Supplementary Materials:

Excision of expanded GAA repeats alleviates the molecular phenotype of Friedreich's ataxia

Yanjie Li[1,7], Urszula Polak[2,3,7], Angela Bhalla[1], Natalia Rozwadowska[1,6], Jill Sergesketter Butler[1,2], David Lynch[4], Sharon Y.R. Dent[2], and Marek Napierala[1,5*]

[1] University of Alabama at Birmingham
Department of Biochemistry and Molecular Genetics
UAB Stem Cell Institute
1825 University Blvd., Birmingham, Alabama 35294, USA

[2] University of Texas MD Anderson Cancer Center
Department of Molecular Carcinogenesis
Center for Cancer Epigenetics
Science Park, Smithville, Texas 78957, USA

[3] Poznan University of Medical Sciences
Department of Cell Biology
Rokietnicka 5D
60-806 Poznan, Poland

[4] Division of Neurology and Pediatrics
Children's Hospital of Philadelphia,
Abramson Research Center Room 502, Philadelphia, PA 19104, USA

[5] Department of Molecular Biomedicine, Institute of Bioorganic Chemistry,
Polish Academy of Sciences, Poznan, 61-704, Poland

[6] Institute of Human Genetics, Polish Academy of Science
Strzeszynska 32, 60-479, Poznan, Poland

[7] These authors contributed equally to this work

Correspondence should be addressed to: Marek Napierala, University of Alabama at Birmingham, Department of Biochemistry and Molecular Genetics, UAB Stem Cell Institute, 1825 University Blvd., Birmingham, Alabama 35294, USA. (205) 975 5320 phone, (205) 975 3335 fax, mnapiera@uab.edu

Table of Contents:

1. Supplementary Materials and Methods……pg 3 - 14
2. Supplementary Figure S1……...pg 15 - 16
3. Supplementary Figure S2……pg 17 - 18
4. Supplementary Figure S3……pg 19 - 20
5. Supplementary Figure S4……pg 21 - 22
6. Supplementary Figure S5……pg 23 - 24
7. Supplementary Figure S6……pg 25 - 26
8. Supplementary Figure S7……pg 27 - 28
9. Supplementary Figure S8……..pg 29 - 30
10. Supplementary References……..…pg 31

Supplementary Materials and Methods:

Cell lines and culture media

K562 chronic myelogenous leukemia cells were purchased from ATCC (#CCL-243) and cultured according to manufacturer conditions in Iscove's Modified Dulbecco's Medium, (IMDM) (Hyclone, cat. SH30228.01) and 10% FBS (Hyclone, cat. SH30910.03) supplemented with penicillin and streptomycin (Hyclone, cat. SV30010).

FRDA and control B-lymphoblasts GM15850 and GM15851, respectively, were purchased from the Coriell Institute for Medical Research. Cells are homozygous for the GAA expansion in the *FXN* gene with alleles of approximately 630 and 860 repeats. Cells were cultured in Roswell Park Memorial Institute Medium 1640 (RPMI1640, cat. SH30096.01) supplemented 2mM L-glutamine (Hyclone, cat. SH30034.01), 15% FBS, and 1% penicillin and streptomycin.

FRDA68 primary skin fibroblasts were obtained through collaboration with Dr. David Lynch (Children’s Hospital of Philadelphia). All studies were conducted in accordance with approvals of CHOP and UAB Institutional Review Boards (CHOP IRB # 10-007864 and UAB IRB#N131204003). Fibroblast cells were cultured in high glucose DMEM/F12 (Invitrogen, cat. 11320) supplemented with glutamine (Hyclone cat. SH30034.01), 15% FBS, MEM Non-Essential Amino acid solution (Corning Cellgro, cat. 25-025-CL) and penicillin and streptomycin. To increase their proliferative potential and enable clonal selection, FRDA68 fibroblasts were transduced with the pBABE-hygro-hTert retroviral vector (Addgene, cat. 1773) followed by 4 days of selection with hygromycin (300 µg/ml).

Excision of the GAA repeat region using custom ZFNs

Determination of ZFN cleavage efficiency and editing of K562 cell lines harboring short GAAs:

K562 nucleofections were conducted using Amaxa Cell Line Nucleofection Kit V (Lonza, cat. VCA-1003) according to the manufacturer's recommendations. Briefly, $1x10^6$ cells were nucleofected with a mix of 2 µg of each UP-ZFN and DN-ZFN encoding RNAs. Separate nucleofection of 2 µg pMAX-GFP control vector (Lonza) was used to monitor the efficiency of transfection. Genomic DNA was extracted 48-72h post-transfection using a GenElute Mammalian Genomic DNA Miniprep kit (Sigma, cat. G1N70-1kt). DNA fragments surrounding the ZFN cleavage sites were PCR-amplified using UP-F/R primers (for UP-ZFN) or DN-F/R primers (for DN-ZNF). Sequences of all primers used in this work are provided below. Amplifications were conducted using 100 ng of genomic DNA and JumpStart *Taq* ready mix (SIGMA, cat. P2893) in a 25 µl volume with the following conditions: initial denaturation 95°C 2 min followed by 30 cycles of 95°C 30 s, 59°C 30 s (60°C 30 s for DN-F/R primers) and 72°C 30 s. The CEL I assay (SURVEYOR Mutation Detection Assay) was carried out precisely as recommended by the manufacturer (Transgenomic cat. 706025). Briefly, heteroduplexes formed by hybridization of thermally denatured PCR products generated from ZFN-cleaved, NHEJ repaired genomic DNA templates and PCR products amplified from non-cleaved DNA templates were digested by the mismatch-specific CEL I nuclease. Products of the CEL I reactions were analyzed on 10% PAGE-TBE gels followed by ethidium bromide staining and quantitative analyses using ImageJ software (NIH).

To obtain edited, single cell-derived clones, a population of UP- and DN-ZFN transfected cells was sorted by FACSAria (BD) or diluted manually and plated at a density of 0.5 - 1 cell per well of a 96-well plate in 100 µl of IMDM media. Approximately 3 – 5 weeks later, genomic DNA was extracted from the clonal cultures using a Quick Extract DNA Extraction Solution (Epicentre, cat. QE09050). The DNA extraction procedure was performed in a thermocycler by incubating ~10^4 cells for 10 min at 70°C and 95°C for 5 min in 100 µl of Quick Extract DNA Extraction Solution. One to five microliters of each DNA solution was used as the template for PCR in a 25 µl reaction volume using UP-F and DN-R primers under the following conditions: initial denaturation 95°C 2 min followed by 35 cycles of 95°C 30 s, 62°C 30 s and 72°C 30 s, final extension 72°C 2 min. Products were resolved on 1% agarose gels.

Determination of ZFN toxicity and off-target effects in FRDA68 fibroblasts:

A set of 3 different assays were used to address functional toxicity and off-target effects of UP/DN ZFNs: (i) DNA damage detection using γH2AX immunostaining, (ii) determination of ZFNs toxicity using XTT cell toxicity/proliferation assays, and (iii) *in silico* identification of the potential off-target sites followed by CEL I analyses of selected DNA regions. FRDA68 fibroblasts were transfected by nucleofection (Amaxa Normal Human Dermal Fibroblast Nucleofector kit VPD-1001, Lonza) with 2 µg of pMAX-GFP, 4 µg *GFP* mRNA, 4 µg UP/DN ZFN mRNAs or 2 µg pDFFB plasmid. The pDFFB plasmid was generated from pENTR223-DFFB (DNASU HsCD00515521) by Gateway gene cloning (Life Technologies). The number of γH2AX foci was also assessed in untransfected cells, mock transfected (nucleofection without genetic material) and cells

treated with hydrogen peroxide (1h, 100 μM at 37 °C). Cells were plated on gelatin-coated cover slips after nucleofection and cultured for 48h, followed by immunodetection of γH2AX foci. Approximately 80 – 100 randomly selected nuclei were analyzed by a blinded investigator. Based on the number of foci, nuclei were classified into 3 groups: no foci, < 5 foci per nucleus, and > 5 foci per nucleus. To determine toxicity of transfection and expression of the UP/DN ZFNs, FRDA68 cells were mock nucleofected or nucleofected with 2 μg pMAX-GFP, 4 μg UP and DN ZFN mRNA, or 4 μg *GFP* mRNA. Cell viability was determined relative to the untransfected cells at 24, 48 and 72h post-transfection using the XTT Cell Proliferation Assay (ATCC 30-1011K) according to manufacturers' recommendations. The absorbance at 475 and 660 nm was determined using a Synergy H1 Hybrid Reader (BioTek).

In silico analysis of potential specific off-target sites was conducted using 3 different algorithms: (i) The ZFN-Site (Cradick, Ambrosini et al. 2011) allowing for a maximum of 2 mismatches per ZFN arm (4 mismatches per each ZFN); (ii) Predicted Report of Genome-wide Nuclease Off-target Sites (Prognos) (Fine, Cradick et al. 2014). (iii) Sigma-Aldrich algorithm. No potential off-targets with less than 6 mismatches were identified. All 4 possible cleavage combinations formed by UP-ZFN and DN-ZFN simultaneously transfected into cells were predicted to target 22 loci with 6 mismatches. Ten of these 22 loci were located in the vicinity of a gene or an annotated transcript, and were analyzed in FRDA68 fibroblasts for UP/DN ZFN cleavage using the CEL I assay as described above. All primer pairs and the predicted sizes of the CEL I cleavage products are listed in the table below.

ZFN editing of the Friedreich's ataxia GM15850 lymphoblasts:

Transfection of ZFNs, culturing of the cells and screening analyses were conducted as described above for K562 cells with one exception. Obtaining single-cell derived clones from GM15850 cells was possible only in the presence of a normal human fibroblast feeder layer (Coriell Institute for Medical Research, GM08399).

