Generation of Transgene-Free Lung Disease-Specific Human Induced Pluripotent Stem Cells Using a Single Excisable Lentiviral Stem Cell Cassette

Aba Somers, Jyh-Chang Jean, Cesar A. Sommer, Amel Omari, Christopher C. Ford, Jason A. Mills, Lei Ying, Andreia Gianotti Sommer, Jenny M. Jean, Brenden W. Smith, Robert Lafyatis, Marie-France Demierre, Daniel J. Weiss, Deborah L. French, Paul Gadue, George J. Murphy, Gustavo Mostoslavsky, Darrell N. Kotton

Stem Cells®

Embryonic Stem Cells/Induced Pluripotent Stem Cells

Generation of Transgene-Free Lung Disease-Specific Human Induced Pluripotent Stem Cells Using a Single Excisable Lentiviral Stem Cell Cassette

Aba Somers,[a,b] Jyh-Chang Jean,[a,b] Cesar A. Sommer,[c] Amel Omari,[a,b] Christopher C. Ford,[a,b] Jason A. Mills,[d] Lei Ying,[d] Andreia Gianotti Sommer,[c] Jenny M. Jean,[a] Brenden W. Smith,[e] Robert Lafyatis,[f] Marie-France Demierre,[g] Daniel J. Weiss,[h] Deborah L. French,[d] Paul Gadue,[d] George J. Murphy,[e,i] Gustavo Mostoslavsky,[c,i] Darrell N. Kotton[a,b,i]

[a]Boston University Pulmonary Center and; [b]Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA; [c]Section of Gastroenterology, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA; [d]Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; [e]Section of Hematology and Oncology, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA; [f]Section of Rheumatology, Department of Medicine, Boston University School of Medicine, Boston, Massachusets, USA; [g]Department of Dermatology, Boston University School of Medicine, Boston, Massachusets, USA; [h]Department of Medicine, University of Vermont College of Medicine, Burlington, Vermont, USA; [i]Center for Regenerative Medicine (CReM), Boston University School of Medicine, Boston, Massachusets, USA

Abstract

The development of methods to achieve efficient reprogramming of human cells while avoiding the permanent presence of reprogramming transgenes represents a critical step toward the use of induced pluripotent stem cells (iPSC) for clinical purposes, such as disease modeling or reconstituting therapies. Although several methods exist for generating iPSC free of reprogramming transgenes from mouse cells or neonatal normal human tissues, a sufficiently efficient reprogramming system is still needed to achieve the widespread derivation of disease-specific iPSC from humans with inherited or degenerative diseases. Here, we report the use of a humanized version of a single lentiviral "stem cell cassette" vector to accomplish efficient reprogramming of normal or diseased skin fibroblasts obtained from humans of virtually any age. Simultaneous transfer of either three or four reprogramming factors into human target cells using this single vector allows derivation of human iPSC containing a single excisable viral integration that on removal generates human iPSC free of integrated transgenes. As a proof of principle, here we apply this strategy to generate >100 lung disease-specific iPSC lines from individuals with a variety of diseases affecting the epithelial, endothelial, or interstitial compartments of the lung, including cystic fibrosis, α-1 antitrypsin deficiency-related emphysema, scleroderma, and sickle-cell disease. Moreover, we demonstrate that human iPSC generated with this approach have the ability to robustly differentiate into definitive endoderm in vitro, the developmental precursor tissue of lung epithelia. Stem Cells 2010;28:1728–1740

Disclosure of potential conflicts of interest is found at the end of this article.

Introduction

The reprogramming of postnatal cells by defined transcription factors has allowed the derivation of induced pluripotent stem cells (iPSC) with similar functional and molecular phenotypic characteristics to embryonic stem cells (ESC) [1–4]. This seminal advance has considerable implications for the field of regenerative medicine and suggests the prospect of generating autologous pluripotent stem cells from easily accessible

human tissues, such as skin biopsies, hair follicles, or peripheral blood [5–10]. The original method for deriving iPSC as described by Yamanaka and coworkers, employed four integrating retroviral vectors to deliver the four "reprogramming transcription factors," Oct4, Sox2, Klf4, and c-Myc into mouse fibroblasts to generate approximately 150 iPSC clones per 8×10^5 transduced target cells, indicating a reprogramming efficiency of 0.01% [3]. It has since become clear that combinations of alternative genes or chemicals can be used to substitute for some of the original four reprogramming factors, modifying the number of viral vectors required, in some cases at the expense of reprogramming efficiency [7, 9, 11, 12]. More recently, derivation of iPSC with nonintegrating vectors, plasmid transfection, or even direct protein delivery has been achieved, although with exceedingly low efficiencies that prevent reliable application for reprogramming disease-specific adult human somatic cells [13–17].

Regardless of the method used, somatic cells from humans appear to be more difficult to reprogram than murine cells [7]. Moreover, it is becoming clear that the development of methods to achieve efficient reprogramming of cells from adult humans with disease while avoiding the permanent presence of the reprogramming transgenes, represents a critical step toward the use of this technology for clinical purposes [18–20]. Importantly, such methodology should allow for the reliable and consistent reprogramming of human somatic cells, regardless of the age or disease state of the individual from whom they are derived.

Recently, we reported the use of a single lentiviral "stem cell cassette" (STEMCCA) encoding all four reprogramming factors, Oct4, Sox2, Klf4, and c-Myc in a single polycistronic vector [21]. By combining all reprogramming transgenes in a single cassette, STEMCCA accomplished reprogramming of postnatal mouse fibroblasts with high efficiency and allowed the derivation of mouse iPSC containing a single viral integration. Most recently, we generated an excisable version of STEMCCA based on Cre/loxP technology that allowed the derivation of murine iPSC free of exogenous transgenes [22]. Here, we report the use of a humanized version of the single lentiviral stem cell cassette vector flanked by loxP sites (hSTEMCCA-loxP) to achieve highly efficient reprogramming of normal or diseased postnatal human skin fibroblasts. Simultaneous transfer of either three or four reprogramming factors into human target cells using this single vector allows derivation of human iPSC containing a single excisable viral integration that on removal generates human iPSC free of integrated transgenes. In contrast to previously described methods, the high efficiency of reprogramming with this reagent allows minute quantities of viral vector to be used for reprogramming normal and disease-specific somatic cells taken from humans of virtually any age. As a proof of principle, here we apply this strategy to generate >100 of the first known lung disease-specific iPSC lines from individuals with diseases affecting the epithelial, endothelial, or interstitial compartments of the lung.

MATERIALS AND METHODS

Vector Design and Construction

The humanized hSTEMCCA lentiviral vector was constructed by adapting the previously published mouse pHAGE2- elongation factor 1 alpha constitutive promoter (EF1α)-STEMCCA vector [21] as follows: First, human cDNAs encoding the transcription factors OCT4, KLF4, SOX2, and cMYC were amplified using the following primers: hOCT4 5′ NotI (5′TTT TGC GGC CGC CAT GGC GGG ACA CCT

GGC TTC GG-3′); hOCT4 F2A 3′ (5′-CCT GCA AGT TTC AGC AAA TCA AAG TTT AAT GTC TG CTT TAC TGG CGC ACC CGA ACC CGA GTT TGA ATG CAT GGG AGA GCC CAG AGT GGT G-3′); hKLF4 F2A 5′ (5′-GCA GAC ATT AAA CTT TGA TTT GCT GAA ACT TGC AGG TG ATG TAG AG TCA AAT CCA GGT CCA ATG GCT GTC AGC GAC GCG CTG CTC CCA TC-3′); hKLF4 3′ BamHI (5′TGT TGG ATC CTT AAA AAT GCC TCT TCA TGT GTA AGG CG-3′); hSOX2 5′ NdeI (5′-TTT AGT GCA TAT GAT GTA CAA CAT GAT GGA GAC GG AGC TG-3′); hSOX2 P2A 3′-(5′-TTC TCT TCG ACA TCC CCT GCT TGT TTC AAC AGG GA GAA GTT AGT GGC TCC GCT TCC GGA CAT GTG TGA GAG GGG CAG TGT GCC GTT AAT G-3′); hcMYC P2A 5′ (5′-GCC ACT AAC TTC TCC CTG TTG AAA CAA GCA GGG GA TGT CGA AGA GAA TCC CGG GCC AAT GCC CCT CAA CGT TAG CTT CAC CAA CAG GAA C-3′); hcMYC 3′ AccI (5′-TTT AGC AGT GGT ACG TCG ACT TAC GCA CAA GAG TTC CGT AGC TGT TC-3′). Cloning of amplified products into the pHAGE2 vector was performed as described [21].

Next, the loxP sequence flanked by AscI restriction enzyme sites was constructed by annealing together the following complementary oligonucleotides: 5′-CGC GCA GGT ACC ATA ACT TCG TAT AAT GTA TGC TAT ACG AAG TTA TGG-3′ and 5′-CGC GCC ATA ACT TCG TAT AGC ATA CAT TAT ACG AAG TTA TGG TAC CTG-3′. The resulting double-stranded DNA fragment was inserted in the deleted U3 portion of the pHAGE2 lentiviral 3′LTR by ligation of cohesive compatible ends into the vector's AscI restriction site. Cloning of mCherry was performed as previously described [22].

Lentiviruses were produced in 293T packaging cells by five plasmid cotransfection and were concentrated by ultracentrifugation as previously described [21, 22]. Viral titers were calculated based on Southern blots of genomic DNA (gDNA) from FG293 cells transduced with defined volumes of concentrated viral supernatants to determine transducing units per milliliter (TU/ml). On average, viral titers of ~1 × 10^8 TU/ml were employed for reprogramming experiments.

