Stem Cell Reports
Report
OPEN ACCESS

Targeted Correction and Restored Function of the *CFTR* Gene in Cystic Fibrosis Induced Pluripotent Stem Cells

Ana M. Crane,[1,7] Philipp Kramer,[1,7] Jacquelin H. Bui,[1,7] Wook Joon Chung,[2] Xuan Shirley Li,[1] Manuel L. Gonzalez-Garay,[3] Finn Hawkins,[4] Wei Liao,[1] Daniela Mora,[1] Sangbum Choi,[5] Jianbin Wang,[6] Helena C. Sun,[6] David E. Paschon,[6] Dmitry Y. Guschin,[6] Philip D. Gregory,[6] Darrell N. Kotton,[4] Michael C. Holmes,[6] Eric J. Sorscher,[2] and Brian R. Davis[1,*]

[1]Center for Stem Cell and Regenerative Medicine, Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center, Houston, TX 77030, USA
[2]Gregory Fleming James Cystic Fibrosis Research Center, University of Alabama, Birmingham, AL 35294, USA
[3]Center for Molecular Imaging, Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center, Houston, TX 77030, USA
[4]Center for Regenerative Medicine, Boston University and Boston Medical Center, Boston, MA 02118, USA
[5]Division of Clinical and Translational Sciences, Department of Internal Medicine, University of Texas Health Science Center, Houston, TX 77030, USA
[6]Sangamo BioSciences, Inc., Richmond, CA 94804, USA
[7]Co-first author
*Correspondence: brian.r.davis@uth.tmc.edu
http://dx.doi.org/10.1016/j.stemcr.2015.02.005

SUMMARY

Recently developed reprogramming and genome editing technologies make possible the derivation of corrected patient-specific pluripotent stem cell sources—potentially useful for the development of new therapeutic approaches. Starting with skin fibroblasts from patients diagnosed with cystic fibrosis, we derived and characterized induced pluripotent stem cell (iPSC) lines. We then utilized zinc-finger nucleases (ZFNs), designed to target the endogenous *CFTR* gene, to mediate correction of the inherited genetic mutation in these patient-derived lines via homology-directed repair (HDR). We observed an exquisitely sensitive, homology-dependent preference for targeting one *CFTR* allele versus the other. The corrected cystic fibrosis iPSCs, when induced to differentiate in vitro, expressed the corrected *CFTR* gene; importantly, *CFTR* correction resulted in restored expression of the mature CFTR glycoprotein and restoration of CFTR chloride channel function in iPSC-derived epithelial cells.

INTRODUCTION

Cellular transplantation of lung stem/progenitor cells represents a potential therapeutic approach for a variety of inherited monogenic lung diseases. A patient-specific approach would first involve derivation of autologous induced pluripotent stem cells (iPSCs) from skin or blood cells of affected patients. Utilizing site-specific homology-directed repair (HDR), the disease-causing mutation would then be corrected in the endogenous, chromosomal DNA sequence. Finally, a directed differentiation approach would be employed to obtain highly purified populations of the relevant lung stem/progenitor cells from the corrected iPSCs for the purpose of transplantation.

We focused our initial development of this therapeutic approach on cystic fibrosis (CF). The primary defect in CF, an autosomal recessive disorder, is the regulation of epithelial chloride transport by a chloride channel protein encoded by the CF transmembrane conductance regulator (*CFTR*) gene (Kerem et al., 1989). Recurrent pulmonary infections are responsible for 80%–90% of the deaths in CF patients. Therefore, transplantation of *CFTR*-corrected, autologous lung stem/progenitor cells provides an attractive alternative strategy for treating CF.

Here we used zinc-finger nuclease (ZFN)-mediated HDR to edit the endogenous *CFTR* locus and precisely correct mutations responsible for CF in patient-derived iPSCs.

RESULTS

Correction of *CFTR* Mutation via ZFN-Mediated HDR in CF iPSCs

Retroviral vectors encoding reprogramming factors (OCT4, SOX2, KLF4, C-MYC, and NANOG) were utilized to reprogram CF primary fibroblasts to iPSCs; the CF fibroblasts and derived iPSCs were compound heterozygous at the *CFTR* locus, with one allele ΔF508 and the other allele ΔI507 (Figure S1A). CF iPSC clones expressed cellular antigens characteristic of undifferentiated human embryonic stem cells (hESCs), were pluripotent as assayed by teratoma formation, and retained a normal karyotype (Figures S1B–S1D).

The overall strategy for correction of *CFTR* exon 10 (*CFTR* legacy exon notation) mutations consisted of delivering *CFTR*-specific ZFNs together with a selectable *CFTR* donor DNA (Figure 1A). We designed ZFNs targeting *CFTR* exon 10, recognizing DNA sequences approximately 110 bp upstream of either the ΔI507 or ΔF508 deletions, to facilitate

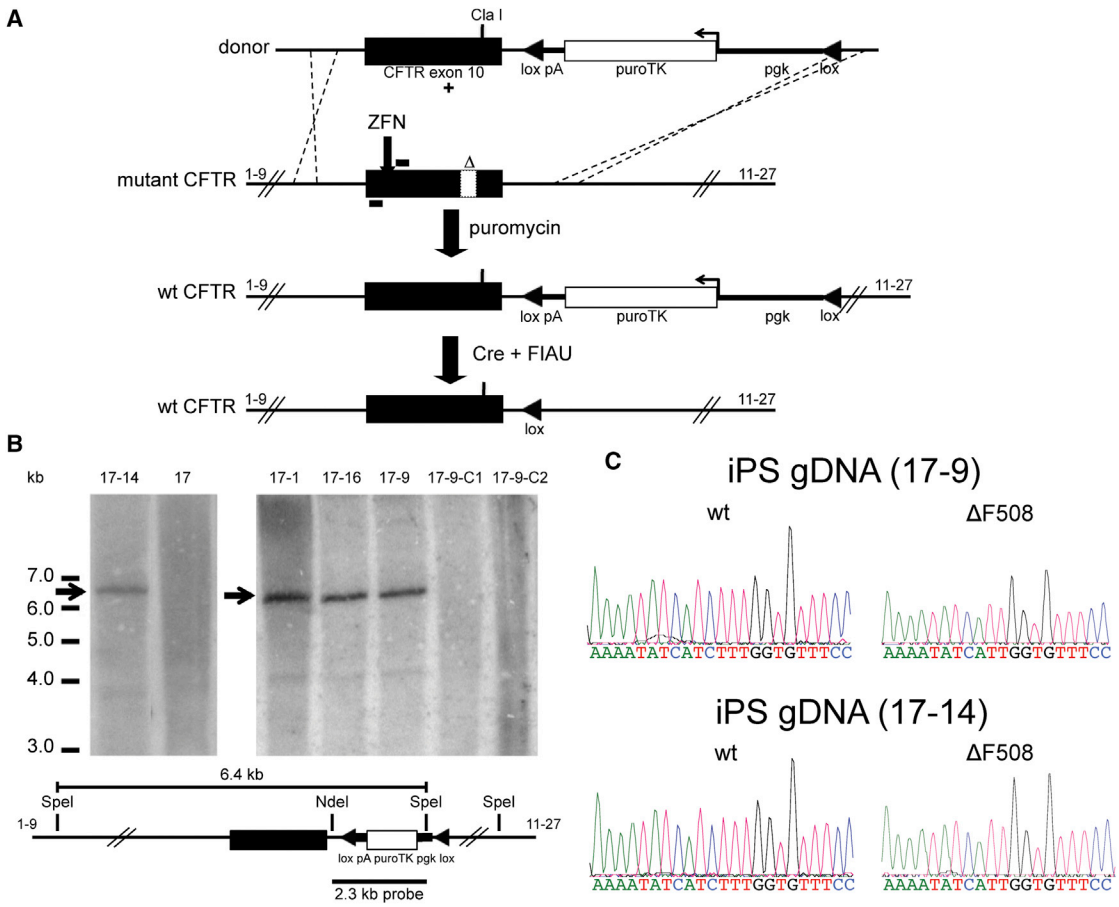

Figure 1. ZFN-Mediated Correction of ΔI507 or ΔF508 *CFTR* Mutations in CF iPSCs

(A) Outline of methodology involving co-delivery of *CFTR*-specific ZFNs together with *CFTR* donor, followed by Cre-recombinase-mediated excision.

(B) The schematic shows the expected genomic organization of a targeted *CFTR* allele including the WT exon 10 (shown in black) together with the pgk-puroTK selection cassette. A unique 6.4-kb hybridizing band is expected for a correctly modified clone and is apparent in the four corrected clones (17-14, 17-1, 17-16, and 17-9), but absent in the Cre-excised clones and the non-targeted clone 17 CF iPSCs.

(C) Sequence chromatograms of the modified WT and unmodified ΔF508 *CFTR* alleles from corrected CF iPSC clones.

See also Figure S1.

the correction of either mutant allele by HDR. The *CFTR* ZFNs were co-delivered with a plasmid encoding the *CFTR* donor to CF iPSCs. Puromycin-resistant colonies were initially screened via PCR and then sequenced to confirm that *CFTR* exon 10 was corrected via HDR. Southern blot analysis confirmed that four clones (17-1, 17-9, 17-14, and 17-16) exhibited the expected genomic organization in the corrected *CFTR* allele without any additional integration of pgk-puroTK sequences (Figure 1B). Sequencing of *CFTR* genomic DNA exon 10 sequences at targeted (wild-type [WT]) and unmodified (ΔF508) alleles for each of the four corrected clones demonstrated correction of one *CFTR* allele (ΔI507) per clone (Figure 1C). Transient delivery of a Cre-recombinase expression plasmid resulted in numerous puroTK-excised clones from each of the four successfully edited clones; successful excision was confirmed via PCR analysis with subsequent Cla I digestion (Figures 2A and 2B) and Southern blot analysis (Figure 1B).

We demonstrated that two of two correctly edited puroTK-excised iPSC clones (17-9-C1 and 17-14-C1) retained pluripotency, as evidenced by teratoma-forming ability (Figures S2A and S2C); furthermore, quantitative transcriptional profiling of 44 genes characteristic of human pluripotent stem cells revealed that both corrected iPSC lines (17-9-C1 and 17-14-C1) exhibited a gene expression pattern highly similar to both the original uncorrected CF iPSCs (clone 17) and hESCs (line WA09) (Figures S2B and S2D).

Karyotypic analysis of 17-9-C1 and 17-14-C1 confirmed retention of a normal karyotype (Figures S2A and S2C).

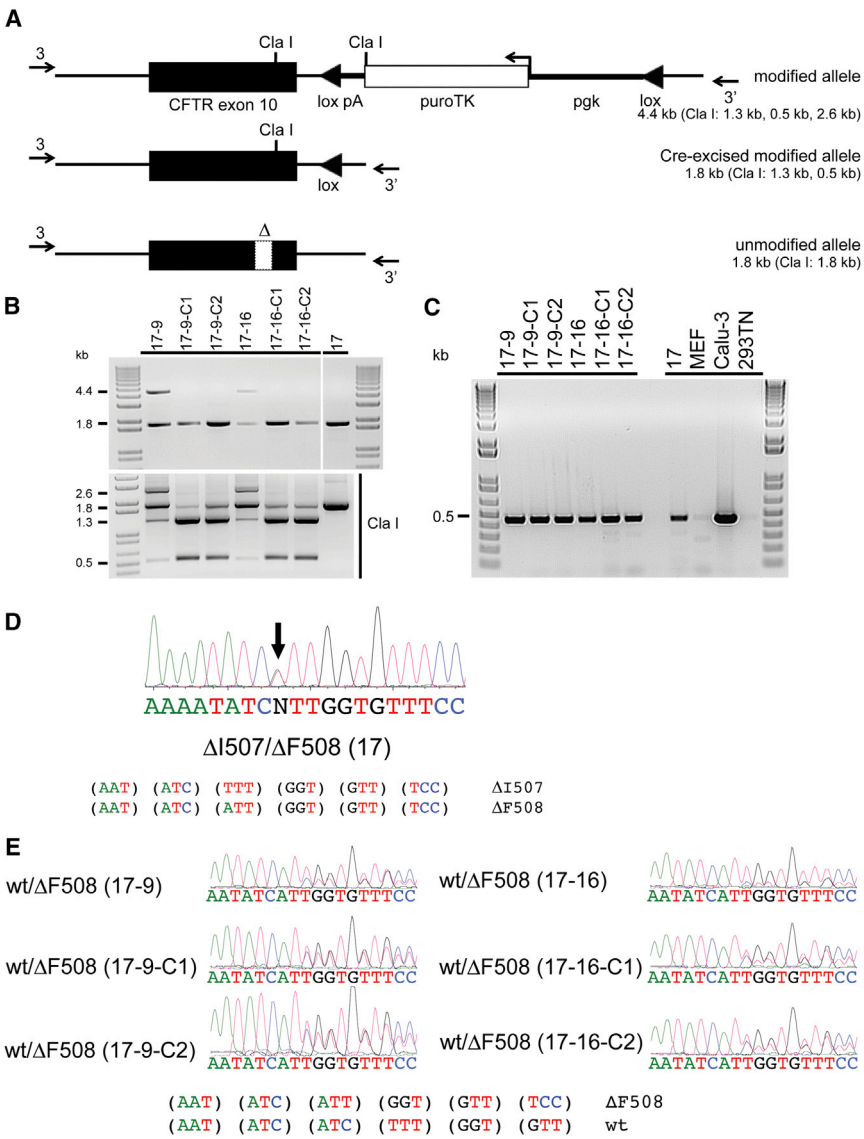

Figure 2. Cre-Mediated Excision of puroTK Cassette from Corrected CF WT/ΔF508 iPSCs

(A) Schematic of the modified allele, before and after Cre-mediated excision, and the unmodified allele. The location of PCR primers (arrows marked 3 and 3′), both located outside of donor sequences, used in verification by amplification are shown. Also indicated are the expected sizes of Cla I digestion products.

(B) (Top) The PCR amplicons for the original targeted clones (17-9 and -16), the Cre-excised clones (17-9-C1 and -C2; 17-16-C1 and -C2), and the original clone 17 CF iPSCs. The presence of only the 1.8-kb band for Cre-excised clones is consistent with successful excision. (Bottom) The results of Cla I digestion of the PCR amplicons. The size of bands for the Cre-excised clones is consistent with successful excision.

