nature biomedical engineering

SUPPLEMENTARY INFORMATION

DOI: 10.1038/s41551-017-0141-6

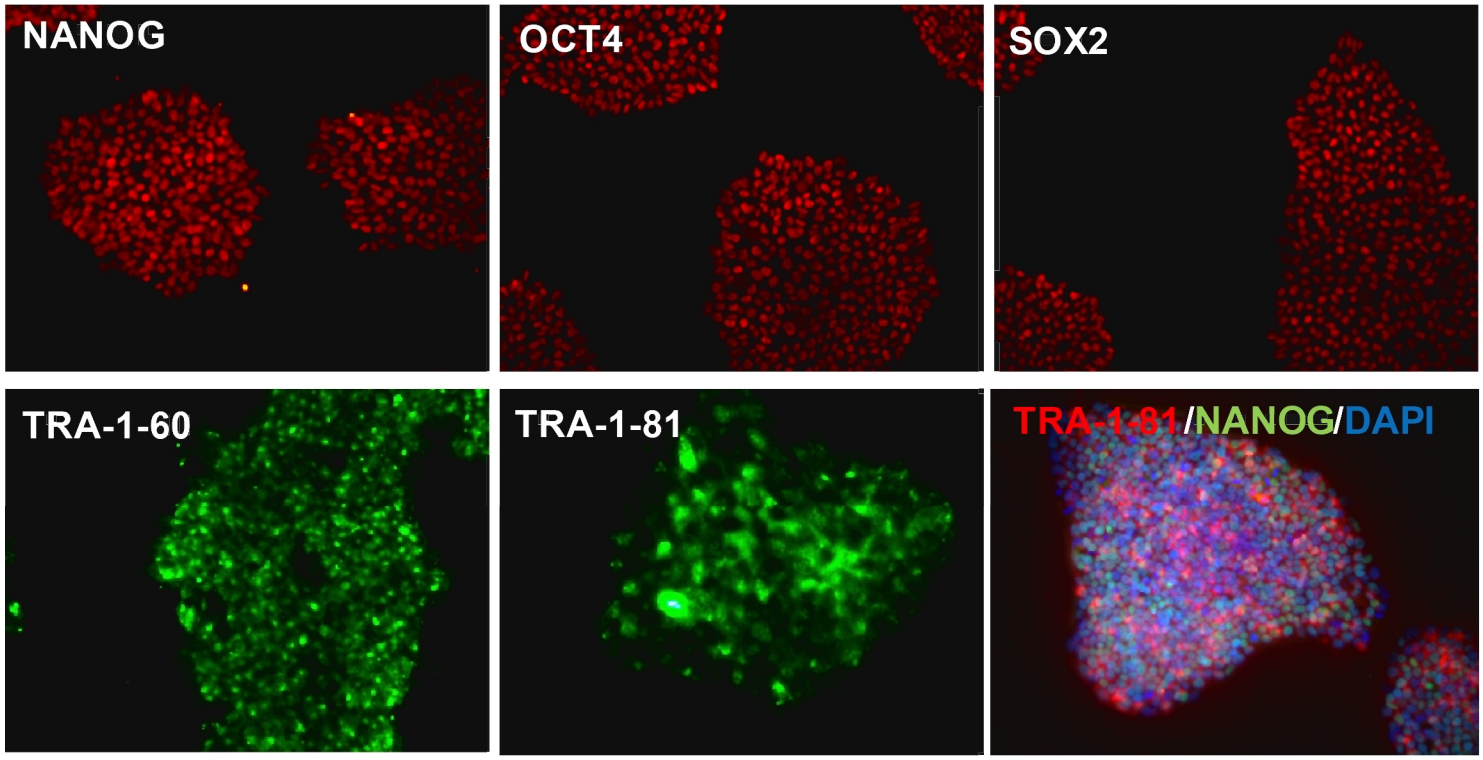
In the format provided by the authors and unedited.

Transcriptomic and epigenomic differences in human induced pluripotent stem cells generated from six reprogramming methods