For transient Cre expression, the plasmid pHAGE2-EF1α-Cre-IRES-PuroR was constructed by ligating the Cre cDNA into compatible NotI-BamHI restriction sites downstream of the EF1α promoter of the pHAGE2 backbone. The PuroR cDNA was ligated downstream of the internal ribosome entry site (IRES) element and upstream of a Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE) sequence using NdeI-ClaI sites.

Skin Biopsies and Expansion of Human Dermal Fibroblasts

Individuals with α-1 antitrypsin (AAT) deficiency because of inheritance of two Z alleles of the AAT protease inhibitor (PiZZ), individuals with cystic fibrosis (CF) caused by inheritance of homozygous ΔF508 cystic fibrosis transmembrane conductance regulator (CFTR) mutations, or individuals with systemic sclerosis (SSc) underwent 6-mm full thickness skin punch biopsies from the arm. All procedures were approved by the Institutional Review Board (IRB) of the Boston University Medical Campus or the University of Vermont College of Medicine, and informed consent was documented from all individuals. Deidentified skin samples were digested overnight at 37°C with 0.25% collagenase I (Worthington-biochem.com; CLS-1) and 0.05% DNase I (Sigma, Worthington Biochemical Corp., Lakewood, NJ) in high-glucose Dulbecco's modified Eagle's medium (DMEM-Aldrich, St. Louis, MO, http://www.sigmaaldrich.com) containing 20% fetal bovine serum

(FBS). Cell suspensions cultured in T75 plates were split 1:3 when 80% confluent, to obtain outgrowth of dermal fibroblasts. Reprogramming was typically initiated on passage 3 or 4.

Human Fibroblast Reprogramming and Characterization of iPSC

A total of 1×10^5 human fibroblasts were plated in DMEM with 10% FBS on a gelatin-coated 35-mm plastic tissue culture dish. The next day polybrene was added to the media (5 μg/ml), and the cells were infected with hSTEMCCA-loxP lentiviruses at a multiplicity of infection (MOI) = 0.1, 1, or 10 where indicated in the text. On day 2, the media was changed to serum-free "iPSC media" (see below), and on day 6 the entire well was trypsinized and passed at a 1:16 split by plating onto two 10-cm gelatin-coated culture dishes which had been preseeded the day before with mitomycin C-inactivated mouse embryonic fibroblast (MEF) feeder cells. iPSC colonies were mechanically isolated 30 days postinfection with the four factor hSTEMCCA-loxP or 45 days postinfection with the three factor hSTEMCCA-RedLight-loxP based on morphology and expanded on MEF feeders in iPSC media. For three factor reprogramming, where indicated, GSK3 inhibitor (Bio; EMD Biosciences, San Diego, CA, http://www.emdchemicals.com, 361550; 10 μM) was added to the culture media on days 7–30 of reprogramming. Reprogramming efficiency was calculated by dividing the number of total colonies obtained on day 30 by the number of starting input fibroblasts. To avoid miscalculating efficiency that can occur when counting the progeny of passaged cells, efficiency was determined from separate experiments in which the input fibroblasts were not passaged onto feeders prior to colony counting.

Candidate iPSC clones were characterized based on staining for expression of alkaline phosphatase (Alkaline Phophatase Substrate Kit I, Vector Laboratories, Burlingame, CA, http://www.vectorlabs.com, SK–5100) or immunostaining of 4% paraformaldehyde-fixed cell colonies with antibodies against stage-specific embryonic antigen 4 (SSEA-4), TRA1–60, and TRA1–81 (ES Cell Characterization Kit, Millipore, Bedford, MA, http://www.millipore.com, SCR001). Primary antibodies were detected with secondary Alexa Fluor 488-conjugated, goat anti-mouse IgG or IgM (Invitrogen, Carlsbad, CA, http://www.invitrogen.com, A10680). In addition, the number of hSTEMCCA lentiviral integrations was determined by Southern blot of gDNA digested with BamHI and probed for the WPRE element as previously published [21]. Reverse transcription polymerase chain reaction (RT-PCR) was performed as previously described [21].

To evaluate the degree of DNA methylation of the human NANOG promoter, gDNA extracts of each indicated sample underwent bisulfite conversion using the EpiTect Bisulfite Kit (Qiagen, Chatsworth, CA, http://www.qiagen.com). Quantitative methylation analyses of 6 CpG islands in the proximal NANOG promoter were performed via pyrosequencing by EpigenDx Inc. (Worcester, MA, www.epigendx.com) using the ADS502/Human NANOG promoter assay, spanning positions −565 to −431 relative to the NANOG ATG start site.

For teratoma formation assays, six wells of a six-well plate of iPSC colonies were harvested with collagenase IV and resuspended in 140 μl of DMEM/F12. Immediately prior to injection, 60 μl of Matrigel (BD Biosciences, San Jose, CA, www.bdbiosciences.com) was added to the cell suspension at 4°C, and the resulting mixture was injected subdermally between the scapulae of each anesthetized Severe combined immunodeficiency (SCID)-Beige mouse (Charles River, Wilmington, MA, www.criver.com, strain 250). Resulting tumors were harvested at 6–8 weeks after injection, fixed in 4% paraformal-

dehyde, and paraffin tissue sections were prepared and stained with hematoxylin and eosin according to standard methods.

Tissue Culture Maintenance of Undifferentiated iPSC

Reprogrammed cells were propagated in "iPSC Media" consisting of DMEM F12 (Sigma-Aldrich) with 20% KnockOut Serum Replacement (Invitrogen), 1 mM nonanimal L-glutamine (Sigma-Aldrich), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 1% nonessential amino acid solution (Invitrogen), and 10 ng/ml of FGF2 (Invitrogen). Culture dishes were coated with sterile gelatin (Millipore) before use. The cells were cultured on a feeder layer of mitomycin-C (Fisher, Pittsburgh, PA, www.fishersci.com)-treated mouse embryonic fibroblasts (MEFs) and were incubated at 37°C at 5% CO_2. iPSC and hESC lines were typically passaged approximately every 5 days at a one-to-three split ratio. Collagenase IV (Invitrogen) was used to loosen the cells from the dish before mechanically scraping to pass. The cells were maintained in the undifferentiated state by scraping off differentiated cells with a glass pipette or alternatively by mechanical passage of individual colonies of undifferentiated cells.

Quantitative RT-PCR

Total RNA was isolated from cells with the RNeasy micro kit (Qiagen) and treated with RNase-free DNase (Qiagen). Hundred nanograms to one microgram of RNA was reverse transcribed into cDNA using random hexamers with Superscript III Reverse Transcriptase (Invitrogen). Real-time quantitative PCR was performed in triplicate for all samples using the LightCycler 480 Real-Time PCR System (Roche, Indianapolis, IN, www.roche.com) with LightCycler 480 SYBR Green I Master (Roche). A 10-fold dilution series of human gDNA ranging from 10 ng to 10 pg per reaction was used to evaluate the efficiency of the PCR and calculate the copy number of each gene relative to the housekeeping gene cyclophilin. Calculated expression levels for each indicated gene were then reported as number of molecules of RNA for that gene per number of molecules of cyclophilin, following a similar method previously described using mouse ESC [23]. Given this approach, all primers were designed not to cross introns; the oligonucleotide sequences for primers were as follows: Cyclophilin F-GAA GAG TGC GAT CAA GAA CCC ATG AC, R-GTC TCT CCT CCT TCT CCT CCT ATC TTT ACT T; DNMT3B F-TAC AGA CGT GTG CAG TTG TAG GCA, R-GTG CAG ACT CCA GCC CTT GTA TTT; NANOG-CCT GAA GAC GTG TGA AGA TGA G, R-GCT GAT TAG GCT CCA ACC ATA C; OCT4 F-AAC CTG GAG TTT GTG CCA GGG TTT, R-TGA ACT TCA CCT TCC CTC CAA CCA; SOX2 F-AGA AGA GGA GAG AGA AAG AAA GGG AGA GA, R-GAG AGA GGC AAA CTG GAA TCA GGA TCA AA; SOX17-AGG AAA TCC TCA GAC TCC TGG GTT, R-CCC AAA CTG TTC AAG TGG CAG ACA; FOXA2 F- GCA TTC CCA ATC TTG ACA CGG TGA; R-GCC CTT GCA GCC AGA ATA CAC ATT; SOX7 F-TGG AGG TTG CAG TGA GCT GAG ATT G; R- TGC ATG AAG TGG GCA TGT GTC TCT; HNF4A F-TCC AAC CCA ACC TCA TCC TCC TTC TT; R- TCC TCT CCA CTC CAA GTT CCT GTT.

Flow Cytometry

iPSC were dissociated by incubation with trypsin for 2–4 minutes and stained for the following cell surface antigens: anti-human CD117-allophycocyanin, anti-human CXCR4-phycoerythrin (Invitrogen), anti-human TRA 1–81-phycoerythrin (Biolegend, San Diego, CA, www.biolegend.com), and anti-human SSEA4–Alexa647 (Biolegend). Intracellular staining was performed as follows: Cells were fixed in 1.6%

paraformaldehyde for 20 minutes at 37°C, washed, then permeabilized, and stained in ×1 saponin buffer (Biolegend). The permeabilized cells were stained with anti-human FOXA1 (Biotechnology, Santa Cruz, CA, http://www.scbt.com, sc-101058) and anti-human OCT4 (Santa Cruz, Biotechnology sc-5279) followed by goat anti-mouse IgG2a-DyLight488 (Jackson ImmunoResearch, West Grove, PA, www.jacksonimmuno. com) and goat anti-mouse IgG2b-DyLight649 (Jackson ImmunoResearch). Cells were analyzed on a FACSCantos II flow cytometer (BD Biosciences) and analysis was performed using FlowJo software (Tree Star Inc., Ashland, OR, www.treestar.com).