(C) RT-PCR analysis of *CFTR* expression for two targeted CF iPSCs (17-9 and -16) as well as their derived Cre-excised clones. Also shown is *CFTR* expression by the original clone 17 CF iPSCs and CFTR-expressing Calu-3 cells; mouse embryo fibroblasts (MEFs) and HEK293TN cells are negative controls.

(D) Sequencing of *CFTR* RT-PCR product from mutant ΔI507/ΔF508 *CFTR* iPSCs (clone 17) revealed a mixture of ΔI507 and ΔF508 *CFTR* sequences.

(E) Sequencing of *CFTR* RT-PCR product from corrected WT/ΔF508 *CFTR* iPSCs (clones 17-9 and 17-16), together with their Cre-excised derivatives, revealed a mixture of WT and ΔF508 *CFTR* sequences.

See also Figure S2 and Tables S1 and S2.

Comparative genomic hybridization (CGH), validated by complete genome sequencing, did not identify any copy-number variations in corrected clones 17-9-C1 and 17-14-C1 not already present in uncorrected clone 17 iPSCs. To assess any off-target effects of ZFN- or Cre-mediated excision methodologies at higher resolution, we also submitted genomic DNA from corrected clones 17-9-C1 and 17-14-C1, as well as from the CF fibroblasts and clone 17 CF iPSCs, for whole-exome and complete genome sequencing. Concordance between exome and whole-genome sequencing was high, validating the variations identified as well as our analysis methodology. We detected only one non-synonymous coding variant (NSCV) unique to clone 17 iPSCs and not present in the original mutant CF fibroblasts. We identified, in comparison with the parental clone 17 iPSCs, two novel NSCVs in 17-9-C1 and eight novel NSCVs in 17-14-C1; these changes consisted entirely of single-base-pair substitutions (Table S1). The DNA sequences flanking these variants were examined to determine whether the DNA variations in the ZFN-corrected iPSCs possibly resulted from non-homologous end joining (NHEJ) at sites of off-target cutting by the ZFNs. Importantly, the NSCV found in the corrected cell lines did not share significant homology to any permutation of ZFN target site (Table S1). Additionally, we did not find any variations that were shared by the two corrected cell lines. Taken together, the type of mutation, the lack of commonality between clones, and the lack of homology to ZFN target sites at/around these changes argue strongly against any ZFN off-target activity in these lines at the whole-exome level. In

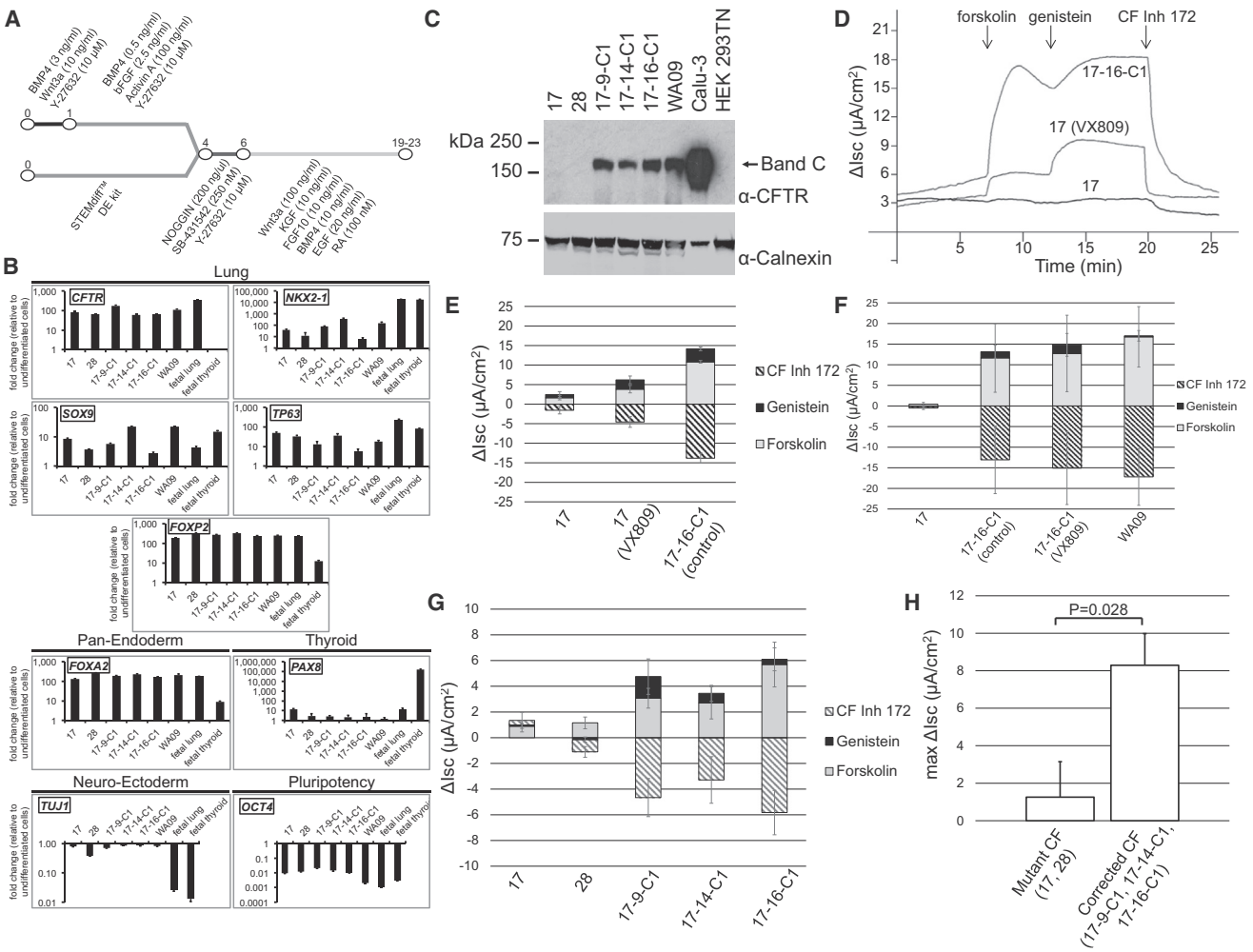

Figure 3. In Vitro Differentiation of Corrected CF WT/ΔF508 iPSCs

(A) Outline of the defined, step-wise differentiation protocol used to generate anterior foregut cell fates.

(B) Gene expression analysis of day 19 differentiated mutant (17 and 28), corrected CF iPSCs (17-9-C1, 17-14-C1, and 17-16-C1), and WA09 hESCs indicates upregulation of lung and pan-endodermal markers. Data (mean ± SD, three well replicates) from a representative experiment further characterized in Figures 3C and 3G. See also Figure S3E.

(C) Western blot analysis of protein lysates from day 19 differentiation cultures probed with a CFTR-specific antibody. Detection of Calnexin demonstrated equal sample loading for differentiated iPSC/hESC samples. See also Figures S3D and S3F.

(D) Representative short-circuit current (Isc) traces of epithelial monolayers differentiated from mutant (17) and corrected (17-16-C1) CF iPSC by Ussing chamber analysis. Cells were treated with either DMSO (0.03%) or VX809 (3 μM) for 24 hr. After establishment of Cl^- gradient and the addition of amiloride, monolayers were treated with forskolin and genistein followed by the administration of CFTR inhibitor 172. The change in Isc (μA/cm^2) for each perturbation is shown. Only corrected clone 17-16-C1 demonstrates the presence of CFTR channels in the cell membrane, as evidenced by activation of cAMP-dependent short-circuit currents. The addition of CF corrector VX809 to differentiated clone 17 sample partially restores CFTR-mediated chloride activity.

(E) Aggregated data of short-circuit current measurements from an experiment with differentiated mutant (17), with or without CF corrector VX809, and corrected (17-16-C1) CF iPSC (three transwell replicates per sample, mean ± SD.

(F) Aggregated data of short-circuit current measurements from an experiment with differentiated mutant (17) and corrected (17-16-C1), with or without CF corrector VX809, and WA09 hESC control (five transwell replicates per sample, mean ± SD).

(G) Aggregated data of short-circuit current measurements from an experiment including all independent mutant (17 and 28) and corrected (17-9-C1, 17-14-C1, and 17-16-C1) differentiated CF iPSC clones (three to six transwell replicates per sample, mean ± SD).

(H) Aggregated data of short-circuit current measurements; graphed is the maximum change in short-circuit current resulting from the addition of forskolin and genistein. Shown is the mean ± SE. The comparison shown is between two mutant CF clones (17 and 28; total of four independent differentiated experimental samples, three to six transwell replicates per sample) and three corrected CF clones (17-9-C1,

(legend continued on next page)

addition, we interrogated the complete genome sequences of the two corrected iPSC clones (17-9-C1 and 17-14-C1) at the top 100 ranked potential off-target ZFN-binding sequences (representing either homo-dimer or hetero-dimer binding) for any evidence of cleavage (see Supplemental Experimental Procedures). No evidence for NHEJ-induced mutation, either in exon or intron sequences, was identified (Table S2), again arguing against any ZFN off-target activity in these lines.

We similarly applied the ZFN-mediated gene correction methodology to two transgene-free CF iPSC lines homozygous for the ΔF508 mutation. These CF iPSC lines (RC202 and RC204) were originally derived by reprogramming of ΔF508/ΔF508 CF fibroblasts with a Cre-excisable polycistronic lentiviral vector (Somers et al., 2010). A total of six corrected iPSC clones (five from RC202 and one from RC204) were obtained, each of which was corrected at one of the two ΔF508 alleles (i.e., of resulting genotype WT/ΔF508) (data not shown).

Expression of the Corrected *CFTR* Gene in Gene-Edited iPSCs and iPSC-Derived Cells

We detected low-level *CFTR* expression in the original, uncorrected ΔI507/ΔF508 clone 17 iPSCs (Figure 2C) and WA09 hESCs (data not shown) by RT-PCR. Similarly RT-PCR analysis for clones 17-1, -9, -14, and -16 yielded a single band of similar size to that seen for clone 17 iPSCs (Figure 2C). Sequencing of the clone 17 iPSC line RT-PCR product demonstrated *CFTR* mRNA expression arising from both the ΔI507 and ΔF508 alleles (Figure 2D). Sequencing of the RT-PCR product from the four corrected clones, both prior to and following pgk-puroTK excision, confirmed the expected cDNA organization (exons 9–13) and demonstrated *CFTR* mRNA expression from both the corrected and mutant alleles (ΔF508) (Figure 2E).

We next examined expression of corrected *CFTR* mRNA and protein under in vitro differentiation conditions previously demonstrated to generate anterior foregut endoderm and primordial lung progenitors from mouse ESC (Longmire et al., 2012) and human ESC/iPSC (Green et al., 2011). In brief, after induction of definitive endoderm, inhibition of BMP/TGF-β signaling with NOGGIN/SB431542 was employed to enrich for anterior foregut endoderm. Subsequent exposure to growth factors implicated in lung development and maturation (WNT3a/KGF/FGF10/BMP4/EGF and retinoic acid) was then used to induce expression of NKX2-1, the earliest marker of commitment of endoderm to either a lung or thyroid epithelial cell fate (Figure 3A). Employing this protocol, mutant CF iPSCs (two independent clones, 17 and 28), corrected CF iPSCs (three clones, 17-9-C1, 17-14-C1, and 17-16-C1, independently obtained from correction of mutant clone 17), and WA09 hESCs efficiently generated definitive endoderm, as evidenced by co-expression of CXCR4 and C-KIT in >90% of cells (Figure S3A) and upregulated expression of endodermal transcriptional regulators, *FOXA2* and *SOX17* mRNA, by qPCR (data not shown). Further endodermal differentiation, for a total of 19 days in this protocol, subsequently upregulated expression of *NKX2-1*, *SOX9*, T*P63*, *FOXP2*, and *FOXA2*, suggesting commitment of at least a sub-population of cells within the endodermal culture to a lung epithelial cell fate (Longmire et al., 2012), and demonstrated increased expression of the *CFTR* target gene (Figure 3B), as expected from differentiated endodermally derived epithelia. Immunostaining confirmed that a subpopulation of cells co-expressed NKX2-1 and FOXA2 in this directed differentiation protocol (Figure S3B). Importantly, *CFTR* mRNA expressed on day 19 from differentiated mutant CF iPSCs (clone 17) reflected expression of both mutant ΔI507 and ΔF508 *CFTR* alleles, whereas differentiation of corrected iPSCs (17-16-C1) revealed co-expression of corrected WT and mutant ΔF508 *CFTR* mRNAs (Figure S3C). These results indicate appropriately regulated gene expression of the corrected WT allele, in comparison to the mutant ΔF508 allele, in the corrected iPSC-derived cells.

CF patients, either compound heterozygous ΔI507/ΔF508 or homozygous ΔF508/ΔF508, fail to express the mature, fully glycosylated CFTR protein. Correction of either allele, corresponding to carrier state WT/ΔF508 or WT/ΔI507, should result in restored expression of the mature CFTR glycoform. Western blotting of protein lysates from day 19 differentiation cultures, probed with a CFTR-specific antibody, identified the mature 170-kDa CFTR protein in differentiated WT CFTR, WA09 hESCs, and in differentiated corrected iPSCs (17-9-C1, 17-14-C1, and 17-16-C1), but not in differentiated ΔI507/ΔF508 iPSCs (clones 17 and 28; Figure 3C). The identity of the mature glycoform is confirmed by its presence in HEK293TN cells transfected with WT CFTR, but not ΔF508, expression constructs (Figure S3D), as well as in WT CFTR-expressing Calu-3 cells (Figures 3C and S3F). In addition, its identity as the mature CFTR protein is verified by its reduction in molecular weight upon treatment with peptide-N-glycosidase (PNGase F; Figure S3F). These results are consistent with maturation of CFTR protein from the corrected *CFTR* allele in cells derived from the ZFN-edited iPSC clones.

