM-infected cells were plated and grown in ESC medium in the presence or absence of 2-deoxy-D-glucose (2-DG, 1 mM) or D-fructose-6-phosphate (F6P, 5 μM). GFP-positive colonies were numerated at ~day 16 after the initial infection. The number of GFP-positive colonies for each condition relative to controls are shown. **(D)** dFib-OCT4^{GFP} cells were treated with 2-deoxy-D-glucose (2-DG, 1 mM) or D-fructose-6-phosphate (F6P, 5 μM) for 48 h, and ECAR and OCR values relative to media only controls determined. All OCR and ECAR values were normalized to cell number. Error bars depict the SEM. *$P < 0.05$.

5D). These collective findings indicate that somatic cells convert from an oxidative to glycolytic state when they are reprogrammed, that the bioenergetic states of somatic cells appear to correlate with their reprogramming efficiencies, and that manipulating these bioenergetic changes can affect reprogramming.

Discussion

To our knowledge, this is the first study to use an untargeted approach to globally analyze the metabolome of

iPSCs, providing mechanistic insight into another layer of regulation that is involved in inducing pluripotency, and in evaluating iPSC and ESC equivalence. Changes in cell metabolism are critical for somatic cell reprogramming, as indicated by the fact that metabolites involved in transmethylation, cellular respiration and energy production were all shown to functionally affect reprogramming. Our findings also suggest an overall correlation between the bioenergetic state of somatic cells and their reprogramming efficiency; somatic cells that demonstrated oxidative:glycolytic energy production ratios closer

to pluripotent cells reprogrammed more quickly and efficiently. Thus, these data suggest that changes in metabolism may play a role in enabling the reprogramming process to occur, rather than simply being a consequence of acquiring a pluripotent state [25]. Future studies, which will rely on the ability to identify a reprogrammed cell before it becomes pluripotent, will help distinguish between these possibilities.

Interestingly, many of the changes in cell metabolism we have identified to be important in regulating somatic cell reprogramming have also been shown to play a role in cancer, in agreement with previous studies demonstrating parallels in the pathways regulating oncogenesis and induced pluripotency [26]. For example, treatment with arachidonic acid has been shown to prevent cancer cell growth [27]. The inhibition of glycolysis as an anti-cancer treatment is a current focus in many areas of cancer research [22, 28]. An interesting study also found that SAM treatment was effective in preventing the formation of primary liver cancer, but was not able to successfully treat hepatocarcinoma once it had been established, thus suggesting an important role for SAM in preventing the conversion of a normal cell to a cancer cell [29]. Therefore, our study not only highlights the importance of metabolism in inducing pluripotency, but suggests that understanding the metabolic changes associated with somatic cell reprogramming may also shed light on the metabolic mechanisms regulating cancer.

Materials and Methods

Materials

All pure standards were purchased from Sigma-Aldrich (St Louis, USA), except UDP-glucose that was purchased from American Custom Chemicals Corp. Ammonium fluoride, formic acid and EDTA were purchased from Sigma-Aldrich (St Louis, USA). LC/MS grade acetonitrile and water were purchased from JT Baker (Phillipsburg, USA).

Metabolite extraction

300 µl of LC/MS-grade water in 1 mM HEPES and 1 mM EDTA (pH 7.2) was added to the ESC, iPSC or somatic cell pellet (~250 000 cells), vortexed for 30 s, and the sample incubated for 1-2 min in boiling water. Next, the sample was incubated for 1 min in liquid nitrogen and thawed at room temperature. The incubation in liquid nitrogen was repeated. Finally, the sample was incubated for 1 h at −20 °C, followed by 15 min centrifugation at 13 000 rpm (4 °C). The resultant supernatant was transferred to an HPLC vial.