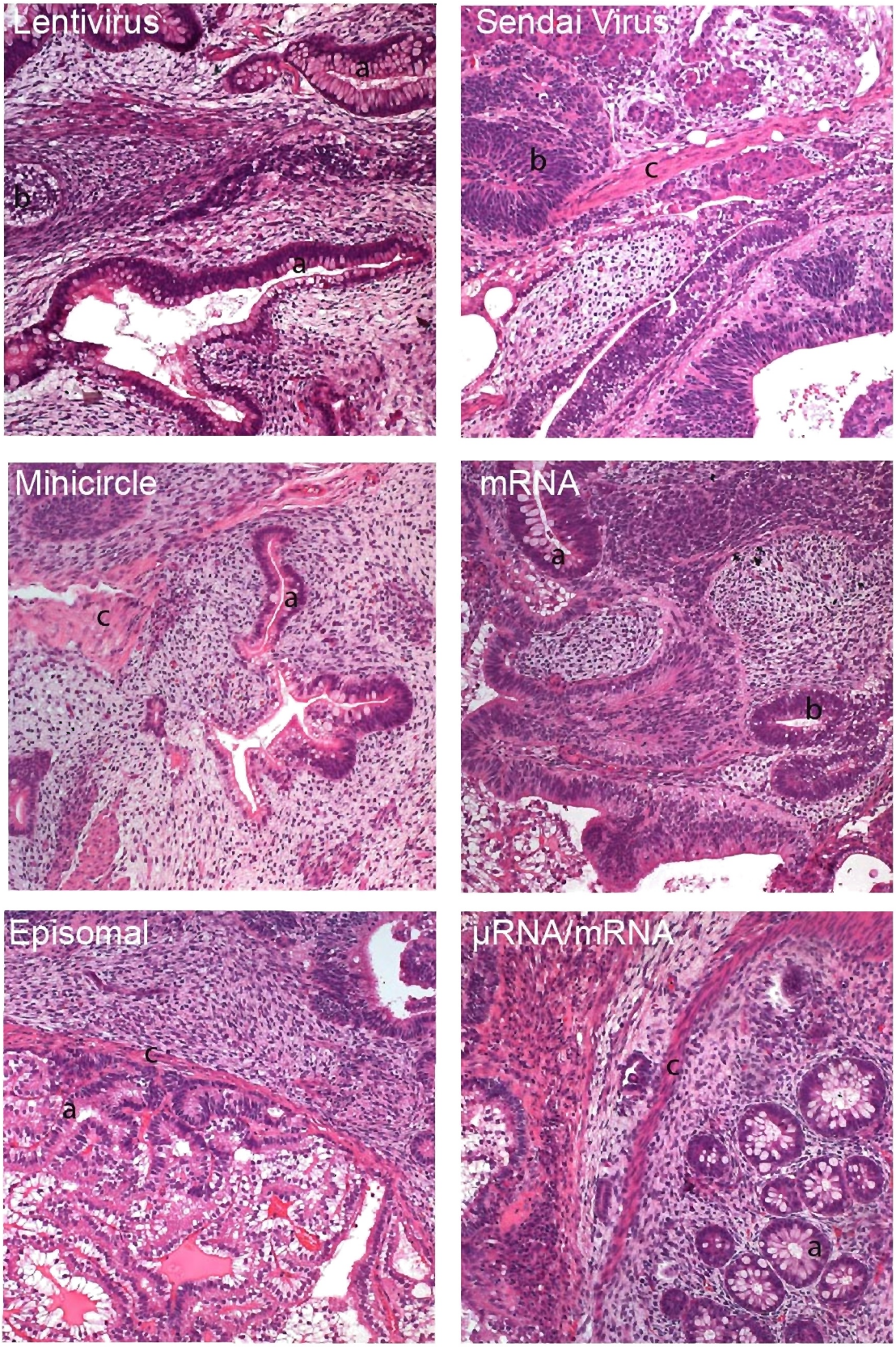
Jared M. Churko[1,2,3], Jaecheol Lee[1,2,3], Mohamed Ameen[1,2,3], Mingxia Gu[1,2,3], Meenakshi Venkatasubramanian[4], Sebastian Diecke[1,2,3], Karim Sallam[1,2,3], Hogune Im[4], Gavin Wang[5], Joseph D. Gold[1], Nathan Salomonis[4], Michael P. Snyder[5] and Joseph C. Wu[1,2,3]*

[1]Institute of Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA. [2]Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA 94305, USA. [3]Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA. [4]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, USA. [5]Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA. *e-mail: joewu@stanford.edu