17-14-C1, and 17-16-C1; total of five independent differentiated experimental samples, three to six transwell replicates per sample). Results were clustered by clonal cell line and experiment number; statistical analysis was performed using a linear mixed-effect model by restricted maximum likelihood to account for correlated replicates within the same experiment.
See also Figure S3.

Restoration of CFTR Chloride Channel Function in Gene-Edited iPSC-Derived Epithelial Monolayers

To examine whether *CFTR* gene correction resulted in functional rescue of CFTR chloride channel activity, we initially selected the mutant (clone 17) and the corrected CF iPSC line (17-16-C1) for analysis. iPSCs were differentiated for 19 to 20 days as outlined above, replated onto permeable supports, and grown until confluent monolayers were established. Assessment of functional CFTR chloride channel activity in iPSC-derived epithelial monolayers was performed by Ussing chamber analysis. Sodium channel activity was first blocked by amiloride to establish a baseline. Stimulation with forskolin and genistein increased CFTR-dependent short-circuit current (Isc) for corrected 17-16-C1, but not for mutant clone 17 (Figures 3D and 3E). The addition of CFTR inhibitor 172 decreased Isc levels back to baseline and verified the CFTR specificity of the assay (Figures 3D and 3E). Mutant clone 17 monolayers, when treated with VX809 (a CFTR modulator capable of partially correcting the defective folding and aberrant maturation of ΔF508 mutant CFTR), exhibited partial restoration of CFTR-mediated chloride transport (Figures 3D and 3E).

To evaluate the extent of functional correction, in a subsequent experiment we demonstrated that 17-16-C1-derived cells functioned similarly to those of control WA09 hESCs, and they showed no further significant increase in CFTR chloride channel activity when treated with VX809 (Figure 3F). These results establish functional expression of WT CFTR in corrected clone 17-16-C1 samples. A second functional measurement, iodide efflux analysis, confirmed successful *CFTR* gene correction in clone 17-16-C1. Mutant and corrected CF-iPSC-derived epithelial monolayers were loaded with radioactive ^{125}Iodide and subsequently stimulated with forskolin and genistein to activate cAMP-sensitive CFTR anion transport. In contrast to mutant clone 17 cells, corrected 17-16-C1 cell samples released higher levels of ^{125}Iodide following stimulation (Figure S3G). The suppression of ^{125}I-efflux with CF inhibitor 172 demonstrated specificity for CFTR channels in corrected clone 17-16-C1 (Figure S3G). To ensure that functional correction seen for 17-16-C1 was not simply an anomalous feature of this specific clone, we assayed CFTR functional activity, in addition to the originally assayed 17 and 17-16-C1, in an additional mutant CF iPSC line (28) and two additional corrected CF iPSC lines, each of which were independently derived from mutant clone 17 (17-9-C1 and 17-14-C1). Whereas both mutant clones showed similar non-responsive behavior, all three corrected iPSC clones yielded epithelial monolayers with CFTR functional activity increased with respect to both mutant clones 17 and 28 (Figure 3G). When considered together, the three corrected CF clones (17-9-C1, 17-14-C1, and 17-16-C1) exhibited maximum short-circuit currents significantly greater than mutant CF clones (17 and 28) (Figure 3H, $p = 0.028$). These functional results exhibit a strong correlation with the expression of fully glycosylated CFTR protein in the corrected clones (Figure 3C).

Allele-Preferred Targeted Correction

In our targeted correction of the ΔI507/ΔF508 clone 17 iPSCs, we observed a strong preference for targeting the ΔI507 allele versus the ΔF508 allele; sequencing of the unmodified allele yielded either the ΔI507 mutant allele (present in 3 of 21 clones) or ΔF508 mutant allele (in 18 of 21 clones) (a ΔI507/ΔF508 targeting ratio of 6:1) (Figure 4A). Although we initially speculated that perhaps a greater level of chromatin accessibility or transcriptional activity for one allele versus the other may have been responsible, our sequencing of *CFTR* cDNA from the original ΔI507/ΔF508 iPSCs, as shown previously, revealed transcriptional activity from both mutant alleles (Figure 2D).

To investigate the cause of this allele-specific correction, we fully sequenced the 1.6-kb endogenous *CFTR* sequences homologous to the donor DNA used in the targeting vector, of each mutant allele (ΔI507 or ΔF508), to determine whether the donor was truly isogenic for both alleles. This analysis revealed a single-base-pair difference (A > G) in intron 9, 76 bp upstream of the ZFN cleavage site, present in the ΔF508 mutant allele, but absent in both the ΔI507 mutant allele and the donor (Figure 4B). In classical homologous recombination, isogenic regions of homology are preferred for donor construction; thus, we speculated that this single-base-pair difference occurring selectively in the ΔF508 allele of the CF iPSCs may have caused this selective behavior. To test this hypothesis, we introduced this A > G mutation into the donor sequences to see whether this donor would now selectively favor targeting of the ΔF508 allele. As shown in Figure 4A, the A > G single-base-pair substitution did in fact skew the allele-specific targeting from the ΔI507 allele to the ΔF508 allele (from a ΔI507:ΔF508 ratio of 6:1 for the A donor to a ΔI507:ΔF508 ratio of 1:3.7 for the G donor, a net 22-fold change resulting from a single-base-pair change). Molecular characterization and sequencing of these ΔF508-targeted, puromycin-resistant iPSC clones now identified three clones with specific correction of the ΔF508 allele. As such, these clones were now of WT/ΔI507 *CFTR* genotype.

DISCUSSION

We utilized ZFN-mediated gene editing (Urnov et al., 2005; Hockemeyer et al., 2009) to correct, in a site-specific manner, the CFTR mutation in iPSCs derived from CF patients. The generation of iPSCs from CF patients has been reported previously, with subsequent differentiation into

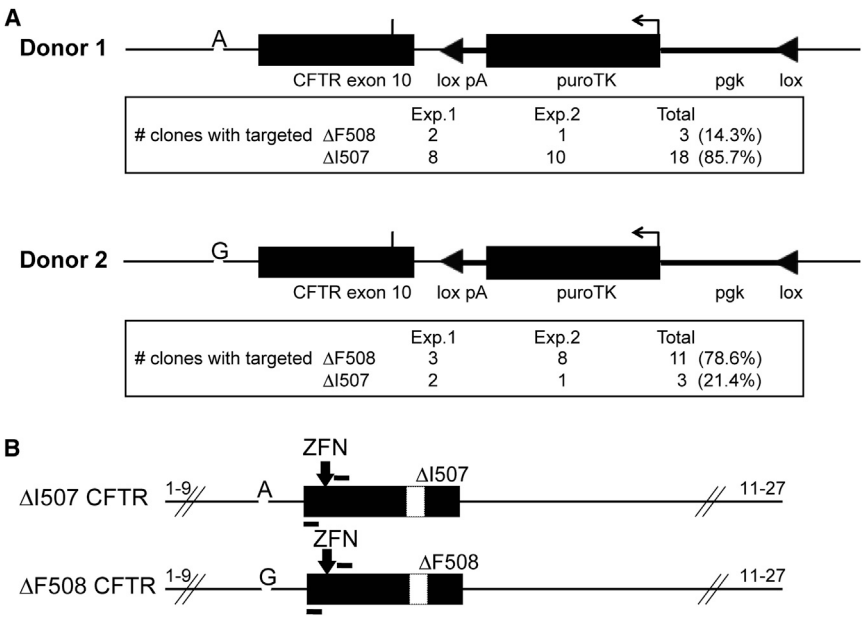

Figure 4. Allele-Preferred Targeting of CF iPSCs

(A) Schematic of donor 1 and donor 2 engineered with respective A or G base change at intron 9; results from a total of two independent targeted-integration experiments with either donor 1 or donor 2 are shown.

(B) Schematic of uncorrected CFTR alleles ΔI507 and ΔF508, highlighting A or G base within intron 9.

epithelial cells (Somers et al., 2010; Mou et al., 2012; Wong et al., 2012; Sargent et al., 2014). The design and assessment of CFTR-specific nucleases also has been reported previously (Maeder et al., 2008; Lee et al., 2012; Sargent et al., 2014), including repair of the mutant CFTR gene. Site-specifically editing the endogenous gene (Garate et al., 2013) offers the potential for physiologically regulated expression of the therapeutic gene, retaining the expression of alternately spliced isoforms and eliminating the potential interfering influence of inherited mutations that may be dominant negative. In vitro differentiation of the mutant CF iPSCs into lung epithelial cells and tissue, controlled for by the parallel differentiation of the otherwise isogenic corrected CF iPSCs, may provide a valuable tool for examining the functional consequence of mutant CFTR expression. Furthermore, corrected CF iPSCs present a potential source of patient-specific cells capable, in vitro, of differentiation into various lung stem/progenitor cells (Weiss et al., 2011), either for transplantation of autologous lung cells or for seeding de-vitalized lung scaffolds ex vivo to generate autologous lungs (Ott et al., 2010).

CF patients, either compound heterozygous ΔI507/ΔF508 or homozygous ΔF508/ΔF508, lacking expression of the normal CFTR protein, manifest various features of CF disease. Individuals bearing either the ΔI507 or ΔF508 *CFTR* mutations at only one allele (i.e., WT/ΔI507 or WT/ΔF508) are CF carriers and exhibit no defects in lung cell function (Kerem et al., 1990). Thus, correction of CF iPSCs derived from either ΔI507/ΔF508 or ΔF508/ΔF508 CF patients would be achieved by converting to WT either one or both mutant alleles. In this paper we first report targeted correction of ΔI507/ΔF508 iPSCs with the resulting iPSCs, corrected at a single allele, of genotype WT/ΔF508; as such, they correspond to the heterozygous ΔF508 carrier state. We demonstrated restored expression of both WT *CFTR* mRNA and mature CFTR protein in cells differentiated from a corrected, puroTK-excised clone 17-16-C1. This was not unexpected since the *CFTR* gene-editing approach simply restores the mutant ΔI507 or ΔF508 sequences in exon 10 to WT, with the only residual footprint being the loxP sequences in intron 10. Starting with the same ΔI507/ΔF508 iPSCs, but utilizing a donor with greater similarity to the ΔF508 *CFTR* allele (Figure 4), we were then able to obtain corrected iPSCs of genotype WT/ΔI507. Additionally, starting with two transgene-free ΔF508/ΔF508 CF iPSC lines, we were able to obtain corrected iPSCs of genotype WT/ΔF508.

The exquisite sensitivity of the allele-preferred targeting, that only a single-base-pair difference (of a total 1.6 kbp donor) could so dramatically specify the targeting for one allele versus the other was perhaps unanticipated. In subsequent correction of another mutant gene in disease-specific human iPSCs (the *PKLR* gene in pyruvate kinase deficiency), we again observed a precise allele-specific targeting caused by a single-base-pair difference in 2.0 kb of homology sequences (Z. Garate, A.M.C., B.R.D., and J.C. Segovia, unpublished data). It is possible that mismatches present in the integrated genome sequences, particularly those in proximity to nuclease cleavage sites, may dramatically influence the efficiency of targeting via decreasing the efficiency of strand invasion, recognition of the donor as a homologous template, and/or synthesis-dependent strand annealing. A preference for donor/target similarity also was seen in ZFN-mediated targeting of the histone variant

H3.3 gene (Goldberg et al., 2010), although to a lesser extent than seen here. A requirement for exquisite matching between donor sequences and target alleles also recently was observed in AAV-mediated targeting, in which even 1 bp mismatch in 1.8 kb of homology sequence decreased the targeting efficiency for the *APP* locus by 4.5-fold (Deyle et al., 2014). Although this last report did not include site-specific donor sequence DNA cleavage, it does support the need for close donor-to-target sequence similarity. Taken together, these results suggest that donor sequences utilized for nuclease-mediated HDR should be finely tuned to the targeted gene locus in the recipient cells. Specific rules for the application of allele-preferred targeting remain to be developed (for example, the quantitative effect of mismatch as a function of distance from nuclease cleavage site). Finally, we note that this allele-specific targeting offers the potential for preferential targeting of specific mutant alleles, for example dominant alleles, by sequence-specific nuclease-mediated gene correction.

Having successfully corrected various mutant *CFTR* alleles in the CF iPSCs, we sought to develop in vitro differentiation conditions allowing us to confirm appropriately regulated expression of the corrected *CFTR* gene. After a total of 19 days of differentiation, upregulated expression of *NKX2-1*, *SOX9*, *FOXP2*, *FOXA2*, and *CFTR* were all consistent with some cells committed to a lung cell fate arising in the culture (Figure 3B). Future work will focus on enriching the NKX2.1-expressing lung progenitors to generate a fully developed lung airway epithelium (Longmire et al., 2012).

We demonstrated that *CFTR* gene correction resulted in restoration of expression of the mature CFTR glycoprotein and CFTR chloride channel function in iPSC-derived epithelial cells. Mutant clone 17 monolayers, when treated with CFTR corrector VX809, exhibited partial restoration of CFTR-mediated chloride transport. This indicates that the failure of non-drug-treated clone 17 cultures to exhibit CFTR chloride transport was not due to incomplete differentiation. Furthermore, in Figure 3F we show that treatment of the corrected line 17-16-C1 with the corrector VX809 yields no further significant increase in chloride channel activity. Since the 17-16-C1 line is WT/ΔF508, this result suggests that the one corrected allele, corresponding to the carrier CF state, produces sufficient WT CFTR protein to respond normally.

Although in our experience there may be some differences, experiment to experiment, in the efficiency of differentiation, these have not affected the primary observation that, once differentiated, mutant CF iPSCs (17 and 28) yield neither mature CFTR protein nor CFTR-specific chloride channel activity; whereas corrected CF iPSCs (17-9-C1, 17-14-C1, and 17-16-C1) and the normal control WA09 hESCs yield mature CFTR protein and CFTR-specific chloride channel activity. Although prior studies of hESC/iPSC-derived epithelial cells documented CFTR functional activity via patch clamping of individual cells (Firth et al., 2014) or iodide efflux (Wong et al., 2012), our demonstrated ability to evaluate functional CFTR chloride channel activity via the Ussing chamber assay in hESC/iPSC-derived polarized epithelial monolayers should be a valuable tool in CFTR drug screening and analysis of various *CFTR* mutations.