LC-MS and MS/MS analysis

Analyses were performed using an HPLC system (1200 series, Agilent Technologies) coupled to a 6538 UHD Accurate-Mass Q-TOF (Agilent Technologies) operated in positive (ESI+) or negative (ESI−) electrospray ionization mode. Vials containing extracted metabolites were kept at −20 °C prior to LC/MS analy-

sis. ESCs, iPSCs and somatic cells extractions were separated using a Cogent Bidentate C18: 4 µm, 100 Å, 150 mm × 2.1 mm ID (Cat No. 40018-15P-2), or a Waters XBridge C18: 3.5 µm, 135Å, 150 mm × 1.0 mm ID (Part No. 186003128) column when the instrument was operated in positive or negative ionization mode, respectively. The solvent system in positive ionization mode was: A = 0.1% formic acid in water, and B = 0.1% formic acid in acetonitrile. When the instrument was operated in negative ionization mode, the solvent system was: A = 1 mM ammonium fluoride in water, B = acetonitrile [30]. The linear gradient elution used started at 100% A (time 0-5 min), and finished at 100% B (35-40 min). The injection volume was 8 µl. ESI conditions were: gas temperature 325 °C, drying gas 11 l/min, nebulizer 30 psig, fragmentor 120 V and skimmer 65 V. The instrument was set to acquire over the *m/z* range 80-1 000 with an acquisition rate of 1.3 spectra/s. MS/MS was performed in targeted mode and the instrument was set to acquire over the *m/z* range 50-1 000, with a default iso width (the width-half maximum of the quadrupole mass bandpass used during MS/MS precursor isolation) of 4 *m/z*. The collision energy was fixed at 20 V.

Metabolomic data processing

LC/MS data from the ESC, iPSC and somatic cell extractions (ESI+ and ESI− modes) were processed using the XCMS software (version 1.24.1) to detect and align features. XCMS analysis of these data provided a matrix containing the retention time, *m/z* value and intensity of each feature for every sample described above. Statistically significant differences (*P*-value < 0.01) are ranked using univariate *t*-test in order to obtain quantitative information between cell populations. In addition, using the statistical software R, each row of feature intensities was normalized such that the highest value was one. A two-dimensional representation of this matrix was calculated using MDS as implemented within the software R.

Metabolite database searching

Metabolites were identified using exact mass (*m/z*), retention time and tandem mass spectrometry data. The databases that helped in these investigations included METLIN (http://metlin.scripps.edu/), Human Metabolome Project (http://redpoll.pharmacy.ualberta.ca/hmdb/HMDB/), and KEGG ligand (http://www.genome.jp/kegg/ligand.html).

Cell culture and cell sample preparation

Human neonatal keratinocytes (Lonza), fibroblasts (IMR90, BJ, from ATCC) and HUVECs (Lonza) were grown according to manufacturer's recommendations. 293T cells were cultured in DMEM (Invitrogen) containing 10% fetal calf serum. Human H1 (WA01) or H9 (WA09) ESC lines (Wicell) and Hues6 (Harvard), and all iPSC lines were cultured as previously described in mTeSR1 before metabolite extraction [16, 31]. For cell sample preparation, keratinocytes, fibroblasts, ESCs and iPSCs (individual iPSC and ESC colonies were mechanically picked) were individualized using TrypLE (Invitrogen) according to manufacturer's recommendations. Equivalent numbers of cells were pelleted, washed 1× in PBS, and stored in −80 °C until metabolites were extracted as described.

iPSC generation and reprogramming experiments

Keratinocytes or fibroblasts were infected with equivalent ra-

tios of retroviruses encoding KLF-4, OCT4, SOX2 and c-MYC (KOSM) by spinfection at 800× *g* for 1 h at RT in the presence of polybrene (8 µg/ml). Cells were replated onto MEFs (Millipore) in their respective medias, and switched to ESC medium for iPSC colony formation. Resulting iPSC colonies were manually picked for iPSC line derivation (~14 days after infection for keratinocytes, and ~3-4 weeks after infection for fibroblasts). For reprogramming experiments, dFibOCT4^{GFP} cells [19] or keratinocytes were KOSM-infected as described. Reprogramming efficiencies were then determined by calculating the number of GFP-positive, or Nanog-positive colonies, respectively, in metabolite treated conditions (2-DG, Sigma; F6P, Sigma; arachidonic acid, Cayman Chemicals; S-adenosyl methionine, New England Biolabs) relative to controls. Metabolite concentrations for use in reprogramming experiments were determined empirically by examining somatic cell survival by TUNEL staining.