ZFN editing of the Friedreich's ataxia FRDA68 fibroblasts:

Nucleofection of $5x10^5$ FRDA68 fibroblasts with ZFN mRNAs was performed using Amaxa Human Dermal Fibroblast Nucleofector Kit (Lonza, cat. VPD-1001). For clonal selection, single cells were plated in individual, 0.2% gelatin covered wells of a 96 well plate in 100 µl of fibroblast media. PCR analysis of individual clones was conducted as described above for K562 cells except for trypsin/EDTA treatment to detach the fibroblasts.

Reprogramming of fibroblasts to iPS cells

Human iPS cells were obtained from non-corrected and ZFN-edited fibroblasts using retroviral transduction of Oct3/4, Sox2, Klf-4 and c-Myc transcription factors as previously described (Ku, Soragni et al. 2010, Polak, Hirsch et al. 2012). After 3-4 weeks of culture, iPSC colonies were manually picked and transferred to matrigel (hESC-qualified Matrix BD, cat. 354277)-coated 24-well plates containing mTeSR1 medium (Stem Cell Technologies, cat. 05850) supplemented with 10 µM ROCK inhibitor Y27632 (Stemgent, cat. 04-0012). Long-term culture of iPSCs was conducted under feeder-free conditions in the mTeSR1 media according to manufacturer

recommendations.

Human iPSCs characterization

Determination of pluripotency:

Characterizing the expression of pluripotency markers in the iPSCs was performed using a previously described immunostaining protocol (Ku, Soragni et al. 2010, Polak, Hirsch et al. 2012). Briefly, exponentially growing iPSCs were washed with PBS buffer and fixed 10 min at 4°C in 4% formaldehyde (Fisher Scientific, cat. BP531). Fixation solution was removed and ice cold methanol was added for 10 min and kept at -20°C. After washing 3 times with PBS at room temperature, the cells were incubated for 30 min with a blocking buffer containing 5% goat or donkey serum and 0.5% Triton X-100. Antibodies for these experiments are listed in the Antibodies section. Alkaline phosphatase and Tra-1-60 staining were performed using Alkaline Phosphatase Live Stain (Life Technologies, cat. A14353) and StainAlive Tra-1-60 antibody. After staining, cells were visualized using Nikon Eclipse Ti-S microscope with NIS-Elements F software (Nikon) or Zeiss Axiovert 40CFL microscope equipped with a Nikon Digital Sight DS-L2 system (Nikon). Determination of retroviral transgene expression in the established iPS cell lines. Retroviral transgene silencing was determined as described in (Wang, Guo et al. 2013).

Expression analysis of pluripotency and differentiation markers using qRT-PCR.

For comprehensive characterization of the expression profiles of pluripotency and differentiation markers in the iPSCs, we used the TaqMan hPSC ScoreCard Panel (Life Technologies, cat. A15872). We conducted ScoreCard analyses of all iPSC lines

generated in this study using Viia 7 (Applied Biosystems). RNA isolation was performed using an RNeasy Mini Kit (Qiagen, cat. 74104) followed by cDNA synthesis using The High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Life Technologies, cat. 4374966). Results were analyzed using the ScoreCard analysis software provided by manufacturer.

Karyotype analysis:

Karyotype analyses of iPSC lines was conducted by Cell Line Genetics Inc. (Madison, WI).

Embryoid bodies (EB) formation and immunostaining:

EBs were cultured from all iPSC lines using AggreWell 400 plates (Stem Cell Technologies, cat. 27845) and APEL media (Stem Cell Technologies, cat. 05210) followed by the spontaneous differentiation protocol described in (Bodnar, Meneses et al. 2004, Son, Kim et al. 2011). Briefly, the iPSCs were cultured until exponential phase and treated with accutase (Stem Cell Technologies, cat. 07920) to obtain a single cell suspension. Cells were plated onto AggreWell 400 at a density of ~ $3x10^6$ cells per well in APEL media supplemented with 10 μM of Y27362 ROCK inhibitor. EBs were cultured on AggreWell plates for 4 days with daily medium changes. Subsequently, EBs were transferred to ultra-low attachment 6-well plates (Corning Costar, cat. 3471) and cultured for 24 h in APEL medium. The next day, 50% of the medium was replaced with EB culture medium containing DMEM high glucose with 20% FBS, L-glutamine, penicillin/streptomycin, non-essential amino acids and β-mercaptoethanol. The expression of SOX17, neuronal class III β-Tubulin (TUJ1) and actin smooth muscle (ASM-1) was analyzed using immunofluorescence as described above to determine the

differentiation potential of the iPSCs into all three germ layers.

RNA isolation, qRT-PCR and western blot analyses.

Total RNA was isolated using an RNeasy Mini kit (Qiagen, cat. 74104) and treated with DNase (Ambion, cat. AM1907). All qRT-PCR analyses, except for the TaqMan ScoreCard panel, were conducted using *Power* SYBR® Green RNA-to-CT™ *1-Step* Kit, according to the manufacturer's protocol (Applied Biosystems, cat. 4389986). Typically, a reaction mixture contained 50 ng RNA in a 10 µl total volume. All reactions were conducted in triplicate with "No RT" control reactions. A standard qRT-PCR reaction protocol included a reverse transcription step at 48°C for 30 min followed by 10 min incubation at 95°C and 40 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 20 s and extension at 60°C for 1 min. The mRNA expression was normalized to the levels of *GAPDH* and/or *β–actin* mRNA. All reactions were performed using a 7500 Fast or StepOnePlus Real-Time PCR System (Applied Biosystems). Expression levels were calculated by $2^{-\Delta\Delta Ct}$ method. All PCR primers are listed below.

For immunoblots, whole cell lysates were prepared using cell lysis buffer (Promega) in the presence of cOmplete Protease Inhibitors (Roche, cat. 11873580001), resolved on 4-12% NuPAGE gels (Invitrogen, cat. NP0322BOX) in MES buffer, transferred to nitrocellulose membranes and probed with appropriate primary and secondary antibodies. Proteins were visualized using ECL Plus Western Blotting Detection Reagents (GE, cat. RPN2133), and the chemiluminescent signal was quantified with the Gel Doc CR+ Imaging System (BioRad).

Amplification of the GAA repeat region

To determine the number of GAA repeats in FRDA and the ZFN-corrected cells, genomic DNA was extracted using a GenElute™ Mammalian Genomic DNA Miniprep Kit (Sigma, cat. G1N70). PCR analyses were carried out with two sets of primers: ZFN-Int amplifying GAA repeats and shorter, 498 bp flanking sequences; and ZFN-Ext amplifying GAAs together with longer, 1370 bp flanking sequences. Primers used for amplification of the GAA region are depicted below and positions indicated in **Supplementary Figure S1**. PCR was conducted as described earlier (Campuzano, Montermini et al. 1996, Kim, Napierala et al. 2011) using the FailSafe PCR System and mix D (Epicentre, cat. FS99100), and amplification products were resolved on 1% agarose gels.

Chromatin immunoprecipitation

Non-corrected and ZFN-edited fibroblasts were cultured in standard DMEM/F12 media. Proteins and DNA were cross-linked with 1% formaldehyde. Chromatin was sheared to ~100-300 bp using a Bioruptor sonicator (Diagenode). Approximately 10^5 cells were used for immunoprecipitation along with 10 µg of each antibody: anti-total histone H3, anti-H3K9K14Ac, and Rabbit IgG. Immunoprecipitates were immobilized using protein A Magna-beads (Millipore, cat. 16-661). After washing, immunoprecipitated DNA was eluted from the beads, reverse cross-linked, and analyzed by qPCR using primers listed below.

Antibodies

Oct ¾ (Santa Cruz Biotechnology, cat. sc-8628, 1:250), Nanog (Cell Signaling Technologies, cat. 4903S, 1:1000), c-Myc Y69 (Abcam, cat. ab32072 1:100), Sox2 (Cell Signaling Technologies, cat. 2748 1:25), Stage-Specific Embryonic Antigen-4, SSEA-4 MC813 (Abcam, cat. Ab16287, 1:100), Tra-1-60 antibody StainAlive Tra-1-60 (Stemgent, cat. 09-0068, 1:50), SOX17 (R&D Systems, cat. AF1924, 1:100), neuronal class III β-Tubulin, TUJ1 (Covance, cat. MMS-435P 1:1000), actin smooth muscle, clone ASM-1 (Millipore, cat. CBL171, 1:100), frataxin H-155 (Santa Cruz Biotechnology, cat. sc-25820, 1:100), GAPDH 6C5 (Millipore, cat. MAB374, 1:40 000), donkey anti-rabbit IgG HRP-linked (GE, cat NA934V, 1:20,000), sheep anti-mouse IgG HRP-linked (GE, cat. NA931V, 1:20,000), goat anti-rabbit with Alexa Fluor 488 (Life Technologies, cat. A11008, 1:1000), donkey anti-mouse with Alexa Fluor 488 (Life Technologies cat. A21202, 1:1000), donkey anti-goat with Alexa Fluor 555 (Life Technologies, cat. A21432, 1:1000), phospho-histone H2A.X (Cell Signaling, cat. 9718, 1:400), anti-total histone 3 (Cell Signaling, cat. 2650), anti-H3K9K14Ac (Cell Signaling, cat. 9677L), and Rabbit IgG (Cell Signaling, cat. 2729),

Primers used in this study:

Primer	Forward 5 → 3	Reverse 5 → 3
To determine ZFN efficiency and editing		
ZFN-UP	ATCTGACCCAGTTACGCCAC	TCTTCAAACACAATGTGGGC
ZFN-DN	TACCTGGTGGCTGTTAAGGG	GAGAAAAGGGTGGGGAAGAG
To determine number of GAA repeats		
GAA-Ext	GGAGGGAACCGTCTGGGCAAAGG	CAATCCAGGACAGTCAGGGCTTT
GAA-Int	GGCTTGAACTTCCCACACGTGTT	AGGACCATCATGGCCACACTT
Quantitative RT-PCR for FXN mRNA		
frataxin	CAGAGGAAACGCTGGACTCT	AGCCAGATTTGCTTGTTTGG
GAPDH	GAAGGTGAAGGTCGGAGTC	GAAGATGGTGATGGGATTTC
ChIP at the FXN locus		
Intron 1	GCATTTACACTGGCTTCTGCTTTCC	AAGTTATCGCAGAGAAGTGACAAGC
Quantitative RT-PCR for FRDA lymphocyte biomarkers		
AGRP	CCTTGCTGTGACCCATGTG	ATGGCAGTACCCAGCTTGC
CBS	AGTCCCCACATCACCACACT	GGGTGTCCCCGATTTTCTTCA
FXYD6	ACCCTGAGGATTGGGGGAC	CATTGGCGGTGATGAGGTT
LEP	CACCAAAACCCTCATCAAGACA	CTTTCTGTTTGGAGGAGACTGACT
LGMN	TTCACCGCAATGGGATTCCTG	GGACTCCCTGATAGACATCTGTG
LILRA2	ATCCTGGGTTAGACGGATACAA	CTGACAGTGATACCGCCCTG
P8	CTATAGCCTGGCCCATTCCT	TCTCTCTTGGTGCGACCTTT
SCARB1	GGAGCCAAAACCGTGTAAAAC	CCTCCTTATCCTTTGAGCCCTTT
TSPAN4	TTTCTCGTGTCCGCTTGACT	GACCAAGGAAGCCCCAAG
PGC1A	AACAGCAGCAGAGACAAATGCACC	TGCAGTTCCAGAGAGTTCCACACT
Quantitative RT-PCR for neuronal markers		
ASCL1	CAACGCCACTGACAAGAAAG	GGAGCTTCTCGACTTCACCA
FABP7	CCCGACCAGGAACATTTTTA	GAAATGGGATGGAAAAGAAA
FOXG1B	CCCTTACTACCGCGAGAACA	CAGCATCCAGTAGTTGCCCT
GAD2	GCACTCACGAGGAAAGGAAC	TGGAACAGACAGCGTGATTC

MAPT	CACACTTGGACTGGACGTTG	CCATGCCAGACCTGAAGAAT
NCAM1	ACTCTCCAACGCTGATCTCC	CAGCCAGCAGATTACAATGC
NEUROD1	CTGTCCAGCTTGGAGGACC	GCCCCAGGGTTATGAGACTA
NKX2	GAGCTTGAGTCCTGAGGGG	CTTCTACGACAGCAGCGACA
NR4A2	CAGGCGTTTTCAGAGGAAAT	GAGACGCGGAGAACTCCTAA
OLIG2	CTGGCGTCCGAGTCCAT	CCTGAGGCTTTTCGGAGC
S100B	TCCACAACCTCCTGCTCTTT	CCACCAATATTCTGGAAGGG
SLC32A1	AGAAACAACCCCAGGTAGCC	CACGACAAGCCCAAAATCA
SLC1A2	GAGCCAAGATGACTGTCGTG	CCGGATAGTGCTGAAGAGGA
SLC1A3	CCATCTTCCCTGATGCCTTA	TTCTCCTTTCCTGGGGAACT
Quantitative RT-PCR for retroviral transgene silencing		
pMXs-F	GTGGTGGTACGGGAAATCAC	–
pMX-Oct3/4	–	TAGCCAGGTTCGAGAATCCA
pMX-Klf4	–	GGGAAGTCGCTTCATGTGAG
pMX-Sox2	–	GGTTCTCCTGGGCCATCTTA
pMXs-Myc	–	AGCAGCTCGAATTTCTTCCA
ZFN off-targets		

Primer	**Forward 5 → 3**	**Reverse 5 → 3**	**CEL I fragments (bp)**
RBFOX3	CTCAGTTCCTGAAGATCAGGCAGT	GGGAAATGGAGGTGGAAACACG	200, 125
RHBG	CATGGTCTTCGTGGGCTTTG	CCTAGAGAAACCAAGTGACGC	201, 99
PXDN	AGGCCACAGCTGGACTTGATG	GTGCTGTGGTCCCCAGTTAG	205, 131
VEPH1	GACCTCCTCAAACACAACCTGC	CATTGCAATCTGCCCTGCTGTG	204, 130
ABCC8	CAGTGCCTGATGCTTATCTTCCC	ATGCTGCTCCTCGTGGAGGT	203, 127
CAMK1G	CAGGATCTATCTCTACTGCTCCC	TTGCCAATTTCCCAGAGTGCTTCC	207, 132
STAT1	GTCCTGGTGTGGAAGGACTG	GGAGTAGCTGTTGTGGGCTT	223, 132
ASB10	CCAGGCAATTTCCCAATGCG	GCTGACCAGGATGGGAAACG	242, 86
CNBD1	TGTGTCCCACACAGAATAGTCC	AAGTCATCTGCAAGCTGAAGG	281, 119
STX2	GCAACCATAAGTTAGCAACTACA	ACATGGAAAGTGTTCAGGCA	247, 105