EXPERIMENTAL PROCEDURES

Human subject and animal subject reviews were performed by University of Texas Health Science Center institutional review committees.

CF iPSC Generation and Characterization

CF fibroblasts were transduced with pMXs retroviruses expressing OCT4, SOX2, KLF4, NANOG, and C-MYC, and reprogrammed iPSC colonies, selected for in hESC media, were subsequently identified based on morphology and live cell staining for Tra-1-60 and Tra-1-81. Pluripotency was assayed by teratoma formation and confirmed by quantitative transcriptional profiling.

ZFN-Mediated Correction

Dissociated iPSCs were nucleofected with ZFNs together with a 1.6-kb donor construct (including WT exon 10, a loxP-flanked pgk-puroTK selection cassette in intron 10, and flanking homology sequences). Colonies exhibiting *CFTR* correction were identified by PCR and confirmed via Southern blot analysis.

In Vitro Differentiation and Functional Analysis

CF iPSCs, either mutant or corrected, and WA09 hESCs (as control) were differentiated based on Green et al. (2011) and Longmire et al. (2012), with minor modifications. CFTR protein was detected by immunoblotting. Ussing chamber analysis and ^{125}I-efflux experiments were performed to demonstrate restoration of CFTR-mediated chloride channel function.

Assessment of Genome Integrity

The genomic integrity of mutant (clone 17) and two corrected CF iPSC lines (17-9-C1 and 17-14-C1) was assessed by karyotyping, whole-exome and whole-genome sequencing, and CGH.

Data Access

Next-generation sequencing data can be downloaded from http://www.ncbi.nlm.nih.gov/Traces/sra/ using the study accession numbers SRA058070 for the Complete Genomics whole-genome dataset and Illumina exome sequencing data (in .bam file format).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.02.005.

AUTHOR CONTRIBUTIONS

A.M.C., P.K., J.H.B., and B.R.D designed the experiments and wrote the manuscript with valuable guidance provided by M.C.H., D.N.K., W.J.C., and E.J.S. J.W, H.C.S., D.E.P., D.Y.G., P.D.G., and M.C.H. designed and generated the ZFNs. A.M.C. and J.H.B. performed the targeted correction. A.M.C., P.K., and F.H performed the in vitro differentiation. A.M.C., W.J.C., and E.J.S. performed the functional analysis. X.S.L., P.K., M.L.G.G., and M.C.H. performed the assessment of genome integrity. S.C. performed the statistical analysis. W.L. and D.M. provided valuable technical assistance.

ACKNOWLEDGMENTS

We thank Naoki Nakayama (Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center) for helpful discussions; Veronica K. Quiceno for editorial assistance; Kathryn Plath, Connie Cepko, Brian Sauer, and Jeffrey Beekman for providing DNA constructs either directly or through Addgene; and Fyodor Urnov (Sangamo BioSciences, Inc.) for ZFN development through funding from the Cystic Fibrosis Foundation (CFF) Folding Consortium. This study was supported by NIH RC1HL099559 (B.R.D. and R. Wetsel, primary investigators [PIs]); CFF DAVIS12GO (B.R.D., PI); NIH P30 DK072482 (E.J.S., PI); and CFF R464 (E.J.S., PI). J.W., H.C.S., D.E.P., D.Y.G, P.D.G., and M.C.H. are employees of Sangamo BioSciences, Inc.

Received: August 7, 2013
Revised: February 9, 2015
Accepted: February 10, 2015
Published: March 12, 2015

REFERENCES

Deyle, D.R., Li, L.B., Ren, G., and Russell, D.W. (2014). The effects of polymorphisms on human gene targeting. Nucleic Acids Res. *42*, 3119–3124.

Firth, A.L., Dargitz, C.T., Qualls, S.J., Menon, T., Wright, R., Singer, O., Gage, F.H., Khanna, A., and Verma, I.M. (2014). Generation of multiciliated cells in functional airway epithelia from human induced pluripotent stem cells. Proc. Natl. Acad. Sci. USA *111*, E1723–E1730.

Garate, Z., Davis, B.R., Quintana-Bustamante, O., and Segovia, J.C. (2013). New frontier in regenerative medicine: site-specific gene correction in patient-specific induced pluripotent stem cells. Hum. Gene Ther. *24*, 571–583.

Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., et al. (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell *140*, 678–691.

Green, M.D., Chen, A., Nostro, M.C., d'Souza, S.L., Schaniel, C., Lemischka, I.R., Gouon-Evans, V., Keller, G., and Snoeck, H.W. (2011). Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. Nat. Biotechnol. *29*, 267–272.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat. Biotechnol. *27*, 851–857.

Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., and Tsui, L.C. (1989). Identification of the cystic fibrosis gene: genetic analysis. Science *245*, 1073–1080.

Kerem, B.S., Zielenski, J., Markiewicz, D., Bozon, D., Gazit, E., Yahav, J., Kennedy, D., Riordan, J.R., Collins, F.S., Rommens, J.M., et al. (1990). Identification of mutations in regions corresponding to the two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene. Proc. Natl. Acad. Sci. USA *87*, 8447–8451.

Lee, C.M., Flynn, R., Hollywood, J.A., Scallan, M.F., and Harrison, P.T. (2012). Correction of the ΔF508 mutation in the cystic fibrosis transmembrane conductance regulator gene by zinc-finger nuclease homology-directed repair. BioRes. Open Access *1*, 99–108.

Longmire, T.A., Ikonomou, L., Hawkins, F., Christodoulou, C., Cao, Y., Jean, J.C., Kwok, L.W., Mou, H., Rajagopal, J., Shen, S.S., et al. (2012). Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. Cell Stem Cell *10*, 398–411.

Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., et al. (2008). Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol. Cell *31*, 294–301.

Mou, H., Zhao, R., Sherwood, R., Ahfeldt, T., Lapey, A., Wain, J., Sicilian, L., Izvolsky, K., Musunuru, K., Cowan, C., and Rajagopal, J. (2012). Generation of multipotent lung and airway progenitors from mouse ESCs and patient-specific cystic fibrosis iPSCs. Cell Stem Cell *10*, 385–397.

Ott, H.C., Clippinger, B., Conrad, C., Schuetz, C., Pomerantseva, I., Ikonomou, L., Kotton, D., and Vacanti, J.P. (2010). Regeneration and orthotopic transplantation of a bioartificial lung. Nat. Med. *16*, 927–933.

Sargent, R.G., Suzuki, S., and Gruenert, D.C. (2014). Nuclease-mediated double-strand break (DSB) enhancement of small fragment homologous recombination (SFHR) gene modification in human-induced pluripotent stem cells (hiPSCs). Methods Mol. Biol. *1114*, 279–290.

Somers, A., Jean, J.C., Sommer, C.A., Omari, A., Ford, C.C., Mills, J.A., Ying, L., Sommer, A.G., Jean, J.M., Smith, B.W., et al. (2010). Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. Stem Cells *28*, 1728–1740.

Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature *435*, 646–651.

Weiss, D.J., Bertoncello, I., Borok, Z., Kim, C., Panoskaltsis-Mortari, A., Reynolds, S., Rojas, M., Stripp, B., Warburton, D., and Prockop, D.J. (2011). Stem cells and cell therapies in lung biology and lung diseases. Proc. Am. Thorac. Soc. *8*, 223–272.

Wong, A.P., Bear, C.E., Chin, S., Pasceri, P., Thompson, T.O., Huan, L.J., Ratjen, F., Ellis, J., and Rossant, J. (2012). Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein. Nat. Biotechnol. *30*, 876–882.

Stem Cell Reports

Supplemental Information

Targeted Correction and Restored Function of the *CFTR* Gene in Cystic Fibrosis Induced Pluripotent Stem Cells

Ana M. Crane, Philipp Kramer, Jacquelin H. Bui, Wook Joon Chung, Xuan Shirley Li, Manuel L. Gonzalez-Garay, Finn Hawkins, Wei Liao, Daniela Mora, Sangbum Choi, Jianbin Wang, Helena C. Sun, David E. Paschon, Dmitry Y. Guschin, Philip D. Gregory, Darrell N. Kotton, Michael C. Holmes, Eric J. Sorscher, and Brian R. Davis

Supplemental Information

Supplemental Data

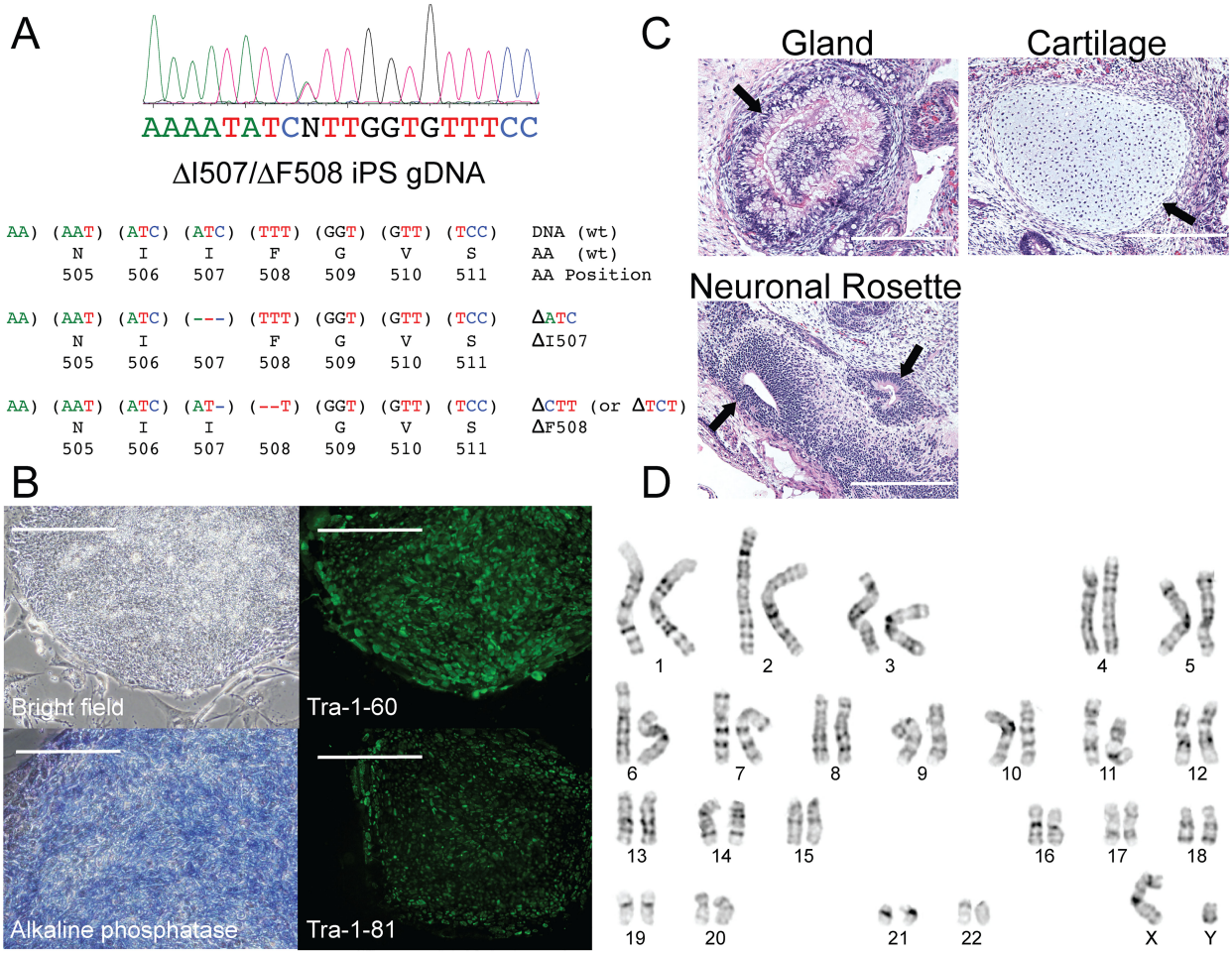

Figure S1. Characterization of mutant CF iPSCs. (related to Figure 1). **A.** DNA sequence of PCR amplicon spanning the *CFTR* exon 10 region from Clone 17 CF iPSCs containing ΔI507 and ΔF508 mutations. The mixture of bases (A and T; marked by the arrow) is consistent with amplification arising from two alleles: one ΔI507 and the other ΔF508. **B.** CF iPSCs express antigens characteristic of human ES cells. Expression of Tra-1-60, Tra-1-81, and non-specific alkaline phosphatase by CF iPSCs, Clone 17. Scale bars = 200 μm. **C.** Hematoxylin and Eosin staining of teratomas formed by CF iPSCs in immunodeficient mice; the three panels include examples of endoderm (gland), ectoderm (neuronal rosette), and mesoderm (cartilage). Scale bars = 200 μm. **D.** CF iPSCs exhibit normal male karyotype.