Excision of hSTEMCCA

The transfection of human iPSC colonies for the excision of viral sequences was performed using the Hela Monster transfection reagent (Mirus, Madison, WI, http://www.mirusbio. com) according to manufacturer's instructions. Briefly, 30% confluent 35-mm tissue culture wells of iPSC colonies growing on puromycin resistant MEFs were exposed to media containing 2 μg of pHAGE2-Cre-IRES-PuroR plasmid DNA along with 6 μl of transfection reagent and 3 μl of the Monster reagent. The media was changed the next day and puromycin selection (1.2 μg/ml) was begun 24 hours following the initial transfection and lasted for 48 hours. The re-emergence of 10–50 colonies was noted within 1 week, and five colonies from each well were picked on day 11–14 for passaging. PCR of gDNA extracted from each subclone was performed to screen for excision of the hSTEMCCA vector using the following primers and conditions: cMYC F5′-GGA ACT CTT GTG CGT AAG TCG ATA G-3′; WPRE R5′-GGA GGC GGC CCA AAG GGA GAT CCG-3′; 95°C for 3 minutes; followed by 33 cycles of 94°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute; followed by a single cycle of 72°C for 5 minutes. Vector excision was then confirmed by Southern blotting of BamHI digested gDNA, probed against WPRE as above.

Directed Differentiation of Human iPSC into Definitive Endoderm

Cells were differentiated in a serum-free media base developed for mouse ES cell differentiation [24]. Briefly, iPSC were plated on Matrigel (BD Biosciences) in iPSC media to deplete feeder cells prior to differentiation. iPSCs were trypsinized for 1–2 minutes with cold 0.25% trypsin/EDTA then scraped off the dish to form cell clusters (5–50 cells per cluster). To induce differentiation, cells were first cultured for 24 hours in serum-free media with the addition of human bone morphogenetic protein 4 (BMP4) at 10 ng/ml (R&D systems, Minneapolis, MN, www.rndsystems.com). The next day to induce differentiation toward definitive endoderm, the cells were changed to media containing Activin A at 100 ng/ml (R&D systems), FGF2 2.5 ng/ml, and BMP4 0.5 ng/ml. The cultures were fed fresh media on day 4 and analyzed for differentiation on day 6. As a negative control, parallel aliquots of cells in each experiment were differentiated into extraembryonic endoderm. This was performed by addition of only BMP4 10 ng/ml in serum-free differentiation media with iPSC plated in adherent cultures on Matrigel and analyzed on day 6.

RESULTS

## Design of a Human STEMCCA Vector for the Efficient Generation of Human iPSC

Using a similar strategy that was employed to generate the mouse STEMCCA vector [21, 22], we constructed a single

lentiviral vector expressing a constitutive polycistronic message encoding the four human transcription factors, OCT4, KLF4, SOX2, and cMYC (Fig. 1A). In addition, we inserted a loxP site at the 3′ LTR of the vector, for future Cre-mediated excision, and named this vector hSTEMCCA-loxP. A combination of 2A peptide and IRES elements allowed for the production of the four individual transcription factors. The simultaneous expression of all four factors resulted in a high efficiency of reprogramming human cells, as evidenced by the large number of putative iPSC colonies obtained after infection of human foreskin fibroblasts (HFF) with concentrated hSTEMCCA-loxP viral particles at a MOI of one (Fig. 1B). The efficiency of reprogramming using our vector was ∼1%, however, this result varied with viral MOI. For example, increasing MOIs >10 resulted in declining reprogramming efficiencies and observable death of HFF during the first week postinfection, suggesting some toxicity associated with either viral transduction or transgene overexpression. The dynamics of reprogramming with the hSTEMCCA-loxP vector was evident morphologically by microscopy beginning 6–8 days post-transduction, followed by formation of early colonies 12–15 days post-transduction. Outgrowth of mature candidate iPSC colonies appeared at 25 days, which were picked at 30 days after initial infection. Using this method, >70% of picked colonies expanded in culture after passaging with a morphology indistinguishable from control H9 human ESC (Fig. 1B). Passaged colonies expressed a broad panel of "pluripotent marker genes" by RT-PCR and immunostaining, including GDF3, NANOG, hTERT, REX1, TRA-1–81, TRA-1–60, and SSEA-4 (Fig. 1C, 1D). G-banding analysis of a representative clone (Fig. 1E) revealed a normal 46XY karyotype with no apparent chromosomal abnormalities.

iPSC Generated with the hSTEMCCA-LoxP Vector Contain a Single Excisable Viral Integrant

An important attribute of a human reprogramming vector should be its ability to reliably produce iPSC clones containing a single viral integration that could be simply excised allowing the generation of iPSC free of exogenous transgenes. Although this approach has been achieved previously with a piggyBAC transposon-transposase method [19, 25] in mouse embryonic fibroblasts, the method has yet to be successfully applied for generating transgene-free human iPSC. Importantly, an excisable multivector lentiviral system has been utilized to reprogram fibroblasts from individuals with Parkinson's disease at low efficiency [18], however, multiple integrated copies of the four dox-inducible vectors, along with additional integrated copies of reverse tetracycline-controlled transactivator (RTTA)-expressing vectors were required. To achieve reprogramming of human cells with a single viral integration, we reasoned that reducing the number of infectious particles would increase the chance of obtaining single copy integrations for reprogramming. Surprisingly, as evidenced by Southern blot analysis of gDNA, we observed that regardless of the MOI (ranging from 0.1 to 10), we consistently obtained a high percentage (94% ± 9%; avg ± SD) of iPSC colonies containing a single viral integration (Figs. 1F, 3). This suggests that a specific range of reprogramming transgene expression, in this case depending on the number of viral copies per cell, must be obtained for a transduced cell to complete reprogramming to the point of colony expansion postpicking.

We have previously reported the use in mice of a modified version of the floxed STEMCCA vector, where cMyc is replaced with the red fluorescent reporter mCherry (STEMCCA-RedLight-loxP), allowing monitoring in real time of Cre/loxP-mediated vector excision during the derivation of transgene-free murine iPSC. Because the transgene cMYC has
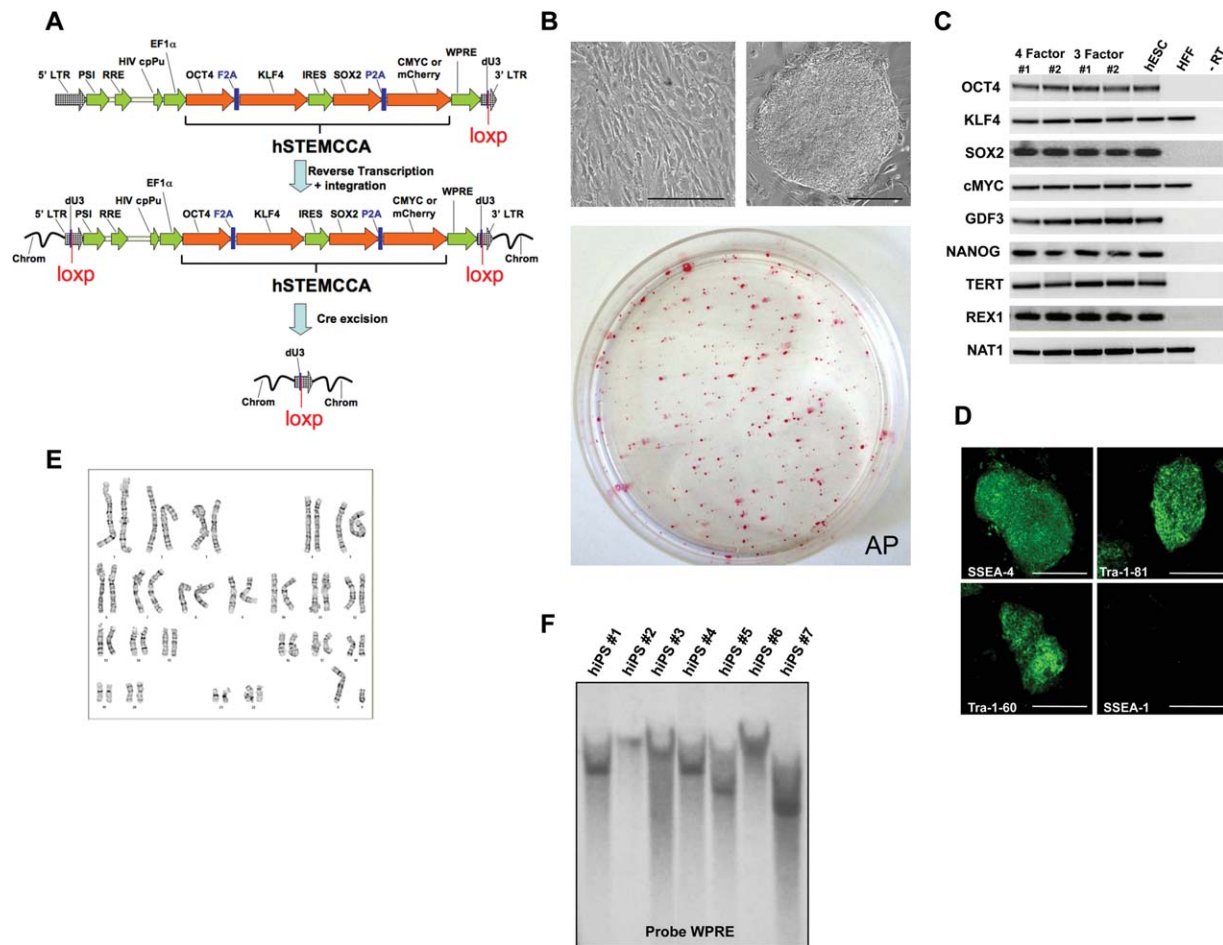
Figure 1. Human induced pluripotent stem cell (iPSC) generation using a humanized, floxed single lentiviral stem cell cassette (hSTEMCCA-loxP). (**A**): Vector schematic illustrating the polycistronic lentiviral backbone encoding either four reprogramming factors or three factors plus mCherry. A loxP site inserted in the viral 3′LTR is duplicated to the 5′LTR during viral infection and reverse transcription. The resulting floxed vector integrated in the host mammalian genome can then be excised on exposure to Cre recombinase. (**B**): Representative micrographs of HFF in culture (left panel) and reprogrammed human iPSC colonies (right panel). Multiple alkaline phosphatase positive colonies are observed 30 days after reprogramming 50,000 human fibroblasts with the hSTEMCCA-loxP virus. (**C, D**): Characterization of four independent iPSC clones, generated with either four factor or three factor hSTEMCCA-loxP, showing expression by reverse transcription polymerase chain reaction (**C**) and immunostaining (**D**) of typical pluripotent stem cell markers. (**E**): Representative normal 46XY karyotype of HFF-derived iPSC clone. (**F**): Southern blot of BamHI digested genomic DNA from representative HFF-derived iPSC clones, probed against WPRE, demonstrates a single viral integration in all clones. Scale bars = 200 μm (**B**); 250 μm (**D**). Abbreviations: CMYC, cellular myelocytomatosis oncogene; dU3, deleted U3 region of viral LTR; EF1α, elongation factor 1 alpha constitutive promoter; hESC, H9 human embryonic stem cell; HFF, human foreskin fibroblasts; hiPS, human induced pluripotent stem cell; HIV, human immunodeficiency virus; hSTEMCCA, humanized version of the single lentiviral "stem cell cassette"; IRES, internal ribosome entry site; LTR, long terminal repeats; PRE, post-transcriptional regulatory element; PSI, packaging signal; RT, reverse transcription; SSEA, stage-specific embryonic antigen; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