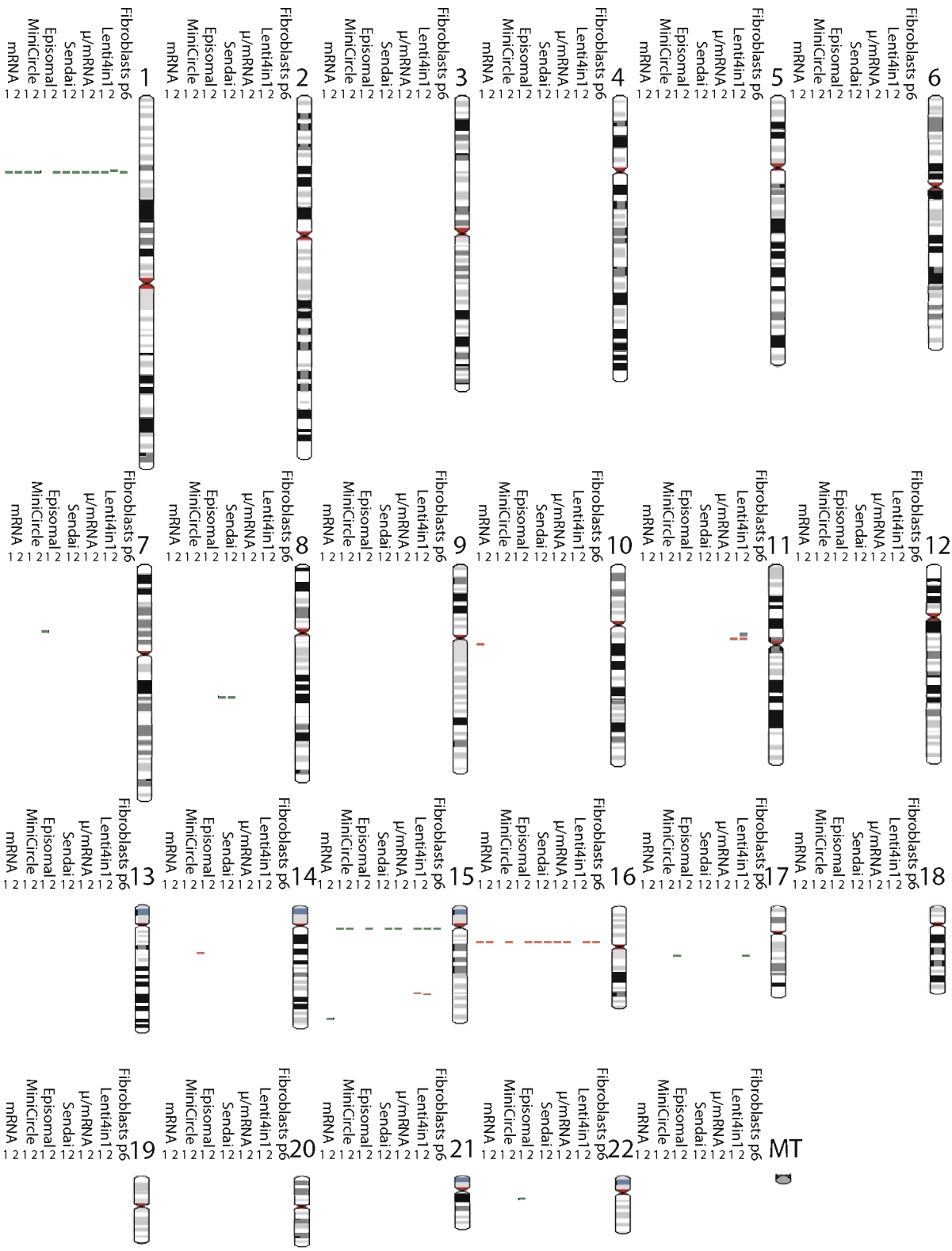
Table of content **page**



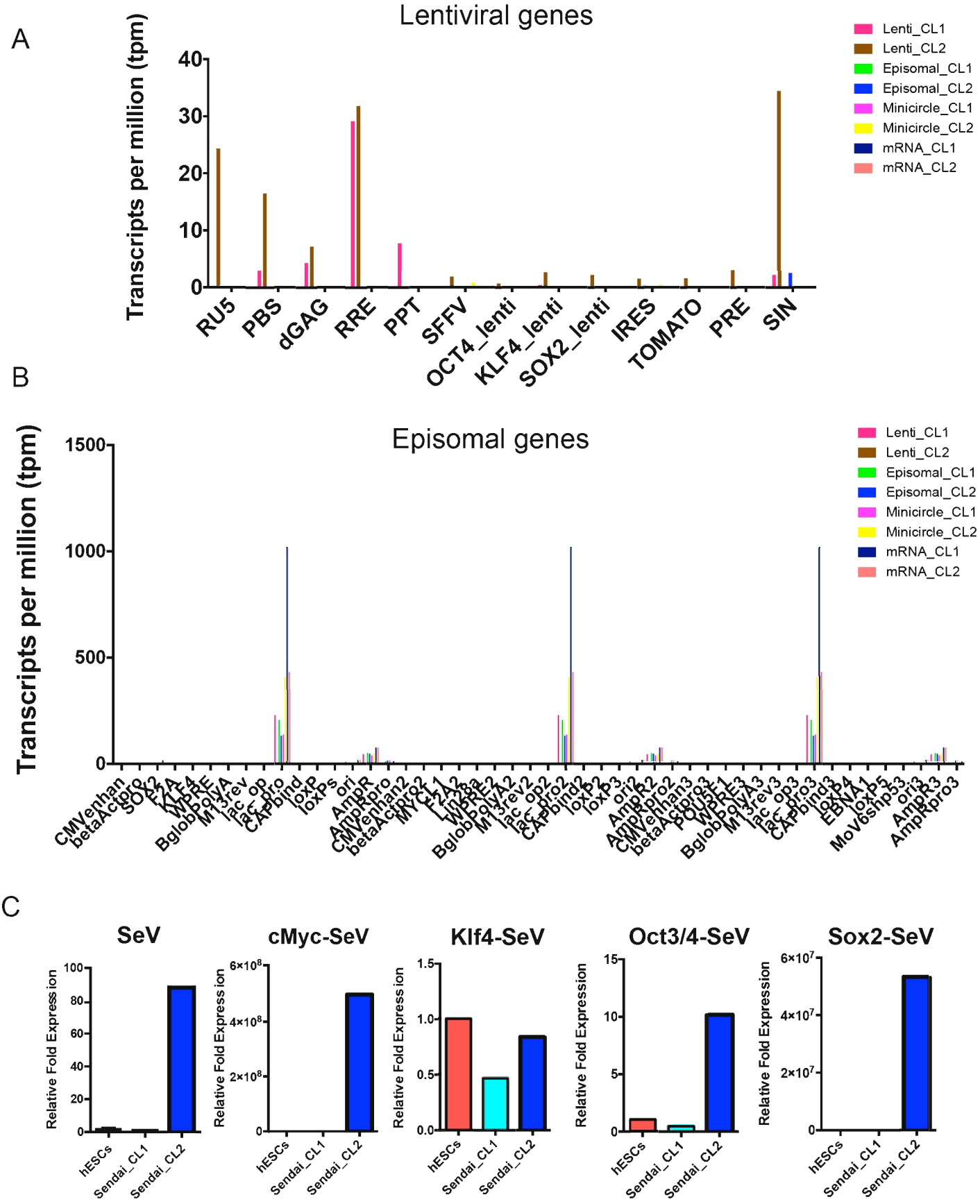
Supplementary Figure 1. Immunofluorescent labeling of stem cell markers. Representative images demonstrating the expression of hiPSC markers by immunofluorescent labeling of hiPSC colonies for Oct4, Nanog, Sox2 in red, Tra-1-81 and Tra1-60 in green, and nuclei labeled with DAPI in blue. All hiPSC lines were positive for these markers.