Supplementary Figure 1

Sequence from: NG_008845.2, 71616 bp, human frataxin (FXN)

```
5703    TCTCCATGCT TGTCACTTCT CTGCGATAAC TTGTTTCAGT AATATTAATA GATGGTATCT GCTAGTATAT ACATACACAT AATGTGTGTG TCTGTGTGTA
        AGAGGTACGA ACAGTGAAGA GACGCTATTG AACAAAGTCA TTATAATTAT CTACCATAGA CGATCATATA TGTATGTGTA TTACACACAC AGACACACAT

5803    TCTGTATATA GCGTGTGTGT TGTGTGTGTG TGTTTGCGCG CACGGGCGCG CGCACACCTA ATATTTTCAA GGCTGGATTT TTTTGAACGA AATGCTTTCC
        AGACATATAT CGCACACACA ACACACACAC ACAAACGCGC GTGCCCGCGC GCGTGTGGAT TATAAAAGTT CCGACCTAAA AAAACTTGCT TTACGAAAGG

5903    TGGAACGAGG TGAAACTTTC AGAGCTGCAG AATAGCTAGA GCAGCAGGGG CCCTGGCTTT TGGAAACTGA CCCGACCTTT ATTCCAGATT CTGCCCCACT
        ACCTTGCTCC ACTTTGAAAG TCTCGACGTC TTATCGATCT CGTCGTCCCC GGGACCGAAA ACCTTTGACT GGGCTGGAAA TAAGGTCTAA GACGGGGTGA

6003    CCGCAGAGCT GTGTGACCTT GGGGGATTCC CCTAACCTCT CTGAGACGTG GCTTTGTTTT CTGTAGGGAG AAGATAAAGG TGACGCCCAT TTTGCGGACC
        GGCGTCTCGA CACACTGGAA CCCCCTAAGG GGATTGGAGA GACTCTGCAC CGAAACAAAA GACATCCCTC TTCTATTTCC ACTGCGGGTA AAACGCCTGG

6103    TGGTGTGAGG ATTAAATGGG AATAACATAG ATAAAGTCTT CAGAACTTCA AATTAGTTCC CCTTTCTTCC TTTGGGGGGT ACAAAGAAAT ATCTGACCCA
        ACCACACTCC TAATTTACCC TTATTGTATC TATTTCAGAA GTCTTGAAGT TTAATCAAGG GGAAAGAAGG AAACCCCCCA TGTTTCTTTA TAGACTGGGT

6203    GTTACGCCAC GGCTTGAAAG GAGGAAACCC AAAGAATGGC TGTGGGGATG AGGAAGATTC CTCAAGGGGA GGACATGGTA TTTAATGAGG GTCTTGAAGA
        CAATGCGGTG CCGAACTTTC CTCCTTTGGG TTTCTTACCG ACACCCCTAC TCCTTCTAAG GAGTTCCCCT CCTGTACCAT AAATTACTCC CAGAACTTCT

6303    TGCCAAGGAA GTGGTAGAGG GTGTTTCACG AGGAGGGAAC CGTCTGGGCA AAGGCCAGGA AGGCGGAAGG GGATCCCTTC AGAGTGGCTG GTACGCCGCA
        ACGGTTCCTT CACCATCTCC CACAAAGTGC TCCTCCCTTG GCAGACCCGT TTCCGGTCCT TCCGCCTTCC CCTAGGGAAG TCTCACCGAC CATGCGGCGT

6403    TGTATTAGGG GAGATGAAAG AGGCAGGCCA CGTCCAAGCC ATATTTGTGT TGCTCTCCGG AGTTTGTACT TTAGGCTTGA ACTTCCCACA CGTGTTATTT
        ACATAATCCC CTCTACTTTC TCCGTCCGGT GCAGGTTCGG TATAAACACA ACGAGAGGCC TCAAACATGA AATCCGAACT TGAAGGGTGT GCACAATAAA

6503    GGCCCACATT GTGTTTGAAG AAACTTTGGG ATTGGTTGCC AGTGCTTAAA AGTTAGGACT TAGAAAATGG ATTTCCTGGC AGGACGCGGT GGCTCATGCC
        CCGGGTGTAA CACAAACTTC TTTGAAACCC TAACCAACGG TCACGAATTT TCAATCCTGA ATCTTTTACC TAAAGGACCG TCCTGCGCCA CCGAGTACGG

6603    CATAATCTCA GCACTTTGGG AGGCCTAGGA AGGTGGATCA CCTGAGGTCC GGAGTTCAAG ACTAACCTGG CCAACATGGT GAAACCCAGT ATCTACTAAA
        GTATTAGAGT CGTGAAACCC TCCGGATCCT TCCACCTAGT GGACTCCAGG CCTCAAGTTC TGATTGGACC GGTTGTACCA CTTTGGGTCA TAGATGATTT

6703    AAATACAAAA AAAAAAAAAA AAGAAGAAGA AGAAGAAGAA AATAAAGAAA AGTTAGCCGG GCGTGGTGTC GCGCGCCTGT AATCCCAGCT ACTCCAGAGG
        TTTATGTTTT TTTTTTTTTT TTCTTCTTCT TCTTCTTCTT TTATTTCTTT TCAATCGGCC CGCACCACAG CGCGCGGACA TTAGGGTCGA TGAGGTCTCC

6803    CTGCGGCAGG AGAATCGCTT GAGCCCGGGA GGCAGAGGTT GCATTAAGCC AAGATCGCCC AATGCACTCC GGCCTGGGCG ACAGAGCAAG ACTCCGTCTC
        GACGCCGTCC TCTTAGCGAA CTCGGGCCCT CCGTCTCCAA CGTAATTCGG TTCTAGCGGG TTACGTGAGG CCGGACCCGC TGTCTCGTTC TGAGGCAGAG

6903    AAAAAATAAT AATAATAAAT AAAAATAAAA AATAAAATGG ATTTCCCAGC ATCTCTGGAA AAATAGGCAA GTGTGGCCAT GATGGTCCTT AGATCTCCTC
        TTTTTTATTA TTATTATTTA TTTTTATTTT TTATTTTACC TAAAGGGTCG TAGAGACCTT TTTATCCGTT CACACCGGTA CTACCAGGAA TCTAGAGGAG

7003    TAGGAAAGCA GACATTTATT ACTTGGCTTC TGTGCACTAT CTGAGCTGCC ACGTATTGGG CTTCCACCCC TGCCTGTGTG GACAGCATGG GTTGTCAGCA
        ATCCTTTCGT CTGTAAATAA TGAACCGAAG ACACGTGATA GACTCGACGG TGCATAACCC GAAGGTGGGG ACGGACACAC CTGTCGTACC CAACAGTCGT

7103    GAGTTGTGTT TTGTTTTGTT TTTTTGAGAC AGAGTTTCCC TCTTGTTGCC CAGGCTGGAG TGCAGTGGCT CAGTCTCAGC TCACTGCAAC CTCTGCCTCC
        CTCAACACAA AACAAAACAA AAAAACTCTG TCTCAAAGGG AGAACAACGG GTCCGACCTC ACGTCACCGA GTCAGAGTCG AGTGACGTTG GAGACGGAGG

7203    TGGGTTCAAG TGATTCTCCT GCCTCAGCCT CCCGAGTAGC TGGGATTATC GGCTAATTTT GTATTTTTAG TAGAGACAGA TTTCTCCATG TTGGTCAGGC
        ACCCAAGTTC ACTAAGAGGA CGGAGTCGGA GGGCTCATCG ACCCTAATAG CCGATTAAAA CATAAAAATC ATCTCTGTCT AAAGAGGTAC AACCAGTCCG

7303    TGGTCTCGAA CTCCCAACCT CAGGTGATCC GCCCACCTCG CCCTCCCAAA GTGCTGGAAT TACAGGCGTG AGCCACCGCG TCTGGCCATC AGCAGAGTTT
        ACCAGAGCTT GAGGGTTGGA GTCCACTAGG CGGGTGGAGC GGGAGGGTTT CACGACCTTA ATGTCCGCAC TCGGTGGCGC AGACCGGTAG TCGTCTCAAA

7403    TTAATTTAGG AGAATGACAA GAGGTGGTAC AGTTTTTTAG ATGGTACCTG GTGGCTGTTA AGGGCTATTG ACTGACAAAC ACACCCAACT TGGCGCTGCC
        AATTAAATCC TCTTACTGTT CTCCACCATG TCAAAAAATC TACCATGGAC CACCGACAAT TCCCGATAAC TGACTGTTTG TGTGGGTTGA ACCGCGACGG

7503    GCCCAGGAGG TGGACACTGG GTTTCTGGAT AGATGGTTAG CAACCTCTGT CACCAGCTGG GCCTCTTTTT TTCTATACTG AATTAATCAC ATTTGTTTAA
        CGGGTCCTCC ACCTGTGACC CAAAGACCTA TCTACCAATC GTTGGAGACA GTGGTCGACC CGGAGAAAAA AAGATATGAC TTAATTAGTG TAAACAAATT

7603    CCTGTCTGTT CCATAGTTCC CTTGCACATC TTGGGTATTT GAGGAGTTGG GTGGGTGGCA GTGGCAACTG GGGCCACCAT CCTGTTTAAT TATTTTAAAG
        GGACAGACAA GGTATCAAGG GAACGTGTAG AACCCATAAA CTCCTCAACC CACCCACCGT CACCGTTGAC CCCGGTGGTA GGACAAATTA ATAAAATTTC

7703    CCCTGACTGT CCTGGATTGA CCCTAAGCTC CCCCTGGTCT CCAAAATTCA TCAGAAACTG AGTTCACTTG AAGGCCTCTT CCCCACCCTT TTCTCCACCC
        GGGACTGACA GGACCTAACT GGGATTCGAG GGGGACCAGA GGTTTTAAGT AGTCTTTGAC TCAAGTGAAC TTCCGGAGAA GGGGTGGGAA AAGAGGTGGG

7803    CTTGCATCTA CTTCTAAAGC AGCTGTTCAA CAGAAACAGA ATGGGAGCCA CACACATAAT TCTACATTTT CTAGTTAAAA AGAAAAAAAA ATCATTTTCA
        GAACGTAGAT GAAGATTTCG TCGACAAGTT GTCTTTGTCT TACCCTCGGT GTGTGTATTA AGATGTAAAA GATCAATTTT TCTTTTTTTT TAGTAAAAGT

7903    ACAATATATT TATTCAACCT AGTACATACA AAATATTATC ATTCCAACAT GTAATCAGTA TTTTAAAAAT CAGTAATGAG ACCAGGCACG GTGGCTCACG
        TGTTATATAA ATAAGTTGGA TCATGTATGT TTTATAATAG TAAGGTTGTA CATTAGTCAT AAAATTTTTA GTCATTACTC TGGTCCGTGC CACCGAGTGC
```

UP-F primer

GAA-Ext F primer

UP-ZFN

E-box

GAA-Int F primer

UP-R

GAA repeats (only 6 repeats are presented)

GAA-Int R primer

DN-F

DN-ZFN

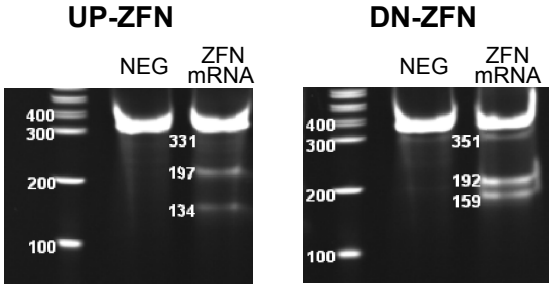
GAA-Ext F primer

DN-R

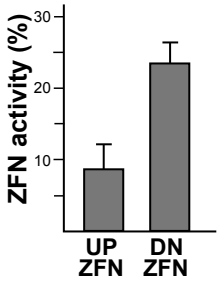
Supplementary Figure S1. Sequence of an *FXN* intron 1 fragment based on NCBI Reference Sequence: NG_008845.2. Locations of UP and DN ZFN binding sites are highlighted in green with the nuclease cleavage sequences in bold between. The GAA repeats (only 6 GAAs are shown for clarity) are highlighted in yellow, and the E-box motif is highlighted in blue. The primers used to determine editing of the GAA repeat region are marked in red, while the primers used for amplification of the GAA region are underlined. See **Supplementary Materials and Methods** for primer sequences.

Supplementary Figure S2

a



b



c

UP-ZFN

>FXN-up-NM000144-r2388a1-FokI-ELD
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAMAERPFQCRICMRNFSQSSHLTRHIRTHTGEKPFACDICGRK
FARLDNRTAHTKIHTGSQKPFQCRICMRNFSRSDTLSEHIRTHTGEKPFACDICGRKFAARSTRTNHTKIHLRGSQLVKSEL
EEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYS
GGYNLPIGQADEMERYVEENQTRDKHLNPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHITNCNGAVLSVEELLIG
GEMIKAGTLTLEEVRRKFNNGEINFRS

>FXN-up- NM000144-2395a1-FokI-KKR
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAMAERPFQCRICMRNFSRSDHLSRHIRTHTGEKPFACDICGRK
FAQWNGRFQHTKIHTGSQKPFQCRICMRNFSQSGSLTRHIRTHTGEKPFACDICGRKFATSSNRKTHTKIHTHPRAPIPKPF
QCRICMRNFSDRSDLSRHIRTHTGEKPFACDICGRKFADRSNRNKHTKIHLRGSQLVKSELEEKKSELRHKLKYVPHEYIEL
IEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQ
TRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNRKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNG
EINF