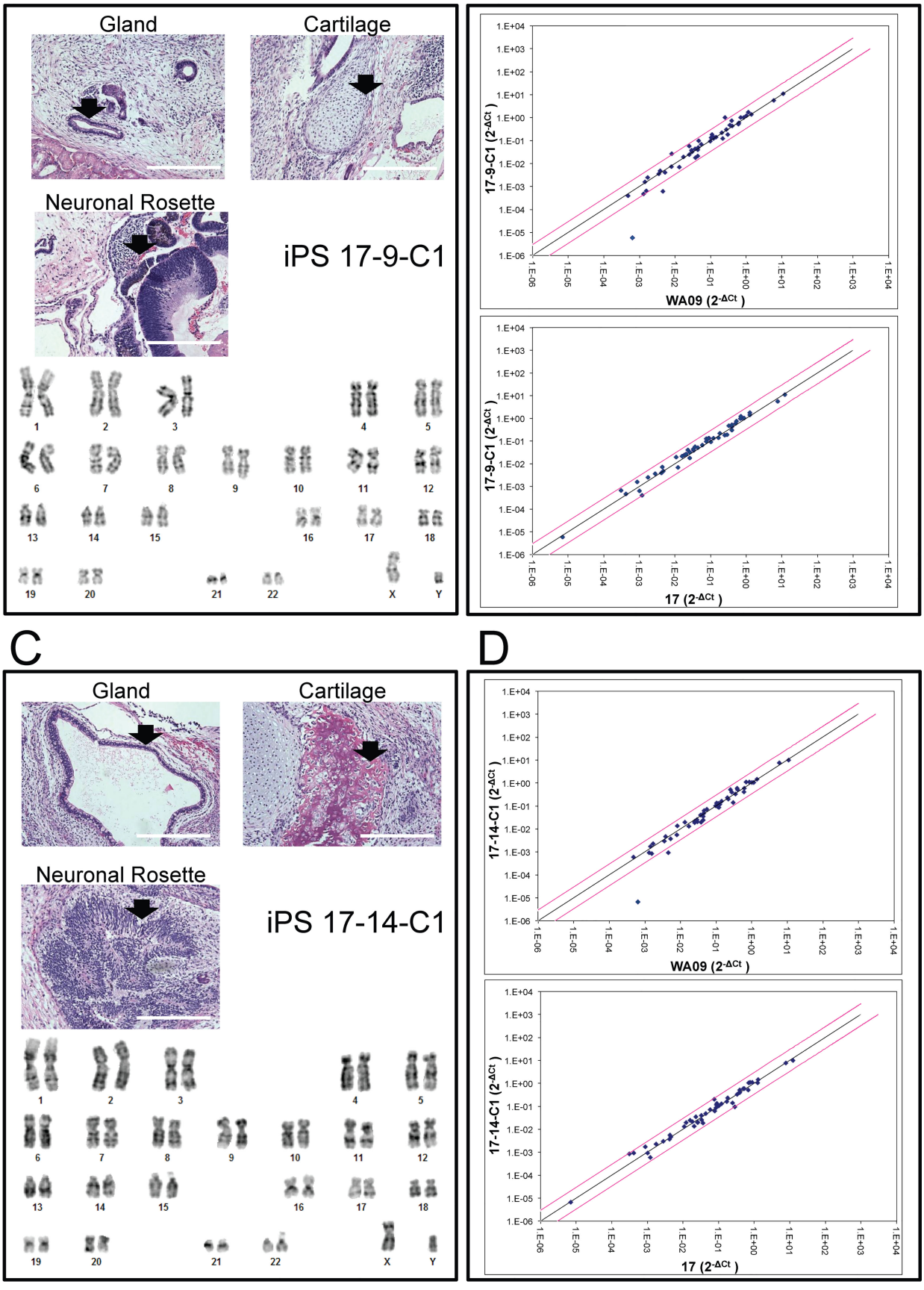
A
Gland
Cartilage
Neuronal Rosette
iPS 17-9-C1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
B
17-9-C1 (2-ΔCt)
WA09 (2-ΔCt)
17 (2-ΔCt)
C
Gland
Cartilage
Neuronal Rosette
iPS 17-14-C1
D
17-14-C1 (2-ΔCt)
WA09 (2-ΔCt)
17 (2-ΔCt)

Figure S2. Characterization of corrected CF iPSCs. (related to Figure 2). Hematoxylin and Eosin staining of teratomas formed by corrected (wt/ΔF508) CF iPSCs (**A.** 17-9-C1, **C.** 17-14-C1) in immunodeficient mice; the three panels include examples of endoderm (gland), ectoderm (neuronal rosette), and mesoderm (cartilage). Corrected CF iPSCs exhibit normal male karyotype. Scale bars = 200 μm. Transcriptional RT^2 Profiler PCR array of corrected CF iPSCs (**B.** 17-9-C1 and **D.** 17-14-C1)**.** The data were analyzed by the ΔCt method and the black line represents fold changes ((2^(-ΔCt)) of 1. Pink lines indicate a three-fold change in gene expression. Top scatter plot: Comparison of corrected CF iPSCs vs. WA09 hES cells. *FGF-5* is the only significant outlier in the corrected 17-14-C1 and 17-9-C1 CF iPSCs with approximately 100-fold down-regulation in comparison to WA09 ES cells. Lower scatter plot: Comparison of corrected CF iPSCs vs. initial uncorrected CF iPSCs (clone 17). Data shown reflect single assays for samples 17-9-C1 and WA09 and average of duplicate assays for samples 17-14-C1 and Clone 17.

Table S1. Non-Synonymous Coding Variants (NSCVs) Identified in ZFN-Corrected iPSCs

CFTR chromosome 7:117199532 CCAGACTTCACTTCTAatggtgATTATGGGAGAACTG

Gene_ID	Chr	Position	Variant type	Ref	Mut	AA Change	Variant DNA Sequence	17-9-C1	17-14-C1
PRG4	1	186276439	snp	G	T	p.A396S	cacccaccaccaccaagtct**T**cacccaccactaccaaggag		
CYTIP	2	158272220	snp	C	T	p.R350H	tttcttcctcttccacagca**T**gatgaaggccagggataaat		
MTHFD1L	6	151243396	snp	G	A	p.R347K	acagcagcacaggcggtgga**A**acttcactgcttgaaacttc		
SPON1	11	14063077	snp	G	A	p.E119K	gttttccgtagatcatagac**A**aagaagaaactcagtttatg		
CLPB	11	72141383	snp	C	T	p.R143H	cttcttgcatattgttggca**T**gggcagcttccaacagggct		
APOBEC1	12	7802196	snp	G	T	p.P220T	tgtagctaaaaggatgtgtg**T**cggaatcgtttggtaatggc		
OR6C68	12	55886587	snp	G	C	p.M147I	aacaaagtgtgcaaaacaat**C**gttatttgttgttggatggc		
ITGBL1	13	102366883	snp	G	T	p.D459Y	acagagactgcgacaaacat**T**atggtctcatttgtacaggt		
SHPK	17	3527384	snp	G	A	p.T151M	tggttgcacagccgaagccc**A**tggccacactgagatgagac		
PHACTR3	20	58322854	snp	C	A	p.L108I	gcaggcaaggccgagaggag**A**tcatcaagaaggggctgctg		

NSCVs found in corrected clones 17-9-C1 or 17-14-C1 consist of single base pair substitutions (snp) with the reference (Ref) changed to Mut. Mut changes are in bold and capitalized shown with surrounding sequence under Variant DNA Sequence. Shown above are the CFTR ZFN recognition sequences with ZFN-L/ZFN-R sequences in capitals and the spacer sequences in lowercase.

Table S2. Bioinformatic analysis to predict the top 100 potential ZFN off-target sites in the human genome.

Rank	SELEX Score	Chrom.	Target Site	Mismatch	Gene
homo-dimer: ZFN pair 9940/ 9940 (recognizing: ATTATGGGAGAACTG/ ATTATGGGAGAACTG)					
1	1.37E-12	chr4	CCAGTTCaCCCGTgTACGTGCGTAgATGGatGAACacA	7	None
2	8.36E-13	chr2	CaGaTaCTCCCAgATACAGAGTATcTGGGAttACaGG	8	ARHGAP15
3	4.36E-14	chrX	CCGGTgCTCCgAgAgcCCCCGgATcacGGcGAACTGG	10	POU3F4
4	3.77E-14	chr15	CaGtTTCaCCCATgTAAGAAGTAgATGGGAGAACTtG	6	None
5	3.40E-14	chr8	ACAGTgCaCCCAgcTcCTAGTCTcTgTGGGtGAtCaGG	10	None
6	2.47E-14	chr11	CCtGTgCTCCCAgggcCCTGGgATggGGGtGAtCTGC	11	ST14
7	1.87E-14	chr5	CtGGTTCTtCCAgATAATTGGTAaATGGGAGtAgTtG	7	C5orf58
8	1.44E-14	chr11	CCtGTagTCCCAgATACTCAGgAcAgGGGtGAggTGG	10	LGR4
9	1.11E-14	chr12	CCAGcaCcCCCAccTcCTGCAGTATcTGGGAccACaGG	10	ATF1
10	9.72E-15	chr13	TCAGTaCaCCtGTATACGGCATtTcTGGGtGAAtCaG	8	None
11	9.62E-15	chrX	GCAGTgCTCCtGTAgcCCCAGCTATgCGGGAGgAtTGT	7	TBL1X
12	8.81E-15	chr3	CCtGTaCTCCCtgcTACTCAGgAggCGGGAGAAtTGT	9	None
13	8.79E-15	chrX	CCAGagCaCCtGTgTcCAGCAGacTcaGGGAGcACTGT	11	DHRSX
14	8.79E-15	chrX	CCAGagCaCCtGTgTcCAGCAGacTcaGGGAGcACTGT	11	DHRSX
15	8.79E-15	chrX	CCAGagCaCCtGTgTcCAGCAGacTcaGGGAGcACTGT	11	DHRSX
16	8.79E-15	chrY	CCAGagCaCCtGTgTcCAGCAGacTcaGGGAGcACTGT	11	DHRSX
17	8.79E-15	chrY	CCAGagCaCCtGTgTcCAGCAGacTcaGGGAGcACTGT	11	DHRSX
18	8.79E-15	chrY	CCAGagCaCCtGTgTcCAGCAGacTcaGGGAGcACTGT	11	DHRSX
19	8.21E-15	chr17	CCGaTaCTgCCATgTATCAAGgcTcgGGGAGAACTGT	8	TRIM37
20	6.05E-15	chr11	CCAGaaCTCCCAgggACTTAGgATgCaGGAGctCCtG	11	PRDM10
hetero-dimer: ZFN pair 9940/ 12897 (recognizing: ATTATGGGAGAACTG/ TAGAAG**t**GAAGTCTGG [Linker Collapse*])					
1	2.52E-13	chr1	CCAGTTCTCCCGgcTAAGTAACAAaaAGGAAtTCTGTC	5	MIER1
2	2.01E-13	chr4	ACAGaaCTCCCGgAaACCAGAGAcaaAGGAAGTtTGTA	8	PDLIM3
3	1.96E-13	chr3	CCtGgTCTCCCATATAACATACCAaaAGGAAGTtGGTG	5	None
4	1.15E-13	chr9	TAtAaACgTaCTGaTGCATACGTATATGGGAtcACTGG	7	None
5	9.26E-14	chr1	CCAGTTCTtCtGcAgACAAAAGCAtCAGGAAGTaCcCA	7	None
6	8.99E-14	chr3	CTCAcACaTaCTGtcTTTGGAAGgATATGGaAGAACTGT	7	PTPRG
7	6.21E-14	chr5	GCCAGcCTTCCTGGaTCTTGAGGTtTATGGGAatACTtG	6	None
8	4.90E-14	chr16	CCAGTTCTCtCATtTACTCCTCTgtCAGGAAcTCTGCC	5	LOC283867
9	4.18E-14	chr8	TCAGTTCTCCCcgATACCCCAGTTAaCAGGcAGcCAaCC	6	None
10	3.66E-14	chr2	CCAGTgCTtCCAgATcCTCATGTAgaaAGGctGTCCtGA	10	ARHGEF4
11	3.50E-14	chr10	AGaGGgCTTgCTGtcCACAGGAGgATtTGGGAGAACTGG	7	None
12	3.23E-14	chr11	CCAGTTtcCaCATtTACTTTGGCAACCAGGAAGTgTGGC	5	DAGLA
13	3.03E-14	chr16	CCAcTgCaCCCAgccACATTATCttCAGGAAGTCTGTG	8	TMCO7
14	2.90E-14	chrX	ACAGagCaCCCAgAgACTAGCAGCACCtGGAAGctCtTG	9	None
15	2.85E-14	chr17	CCAGTTCTCCCAgATACGCTAGGctCAaGcccTCAGCC	7	SPNS2
16	2.66E-14	chr17	ACCTGgCTTCCTGGaTCCAGGGGTgTcTGGGtGAAaTGA	6	AARSD1
17	2.62E-14	chr11	CCtAGgCTTCCTaGcACTGGGAGgcTATGGGAGAACCaG	7	TRIM68
18	2.59E-14	chr18	ATCCcACTTCCctGgGCCTAACTATgTGGGtGAACTGT	6	None
19	2.48E-14	chr12	AAgTGtCTTCaTGacCATTTTGTATATGGGgGAACTGG	6	PLXNC1
20	2.48E-14	chr15	CCtGTgCTCCCAcATcTCTCAGAAgCaAGGAAGTgGGTG	7	RASGRF1
homo-dimer: ZFN pair 12897/ 12897 (recognizing: TAGAAG**t**GAAGTCTGG/ TAGAAG**t**GAAGTCTGG [Linker Collapse*])					
1	1.20E-13	chr1	AGgTGACTTCCTGGgCCCATTAAAAaCAGGAAGTCAGGC	3	None
2	2.32E-14	chr14	CCCAGAggTCCTGGTTGCATGAATGACCAGGAAGgCCaGG	4	PPP1R13B
3	2.14E-14	chr18	CACAGACTTCCTGtaGGTCACTGAACaAGGAAGcCGtTT	5	None
4	1.96E-14	chr11	TGaAtACTTCCTGGTGTAAGTTAGAACCAGaAAtTCTGGA	4	None
5	1.84E-14	chr1	AACAacCTTCCTGtcTGAATCAAGAAgCAGGAAGTtTGGG	6	AXDND1
6	1.84E-14	chrX	GTCCaACTTCCTGtTTCCATCATGgCCAGGAAcTCAGTT	4	None
7	1.64E-14	chr1	TCaTGACaTCCTtGTAAGAACCTTAAtCAGGAAGTCCtAT	5	None
8	1.53E-14	chr1	TCaAGgCaTCCTGGTACTGTCAAGTACCAGGAAcTtGGGG	5	KCNH1
9	1.45E-14	chr5	TCaAtACTTCCTtGcTCTGCGTGTCtCCAGGAAGTtCtGT	7	None
10	1.41E-14	chr17	ACaGaACaTCCTGGgTCTAATTTGAACtgGGAAGTCTGGC	6	None
11	1.26E-14	chr3	AGgGGgCTTCCTGGTCCCAGAGAGCACagGGAtGcCTGGG	6	None
12	1.20E-14	chr7	CCaCaACTTCCTGcTTTCCTAAATGgaCAGGAAGTCCtTG	6	None
13	1.10E-14	chr5	GGCAGAtTTCCTGGTTCTCAAGAGAAaCtGGAAaTCTtGA	5	ADAMTS2
14	1.03E-14	chr12	CACAcACTTCCTGGTGACCAGAGGtCCcaGAAGTCTtCT	5	PRICKLE1
15	1.02E-14	chr17	AAaAGACcTtCTGtTTCTGCCAGAAtCAGGAAGTCTtTC	6	None
16	8.89E-15	chr17	AAaAGACTTgCTttTTACGCAGATAAaCAGGAAGTtTGAA	6	YPEL2
17	8.57E-15	chr14	AGaGaACTTtCTGGTCCTTGAGCTAcCCAGGAAGcCAGCA	5	TMEM63C
18	7.86E-15	chr21	CTCCacCTTCtTGtTCTTCTTGTGCACCAGGAAGTtCtTG	6	TIAM1
19	7.61E-15	chr20	CCaGGcCcTCCTGaTTTGCTCCCCACCAGGAAGTtTtTCT	6	None
20	7.25E-15	chr7	GTtTGACTTCCTtaTGAAGTAGAGACCAGGtAGTtTGGA	5	None
hetero-dimer: ZFN pair 9940/ 12897 (recognizing: ATTATGGGAGAACTG/ TAGAAG**T**GAAGTCTGG [Linker Gapped*])					
1	4.85E-12	chr5	TCAGTTCTCCCAgAgcCTTGGAAaCAGTGAAGTtAcCT	6	None
2	4.94E-13	chr4	CCAGTTCTCCCAggTcTCTGGCAtCAGTGAtGTtTtAG	7	None
3	4.56E-13	chr18	ATCCcACTTCCCTGGgCCTAACTATgTGGGtGAACTGT	4	None
4	2.88E-13	chr11	CCtGTaCTCCCAgATACTTGGGAAgCAGAGAtGggAGAA	7	SBF2
5	1.16E-13	chr4	ACAGaaCTCCCGgAaACCAGAGACaAaGGAAGTtTGTA	7	PDLIM3
6	9.14E-14	chr4	GAgTGACTTCTgTGGTTCAAAATATATGaGAGcACTGG	4	None
7	8.01E-14	chr22	CCAGTTCTCCtGgATcCATGGCtCCAGGGAtGgaGGAT	7	None
8	7.25E-14	chr8	CCCCacCTTCCCTGGTGAAGGGGgcTcTGGGAGgACaGA	7	MMP16
9	6.87E-14	chr7	GCGGTgCTCCCATATtGGGATAgCCAGAGAAGTgTGTA	4	CHN2
10	6.85E-14	chr9	CCAGTTCTCCCATtTcCTCCACCACagGAGAAtaCTGCT	6	TMOD1
11	6.36E-14	chr13	CCAGgaCTCCCATAacGATGTGGcCCAGGGAAGTCTtCC	6	None
12	5.72E-14	chr11	CCAGTTtTCCCAgATACCAGTGAaaAGTGcAGctAtGT	8	None
13	5.20E-14	chr8	ACAGTTCTtCCAcAgACCAGGAgCaAGGGAtGgtTtCA	9	RIMS2
14	5.13E-14	chr1	CCAGTTCaCCCAgAcAGCGAGTAgaCAGAGgAGTCTGCT	6	DAB1
15	4.00E-14	chr3	ACAGTTCTCCCATATACCACATtCCAGGGcAtcCTaTG	5	None
16	3.60E-14	chr9	CCAGTTCTCCCATATACTTGTGAgCaAtAaAAaTgCaTA	7	None
17	3.50E-14	chr6	CCAGTTCTCCCAgcacCATATAAAaCAGGGAAtTtTtTC	8	BCKDHB
18	3.48E-14	chr6	CTCCGAgTTCACTGtTGTCTATTcTcTGGGtGcACCGT	6	None
19	3.35E-14	chr5	TCAGTcCTCCCGgATcCCTGCTctCAGCtAAGgCTGGT	7	N4BP3
20	3.12E-14	chr1	CCAGTgCTCCCcTATATCTCCTCcCaAGGGAtcTCAGCC	6	SYT2
homo-dimer: ZFN pair 12897/ 12897 (recognizing: TAGAAG**T**GAAGTCTGG/ TAGAAG**T**GAAGTCTGG [Linker Gapped*])					
1	4.24E-13	chr21	TCCAGgCTTCCCTGtTGTCAGTGAACCAGGGAAGcCTGGT	3	None
2	1.15E-13	chr7	ACCCGACTTCTCTtGTGCTGGTCCACaAGAGAAGTtCaTT	4	DGKI
3	5.71E-14	chr9	TGaAGgCTTCCCTaGcTTCCTATGACCAGAGAAGTttTGTA	5	ENG
4	3.38E-14	chr4	ACCTcACTTCTCTGGTGGGGAGAACaAGTGAAcTgTGAA	4	WHSC1
5	3.34E-14	chr2	GCtGGACTTCACTGaTGCTCCAGGcCCAGAGAgGTCAGTG	4	None
6	1.66E-14	chr15	AGgGGACTTgGCTaGTTCTAGGAACCAGGGAAGTCTtTT	4	USP8
7	1.63E-14	chr9	TCCAaACTTCGCTGGTCTTTGGAACCAGTGAAaaCCcTT	4	None
8	1.56E-14	chr18	TCaAGACTTCACTGtgGTCTGTggtCAGAGAAGgCTGCA	6	KCNG2
9	1.27E-14	chr5	TCgGGcCaTCACTGaTCTTCACCCACCAGAGAAGTtTtTC	6	None
10	1.16E-14	chr4	TGCTGcCTTCACTGGgCACATGAAaCAGGGgAGTCAGCA	4	None
11	1.05E-14	chr1	CCaAaACTTCTCTGGaCCTCAATGtCCAGAGAAGcCAtGT	6	LGR6
12	1.05E-14	chr9	CCaAaACTTCCCTGGaCCTCAATGtCCAGAGAAGcCAtGT	6	C9orf44
13	7.52E-15	chr3	GAaGcACTTtGCTGaTCATAATTAAtCAGGGAAGTtTGCT	6	None
14	6.35E-15	chr15	AGaGGACTTCACTGaaTATGCTCAtCAGTGAtGTtGGCT	6	
15	6.34E-15	chr4	CAaCGACTTCACTGGaGGGACCActCAGAGAAGTgTtGA	6	None
16	5.53E-15	chr17	GGCTGACTTCCCTGaTGAAGAGGAtCaAGTGAttTCTGTG	5	None
17	5.42E-15	chr3	CACAGACTTCACTccaCCTTTGAAaaAGAGAAGTCTGCG	5	DNAH12
18	5.21E-15	chr4	GGgAGACTTCTtTGGaCCAGGTGgCCAGGGAAGTCAtCT	5	ODZ3
19	5.20E-15	chr1	TGCAGcCTTCCCTGaTTCCCAGCCAgCAGGGAtcTCTGAC	5	IL28RA
20	4.75E-15	chr9	TCaGagCcTCTCTGGTTCTTTCAAtCgGGGAtGTCTGCA	7	C9orf53