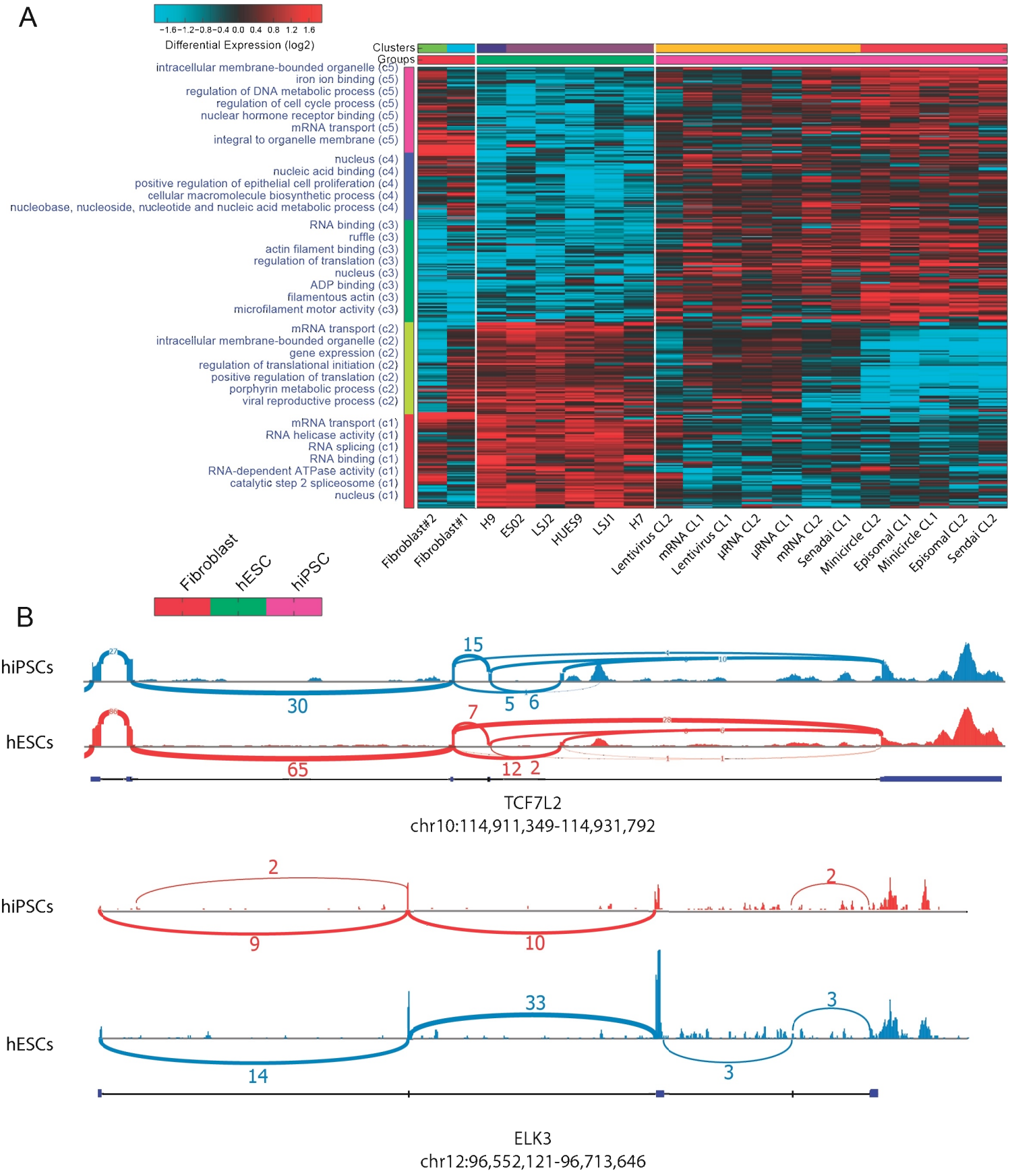
Supplementary Figure 2. Teratoma formation characterization of hiPSC lines. hiPSCs were injected into the hindlimb of NOD/SCID mice and were extracted approximately 45 days later. Extracted teratomas were fixed in 3.7% paraformaldehyde overnight, tissue processed and stained for hematoxylin and eosin (H&E). Representative staining demonstrated that each method of reprogramming (lentivirus, Sendai virus, minicircle, mRNA, episomal, and µRNA/mRNA) generated teratomas containing each germ layer. a = endoderm derivatives, b = ectoderm derivatives, and c = mesoderm derivatives.



Supplementary Figure 3. SNP karyotype. SNP karyotyping was performed on each cell line (12 hiPSC lines and the parental fibroblast line) using the HuCytoSNP-12 chip. No major karyotypic abnormalities were present. Green bars represent an overrepresentation and red bars represent an underrepresentation of genetic material found in some chromosomes.



Supplementary Figure 4. Transgene expression quantification within lentiviral and episomal lines. (A) RNA-seq reads from Lentiviral (Lenti_CL1, Lenti_CL2), Episomal (Episomal_CL1, Episomal_CL2), Minicircle (Minicircle_CL1, Minicircle_CL2) and mRNA (mRNA_CL1, mRNA_CL2) lines were mapped to either the Lentiviral or (B) Episomal vector constructs. In both Lentivrial clones, lentiviral transgene expression could be observed however no specific expression could be observed within the episomal lines when mapped to the episomal specific sequences. (C) Sendai virus could be detected in the Sendai_CL2 hiPSC line but not the Sendai_CL1 at passage 12 by real-time PCR (using hESC lines as control). Sendai_CL2 also expressed high levels of cMyc-SeV, Oct3-SeV, and Sox2-SeV Sendai viral specific genes.