Plasmids and virus preparation

The following moloney-based retroviral vectors were obtained from Addgene: pMX-OCT4, pMX-SOX2, pMX-KLF-4 and pMX-c-MYC (plasmids 17217, 17218, 17219 and 17220, respectively). Packaging plasmids (pCMV-gag-pol-PA and pCMV-VSVg) were kindly provided by Dr Gerald Pao (Laboratory of Genetics, The Salk Institute, La Jolla, CA). Retrovirus was collected 24 h following cotransfection of plasmids in 293T cells using Lipofectamine (Invitrogen) in accordance with manufacturer's recommendations.

Extracellular flux bioenergetic assay

Assays were performed in accordance with manufacturer's instructions (Seahorse Bioscience). Briefly, cells were seeded in extracellular flux 96-well cell culture microplates (Seahorse Bioscience) in 80 µl of corresponding cell growth medium (ESCs and iPSCs were grown in mTeSR1, Stem Cell Technologies) and then incubated at 37 °C/5% CO_2 for ~24 h. Assays were initiated by replacing the cell growth medium from each well and replacing it with 200 µl of assay medium. The microplates were incubated at 37 °C for 60 min to equilibrate the temperature and pH of the media before measurement. A Seahorse Bioscience instrument (model XF96) was used to measure the rate of change of dissolved O_2 and pH in the media. OCR and ECAR were measured simultaneosly for ~3 min in repeated cycles, to obtain a basal average. All OCR and ECAR values were normalized by cell number.

Bisulfite sequencing

Genomic DNA was extracted using the ALLPrep DNA/RNA Mini kit (Qiagen). The bisulfite conversion and capture reactions was carried out on each gDNA sample of somatic or iPSCs [10, 32] using previously established protocols [18]. Briefly, DNA was bisulfite converted using the Zymo Methyl-Direct Bisulfite Conversion kit. Bisulfite-converted DNA was then captured using specially designed padlock probe oligonucleotides, resulting in a circular DNA library of targeted CpG sites. The padlock capture and DNA sequencing was carried out using a previously described protocol [33]. Bisulfite converted data was processed as previously described [18]. Heatmaps and dendrograms were created from the Pearson's correlation matrices of (1) the relative change in methylation level between each iPSC line and its somatic progenitor, and (2) the absolute methylation level at each site in each iPSC line.

Identification of differentially methylated sites in oxidative phosphorylation and in glycolysis

To identify sites showing a change in methylation after reprogramming, a Chi-squared test with Yates' correction was carried out on each CpG site characterized in each iPSC line and corresponding paired-somatic cell line. The Benjamini-Hochberg method was used to correct for multiple testing errors; the false discovery rate was set at 1%. This resulted in a set of differentially methylated sites (DMSs) for each iPSC line; at each site, the methylation level was statistically significantly different from the somatic progenitor line. The DMS list was uploaded to GREAT (http://great.stanford.edu) to find possible gene to DMS interactions based on distance from transcription start sites and curated regulatory regions of genes. Genes were then annotated by using the KEGG database for involvement in oxidative phosphorylation and in glycolysis.

Statistical analysis

Results are shown as mean values ± SEM. Statistics were performed using either an unpaired two-tailed Student's *t*-test or one-way ANOVA with a *post-hoc* test as appropriate. *P*-values < 0.05 were considered statistically significant.

References

1 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; **126**:663-676.

2 Takahashi K, Tanabe K, Ohnuki M, *et al*. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007; **131**:861-872.

3 Yu J, Vodyanik MA, Smuga-Otto K, *et al*. Induced pluripotent

stem cell lines derived from human somatic cells. *Science* 2007; **318**:1917-1920.

4   Stadtfeld M, Hochedlinger K. Induced pluripotency: history, mechanisms, and applications. *Genes Dev* 2010; **24**:2239-2263.

5   Panopoulos AD, Ruiz S, Izpisua Belmonte JC. iPSCs: induced back to controversy. *Cell Stem Cell* 2011; **8**:347-348.

6   Kwon YK, Lu W, Melamud E, *et al*. A domino effect in anti-folate drug action in *Escherichia coli*. *Nat Chem Biol* 2008; **4**:602-608.

7   Sreekumar A, Poisson LM, Rajendiran TM, *et al*. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* 2009; **457**:910-914.