*ZFN 12897 was designed to skip a base (t) in the middle of the binding site (TAGAAGtGAAGTCTGG). Therefore, for the top 20 potential off-target searches we generated two sets for homo-/ heterodimers including ZFN 12897 with either the intended base pair position skipped (Linker Collapse) or where the intended base was not skipped (Linker Gapped).

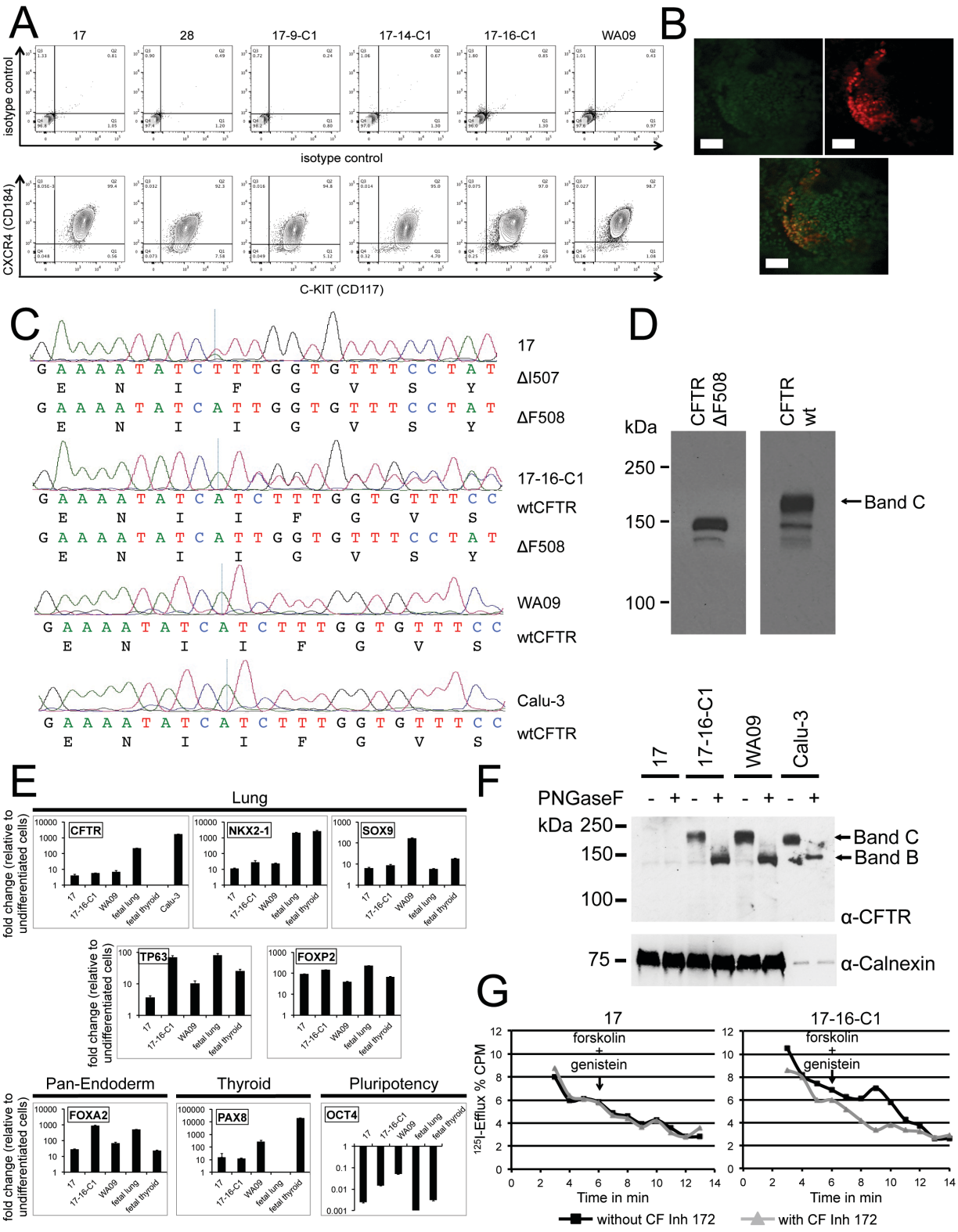
A
17
28
17-9-C1
17-14-C1
17-16-C1
WA09
isotype control
isotype control
CXCR4 (CD184)
C-KIT (CD117)
B
C
17
ΔI507
ΔF508
17-16-C1
wtCFTR
ΔF508
WA09
wtCFTR
Calu-3
wtCFTR
D
CFTR ΔF508
CFTR wt
kDa
250
150
100
Band C
E
Lung
fold change (relative to undifferentiated cells)
CFTR
NKX2-1
SOX9
TP63
FOXP2
Pan-Endoderm
Thyroid
Pluripotency
FOXA2
PAX8
OCT4
17
17-16-C1
WA09
fetal lung
fetal thyroid
Calu-3
F
17
17-16-C1
WA09
Calu-3
PNGaseF
kDa 250
150
100
75
Band C
Band B
α-CFTR
α-Calnexin
G
17
17-16-C1
forskolin
+
genistein
125I-Efflux % CPM
Time in min
without CF Inh 172
with CF Inh 172

Figure S3. *In vitro* differentiation of corrected CF wt/ΔF508 iPSCs. (related to Figure 3). **A.** Representative flow cytometry analysis showing CXCR4 vs. CKIT co expression after 4 days of differentiation to definitive endoderm of all mutant and corrected CF iPSC clones studied. Isotype control (top panels) were used to set gate for experimental clones (bottom panels) **B.** Corrected CF wt/ΔF508 iPSCs (RC 202) were differentiated towards anterior foregut endoderm/NKX2-1 according to the directed differentiation protocol shown in Fig. 3A. Immunostaining for FOXA2 (green) and NKX2-1 (red) was performed on day 14 of the differentiation and demonstrates that the majority of cells are endodermal (FOXA2+) and a subpopulation is FOXA2+/NKX2-1+. Scale bar = 100 μm. **C.** Sequencing of *CFTR* mRNA from 19 day differentiated mutant Clone 17 CF (ΔI507 and ΔF508) and corrected 17-16-C1 (corrected wt and mutant ΔF508) iPSCs confirm the expected expression pattern. Calu-3 and differentiated WA09 cells express wt *CFTR* mRNA. **D.** HEK 293 cells transiently transfected with ΔF508 or wt CFTR expression constructs produce immature or mature glycoforms, respectively. **E.** Gene expression analysis of day 19 differentiated mutant (17), corrected CF iPSCs (17-16-C1), and WA09 hESCs. Data (mean +/- standard deviation, 2-3 well replicates) from a representative experiment (this experiment corresponds to the Western blot in Fig. S3F). **F.** Western blot analysis of 50 μg protein lysates from day 19 differentiation cultures treated with (+) or without (-) PNGaseF and probed with a CFTR-specific antibody. Whereas the mature, fully glycosylated CFTR protein (the 170 kDa protein band C) is present in Calu-3 cells and differentiated WA09 cells, Clone 17 iPSCs are deficient in expression of this glycoform. Differentiation of the corrected Clone 17-16-C1 iPSCs demonstrates restored expression of the mature CFTR glycoform. PNGaseF treated samples showed a shift of the mature 170 kDa band C to the core glycosylated, immature CFTR band B (~ 130 kDa). Detection of Calnexin demonstrated equal sample loading for differentiated iPSC/hESC samples, with less sample (5μg) loaded for the positive control Calu-3 cells. See also Fig. 3C.