also been shown to be dispensable for reprogramming human cells [26], we generated a humanized version of the STEMCCA-RedLight-loxP vector by similarly replacing cMYC with mCherry in the hSTEMCCA vector. As in previous reports, the absence of cMYC diminished the overall efficiency of reprogramming and extended the time required to detect reprogrammed colonies to at least 6 weeks postinfection (data not shown). We found three factor reprogramming with this approach to be unreliable, however, as dermal fibroblasts obtained from two different adults failed to produce any colonies even after 8 weeks postinfection with hSTEMCCA-RedLight-loxP despite >90% transduction efficiency monitored by mCherry transduction. In contrast, the same fibroblasts from these two individuals yielded >100 colonies within 4 weeks of reprogramming with the four-factor hSTEMCCA-loxP vector. Consequently, we used a known GSK3 inhibitor (BIO) that has been suggested to enhance the efficiency of reprogramming

[27]. Indeed, the presence of BIO allowed for the generation of iPSC using the hSTEMCCA-RedLight-loxP vector with an efficiency of ∼0.01%. However, we noted the mCherry fluorochrome, which was easily visible during the first 3 weeks of reprogramming, became undetectable by fluorescence microscopy in seven of seven picked iPSC clones within two passages (Fig. 2A). Absence of mCherry expression following reprogramming of human cells was confirmed by fluorescence-activated cell sorting (FACS), suggesting some degree of silencing of the lentiviral vector (Fig. 2B), a result that sharply contrasted with our previous observation that the STEMCCA-RedLight-loxP vector is not silenced in mouse iPSC [22]. As a result of the observation that lentiviral silencing followed human reprogramming, the application of the mCherry-containing vector to visually monitor vector cre-excision in human cells was not possible. Thus, in all subsequent studies, we elected to employ only the four-factor hSTEMCCA-loxP vector.
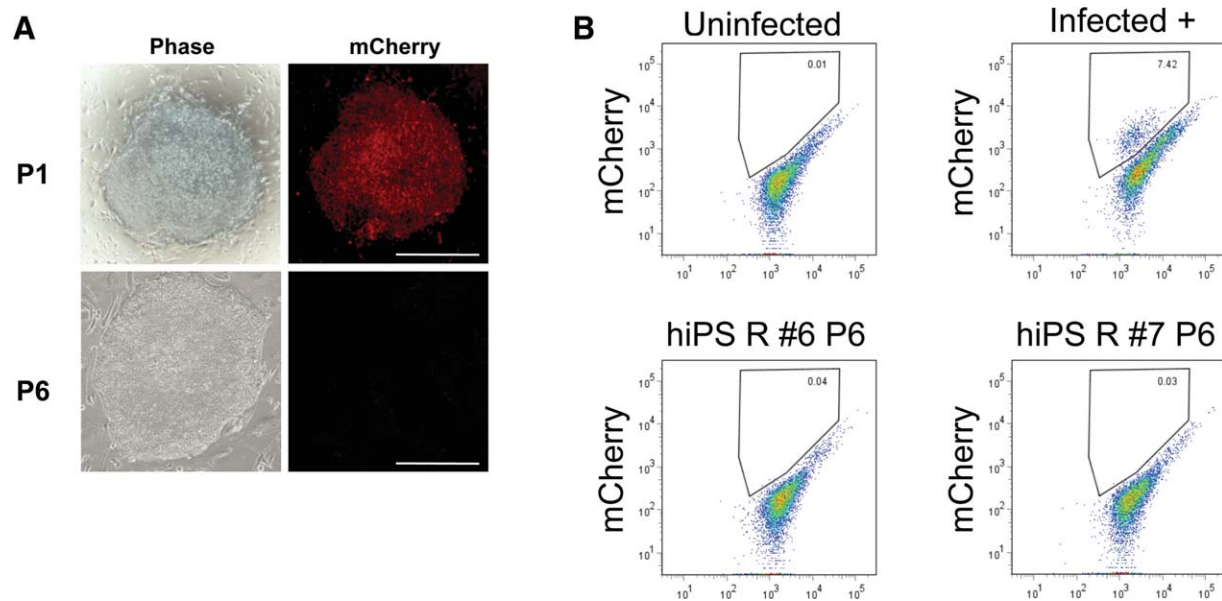
Figure 2. Human three factor reprogramming with humanized version of the single lentiviral "stem cell cassette" vector (hSTEMCCA)-Red-Light-loxP vector. **(A)**: Flourescence microscopy shows mCherry fluorescence in an early passage (P1) induced pluripotent stem cell (iPSC) clone, which becomes undetectable in later passages (P6). **(B)**: Flow cytometry of uninfected control human embryonic stem cell (hESC) as well as hESC 72 hours after hSTEMCCA-RedLight-loxP lentiviral infection (infected+), compared with two individual late passage (P6) iPSC clones (hiPS R #6 and #7) showing absence of detectable mCherry gene expression. Scale bar = 200 μm. Abbreviation: hiPS, human induced pluripotent stem cell.
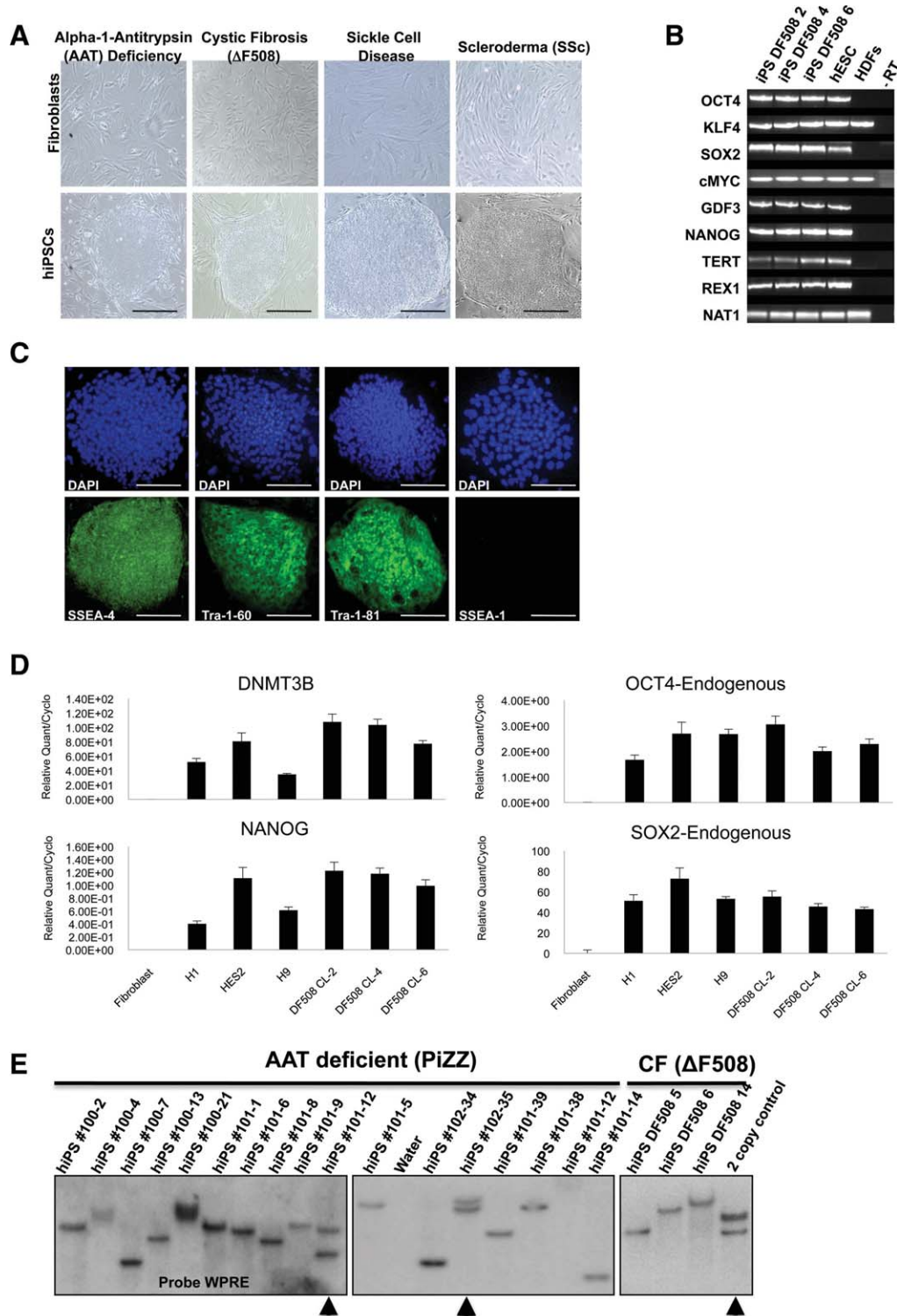
Generation of Human Lung Disease-Specific iPSC