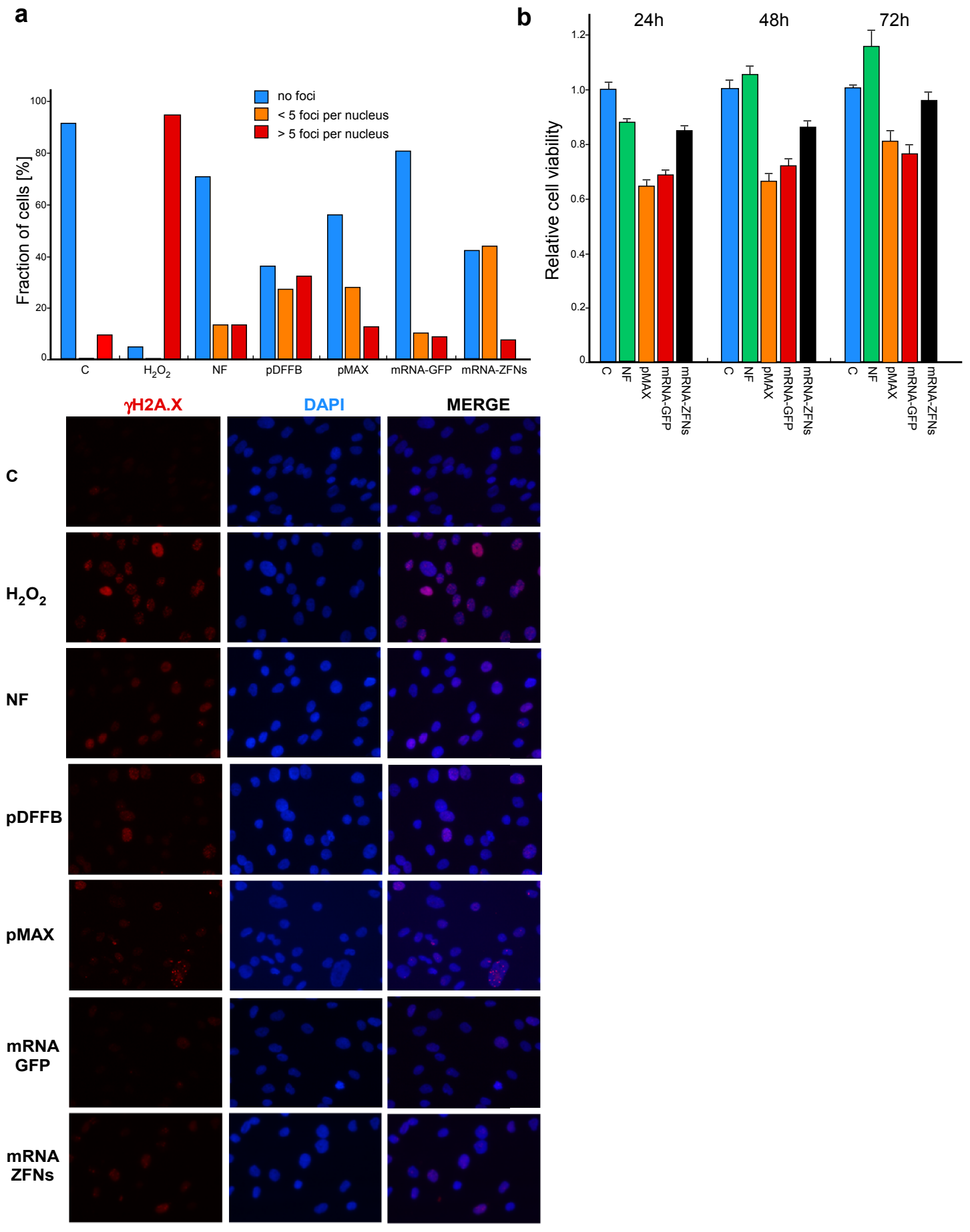
DN-ZFN

>FXN-down-NM000144-r3636a1-FokI-ELD
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAMAERPFQCRICMRNFSRSDHLSRHIRTHTGEKPFACDICGRK
FADRSNRKTHTKIHTGSQKPFQCRICMRKFARSDALARHTKIHTGEKPFQCRICMRNFSTSGNLTRHIRTHTGEKPFACDIC
GRKFAQAENLKSHTKIHLRGSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLG
GSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDKHLNPNEWWKVYPSSVTEFKFLFVSGHFKG
NYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFRS

>FXN-down- NM000144-3643a1-FokI-KKR
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAMAERPFQCRICMRNFSRSDVLSAHIRTHTGEKPFACDICGRK
FARNDHRINHTKIHTGSQKPFQCRICMRNFSRSAHLSRHIRTHTGEKPFACDICGRKFATSGSLTRHTKIHTGSQKPFQCRI
CMRNFSQSGHLARHIRTHTGEKPFACDICGRKFAQRTHLNSHTKIHLRGSQLVKSELEEKKSELRHKLKYVPHEYIELIEIA
RNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNK
HINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNRKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

Supplementary Figure S2. Characterization of UP and DN ZFNs. (**a**) Efficiency of DNA cleavage by UP and DN ZFNs in K562 cells. Cells were nucleofected with ZFN mRNAs (ZFN mRNA lane) or with *GFP* mRNA (NEG lane). PCR products amplified using the UP-F/UP-R and DN-F/DN-R primers were subjected to the CEL I assay. (**b**) ZFN activity was quantitated using ImageJ software. (**c**) Protein sequences of ZFNs used in this study. CompoZr ZFNs UP and DN were designed by Sigma-Aldrich and are available as: Lots 09011112MN (UP-ZFN) and 09011111MN (DN-ZFN). To prevent homodimerization, minimize off-target cleavage events, and increase activity, the ELD:KKR variants of FokI nuclease were used (Doyon, Vo et al. 2011).

Supplementary Figure S3



Supplementary Figure S3. Assessment of UP and DN ZFN toxicity. (**a**) Immunostaining of γH2AX was carried out in FRDA68 fibroblasts. Label designations are as follows: C - untransfected control, H_2O_2 - cells treated with hydrogen peroxide, NF - mock nucleofection, pMAX – cells transfected with pMAX GFP, mRNA-GFP - cells transfected with *GFP* mRNA, and mRNA-ZFNs – cells transfected with UP and DN ZFN mRNAs. Identified foci were counted and nuclei were divided into 3 groups: no foci (blue bars), < 5 foci (orange bars), > 5 foci (red bars). Results are based on analysis of 80 – 100 nuclei. Representative images of γH2AX immunostaining are shown below. (**b**) Relative cell viability was assessed using the XTT assay performed 24, 48 and 72h post-nucleofection. The results are shown with bar designations as described in panel (**a**).

Supplementary Figure S4

a

					No. of mismatches for ZFN heterodimers						
ZFN right arm	target seq	ZFN left arm	target seq		0	1	2	3	4	5	6
NM000144-2395a1	ggTACGCCgCATGTATTAGGGgagatga	NM000144-r2388a1	ccACTCTGAAGGGAtccccttccgcctt	FXN	1	0	0	0	0	0	8
NM000144-3643a1	ttTGAGGAGTTGGGTGGGTGgcagtggc	NM000144-r3636a1	ccCAAGATGTGCAAGGGaactatggaac	FXN	1	0	0	0	0	0	0

				No. of mismatches for ZFN heterodimers						
ZFN right arm	target seq	ZFN left arm	target seq	0	1	2	3	4	5	6
NM000144CROSSPAIR-2935a1	ggTACGCCgCATGTATTAGGGgagatga	NM000144CROSSPAIR-r3636a1	ccCAAGATGTGCAAGGGaactatggaac	0	0	0	0	0	0	0
NM000144CROSSPAIR-3643a1	ttTGAGGAGTTGGGTGGGTGgcagtggc	NM000144CROSSPAIR-r2388a1	ccACTCTGAAGGGAtccccttccgcctt	0	0	0	0	0	0	14