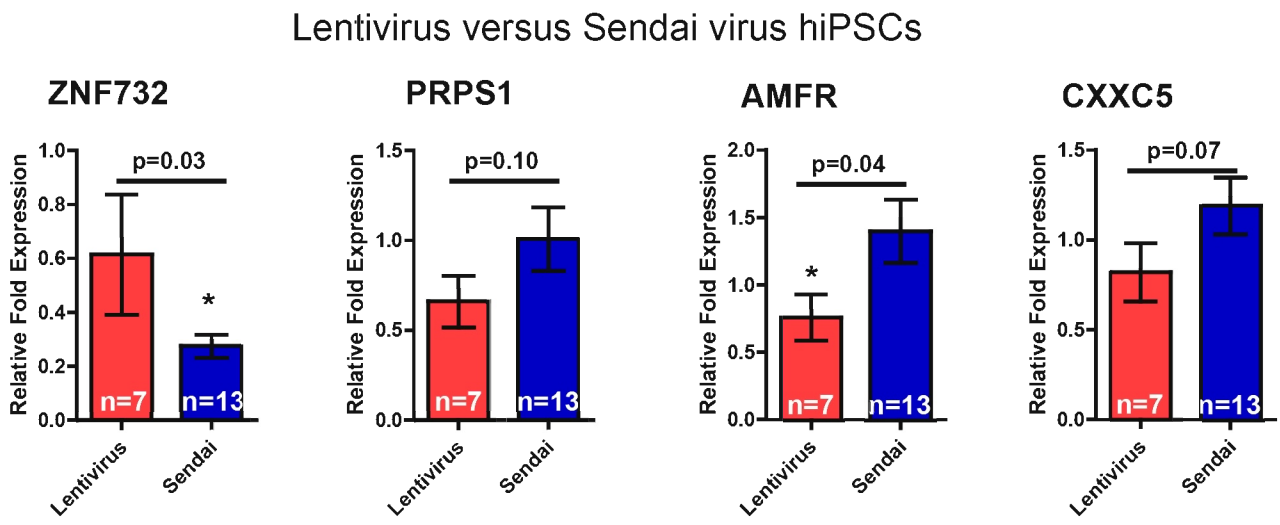


Supplementary Figure 5. Alternative splicing analysis of fibroblasts, hESCs and hiPSCs. hiPSC specific events comprise 1/4th of alternative exon differences compared to fibroblasts. These included a number of transcription factors (n=14), including the WNT signaling regulators TCFL2 (predict gain of microRNA binding site in iPSC) and RBPJ (induced expression of a full-length isoform in iPSC). All together, these genes were enriched in polycomb repressive complex (PRC) targets via EED. In addition, targets associated with mRNA processing and RNA splicing, regulation of cell cycle, DNA recombination and repair, are validated targets of C-MYC transcriptional activation.

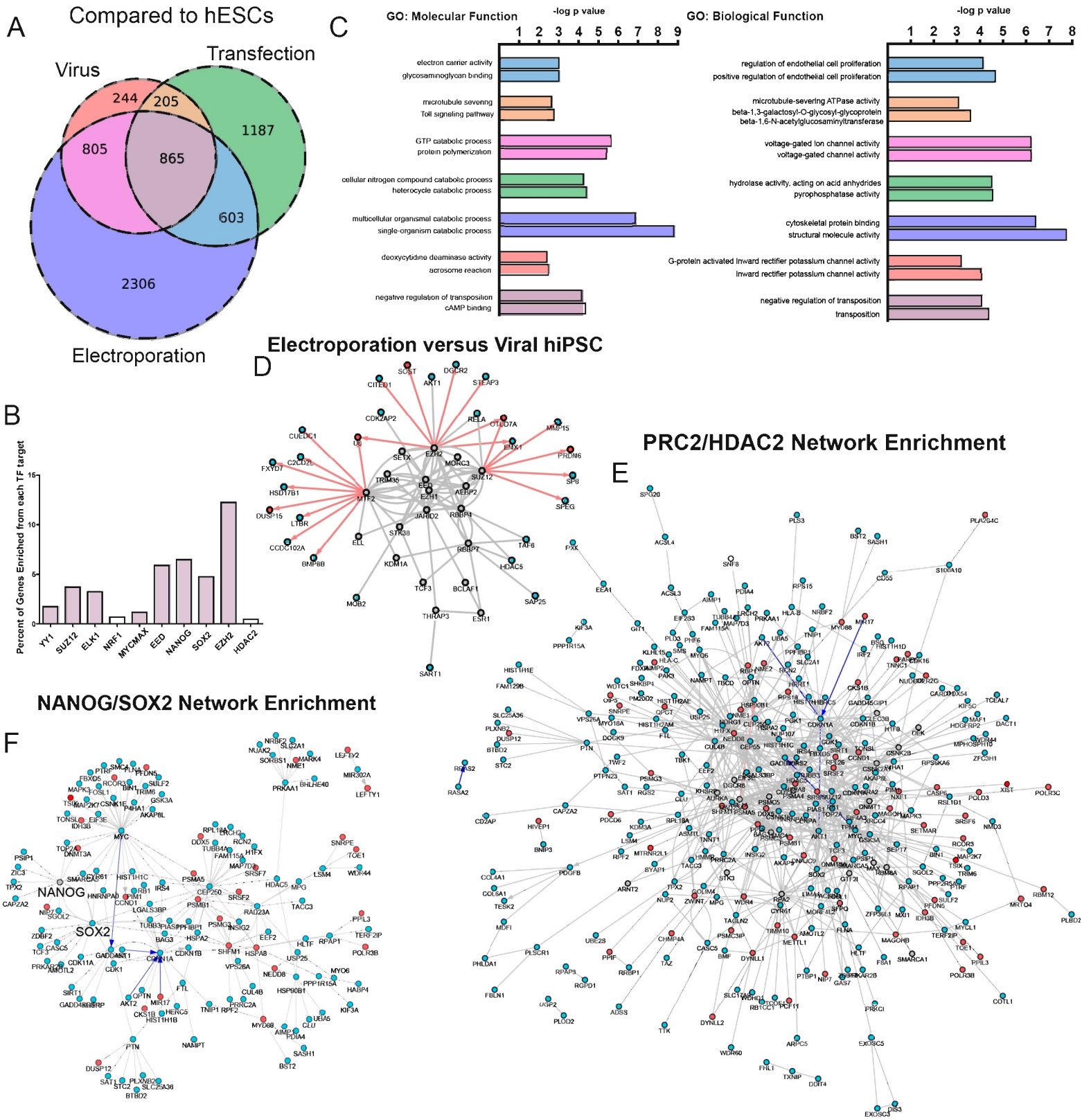
A

Line ID	Passage	Age	Sex	Source	Reprogramming Method
AH018	p47	47	Male	Fibroblasts	Lentivirus
BA037	p12	60	Female	Fibroblasts	Lentivirus
BA057	p8	47	Female	Fibroblasts	Lentivirus
AH022	p14	57	Female	Fibroblasts	Lentivirus
48	p11	60	Female	Fibroblasts	Sendai
149	p11	47	Female	Fibroblasts	Sendai
160	p11	57	Female	Fibroblasts	Sendai
161	p11	49	Male	Fibroblasts	Sendai
23	p12	N/A	Female	Fibroblasts	Sendai
273	p11	42	Female	PBMCs	Sendai
274	p11	31	Female	PBMCs	Sendai
116	p12	33	Male	PBMCs	Sendai
275	p11	72	Female	PBMCs	Sendai
264	p9	3	Female	Fibroblasts	Sendai
267	p11	73	Female	Fibroblasts	Sendai
269	p11	52	Female	Fibroblasts	Sendai
268	p11	51	Female	Fibroblasts	Sendai