We sought to apply our reprogramming vector for the generation of disease-specific iPSC free of reprogramming transgenes. We chose to derive iPSC from individuals with a variety of lung diseases, as there is a significant lack of conventional therapies or human model systems for many inherited or degenerative lung diseases. Pluripotent stem cells or their differentiated progeny have been proposed as attractive candidates for reconstituting injured lung tissues in vivo or modeling lung disease pathogenesis in vitro [28], however, to date no lung-disease-specific iPSC have been available. Hence, we sought to reprogram fibroblasts taken from humans with diseases affecting the three broad cell lineages of the adult lung: lung epithelium, endothelium, and interstitium. Fibroblasts were obtained from humans with either of the two most common inherited lung diseases: CF (which affects the airway epithelium) or AAT deficiency-related emphysema (which affects the lung interstitium and epithelium [29]). In addition, we obtained fibroblasts from individuals with systemic diseases known to frequently affect the lung: sickle-cell disease (which results in pulmonary endothelial injury and pulmonary arterial hypertension [30]), and SSc (scleroderma, which frequently leads to pulmonary arterial hypertension as well as interstitial pneumonitis [31]). We focused on reprogramming dermal fibroblasts that we obtained from 6-mm skin punch biopsies taken from recruited volunteers with either inherited PiZZ phenotype AAT deficiency, CF (homozygous ΔF508 mutant CFTR genotype), or SSc. In some cases, we obtained additional banked, frozen CF or AAT-deficient fibroblasts from the Coriell Cell Repositories. Using the same strategy detailed earlier, the hSTEMCCA-loxP vector allowed efficient generation of iPSC clones from all samples (Fig. 3A), regardless of the age of the individual from which the cells originated (Table 1). As with normal HFF-derived iPSC, the disease-specific iPSC we generated expressed the full complement of stem cell marker genes by RT-PCR, quantitative (q) RT-PCR, and immunostaining (Fig. 3B–3D). In addition, the majority of clones had been reprogrammed with a single integrated vector copy as evidenced by Southern blot (Fig. 3E) and G-banding analysis revealed a normal karyotype (Fig. 5E). The range of reprogramming efficiency in these experiments ranged from 0.1% to 1.5% without significant correlation of efficiency with age, gender, or disease type (Table 1 and data not shown). Finally, to functionally assess pluripotency of the disease-specific iPS cells generated with this method, we subdermally transplanted three representative iPS cell lines into immunodeficient mice and found the cells gave rise to teratomas comprised differentiated tissues characteristic of the three primary germ layers, ectoderm, mesoderm, and endoderm (Fig. 4).

Generation of Transgene-Free Human iPSC by Cre/LoxP-Mediated Excision

Next, we sought to excise the single vector copy to generate lung disease-specific iPSC free of reprogramming transgenes. We have previously shown that constitutive expression of reprogramming transgenes interferes with mouse iPSC differentiation into lineages of all three primary germ layers [22]. Moreover, aberrant expression of some or all of the reprogramming factors could lead to tumorigenesis [2] in vivo and may affect global gene expression [18]. For these reasons, we aimed at generating iPSC free of reprogramming transgenes based on Cre/loxP-mediated excision of the reprogramming cassette (See schematic Fig. 5A). During normal lentiviral reverse transcription, the loxP site present in the deleted U3 region of the vector long terminal repeats (LTR) is duplicated to the 5′ LTR, resulting in a floxed version of the hSTEMCCA (Fig. 1A). To achieve excision, we selected iPSC clones containing a single integration and performed transient transfection with a plasmid expressing Cre and a puromycin resistance gene (Cre-IRES-PuroR). The cells were then exposed to
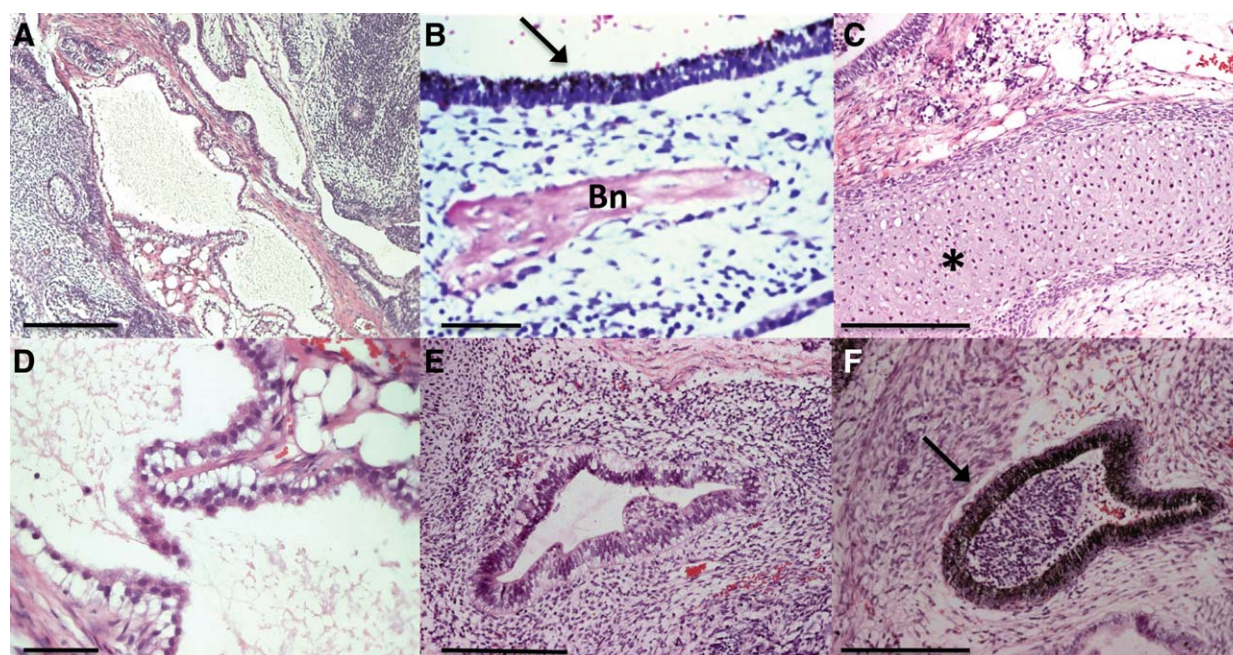
Figure 3. Generation of disease-specific iPSC. **(A)**: Human fibroblasts in culture and their reprogrammed iPSC progeny, derived from individuals with α-1 antitrypsin deficiency, cystic fibrosis, sickle-cell disease, and SSc. **(B, C)**: Characterization of lung disease-specific cystic fibrosis iPSC by reverse transcription polymerase chain reaction (RT-PCR) **(B)**, immunostaining **(C)**, and real time quantitative RT-PCR **(D)** demonstrating expression of stem cell markers, quantified as molecules of RNA of the indicated gene per molecules of RNA of the housekeeping gene, cyclophilin. H1, H9, and HES2 are three control lines of human embryonic stem cells. **(E)**: Southern blot of BamHI digested genomic DNA demonstrating 18 of 20 individual iPSC clones (made from four individuals with either inherited AAT deficiency or CF) have been reprogrammed with a single integrated vector copy. Arrowheads indicate lanes with clones reprogrammed with two vector copies as well as a known two copy clone (far right lane). Blot has been probed against WPRE sequence of the lentiviral vector backbone. "Clone #" represents human volunteer (100–102) followed by clone number after hyphen (e.g., 100-2). Scale bar = 200 μm. Abbreviations: AAT, α-1 antitrypsin; CF, cystic fibrosis; CL, clone number; DAPI, 4′-6- diamidino-2-phenylindole; DF508, ΔF508 homozygous cystic fibrosis iPSC clones; HDFs, human dermal fibroblasts; hESC, human embryonic stem cell; hiPSCs, human induced pluripotent stem cells; iPS, induced pluripotent stem cell; PiZZ, homozygous Z alleles of the AAT protease inhibitor; RT, reverse transcription; SSc, scleroderma; SSEA, stage-specific embryonic antigen; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