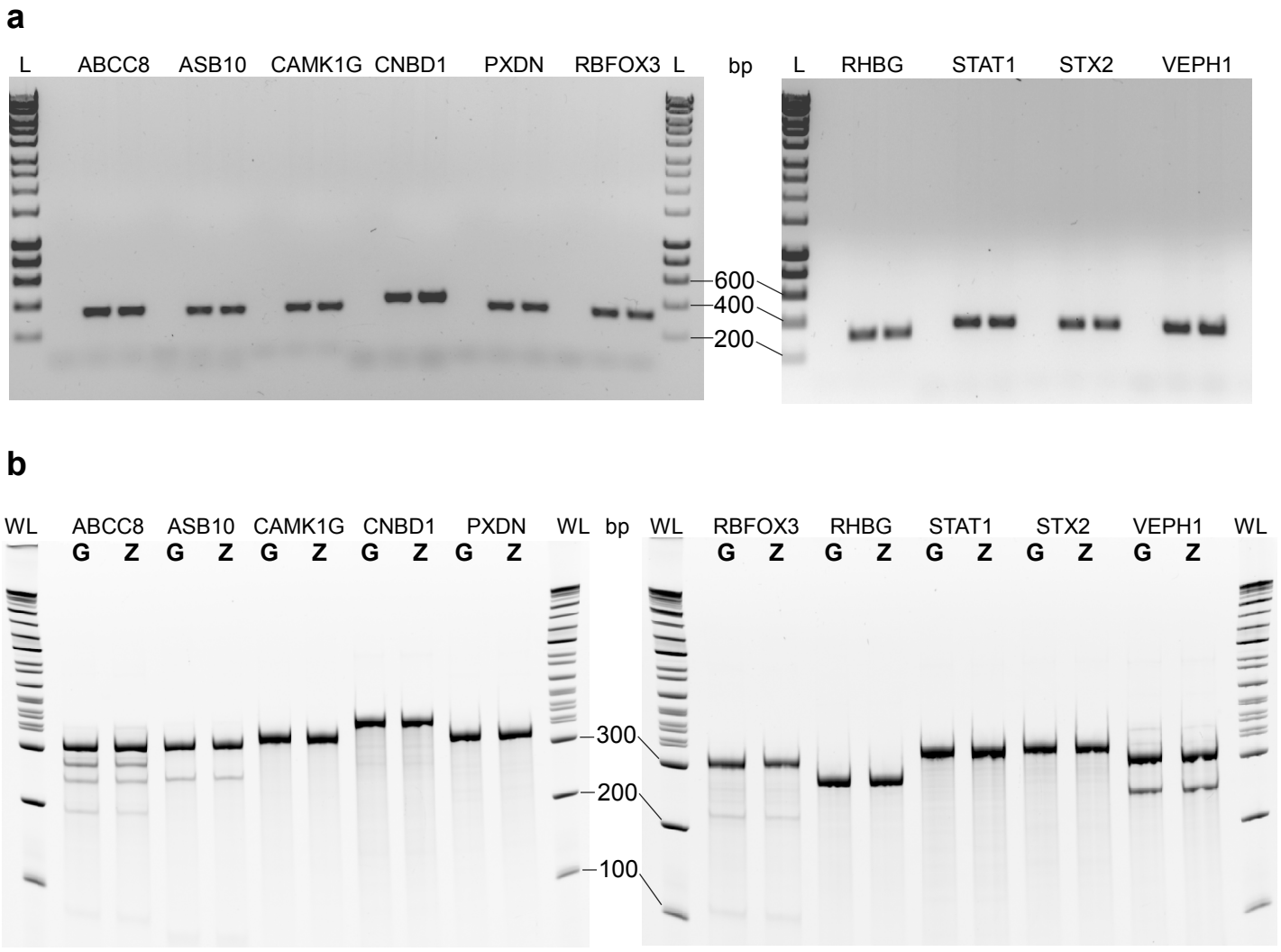
b

chr	pos	seq	# of mismatch	ZFN pair	annotation
chr9	71651853	ATCCCTTCAGAGTGGCTGGTACGCCGCATGTATTAGGGG	0	NM000144-r2388a1_N6_NM000144-2395a1	FXN
chr8	88152588	CTCCCTTCAGAGaTTTTGTACttCTCATtTATTAttGT	6	NM000144-r2388a1_N5_NM000144-2395a1	CNBD1
chr2	147929655	TCCCTAtTACATGTGagGaAATTCAGgCTCTtAAGGGAC	6	NM000144-2395a1_N6_NM000144-r2388a1	
chr2	140366992	TTCCCTTCAGgGTGGCAAGTACcCCCATGctcTAGGcA	6	NM000144-r2388a1_N6_NM000144-2395a1	
chr14	26460996	ATttaTTCAGAGTGTGGTATAtGCCACATGTATTAGtcC	6	NM000144-r2388a1_N6_NM000144-2395a1	
chr12	131286328	TTtCCcTaAGAGTTTAACTcCGCCTCATGTAaaAGGGG	6	NM000144-r2388a1_N5_NM000144-2395a1	STX2
chr11	80449918	ACtCctATACATGaGgaTAAGGTCAACTCTGAAGGGAA	6	NM000144-2395a1_N6_NM000144-r2388a1	
chr11	80449918	ACtCctATACATGAGGataAGGTCAACTCTGAAGGGAA	6	NM000144-2395a1_N5_NM000144-r2388a1	
chr6	14766018	TCCCTAtgACAaGTGGaGTtTTGCGAAaTCTGAAGGGAT	6	NM000144-2395a1_N6_NM000144-r2388a1	

chr	pos	seq	# of mismatch	cross pairs	annotation
chrX	139608453	TTtCCTTCAGAaTTCTTTcaAGcAGTTGGGaGGGTGG	6	NM000144CROSSPAIR-r2388a1_N5_NM000144CROSSPAIR-3643a1	
chr11	17484877	ATCCCTTCtcAGTTTGGCTGAGaAGcaGGGTGGGgGC	6	NM000144CROSSPAIR-r2388a1_N5_NM000144CROSSPAIR-3643a1	ABCC8
chr2	191849865	TCACaaACCCAACTCCTtAAGGATCACgCTGtAGGaAG	6	NM000144CROSSPAIR-3643a1_N6_NM000144CROSSPAIR-r2388a1	STAT1
chr2	1677085	GTCCCTTCAGAtgGCGCCTGgGGgGTgGGGTcGGTGG	6	NM000144CROSSPAIR-r2388a1_N5_NM000144CROSSPAIR-3643a1	PXDN
chr2	213523815	TTCCCTTCAGAGTTTGACTGtGGtGTaaGtTGaGTGT	6	NM000144CROSSPAIR-r2388a1_N5_NM000144CROSSPAIR-3643a1	
chr3	157122068	CTtCCTTCAGAGTGCTTATGgtGgGTTGGGaGtGTGT	6	NM000144CROSSPAIR-r2388a1_N5_NM000144CROSSPAIR-3643a1	VEPH1
chr17	77301424	CCACCCACCCAACTCCcCAGAACAGAaTCaGAtaaGAA	6	NM000144CROSSPAIR-3643a1_N6_NM000144CROSSPAIR-r2388a1	RBFOX3
chr7	64068916	TaACCCACCCAACTCCTgAGATCTCACTggGAAGctAC	6	NM000144CROSSPAIR-3643a1_N6_NM000144CROSSPAIR-r2388a1	
chr7	150883434	TTCCCTTCAGAGTGGAACTGgGGgtTgGGGTGaGgGG	6	NM000144CROSSPAIR-r2388a1_N5_NM000144CROSSPAIR-3643a1	ASB10
chr7	46801554	TgCCCcTCAGAGTGTGTCTTcAGGAGTTcccTGGGTGT	6	NM000144CROSSPAIR-r2388a1_N6_NM000144CROSSPAIR-3643a1	
chr12	30766252	AgACCCACCtAtCTCtgaACTAAAAACTCTGAAGGGAC	6	NM000144CROSSPAIR-3643a1_N6_NM000144CROSSPAIR-r2388a1	
chr1	156347290	CCACCCACCCAgCTCCcCAAGGTTCACTCgGgAGGccC	6	NM000144CROSSPAIR-3643a1_N6_NM000144CROSSPAIR-r2388a1	RHBG
chr1	209780138	GCACCaACCCtACTCCTttCTCCCACTCTGAAGGagA	6	NM000144CROSSPAIR-3643a1_N5_NM000144CROSSPAIR-r2388a1	CAMK1G
chr9	15121361	TCtCCCACCaAACTCCTCAACACTTtCTCTGAAGtacT	6	NM000144CROSSPAIR-3643a1_N6_NM000144CROSSPAIR-r2388a1	

Supplementary Figure S4*. In silico* identification of potential off-target sites of UP and DN ZFNs using the Sigma-Aldrich algorithm. (**a**) Analyses were conducted for all 4 possible ZFN heterodimer combinations (UP-ELD and UP-KKR; DN-ELD and DN-KKR; UP-ELD and DN-KKR and DN-ELD and UP-KKR). (**b**) The list of potential off-target loci with their exact nucleotide position in the genome is given. Ten of the 22 loci were located in the vicinity of a gene or an annotated transcript (highlighted in blue). The target site in the *FXN* gene with no mismatches is highlighted in yellow.

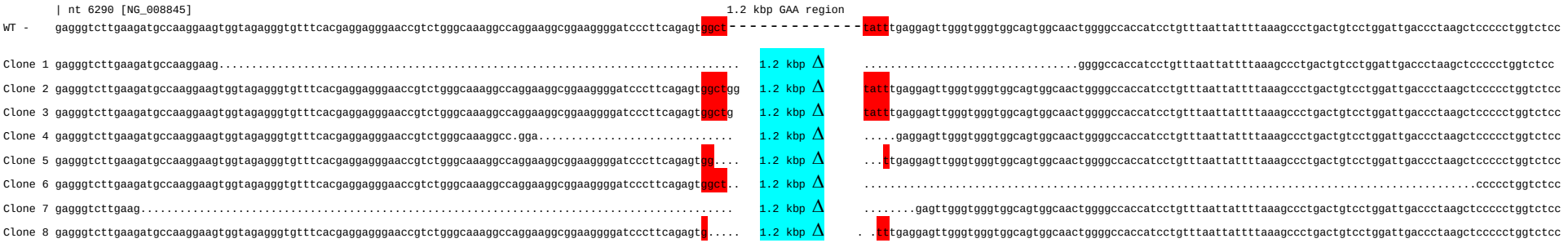
Supplementary Figure S5



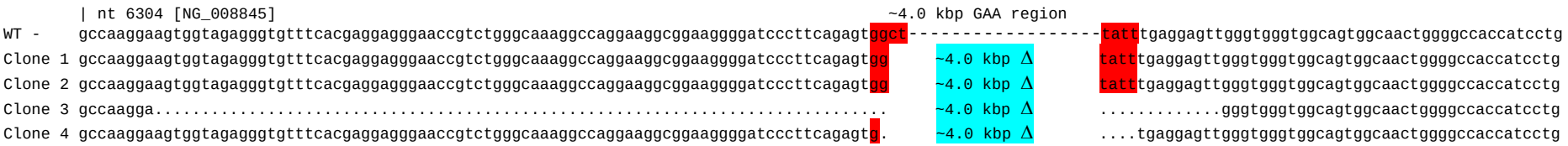
Supplementary Figure S5. CEL I analysis of *in silico* identified potential off-target sites. (**a**) Agarose gel analysis of the PCR products of 10 predicted potential off-target sites of the ZFN pairs (**Supplementary Figure S4**). FRDA68 fibroblasts were nucleofected with UP and DN ZFN mRNAs. Genomic DNA was isolated 48h post-transfection and amplified with primers specific for the loci indicated above the gel lanes. Hyperladder 1 (BioLine) is designated by L. (**b**) Polyacryalamide gel electrophoresis analysis of the CEL I assay products. Reactions conducted using genomic DNA isolated from control (GFP) and ZFN transfected cells are designated G and Z, respectively.