G. cAMP-dependent ^{125}I-efflux from confluent monolayers of differentiated mutant Clone 17 and corrected 17-16-C1 CF iPSCs results in the release of radioactive ^{125}I exclusively in corrected CF cells after addition of forskolin and genestein at 6 minutes. Treatment with CF Inhibitor 172 (grey graph) from minute one completely blocks CFTR channels and iodide release from corrected 17-16-C1 CF cells. Results were expressed as ^{125}I-efflux % CPM released per minute into to the media; the average of in duplicate measurements is shown with and without CF inhibitor 172.

Supplemental Experimental Procedures

Cystic fibrosis primary fibroblasts:

CF primary fibroblast line GM04320 was obtained (Coriell Repository, Camden, NJ) from a patient (17 year old male) reported homozygous for the ΔF508 mutation. Clinical symptoms for this patient were reported as advanced pulmonary disease and pancreatic insufficiency; in addition, defective cAMP stimulated chloride channel activity was demonstrated in fibroblasts from this patient (Lin and Gruenstein 1987). Our sequencing of the *CFTR* alleles in genomic DNA isolated from the GM04320 fibroblasts demonstrated that the patient was, in fact, a compound heterozygote with one allele being ΔF508 and the other ΔI507. ΔF508/ΔI507 compound heterozygosity has previously been reported in CF patients (Kerem et al., 1990).

CF iPSC generation and characterization:

pMXs retroviral vectors encoding human reprogramming factors (OCT4 [17964], SOX2 [17965], KLF4 [17967], C-MYC [17966], NANOG [18115]) were graciously provided by K. Plath through Addgene (Lowry et al., 2008). VSV-G enveloped viral stocks were prepared by transfection of Plat-GP cells (Cell Biolabs) with vector DNA and VSV-G expression plasmids (pCMV-VSV-G [8454], kindly provided by B. Weinberg through Addgene) and concentrated 100 fold by ultracentrifugation. Parallel production of pMXs-GFP vector stocks was performed; titration of the pMXs-GFP virus was performed by infection of primary human fibroblasts and subsequent

FACS analysis for GFP-expressing cells. CF fibroblasts, plated at 10^5 cells per well of a 6-well plate on day 0, were transduced on days 1 and 2 by spin-fection (200 g for 30 minutes) at a multiplicity of infection of 21.5, in the presence of 10 μg/ml protamine sulfate. On day 4, fibroblasts were transferred onto irradiated mouse embryo fibroblasts (MEFs; CF-1 mouse strain, Charles River), and one day later media was switched to human embryonic stem (ES) cell media (per National Stem Cell Bank protocol SOP-CC-001C; http://nationalstemcellbank.org) containing 40 ng/ml basic Fibroblast growth factor (bFGF) and re-fed daily. Starting on day 12, cells were re-fed daily with human ES cell media pre-conditioned on irradiated MEFs. Beginning at 16 days post transduction, iPS-like colonies were first identified based on morphological criteria. Live-cell staining with either Alexa 488-conjugated anti-Tra-1-60 monoclonal antibody (Stemgent, 09-0068), or anti-Tra-1-81 monoclonal antibody (Millipore, MAB4381) followed by Alexa 488 goat anti-mouse IgM (Life Technologies, A-21042), was then used to identify reprogrammed colonies for subsequent expansion and characterization. Of 32 colonies originally picked (all of which stained positive for Tra-1-60 and/or Tra-1-81), 9 colonies were subsequently expanded and cryopreserved and two iPS clones (17 and 28) were selected for more extensive characterization.

Teratoma assay: 2-3 million uncorrected (clone 17) or corrected (17-9-C1, 17-14-C1) CF iPSCs in 30% matrigel were injected into the kidney capsule or testis of six week old Fox Chase SCID beige mice (3 mice per cell line; Charles River) and monitored weekly for the appearance of tumor growth. Six to eight weeks post injection, tumors were removed, paraffin embedded, prepared for histological examination by hematoxylin and eosin, and analyzed (Applied Stem Cell Inc.).

ZFN-mediated targeting:

Potential ZFN target sequences in the vicinity of *CFTR* exon 10 were first examined against a database of ZF DNA binding specificities. ZFN pair 12897 (recognizing TAGAAGtGAAGTCTGG) /9940 (recognizing ATTATGGGAGAACTG) was constructed by fusing

the desired DNA binding motifs to the cleavage domain of the Fok I endonuclease. ZFNs were delivered to cells either in the form of a DNA expression plasmid or *in vitro* generated RNA (MessageMAX T7 ARCA-Capped Message Transcription Kit, A-PlusPolyA Polymerase Tailing from Cell Script Inc. and MegaClear kit from Ambion). A 1.6 kb donor plasmid construct to facilitate HDR containing wt exon 10 sequences (approximately 860 bp and 290 bp of homology sequences upstream and downstream of exon 10, respectively) was originally constructed by PCR amplification of genomic DNA sequences from BAC clone RP11-1152A23. Two silent single base pair substitutions were introduced into the right ZFN binding site with the goal of interfering the ability of the introduced ZFNs to cleave the donor either prior or subsequent to homology-directed repair; an additional silent single base pair substitution was introduced into the wt exon 10 donor sequences in order to create a novel Cla I restriction enzyme site to facilitate identification of target integrated clones. The 1.6 kb amplicon was cloned into pSC-B as per StrataClone Blunt PCR Cloning Kit (Stratagene). Three additional single base pair changes were introduced into intron 10 donor sequences 125 bp downstream of exon 10 to create a unique Avr II restriction enzyme site to introduce a selection marker conferring phosphoglycerate kinase promoter-driven puromycin- resistance and thymidine kinase sensitivity (pgk-puro-TK-bpA). All changes in the wt donor were introduced via Quikchange Lightning Site-Directed Mutagenesis (Agilent). PCR amplification of pgk-puroTK-bpA sequences from plasmid pPthC-Oct3/4 (Masui et al., 2007)(kindly provided by N. Nakayama) with primers including loxP recognition sequences and Avr II sites generated material was cloned into the introduced Avr II site of *CFTR* donor by Phusion Hot Start II high fidelity DNA polymerase (New England Biolabs) in antisense orientation to CFTR gene. Donor plasmid was confirmed by sequencing.

ZFNs, either in the form of DNA expression plasmids (1 or 2 μg) or *in vitro* transcribed RNA (1.5 or 3 μg were delivered together with donor DNA (4 or 8 μg) to CF iPSCs (2 million single cells in suspension obtained via Accutase treatment; clone 17) via nucleofection (Lonza

Amaxa hStemCell kit1 program A23) and cells were plated in the presence of 10 μM Rock-inhibitor (Alexis Biochemicals, Y27632) onto irradiated puromycin-resistant MEFs (Stem Cell Technology). Puromycin selection (0.5 μg/ml) was initiated 4 days post transfection, and puromycin-resistant colonies were picked starting 5-9 days later and expanded, in the presence of puromycin, to establish clonal cell lines.

The donors utilized included either "A" or "G" in intron 9, 76 bp upstream of the ZFN cleavage site. We did not observe a significant difference in the ratio of CFTR-targeted vs. non-targeted puro^R colonies as a function of donor. For example, in a total of two experiments with the "A" donor, we analyzed a total of 63 puro^R colonies, of which 21 exhibited targeting within the *CFTR* locus as evidenced by PCR amplification at the upstream end of the donor. For the "G" donor, we analyzed a total of 57 puro^R colonies, of which 15 gave evidence for targeting within the *CFTR* locus (one of the 15 colonies did not subsequently pass one of our further criteria for targeting and was thus not included in the total in Fig. 4).

Molecular analysis of targeted iPS clones:

Genomic DNA was isolated from puromycin-resistant clones beginning at passage 2 by QIAprep Spin Miniprep Kit (Qiagen) or ArchivePure DNA Cell Tissue Kit (5 Prime). PCR amplification utilizing various primers (e.g. Fig. 2) was performed according to manufacturer protocols. Sequencing was performed on an ABI 3730XL sequencer. Sequences of oligonucleotide primers utilized for PCR amplification and sequencing are available upon request.

Southern blotting:

In order to generate a radio-labeled DNA for probing Southern blotted genomic DNAs, the donor plasmid was digested with Nde I + Spe I, separated on 0.8% agarose gel, and then the 2.3 kb fragment was cut out and gel-purified (Qiagen). The 2.3 kb fragment was labeled with [alpha-^{32}P]dCTP using Prime-It II Random Primer Labeling kit (Agilent Technologies) following manufacturer's instruction. 25 μg of genomic DNAs (gDNAs) were digested with Spe I overnight

and purified by phenol/chloroform extraction. The gDNAs were then resolved on 1% agarose gell and transferred to a Nytran Super Charge membrane (Schleicher and Schuell) and hybridized with ^{32}P-labeled probe. The membrane was exposed and image scanned using a phosphorimager system (Molecular Dynamics).

Cre-mediated excision of selectable marker:

Cre-expression plasmids (pBS513 EF1alpha-cre [B. Sauer, 11918](Le et al., 1999) , pCAG-Cre [C. Cepko, 13775])(Matsuda and Cepko 2007) were delivered to Accutase-treated single cells via Amaxa nucleofection (same conditions as ZFN mediated targeting) and plated onto irradiated puromycin-resistant MEFs. Individual colonies were picked and expanded, and then plated in replicate to identify those clones that had become sensitive to puromycin. Alternatively, some clones were first identified based on their resistance to FIAU (1 μM, Moravek Biochemicals), expanded, and then plated in replicate to identify puromycin sensitive clones.

In vitro differentiation:

CF iPSCs, either mutant or corrected, and WA09 ESCs were differentiated based on (Green et al., 2011) and (Longmire et al., 2012) with minor modifications. Single cell adapted and mTeSR1 (StemCell Technologies) cultured clones were lifted from the plate by using 1mg/ml Dispase (StemCell Technologies). To form embryoid bodies (EBs), cells were transferred into serum-free differentiation medium (SFDM, previously described in (Longmire et al., 2012)) supplemented with 10 μM Y-27632 (Reagent Direct), 3 ng/ml BMP4 and 10 ng/ml Wnt3a onto low attachment six well plates (Corning). All differentiation factors were purchased from R&D Systems if not otherwise indicated. Next day the medium was changed to SFDM containing 10 μM Y-27632, 0.5 ng/ml BMP4, 2.5 ng/ml b-FGF and 100 ng/ml Activin A to induce definitive endoderm over 72 hours. Depending on EB number the medium was refreshed on the second day of Activin A treatment. Alternatively, we derived definitive endoderm from human iPSCs/ESCs cultured as a monolayer, using the StemDiff kit (Stem Cell Technology) according to manufacturer's protocol. On day four or five of differentiation cells were analyzed by flow

cytometry for expression of endodermal markers using CXCR4- (BioLegend, 306505) and C-KIT-antibodies (Life Technologies, CD11705). Subsequently 50,000- 100,000 single cells/cm^2 were plated either into each well of a gelatin (Millipore) or fibronectin (3ug/cm^2,Sigma) coated 24 well plate or a 12.5 cm^2 flask (Falcon) and cultured for 2 days in SFDM supplemented with 10 μM Y-27632, 200 ng/ml NOGGIN and 250 nM SB-431542 (Sigma-Aldrich) to generate anterior foregut endoderm. On day 7 of differentiation the medium was switched to ventralization/ NKX2-1 induction medium containing 100 ng/ml WNT3a, 10 ng/ml KGF, 10 ng/ml FGF10, 10 ng/ml BMP4, 20 ng/ml EGF and 100 nM Retinoic Acid (RA) and refreshed every day until day 19.

Analysis of mRNA:

RNA isolation from iPS and iPS-derived cells with the RNeasy kit (Qiagen), cDNA synthesis was performed with Improm-II Reverse Transcriptase oligo dT kit (Promega), and RT-PCR was performed with Gotaq Hot Start polymerase (Promega) or Universal TaqMan PCR (Applied Biosystems). For detection of *CFTR* expression, two sets of PCR primers were utilized: one set with primers in exons 9 and 13 (yielding a 0.77 kb product; data not shown), and an alternative set with one primer spanning exons 8 and 9 and the other primer in exon 11 (yielding a 0.5 kb product; Fig. 2C).

Quantitative transcriptional profiling: The RT^2 profiler human ES cell PCR array (Qiagen PAHS-081Z), designed to assay the expression of 44 genes involved in maintenance of pluripotency and self-renewal of human ES cells, was used to profile corrected (17-9-C1, 17-14-C1), uncorrected (Clone 17) CF iPSCs, and WA09 hES cells.

In profiling the differentiated cells, gene expression levels were determined in 2-3 well replicates, normalized against housekeeping gene 18S, and compared to undifferentiated samples. All TaqMan gene expression assays were obtained from Life Technologies.

CFTR constructs and transfection:

CFTR wt and ΔF508 over-expression plasmids were previously described in (Hoelen et al., 2010) and a kind gift from Jeffrey Beekman. HEK293 cells grown in six well plates were

transfected with 2 µg of pcDNA3-CFTRwt and pcDNA-F508del plasmids using X-treme Gene HP DNA Transfection Reagent (Roche). 48 hours post transfection cells were lysed in 500ul lysis buffer (1% Triton X-100, 50 mM Tris HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, supplemented with protease inhibitor mixture [Roche]). CFTR protein analysis was performed after removal of cell debris.