Table 1. Human iPSC lines derived from somatic parental cell lines of varying disease types using one of two reprogramming vectors

Disease type	Parental cell line	Source	Sex	Age	Diseased allele(s)	Vector	Number of iPSC lines
Wild-type	HFF	ATCC, CRL2097	M	Neonatal	N/A	EF1α-hSTEMCCA-loxP	8
						EF1α-hSTEMCCA-RedLight-loxP	11
Emphysema (AAT-deficient)	Dermal Fibroblast	Patient biopsies (4)	F	64 yr	PiZZ	EF1α-hSTEMCCA-loxP	15
				61 yr			15
				57 yr			15
				47 yr			20
	Liver Fibroblast	Coriell, GM11423	M	16 wk			9
Cystic fibrosis	Fibroblast	Coriell, GM1012A	M	8 yr	Homozygous	EF1α-hSTEMCCA-loxP	15
	Dermal fibroblast	Patient biopsies (4)	F	21 yr	ΔF508		10
			F	31 yr			10
			M	33 yr			4
			M	29 yr			2
Scleroderma	Dermal fibroblast	Patient biopsy	F	47 yr	Unknown	EF1α-hSTEMCCA-loxP	3
Sickle cell	Fibroblast	Coriell, GM02340	F	20 yr	SS	EF1α-hSTEMCCA-loxP	26

Abbreviations: AAT, α-1 antitrypsin; iPSC, induced pluripotent stem cell; PiZZ, Z alleles of the AAT protease inhibitor; SS, sickle cell alleles.


Figure 4. Functional in vivo pluripotency of disease-specific human iPSC assessed by teratoma formation assay. Transplantation of a representative cystic fibrosis-specific induced pluripotent stem cell (clone 202) into immunodeficient mice gave rise to a teratoma containing differentiated lineages representing the three primary germ layers, mesoderm, ectoderm, and endoderm. **(A):** Low-power magnification (×4) of hematoxylin and eosin-stained tissue section of teratoma, illustrating formation of multiple glandular and cystic structures. **(B):** Mesodermal differentiation demonstrated by Bn and ectoderm differentiation demonstrated by pigmented neuroepithelium (arrow), ×40 magnification. **(C):** Cartilage formation (*) indicative of mesodermal differentiation. **(D, E):** Glandular (endodermal) epithelium. **(F):** Pigmented ectodermal differentiation reminiscent of retinal pigmented epithelium (arrow). Scale bars = 200 μm **(A)**, 100 μm **(B, C, E, F)**, and 50 μm **(D)**. Abbreviation: Bn, bone formation.

puromycin antibiotic selection for 48 hours, followed by a 1–2-week recovery period prior to picking any surviving colonies (Fig. 5D). Employing this method, we were able to recover transgene free subclones within 2 weeks of transfection, as demonstrated by both PCR and Southern blotting (Fig. 5B, 5C). In three repeated experiments, five subclones were derived from each of 10 disease-specific iPSC lines (50 subclones total; CF and AAT deficient) following Cre exposure. Deletion of the hSTEMCCA vector was found in 100% of these screened subclones (Fig. 5 and Supporting Information Fig. 1). In subsequent experiments, occasionally we found iPSC lines (n = 2 to date) that were resistant to successful Cre-mediated excision for

unclear reasons. Our inclusion in the Cre-IRES-PuroR plasmid of a WPRE sequence allowed us to exclude the presence of any residual integrated STEMCCA or Cre-IRES-PuroR events in the successfully excised clones, using a Southern blot probe able to detect this element common to both vectors (Fig. 5). To further exclude the possibility of integration of the Cre-IRES-PuroR plasmid in the excised clones, we demonstrated the return of puromycin susceptibility in all subclones 2 weeks after picking by re-exposure to puromycin antibiotic for 48 hours with resultant death of >99% of cells.

Following Cre-mediated excision, we assessed the self-renewal capacity, stability of epigenetic reprogramming, and
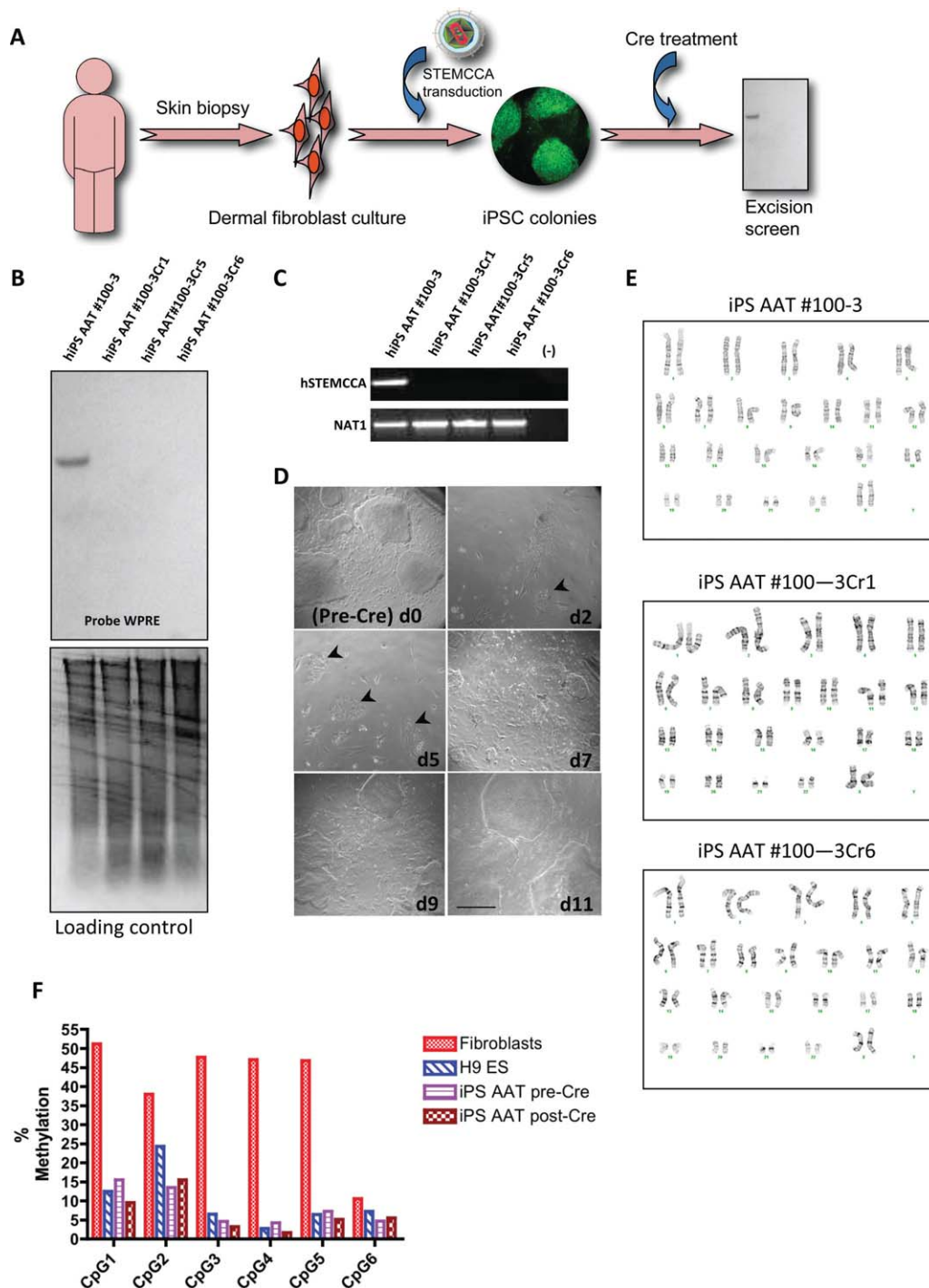
Figure 5. Cre-mediated excision of hSTEMCCA-loxP to generate transgene-free lung disease-specific iPSC. **(A)**: Schematic summarizing experimental approach for generating transgene-free iPSC. **(B)**: Southern blot of BamHI-digested genomic DNA (gDNA) probed against the lentiviral WPRE fragment (top), demonstrating a representative parental iPSC line (AAT #100-3) and successful Cre-mediated excision of the single copy hSTEMCCA vector in three iPSC subclones (AAT #100-3-Cr1, -Cr5, and -Cr6). Equal DNA loading in each lane is evident by ethidium bromide (bottom) staining of the gel prior to transfer. **(C)**: Polymerase chain reaction of gDNA confirming vector excision of the clones shown in **(B)**. **(D)**: Micrographs of iPSC showing colony morphologies after transfection of Cre-IRES-PuroR plasmid and 48 hours of antibiotic selection of excised colonies, at different time points after selection. Arrowheads denote emerging iPSC colonies. **(E)**: Karyotypic stability (46XX) by G-banding analysis of iPSC clone AAT #100-3 both before Cre-mediated vector excision, as well as in the subclones after vector excision. **(F)**: Pyrosequencing analysis of bisulfite-treated gDNA provides quantitation of methylation of the human NANOG promoter across 6 CpG islands in dermal fibroblasts prior to reprogramming, as well as in the AAT iPSC clone before and after Cre-mediated excision (pre-Cre vs. post-Cre) of the hSTEMCCA vector. Scale bar = 200 μm. Abbreviations: AAT, α-1 antitrypsin; H9 ES, control human embryonic stem cells; hiPS, human induced pluripotent stem cell; hSTEMCCA, humanized version of the single lentiviral "stem cell cassette"; iPSC, induced pluripotent stem cell; STEMCCA, stem cell cassette; WPRE, Woodchuck hepatitis virus post-transcriptional regulatory element.

karyotypic stability of the transgene-free iPSC as follows: AAT-deficient iPSC (prevector and postvector excision) were maintained in continuous culture for 3 months (equivalent to passage 20) and karyotypes were evaluated by G-banding analyses (Fig. 5E). Even after this prolonged time in culture, the iPSC showed stability of normal karyotype, stable colony morphology, and stable expression of stem cell marker genes (SSEA4, TRA-1–81, and TRA-1–60 by FACS and immunostaining; SOX2, DNMT3B, OCT4, NANOG, and REX1 by quantitative RT-PCR; Fig. 5 and Supporting Information Fig. 2). As expected, quantitative RT-PCR evaluation of hSTEMCCA transgene mRNA expression revealed removal of all traces of transgene expression in all subclones after Cre-mediated excision of the hSTEMCCA vector (Supporting Information Fig. 3).