Supplementary Figure S6

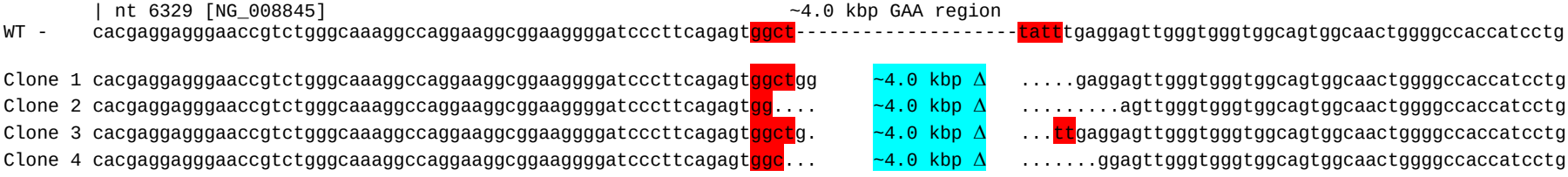
a K562 cells



b GM15850 lymphoblasts

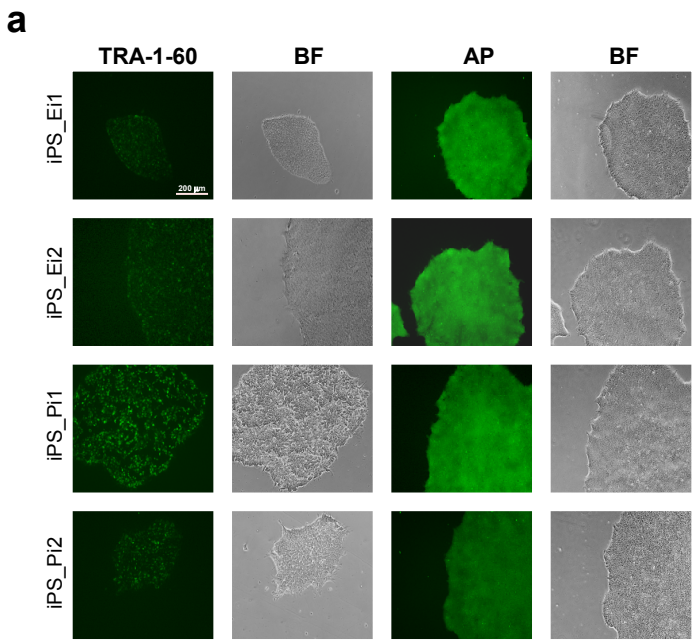


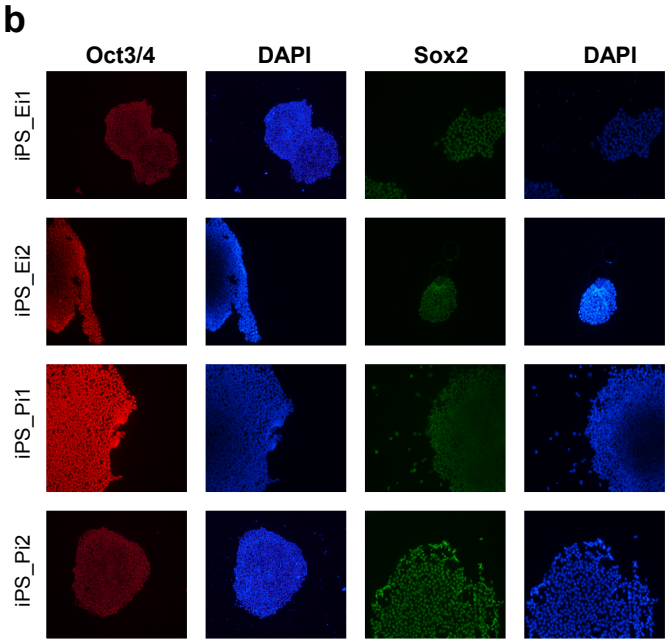
c FRDA68 fibroblasts

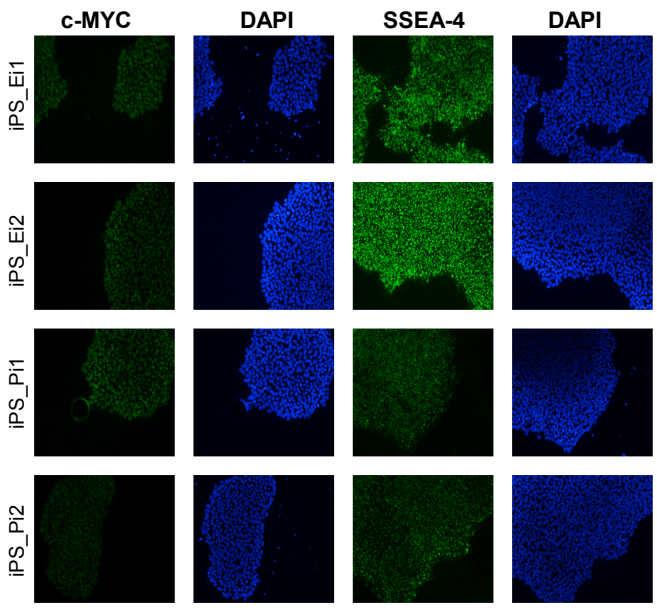


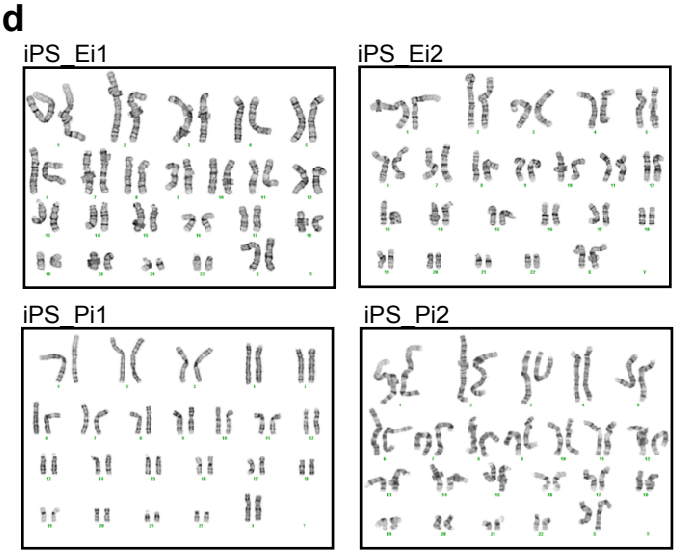
Supplementary Figure S6. DNA sequence analyses of the cell lines edited by UP-ZFN and DN-ZFN. Clones were obtained by editing the *FXN* locus in K562 cells (**a**), GM15850 FRDA lymphoblasts (**b**) and FRDA68 fibroblasts (**c**). PCR products obtained using UP-F and DN-R primers were cloned using a TOPO PCR cloning kit and sequenced. WT designates the initial, non-edited DNA sequence. The location of the edited DNA region within the human *FXN* gene is indicated by nucleotide number (nt) according to the NCBI Reference Sequence NG_008845. Short GAA repeats together with their flanking sequences were edited in K562 cells (~1.2 kbp deletion indicated in blue). Expanded GAAs ranging from ~540 to 1400 repeats depending on the cell line and edited allele (~4 kbp deletion indicated in blue) were cut out in the GM15850 and FRDA68 cells. The remaining sequence of the ZFN cleavage site is highlighted in red, with dotted fragments representing DNA sequences resected during NHEJ repair of the ZFN-cleaved DNA ends.

Supplementary Figure S7







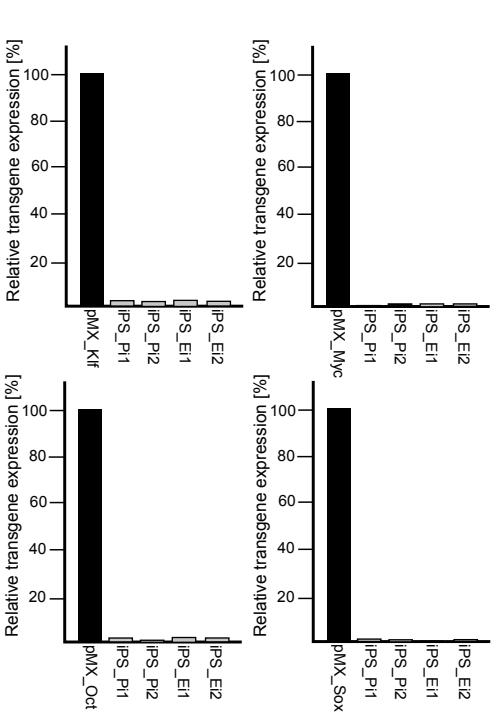


e

Sample Name	Cell Type	Pluripotent	Ectoderm	Mesoderm	Endoderm
iPS_Ei1	Pluripotent	0.42	-1.50	-0.27	-1.31
iPS_Ei2	Pluripotent	0.56	-1.53	-0.10	-1.15
iPS_Pi1	Pluripotent	0.89	-2.14	-0.04	-1.29
iPS_Pi2	Pluripotent	0.60	-3.16	-0.48	-1.57

Gene expression relative to the reference standard

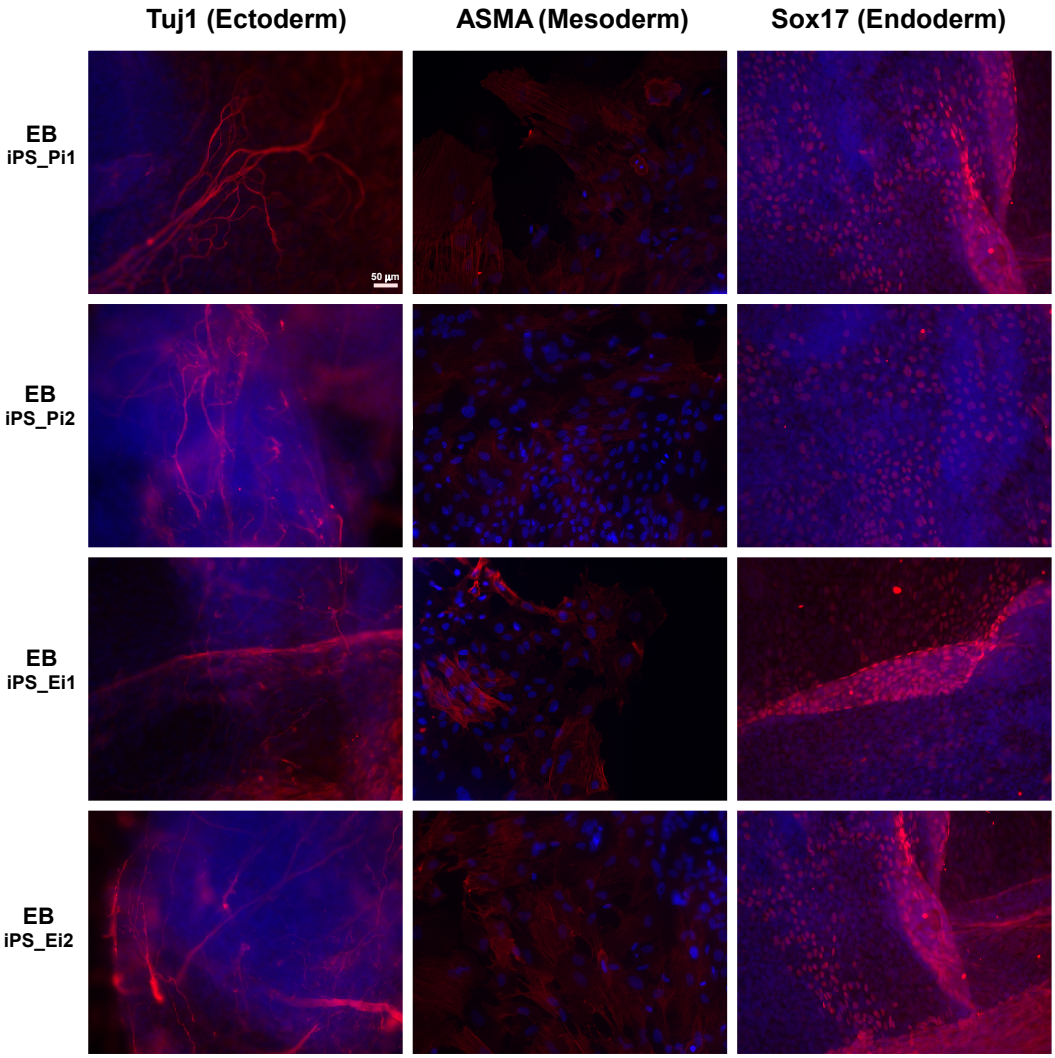
Upregulated			Comparable			Downregulated
x>1.5	1.0<x<=1.5	0.5<x<=1.0	-0.5<=x<=0.5	-1.0<=x<-0.5	-1.5<=x<-1.0	x<-1.5