8   Smith CA, Want EJ, O'Maille G, *et al*. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 2006; **78**:779-787.

9   Yanes O, Clark J, Wong DM, *et al*. Metabolic oxidation regulates embryonic stem cell differentiation. *Nat Chem Biol* 2010; **6**:411-417.

10 Gore A, Li Z, Fung HL, *et al*. Somatic coding mutations in human induced pluripotent stem cells. *Nature* 2011; **471**:63-67.

11 Aasen T, Raya A, Barrero MJ, *et al*. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol* 2008; **26**:1276-1284.

12 Liu GH, Barkho BZ, Ruiz S, *et al*. Recapitulation of premature ageing with iPSCs from Hutchinson Gilford progeria syndrome. *Nature* 2011; **472**:221-225.

13 Chin MH, Mason MJ, Xie W, *et al*. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 2009; **5**:111-123.

14 Kim K, Doi A, Wen B, *et al*. Epigenetic memory in induced pluripotent stem cells. *Nature* 2010; **467**:285-290.

15 Polo JM, Liu S, Figueroa ME, *et al*. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotechnol* 2010; **28**:848-855.

16 Ludwig TE, Bergendahl V, Levenstein ME, *et al*. Feeder-independent culture of human embryonic stem cells. *Nat Methods* 2006; **3**:637-646.

17 Stadtfeld M, Apostolou E, Akutsu H, *et al*. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 2010; **465**:175-181.

18 Deng J, Shoemaker R, Xie B, *et al*. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat Biotechnol* 2009; **27**:353-

360.

19 Ruiz S, Panopoulos AD, Herrerias A, *et al*. Active proliferation is required for somatic cell reprogramming and maintenance of human embryonic stem cell identity. *Curr Biol* 2011; **21**:45-52.

20 Kondoh H, Lleonart ME, Nakashima Y, *et al*. A high glycolytic flux supports the proliferative potential of murine embryonic stem cells. *Antioxid Redox Signal* 2007; **9**:293-299.

21 Prigione A, Fauler B, Lurz R, *et al*. The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells. *Stem Cells* 2010; **28**:721-733.

22 Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* 2009; **324**:1029-1033.

23 Panopoulos AD, Ruiz S, Yi F, *et al*. Rapid and highly efficient generation of induced pluripotent stem cells from human umbilical vein endothelial cells. *PLoS One* 2011; **6**:e19743.

24 Zhu S, Li W, Zhou H, *et al*. Reprogramming of human primary somatic cells by OCT4 and chemical compounds. *Cell Stem Cell* 2010; **7**:651-655.

25 Folmes CD, Nelson TJ, Martinez-Fernandez A, *et al*. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab* 2011; **14**:264-271.

26 Krizhanovsky V, Lowe SW. Stem cells: the promises and perils of p53. *Nature* 2009; **460**:1085-1086.

27 Monjazeb AM, High KP, Connoy A, *et al*. Arachidonic acid-induced gene expression in colon cancer cells. *Carcinogenesis* 2006; **27**:1950-1960.

28 Pelicano, Martin DS, Xu RH, *et al*. Glycolysis inhibition for anticancer treatment. *Oncogene* 2006; **25**:4633-4646.

29 Lu SC, Ramani K, Ou X, *et al*. S-adenosylmethionine in the chemoprevention and treatment of hepatocellular carcinoma in a rat model. *Hepatology* 2009; **50**:462-471.

30 Yanes O, Tautenhahn R, Patti GJ, Siuzdak G. Expanding coverage of the metabolome for global metabolite profiling. *Anal Chem* 2011; **83**:2152-2161.

31 Thomson JA, Itskovitz-Eldor J, Shapiro SS, *et al*. Embryonic stem cell lines derived from human blastocysts. *Science* 1998; **282**:1145-1147.

32 Eguizabal C, Montserrat N, Vassena R, *et al*. Complete meiosis from human induced pluripotent stem cells. *Stem Cells* 2011; **29**:1186-1195.

33 Zhang K, Li JB, Gao Y, *et al*. Digital RNA allelotyping reveals tissue-specific and allele-specific gene expression in human. *Nat Methods* 2009; **6**:613-618.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)