To assess stable reprogramming of the epigenetic state of the NANOG promoter in disease-specific iPS cells before versus after vector excision, we quantified the degree of DNA methylation of 6 CpG islands in the NANOG promoter by pyrosequencing. We found five of six CpG islands evaluated were mostly methylated in dermal fibroblasts prior to reprogramming, but six of six islands were unmethylated in iPSC both before and after vector excision, similar to control H9 ESC (Fig. 5F).

Differentiation of iPSC into Definitive Endoderm, the Developmental Precursor Lineage of Lung and Liver Epithelia

The lung epithelium develops in the embryo from multipotent definitive endoderm progenitors of the anterior foregut [32]. Only recently, protocols have been developed for the reliable derivation of definitive endoderm progenitors from ESC [33–35]. Deriving definitive endoderm from pluripotent stem cells generated from AAT-deficient individuals is also a prerequisite for obtaining endoderm-derived hepatic cells, an important goal as hepatocytes are the main secretors of circulating AAT protein in mammals. Hence, we attempted to direct the differentiation of our iPSC toward definitive endoderm, the direct precursor population of lung and liver epithelium. As shown in Figure 6A and 6B, using a previously described method for activating nodal/activin signaling [33, 34], iPSC were differentiated in serum-free culture conditions into definitive endoderm cells expressing the endodermal transcriptional regulators, FOXA1, FOXA2, SOX17, and HNF4A and the cell surface markers CXCR4 and CD117. Induction of endodermal markers was accompanied by expression of the known nodal target CER1 [36] together with loss of expression of the pluripotent markers SSEA3 and OCT4 (Supporting Information Fig. 4 and data not shown). To confirm directed differentiation into definitive endoderm rather than extraembryonic/visceral endoderm, we assessed expression levels of the early extraembryonic endodermal marker, SOX7, and found no evidence of induction of this lineage. In contrast, SOX7 expression was robustly induced in negative control aliquots of cells differentiated in parallel in conditions designed to induce extraembryonic endoderm (in the presence of BMP4 alone). The efficiency of definitive endodermal differentiation of the iPSC (70%–80% by day 6; Fig. 6) was robust in disease-specific iPSC both before or after vector excision, a finding of particular importance if iPSC are to be used to generate endodermal progenitors for lung and liver disease modeling or future clinical use.

DISCUSSION

Our results demonstrate a complete platform for reliably reprogramming disease-specific somatic cells from fresh or

banked clinical samples into iPSC that are free of reprogramming transgenes. The reliability of the platform rests on the efficiency of the hSTEMCCA-loxP vector to obtain hundreds of putative iPSC clones from a single starting 35-mm plate of human dermal fibroblasts. The vast majority of clones obtained with this method contain a single integrated vector copy, which can be excised to obtain transgene-free iPSC clones. We have illustrated how this method may be widely applied for clinical studies by recruiting volunteers and producing >100 of the first known iPSC relevant for the study of a variety of lung diseases (summarized in Table 1). At this point the entire process from skin biopsy to banking transgene-free iPSC requires ~3–4 months in a protocol that will need further optimization and high-throughput adaptation if it is to be properly applied for clinical trials or epidemiologic and genome-wide association studies of disease.

Although several methods have been published previously for deriving iPSC free of any reprogramming transgenes, most methods to date have been employed either to reprogram only murine cells or to reprogram only neonatal normal human cells with an efficiency too low to ensure reliable reprogramming of valuable clinical disease-specific samples. In contrast, we show that similar to our original description in the mouse system, we can consistently and reliably reprogram human somatic cells with high efficiency by using a single polycistronic excisable lentivirus expressing the four transcription factors, OCT4, KLF4, SOX2, and cMYC. The high efficiency of reprogramming likely derives in part from favorable stoichiometry of the four reprogramming transcription factors when expressed in this particular order by a polycistronic system, as suggested by previous publications [14]. More importantly, reprogramming with our system can be achieved regardless of the age or disease state of the donor individual. We believe that these are important practical considerations that will have a significant impact on the use of this technology in the clinical arena.

Several teams have noted that vector-based reprogramming is often "incomplete" resulting in a significant number of "partially reprogrammed" iPSC clones [37, 38], which can be distinguished from fully reprogrammed clones on the basis of marker gene expression [37]. We found all clones generated with the hSTEMCCA-loxP vector expressed a broad complement of "stem cell markers" including those previously reported to be absent in "partially reprogrammed" clones, such as TRA-1–60, REX1, and DNMT3B [37]. We speculate that the majority of those clones reported as partially reprogrammed in other studies may have arisen from cells that either did not receive all reprogramming factors or expressed the factors with stoichiometries or expression levels that did not allow for complete reprogramming. As most clones generated with the hSTEMCCA-loxP vector received a single copy of all four reprogramming factors, partial reprogramming may be minimized, potentially explaining why our results contrast with those of prior studies using multivector approaches.

It is often claimed that iPSC will be utilized for regenerative medicine applications, either as in vitro models of disease or as in vivo vehicles for gene or cell therapies. For this promise to be realized, iPSC will need to be generated from individuals with inherited or degenerative diseases and differentiated into the relevant cell lineages required to model or treat those diseases. Although we predict there will be variability in the differentiation capacity of each iPSC clone generated by any method, we have demonstrated that iPSC derived with hSTEMCCA-loxP have robust potential to differentiate into definitive endoderm, the developmental precursor lineage of lung epithelium. In contrast to our previous studies
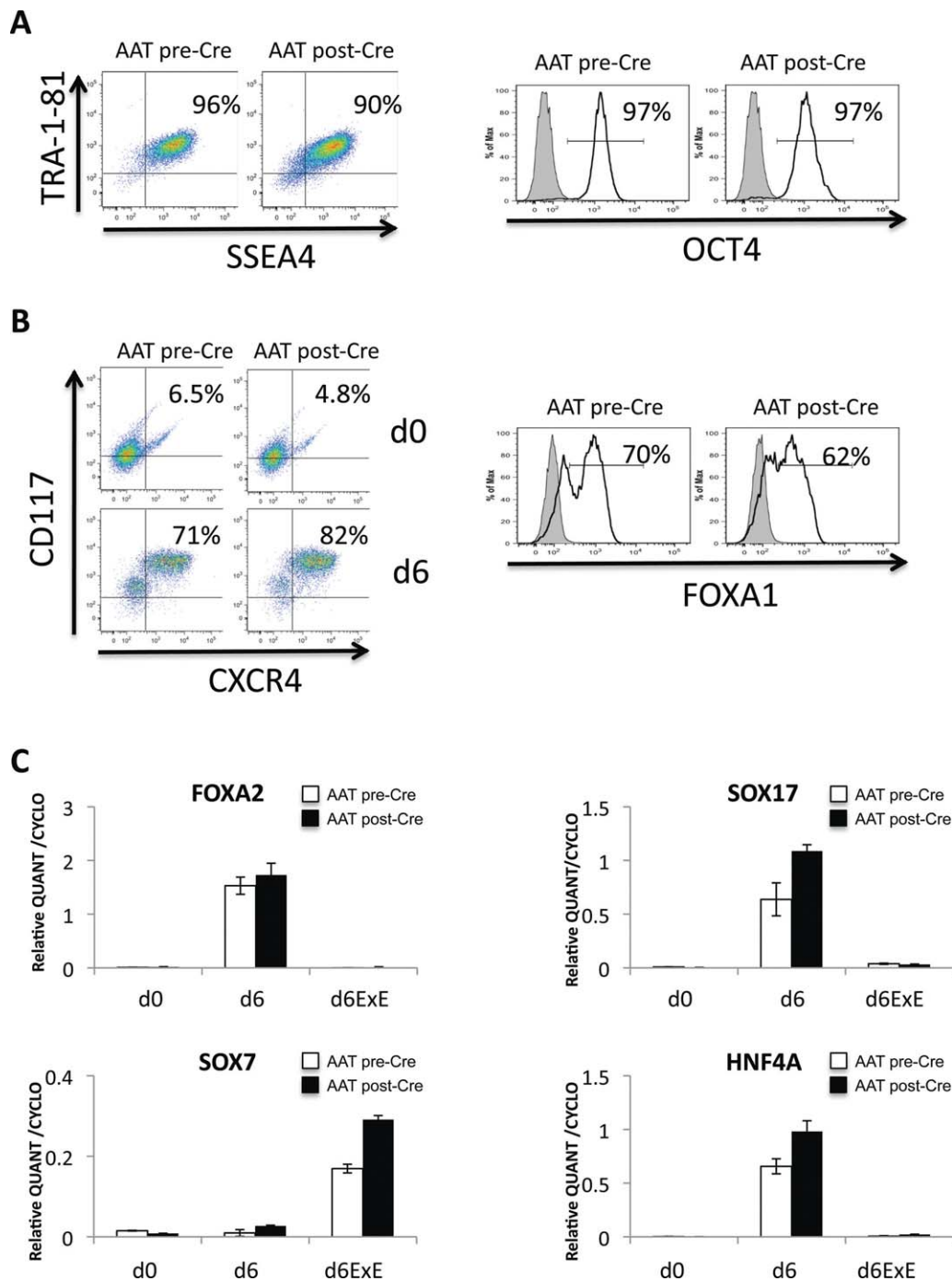
Figure 6. Directed differentiation of disease-specific induced pluripotent stem cell (iPSC) into definitive endoderm after activin stimulation. **(A):** Representative undifferentiated iPSC clone from an individual with AAT deficiency, assessed both before (AAT pre-Cre) and after (AAT post-Cre) excision of the humanized version of the single lentiviral "stem cell cassette" reprogramming vector. Left panel shows fluorescence-activated cell sorting (FACS) dot plot analyses of stage-specific embryonic antigen 4 (SSEA4) and TRA 1–81 cell surface markers. Right panel shows FACS analysis after intracellular staining for the stem cell marker, OCT4 (black line) versus isotype control staining (gray histogram). **(B, C):** Directed differentiation of the iPSC clones shown in **(A)**, in cultures designed to induce definitive endoderm versus extraembryonic endoderm over 6 days. FACS **(B)** was used to quantify induction of the definitive endodermal markers CXCR4, CD117 (C-KIT), and FOXA1. Gray histogram indicates undifferentiated cells on day 0 (d0). Quantitative reverse transcription polymerase chain reaction **(C)** shows induction of endodermal transcriptional regulators FOXA2, SOX17, and HNF4A on day 6 (d6), or the extraembryonic marker SOX7. Gene expression levels are represented as molecules of RNA of the indicated gene per molecule of RNA of the housekeeping gene, Cyclophilin. Abbreviation: AAT, α-1 antitrypsin.