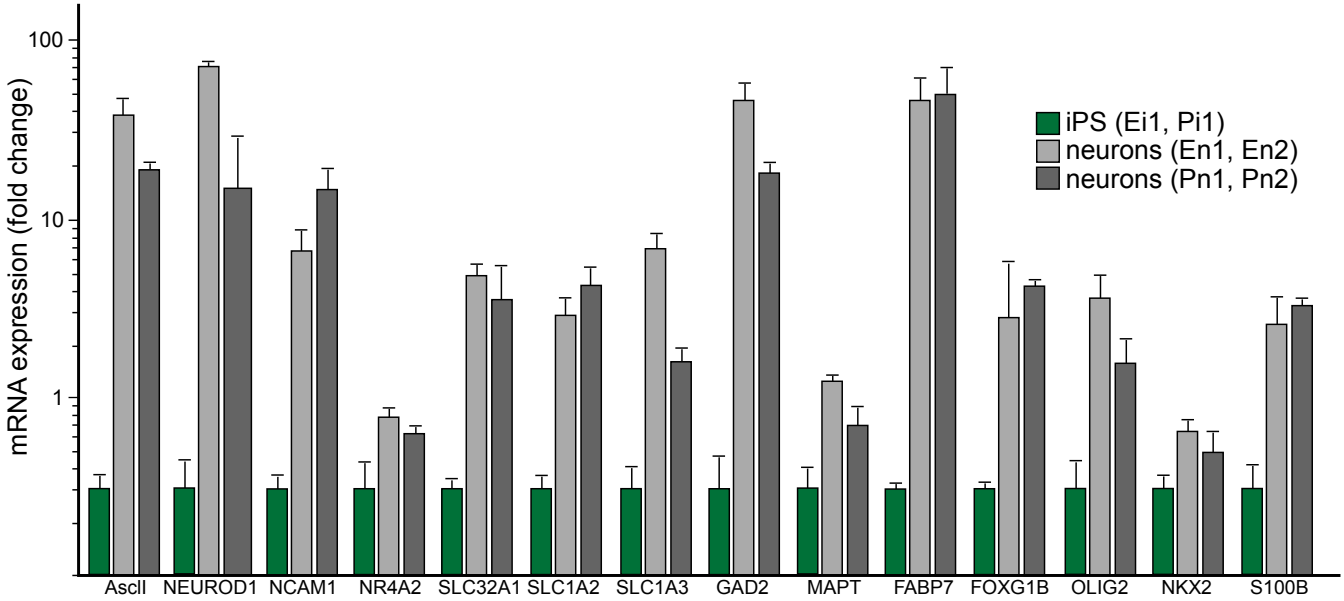
Supplementary Figure S7. Characterization of non-GAA edited FRDA iPS cells (clones Pi1 and Pi2) and ZFN-corrected FRDA iPS cells (clones Ei1 and Ei2). (**a**) Expression of Tra-1-60 and alkaline phosphatase (AP); BF – bright field. (**b**) Expression of Oct3/4 and Sox2 markers; DAPI – nuclear staining. (**c**) Expression of c-MYC and SSEA-4 proteins; DAPI – nuclear staining. The scale bar represents 200 μm. (**d**) Cytogenetic analyses of iPS cells. Karyotyping was performed by Cell Line Genetics (Madison, WI) on 20 G-banded metaphase cells. Typically, 20 metaphases were counted and 5 - 7 metaphases were analyzed. All iPSCs generated in this study demonstrated a correct 46,XX female karyotype. (**e**) Determination of the pluripotency and differentiation score using hPSC ScoreCard. Results of qRT-PCR mRNA expression analysis of 93 different markers of pluripotency, ectoderm, endoderm and mesoderm in four iPSC lines (Pi1, Pi2, Ei1 and Ei2) using hPSC ScoreCard were evaluated using the hPSC ScoreCard Software Analysis Tool (Life Technologies). Scores are a statistical representation of differences between the level of gene expression in iPSCs and an undifferentiated reference standard. (**f**) Determination of retroviral transgene expression in the established iPS cell lines. The expression level of the OSKM transcription factors in all iPSCs was determined using qRT-PCR. Transgene specific expression of OSKM mRNA in fibroblasts 7 days after transduction with retroviruses was used as a reference (lanes pMX_Oct, pMX_Sox, pMX_Klf and pMX_Myc). The qRT-PCR was performed using primers described in (Wang, Guo et al. 2013). Results are an average of two independent experiments.

Supplementary Figure S8

a



b



Supplementary Figure S8. Differentiation potential of the ZFN-edited and non-edited iPSC lines. (**a**) Immunostaining with antibodies specific for ectoderm (Tuj1), mesoderm (ASMA) and endoderm (Sox17) markers in embryoid bodies differentiated from Ei1, Ei2, Pi1 and Pi2 cell lines. The scale bar represents 50 μm. (**b**) Characterization of iPSC derived neurons. qRT-PCR analysis was performed to measure the expression of neuronal markers in iPSC-derived neurons (En1, En2, Pn1 and Pn2; gray bars) relative to the parental iPS cells (average of Ei1 and Pi1; green bars) after 14 days of differentiation. Error bars represent SD of 3 separate analyses.

Supplementary References:

1. Bodnar, M. S., J. J. Meneses, R. T. Rodriguez and M. T. Firpo (2004). "Propagation and maintenance of undifferentiated human embryonic stem cells." Stem Cells Dev **13**(3): 243-253.

2. Campuzano, V., L. Montermini, M. D. Molto, L. Pianese, M. Cossee, F. Cavalcanti, E. Monros, F. Rodius, F. Duclos, A. Monticelli, F. Zara, J. Canizares, H. Koutnikova, S. I. Bidichandani, C. Gellera, A. Brice, P. Trouillas, G. De Michele, A. Filla, R. De Frutos, F. Palau, P. I. Patel, S. De Donato, J. L. Mandel, S. Cocozza, M. Koenig and M. Pandolfo (1996). "Friedreich's ataxia: Autosomal recessive disease caused by an intronic GAA triplet repeat expansion." Science **271**(5254): 1423-1427.

3. Cradick, T. J., G. Ambrosini, C. Iseli, P. Bucher and A. P. McCaffrey (2011). "ZFN-site searches genomes for zinc finger nuclease target sites and off-target sites." BMC Bioinformatics **12**: 152.

4. Doyon, Y., T. D. Vo, M. C. Mendel, S. G. Greenberg, J. Wang, D. F. Xia, J. C. Miller, F. D. Urnov, P. D. Gregory and M. C. Holmes (2011). "Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures." Nat Methods **8**(1): 74-79.

5. Fine, E. J., T. J. Cradick, C. L. Zhao, Y. Lin and G. Bao (2014). "An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage." Nucleic Acids Res **42**(6): e42.

6. Kim, E., M. Napierala and S. Y. Dent (2011). "Hyperexpansion of GAA repeats affects post-initiation steps of FXN transcription in Friedreich's ataxia." Nucleic Acids Res **39**(19): 8366-8377.

7. Ku, S., E. Soragni, E. Campau, E. A. Thomas, G. Altun, L. C. Laurent, J. F. Loring, M. Napierala and J. M. Gottesfeld (2010). "Friedreich's ataxia induced pluripotent stem cells model intergenerational GAATTC triplet repeat instability." Cell Stem Cell **7**(5): 631-637.

8. Polak, U., C. Hirsch, S. Ku, J. M. Gottesfeld, S. Y. R. Dent and M. Napierala (2012). "Selecting and isolating colonies of human induced pluripotent stem cells reprogrammed from adult fibroblasts." J. Vis. Exp. **60**(, DOI: 10.3791): e3416.

9. Son, M. Y., H. J. Kim, M. J. Kim and Y. S. Cho (2011). "Physical passaging of embryoid bodies generated from human pluripotent stem cells." PLoS One **6**(5): e19134.

10. Wang, G., X. Guo, W. Hong, Q. Liu, T. Wei, C. Lu, L. Gao, D. Ye, Y. Zhou, J. Chen, J. Wang, M. Wu, H. Liu and J. Kang (2013). "Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation." Proc Natl Acad Sci U S A **110**(8): 2858-2863.