Western blot analysis:

Day 19 differentiated cells and Calu-3 controls were lysed using 1% Triton X-100 containing lysis buffer in the presence of protease inhibitor. Cell debris was removed by centrifugation and protein amount quantified by BCA protein assay (Thermo Fisher). 100 µg of total protein from differentiated samples and 10 µg from Calu-3 cells were treated with 10X glycoprotein denaturation buffer, 10X G7 reaction buffer, 10% NP-40 and finally supplemented with and without 3,000 units of PNGaseF (New England Biolabs). Protein samples were incubated for 1 h at 37 deg C until deglycosylation reaction was stopped by adding 4X Laemmli sample buffer (Bio-Rad). 50 µg of PNGaseF-treated and untreated protein samples (or 5 µg for Calu-3 samples) were separated on a 7% NuPAGE Tris-Acetate gel (Life Technologies) and transferred onto a nitrocellulose membrane. Membranes were blocked for 1 hour in 5% nonfat dry milk (Bio-Rad) in PBS and probed with a monoclonal 596 anti-CFTR antibody (1:1000, Cystic Fibrosis Foundation Therapeutics, A4 596). A secondary HRP conjugated anti-mouse IgG (1:5000, Cell Signaling Technology, 7076S) and enhanced chemiluminescence reagent (Amersham) was used to visualize the bands. As loading control an anti-Calnexin polyclonal antibody (1:300, abcam, ab22595) was used together with a secondary anti-rabbit IgG IRDye 800CW (LI-COR, 925-32211). An Odyssey Quantitative Infrared Imaging System measured the fluorescence intensities of the protein bands.

Immunostaining:

Cultured cells were fixed with fresh 4% paraformaldehyde for 30 minutes, rinsed with PBS, permeabilized with 0.2% Triton X-100 for 15 minutes, blocked in 4% goat serum for 30 minutes

and then incubated in primary antibody overnight at 4 degrees. After 16 hours the cells were washed twice with PBS and incubated in secondary antibody (1:100) for 2 hours at room temperature. The antibodies used were FOXA2 (1:100, Santa Cruz Biotechnology, sc-101060) and NKX2-1 (1:100, abcam, ab76013). Parallel wells stained with secondary antibodies alone were included as negative controls. Goat anti-mouse Alexa 488 (Life Technologies, A-11001) and goat anti-rabbit Alexa 546 (Life Technologies, A-11010) were used as secondaries.

Functional CFTR chloride channel assays:

Short circuit current: Inserts were mounted in Ussing chambers and short circuit currents (Isc) were tested under voltage clamp conditions as previously described (Bates et al., 2007; McClure et al., 2014). Briefly, iPS cells differentiated in flasks for 19-20 days under the conditions described above were seeded on permeable supports (3×10^5 cells/cm2 filter (Costar 3470),) coated with fibronectin (3ug/cm^2,Sigma) and grown until confluent monolayers established. At time of seeding, differentiated cells expressed, by Western Blot, either corrected (e.g. Clones 17-9-C1, 17-14-C1, 17-16-C1) or mutant (e.g. Clones 17, 28) CFTR proteins, respectively. Prior to short circuit measurements, cells were treated for 24 hours with 0.03% DMSO or 3 uM VX-809 (Selleckchem). Filters in modified Ussing chambers (Jim's Instruments,Iowa City, IA) were bathed on both sides with identical Ringers buffer containing 120 mM NaCl, 25 mM, $NaHCO_3$, 2.4 mM KH_2PO_4, 1.24 mM K_2HPO_4, 1.2 mM $CaCl_2$, 1.2 mM $MgCl_2$, 10 mM D-glucose (pH 7.4). Bath solutions were vigorously stirred and gassed with 95% O_2: 5% CO_2 and maintained at 37°C. Isc expressed as μA/cm^2, was monitored using an epithelial voltage clamp (University of Iowa Bioengineering). A 3-mV pulse of 1 second duration was imposed every 100 seconds to track resistance by Ohm's law. To measure stimulated Isc, the mucosal bathing solution was changed to a low Cl^- buffer containing 1.2 mM NaCl, 115 mM Na gluconate, and other components as above plus 100 μM amiloride to block residual Na^+ channel activity. Agonists (20 μM forskolin and 50 μM genistein) were included in the bath solutions (minimum 5 min observation at each concentration). CFTRInh-172 (10 μM) was added

to the mucosal bathing solution at the conclusion of each experiment to block CFTR-dependent Isc. Statistical Analysis: Study subjects were clustered by clonal cell line and experiment number. To conduct statistical analysis, while accounting for correlated replicates, we applied a linear mixed-effect model (LMM) by restricted maximum likelihood (REML) (Laird and Ware 1982; Pinheiro and Bates 2000) and the corresponding ANOVA (Analysis of Variance) that incorporate random-effects to explain within-subject variations. For independent observations, we use a t-test or Mann-Whitney U test for normally-distributed and non-normally distributed variables, respectively. All statistical analyses were performed using the computing environment R (R Development Core Team, 2014; www.r-project.org).

cAMP-dependent ^{125}I-efflux: Corrected (17-16-C1) and mutant (Clone 17) CF iPSCs were differentiated for 19 to 23 days to carry a sub-population of lung epithelial cell types. CF-iPSC-derived epithelial monolayer was loaded with 0.5 μCi/cm^2 Na^{125}I in HEPES-phosphate buffered Ringer solution (HPBR) and incubated at 37°C in 5 % CO_2 for 1-2 hours (Venglarik et al., 1990). Briefly, cells were washed 4 times with 1 ml aliquots of isotope-free HPBR solution to establish an iodide efflux baseline. From minute one, isotope-free HPBR solution with or without CF inhibitor 172 (10 μM) was added by first removing all the solution from duplicate wells and replacing it with fresh solution at one minute intervals indicated times. At minute six, we added 1 ml cAMP buffer (HPBR with forskolin (10 μM) and genistein (50 μM) in the presence or absence of CF inhibitor 172 (10 μM)). Samples were collected until approximate 14 minutes. Buffer changes at each time point were collected and analyzed for radioactive ^{125}I. Remaining counts were obtained by cell lysis with 1% SDS, 0.1 N NaOH solution. ^{125}I-efflux are expressed in % counts per minute (CPM) as determined by the ratio of counts at a specific time point versus total of measured radioactivity (i.e., released iodide in supernatants versus remaining iodide

from lysed cells). We also assayed Calu-3 cells to establish robustness of the assay (data not shown, (Cantin et al., 2006)).

Genome sequencing and analysis:

Human iPSCs were first depleted of MEFs by passage on matrigel in MEF-conditioned media. Genomic DNA was extracted with the QIAprep Spin Miniprep Kit. DNA sequences were obtained as described previously using DNA nanoball amplification and combinatorial probe-anchor ligation sequencing (Drmanac et al., 2010; Lee et al., 2010). Possible de novo non-synonymous coding variants (NSCV) were analyzed and filtered based on two independent criteria (somaticScore and refScore) and then combined to generate the list in Table S1. Since each cell line was compared to the reference genome (GRCh37/hg19) we used the *calldiff* function with somatic option in CGA Tools to generate pairwise comparisons of our four cell lines. All NSCV with SomaticScore greater than -10 from the 8 pairwise comparisons were combined with NSCV which had a refScore greater than 60 (indicating the likelihood of being homozygous reference) in the mutant CF fibroblast cell line. The compendium of putative NSCV was examined across all four cell lines by loading evidence file based BAM files for each cell line into the Integrative Genomics Viewer (Robinson et al., 2011). We considered false NSCV to consist of variations in areas with poor coverage (<9 reads), seen in the CF mutant fibroblast(Gore et al., 2011), variations present at <10% frequency, and those duplicating nearby repetitive sequences. The 10% frequency is based on observations of loci with variant reads present at 10% of total reads that were not called variant. Since the BAM files were generated from evidence files after *de novo* alignment, coverage was assessed indirectly based on calls in the masterVarbeta files; regions with “no call” or “complex” in any of the cell lines. In other instances where there was no sequence data in the BAM files and a call of “hom ref”, we assumed homozygous reference sequence when coverage scores were positive.

Exome Capture and Sequencing:

Two μg of genomic DNA was submitted to Axeq Technologies for human exome capture

sequencing using TrueSeq 62 Mb target enrichment (http://www.illumina.com/documents/products/datasheets/datasheet_truseq_exome_enrichment_kit.pdf). Axeq Technologies performed sample validation, library preparation, exon enrichment, clustering and sequencing using illumina HiSeq 2000 Sequencer. Approximately 63,000,000 reads of an average size of 107 bp per sample was returned to us as two fastq files (one file per orientation).

Read Mapping:

Each pair of fastq files were aligned to human genome (hg19) using Novoalign (http://www.novocraft.com). All parameters were kept at the default settings, as recommended by Novocraft. SAMtools (http://samtools.sourceforge.net/) (Li et al., 2009) was used to sort the SAM files, create BAM files and generate their index files. Picard (http://picard.sourceforge.net/) was used to remove all the PCR duplicates from the BAM files. The Genome Analysis Toolkit (GATK) (McKenna et al., 2010)was used for local realignments, base quality recalibration, and variant calling. Parameters were set as described in GATK's Best Practices v3. GATK generated standard variant call format files (VCF http://www.1000genomes.org/wiki/Analysis/Variant%20Call%20Format/vcf-variant-call-format-version-41). The VCF files were annotated using snpEff (snpeff.sourceforge.net) (Cingolani et al., 2012) and ANNOVAR(Wang et al., 2010). From this point on we focused our analysis only on the putative coding non-synonymous variants.

NSCV described were independently found in both the whole genome (WGS) and the exome sequencing, which incorporates data from two different sequencing platforms and analysis methods (Lam et al., 2012). Both WGS and exome sequencing identified 4 NSCV in the CF iPS, but only 1 NSCV was present in both data sets. For the 17-9-C1 cells, 1 additional NSCV was identified by exome sequencing while for the 17-14-C1 cells, 1 additional NSCV was found by WGS. Whole genome sequencing was also employed to identify the integration sites of the reprogramming vectors in Clone 17 CF iPSCs; as expected for retroviral vectors, the majority of

integrants mapped to introns but none of the 25 vectors integrated directly into protein coding sequences (data not shown).

SELEX was used to identify the binding site preference for the left and right zinc finger proteins comprising the CFTR ZFNs. Base preferences for the CFTR ZFNS as determined by SELEX were used to guide a genome wide bioinformatics prediction of the top 100 potential off-target sites in the human genome, taking into account either homo-dimer or hetero-dimer binding (Perez et al., 2008). Whole genome sequences for two corrected CF iPSC lines (17-9-C1 and 17-14-C1) were then interrogated at these potential off-target ZFN binding sites for any evidence of mutation.

Supplemental References

Bates E, Miller S, Alexander M, Mazur M, Fortenberry JA, Bebok Z, Sorscher EJ, Rowe SM. 2007. Bioelectric effects of quinine on polarized airway epithelial cells. Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society 6: 351-359.

Cantin AM, Hanrahan JW, Bilodeau G, Ellis L, Dupuis A, Liao J, Zielenski J, Durie P. 2006. Cystic fibrosis transmembrane conductance regulator function is suppressed in cigarette smokers. American journal of respiratory and critical care medicine 173: 1139-1144.

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w (1118) ; iso-2; iso-3. Fly (Austin) 6: 80-92.

Drmanac R, Sparks AB, Callow MJ, Halpern AL, Burns NL, Kermani BG, Carnevali P, Nazarenko I, Nilsen GB, Yeung G et al. 2010. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science 327: 78-81.

Gore A, Li Z, Fung HL, Young JE, Agarwal S, Antosiewicz-Bourget J, Canto I, Giorgetti A, Israel MA, Kiskinis E et al. 2011. Somatic coding mutations in human induced pluripotent stem cells. Nature 471: 63-67.

Green MD, Chen A, Nostro MC, d'Souza SL, Schaniel C, Lemischka IR, Gouon-Evans V, Keller G, Snoeck HW. 2011. Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. Nature biotechnology 29: 267-272.

Hoelen H, Kleizen B, Schmidt A, Richardson J, Charitou P, Thomas PJ, Braakman I. 2010. The primary folding defect and rescue of DeltaF508 CFTR emerge during translation of the mutant domain. PloS one 5: e15458.

Kerem BS, Zielenski J, Markiewicz D, Bozon D, Gazit E, Yahav J, Kennedy D, Riordan JR, Collins FS, Rommens JM et al. 1990. Identification of mutations in regions corresponding to the

two putative nucleotide (ATP)-binding folds of the cystic fibrosis gene. Proceedings of the National Academy of Sciences of the United States of America 87: 8447-8451.

Laird NM, Ware JH. 1982. Random-effects models for longitudinal data. Biometrics 38: 963-974.

Lam HY, Clark MJ, Chen R, Natsoulis G, O'Huallachain M, Dewey FE, Habegger L, Ashley EA, Gerstein MB, Butte AJ et al. 2012. Performance comparison of whole-genome sequencing platforms. Nature biotechnology 30: 78-82.

Le Y, Miller JL, Sauer B. 1999. GFPcre fusion vectors with enhanced expression. Anal Biochem 270: 334-336.

Lee W, Jiang Z, Liu J, Haverty PM, Guan Y, Stinson J, Yue P, Zhang Y, Pant KP, Bhatt D et al. 2010. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature 465: 473-477.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.

Lin PY, Gruenstein E. 1987. Identification of a defective cAMP-stimulated Cl- channel in cystic fibrosis fibroblasts. J Biol Chem 262: 15345-15347.

Longmire TA, Ikonomou L, Hawkins F, Christodoulou C, Cao Y, Jean JC, Kwok LW, Mou H, Rajagopal J, Shen SS et al. 2012. Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. Cell stem cell 10: 398-411.

Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, Sridharan R, Clark AT, Plath K. 2008. Generation of human induced pluripotent stem cells from dermal fibroblasts. Proceedings of the National Academy of Sciences of the United States of America 105: 2883-2888.

Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA et al. 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol 9: 625-635.

Matsuda T, Cepko CL. 2007. Controlled expression of transgenes introduced by in vivo electroporation. Proceedings of the National Academy of Sciences of the United States of America 104: 1027-1032.

McClure ML, Wen H, Fortenberry J, Hong JS, Sorscher EJ. 2014. S-palmitoylation regulates biogenesis of core glycosylated wild-type and F508del CFTR in a post-ER compartment. The Biochemical journal 459: 417-425.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297-1303.

Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Bartsevich VV, Lee YL et al. 2008. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nature biotechnology 26: 808-816.

Pinheiro JC, Bates DM. 2000. Mixed-effects models in S and S-PLUS. Springer, New York.

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nature biotechnology 29: 24-26.

Venglarik CJ, Bridges RJ, Frizzell RA. 1990. A simple assay for agonist-regulated Cl and K conductances in salt-secreting epithelial cells. The American journal of physiology 259: C358-364.

Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 38: e164.