in mouse iPSC, where excision of the STEMCCA vector clearly improved endodermal differentiation capacity [22], human iPSC appeared to differentiate efficiently to endoderm, both before or after vector excision, as quantified by the per-centage of cells expressing CXCR4, CD117, or FOXA1. However, qRT-PCR suggested improved expression of some endodermal marker genes following vector excision (Fig. 6B, 6C). Further assessment of the differentiation capacity of

many more clones prevector versus postvector excision, however, will be required to assess whether any significant effect on differentiation potential occurs following excision of the reprogramming vector in human iPSC.

For the two most common inherited lung diseases, CF and AAT deficiency-related emphysema, disease-specific iPSCs are particularly promising research tools. Both diseases result from single gene defects that typically cause misfolding of CFTR or AAT proteins in lung or liver epithelial cells, respectively. Primary cells from individuals with these diseases are difficult to expand in culture for studies of disease pathogenesis, drug therapy, or targeted correction of their diseased loci. Hence, the prospect of deriving an inexhaustible supply of disease-specific lung and liver epithelia from iPSC is attractive for modeling or treating these diseases.

It should be emphasized that clinical application of iPSC for therapy will need to take into account the benefit to risk ratio of employing this novel stem cell population with known potential to differentiate into any cell type and the capacity for tumorigenesis in vivo. Although we have minimized tumorigenic risk by removing exogenous transgenes, it should be noted, however, that ~200 bp of an inactive viral LTR remains in the host genome after excision, and hence, the theoretical risk of insertional mutagenesis that may arise from genomically integrated exogenous DNA is not completely eliminated. Although the statistical probability is minimal that a single integration event can induce gene disruption and dysregulation, this minimal risk could be further reduced in the future by targeting the STEMCCA into a safe genomic locus, as has been described [39]. Alternatively, improvement in the efficiency of reported reprogramming methods that do not modify the host genome would allow the future derivation of clinically relevant disease-specific iPSC, whereas avoiding the need for integrating sequences.

CONCLUSION

We demonstrate a system for the efficient reprogramming of human somatic cells utilizing a single integrated copy of a floxed polycistronic lentiviral vector expressing a humanized stem cell cassette (hSTEMCCA-loxP). Following the completion of reprogramming, the vector can be excised from the resulting human induced pluripotent stem cells (iPSC) via transient expression of Cre recombinase, generating iPSC free of residual reprogramming transgenes. We employ the system to generate >100 lines of the first known lung disease-specific human iPSC, and demonstrate robust capacity of the resulting pluripotent cells to undergo directed differentiation into definitive endoderm, the developmental precursor tissue of lung epithelia.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1   Maherali N, Sridharan R, Xie W et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell 2007;1:55–70.
2   Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. Nature 2007;448:313–317.
3   Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126:663–676.
4   Wernig M, Meissner A, Foreman R et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 2007;448: 318–324.
5   Staerk J, Dawlaty MM, Gao Q et al. Reprogramming of human peripheral blood cells to induced pluripotent stem cells. Cell Stem Cell 2010;7:20–24.
6   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.
7   Park IH, Zhao R, West JA et al. Reprogramming of human somatic cells to pluripotency with defined factors. Nature 2008;451: 141–146.
8   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.
9   Yu J, Vodyanik MA, Smuga-Otto K et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007;318: 1917–1920.
10 Lowry WE, Richter L, Yachechko R, et al. Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci USA 2008;105:2883–2888.
11 Huangfu D, Osafune K, Maehr R et al. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 2008;26:1269–1275.
12 Han J, Yuan P, Yang H et al. Tbx3 improves the germ-line competency of induced pluripotent stem cells. Nature 2010;463: 1096–1100.
13 Zhou H, Wu S, Joo JY et al. Generation of induced pluripotent stem cells using recombinant proteins. Cell Stem Cell 2009;4: 381–384.
14 Okita K, Nakagawa M, Hyenjong H et al. Generation of mouse induced pluripotent stem cells without viral vectors. Science 2008; 322:949–953.
15 Stadtfeld M, Nagaya M, Utikal J et al. Induced pluripotent stem cells generated without viral integration. Science 2008;322: 945–949.
16 Heng JC, Feng B, Han J et al. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. Cell Stem Cell 2010;6:167–174.
17 Kaji K, Norrby K, Paca A et al. Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature 2009;458: 771–775.
18 Soldner F, Hockemeyer D, Beard C et al. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell 2009;136:964–977.
19 Woltjen K, Michael IP, Mohseni P et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 2009;458: 766–770.

20 Yamanaka S. Strategies and new developments in the generation of patient-specific pluripotent stem cells. Cell Stem Cell 2007;1:39–49.
21 Sommer CA, Stadtfeld M, Murphy GJ et al. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells 2009;27:543–549.
22 Sommer CA, Sommer AG, Longmire TA et al. Excision of reprogramming transgenes improves the differentiation potential of iPS cells generated with a single excisable vector. Stem Cells 2010;28:64–74.
23 Nostro MC, Cheng X, Keller GM et al. Wnt, activin, and BMP signaling regulate distinct stages in the developmental pathway from embryonic stem cells to blood. Cell Stem Cell 2008;2:60–71.
24 Gadue P, Huber TL, Paddison PJ et al. Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. Proc Natl Acad Sci USA 2006;103:16806–16811.
25 Yusa K, Rad R, Takeda J et al. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. Nat Methods 2009;6:363–369.
26 Nakagawa M, Koyanagi M, Tanabe K et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 2008;26:101–106.
27 Silva J, Barrandon O, Nichols J et al. Promotion of reprogramming to ground state pluripotency by signal inhibition. Plos Biol 2008;6:e253.
28 Weiss DJ, Kolls JK, Ortiz LA et al. Stem cells and cell therapies in lung biology and lung diseases. Proc Am Thorac Soc 2008;5:637–667.
29 Davies JC, Alton EW. Airway gene therapy. Adv Genet 2005;54:291–314.

30 Klings ES, Wyszynski DF, Nolan VG et al. Abnormal pulmonary function in adults with sickle cell anemia. Am J Respir Crit Care Med 2006;173:1264–1269.
31 Lynch DA. Lung disease related to collagen vascular disease. J Thorac Imaging 2009;24:299–309.
32 Cardoso WV, Lu J. Regulation of early lung morphogenesis: Questions, facts and controversies. Development 2006;133:1611–1624.
33 Gouon-Evans V, Boussemart L, Gadue P et al. BMP-4 is required for hepatic specification of mouse embryonic stem cell-derived definitive endoderm. Nat Biotechnol 2006;24:1402–1411.
34 Kubo A, Shinozaki K, Shannon JM et al. Development of definitive endoderm from embryonic stem cells in culture. Development 2004;131:1651–1662.
35 D'Amour KA, Agulnick AD, Eliazer S et al. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat Biotechnol 2005;23:1534–1541.
36 Yamamoto M, Saijoh Y, Perea-Gomez A et al. Nodal antagonists regulate formation of the anteroposterior axis of the mouse embryo. Nature 2004;428:387–392.
37 Chan EM, Ratanasirintrawoot S, Park IH et al. Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat Biotechnol 2009;27:1033–1037.
38 Sridharan R, Tchieu J, Mason MJ et al. Role of the murine reprogramming factors in the induction of pluripotency. Cell 2009;136:364–377.
39 Stadtfeld M, Maherali N, Borkent M et al. A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nat Methods 2010;7:53–55.

See www.StemCells.com for supporting information available online.