B

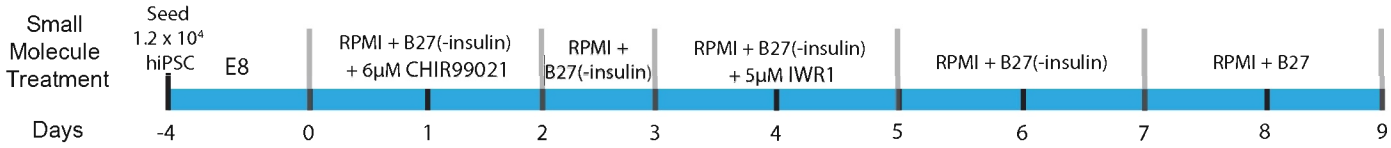


Supplementary Figure 6. PCR quantification of genes called as significantly different between an independent set of Sendai and lentiviral derived lines. (A) Seventeen additional hiPSC lines were obtained representing two reprogramming methods (Sendai (13 lines) or lentiviral (4 lines)) derived from either peripheral blood mononuclear cells (PBMCs) or fibroblasts. (B) Real-time PCR assessment of the relative fold expression of four genes (ZNF732, PRPS1, AMFR, CXXC5) identified to be different between lentiviral and Sendai viral hiPSC lines. (n is total number of lines assessed, p represents t-test p value, * indicates $p < 0.05$).

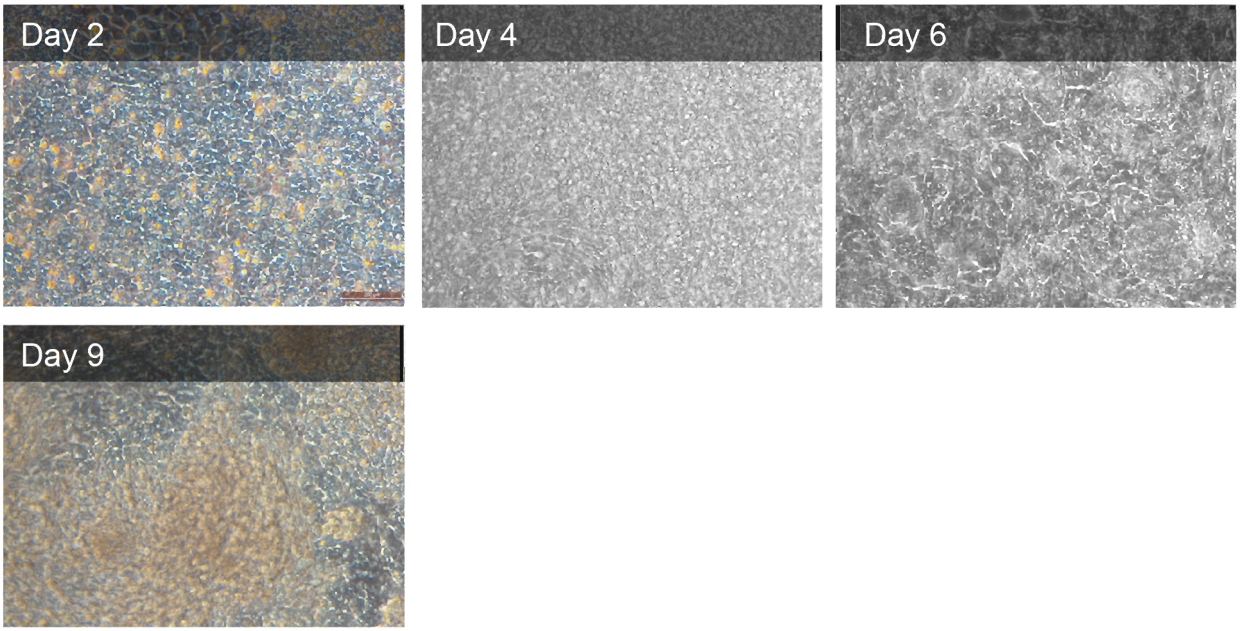


Supplementary Figure 7. Transcription factor targets and GO terms associated with each type of reprogramming. (A) A weighted Venn diagram was used to demonstrate significant gene differences between hESCs and the type of reprogramming (electroporation, viral, or transfection based). The numbers represented in each area of Venn diagram correspond to the number of genes found within each comparison and all bars were color coded to match the gene lists within each comparison (i.e., light purple represent 865 genes that were commonly called as significantly different by each type of reprogramming method to hESCs). (B) The associated transcription factor target representation (as the percent of genes called within each TF) and (C) GO terms correspond to the color of the different unions and intersections from the Venn diagram. (D) Network visualization of the transcriptional targets of PRC2 genes (from WikiPathways WP2916) between electroporation versus viral derived hiPSCs (n=1,177) demonstrates a downregulation of PRC2 direct target genes in the electroporated hiPSCs. Blue = downregulation, red = upregulation. (E,F) Network visualization of the transcriptional targets within NANOG/SOX2 and PRC2/HDAC2 network significantly different between hiPSC lines (red- higher in hiPSCs) when compared to hESCs cells (blue- higher in hESCs).

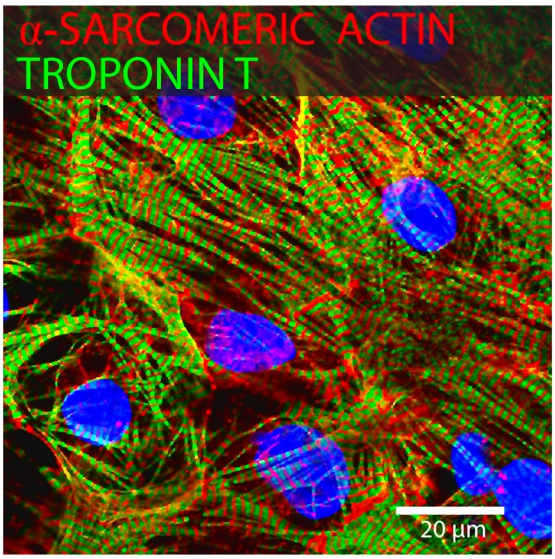
A



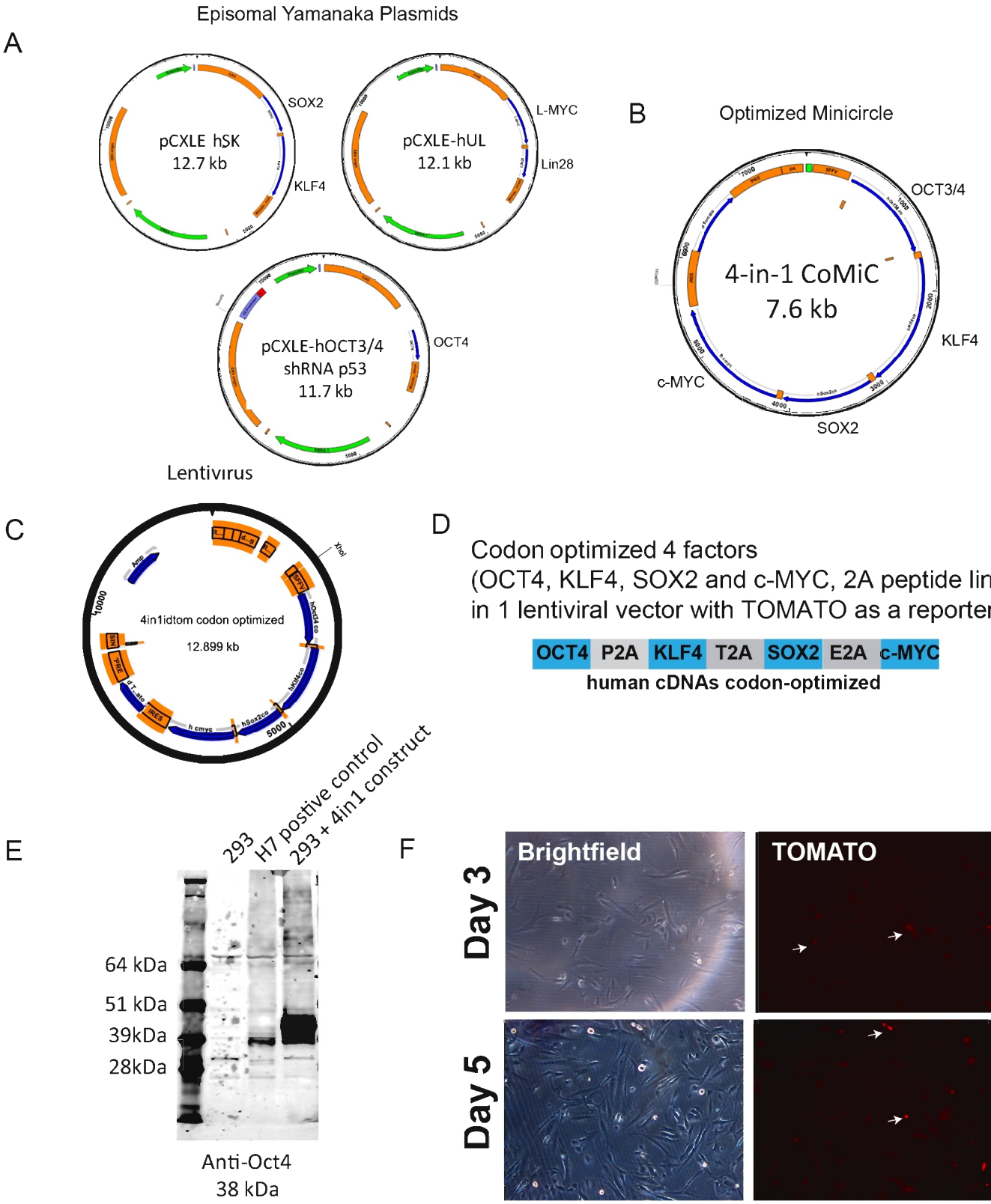
B



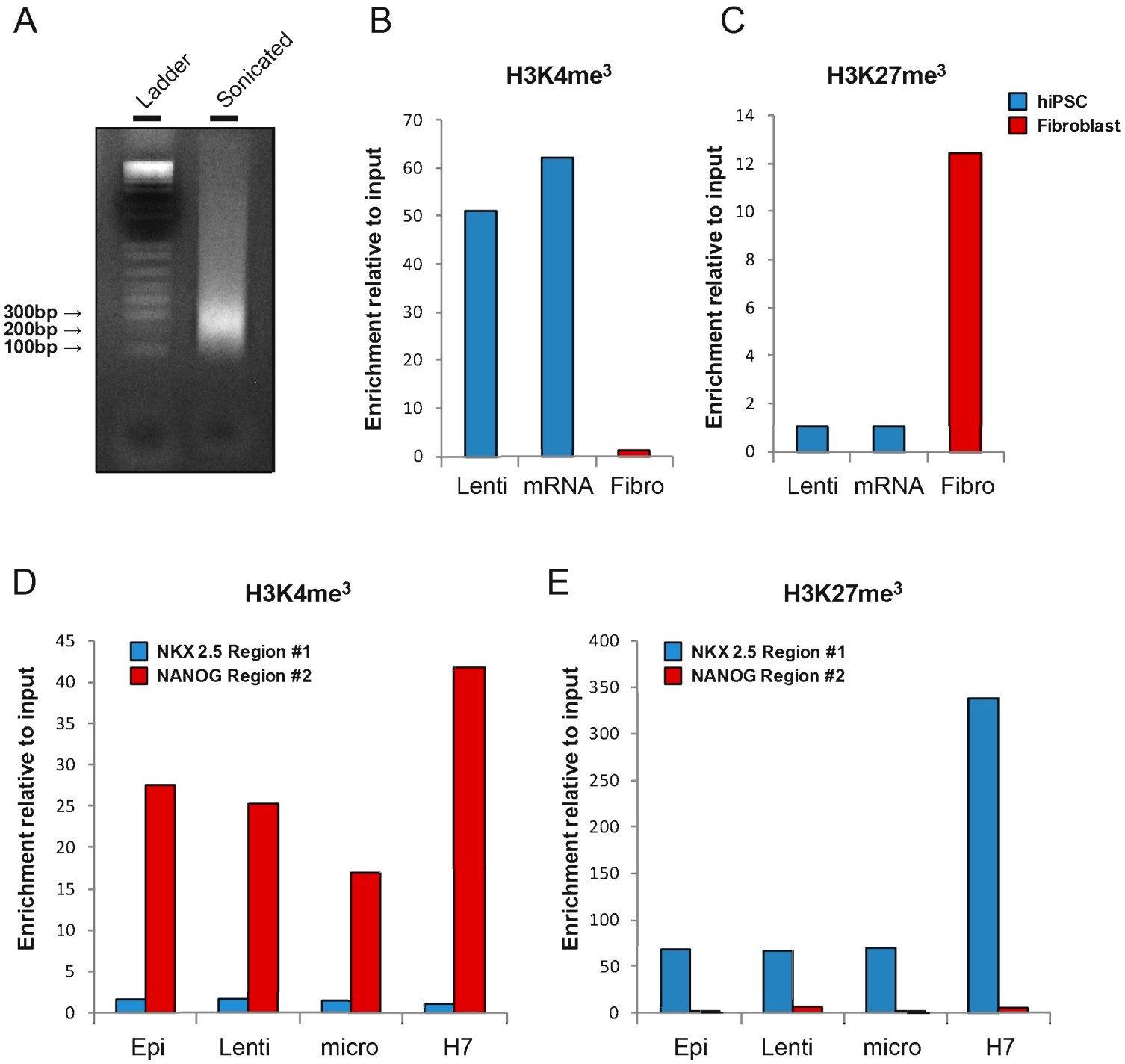
C



Supplementary Figure 8. Cardiomyocyte differentiation from hiPSC lines. (A) Cardiomyocyte differentiation was performed using the outlined treatment schedule. (B) Brightfield images demonstrating the progression of cardiomyocyte differentiation during days 2, 4, 6, 8 and 9 of cardiomyocyte differentiation. (C) Confocal fluorescent image of cardiomyocytes derived from hiPSCs (episomal line) labeled with cardiac specific troponin-T and alpha sarcomeric actin.



Supplementary Figure 9. Plasmids used for episomal, minicircle, and lentiviral hiPSC line generation. (A) Three episomal plasmids were used (pCXLE-hSK, pCXLE-hUL, and pCXLE-hOCT3/4 shRNA p53) containing genes encoding SOX2, KLF4, L-MYC, Lin28, POU5F1 (Oct4), as well as a shRNA for p53 to derive the episomal hiPSC lines. (B) An optimized minicircle was created which contains codon optimized versions of the POU5F1, KLF4, c-MYC, and SOX2. (C) Lentiviral lines were created from a vector also containing codon optimized version of the reprogramming factors called 4-in-1 lentiviral vector. (D) These factors were constructed in the order of POU5F1, KLF4, SOX2, and c-MYC separated by P2A, T2A, and E2A linker DNA. (E) To confirm the expression of the reprogramming factors, Western blots were performed on H7 cells (positive control), 293 cells without the transfected 4-in-1 lentiviral vector, as well as 293 cells transfected with the 4-in-1 lentiviral vector. Western blots demonstrated high expression of Oct4 after transfection with the 4-in-1 lentiviral vector. (F) 4-in-1 lentiviral vector transfection efficiency demonstrated by brightfield and fluorescent (TOMATO (red)) imaging of fibroblasts at three and five days after transfection with the 4-in-1 lentiviral vector.



Supplementary Figure 10. Validation of chromatin immunoprecipitated DNA. (A) Crosslinked DNA was sheared by sonication and separated by gel electroporation to an average size of 100-300 base pairs. The primers (5'-CATTATAACTTGGTGGCACCTG-3'; 5'- GGTTCTCAAACTTACTG CACAC-3') for the promoter of NANOG were used for ChIP-qPCR on sample immunoprecipitated with histone markers H3K4me3 (B) and H3K27me3 (C) in hiPSCs (lentiviral derived lines and mRNA derived lines) and fibroblast. Quantification of the histone markers H3K4me3 (D) and H3K27me3 (E) by ChIP-qPCR was performed for the promoter of NKX2-5 and NANOG using the primers for NKX2-5 [5'-GTAAGAGCGGCTTGACCTAC-3'; 5'-CTGACTTTCTACAAATGCTGAGTG-3'] and NANOG (same as above) in hiPSC (episomal, lentivirus, microRNA) and H7 hESC lines. Both hiPSC and H7 hESC lines expressed high levels of the active H3K4me3 mark at the NANOG promoter as well as high levels of the inactive H3K27me3 for the NKX2-5 promoter.