Supplemental Materials and Methods

Cell culture and maintenance of iPSCs
Human fibroblast lines CHOPWT1 (Coriell GM00969), CHOPWT2 (Coriell GM05565), and BMC (ATCC CRL-2097) were cultured in IMDM (Mediatech) containing 10% FBS at 37°C and 5% CO2 in a humidified incubator. The iPSCs were maintained in HES media: DMEM/F12 containing 2 mM of glutamine, 20% Knockout Serum Replacement (KO), 1X NEAA, 0.1 mM β-mercaptoethanol, and 10 ng/ml of bFGF (R&D Systems, Minneapolis, MN) on 0.1% gelatin (Sigma, Inc.) coated plates with irradiated mouse embryonic fibroblast (MEFs).

Lentivirus production and infection
STEMCCA lentiviral vector production and infection followed the previously published protocol[1]. The STEMCCA viral vector used for Coriell fibroblast GM00969 and GM0565 reprogramming encoded OCT4, KLF4, SOX2, and mCherry cDNAs. The STEMCCA viral vector used for ATCC CRL-2097 reprogramming encoded OCT4, KLF4, SOX2, and cMYC cDNAs. Human fibroblasts were seeded at approximately 3 x 10^4 cells per well of a 6-well plate and incubated with (5-10 μl) STEMCCA virus plus 5 ug/ml of polybrene in fibroblast medium with a change in media 24h after infection. The medium was changed to IMDM with 10% FBS, 1x NEAA, 1x 1-glutamine, 1x-P/S, 20 ug/ml of ascorbic acid, and 4 ng/ml bFGF. GSK3-beta (EMD Biosciences; 5 uM) was added to the culture media on days 2-6 of reprogramming. Six days after infection, cells were replated on 100-mm plates containing MEFs. Cells were replated in IMDM media plus Thiazovivin (Stemgent, Boston, MA, 0.5 uM). The cells were maintained in IMDM with 10% FBS, 1x NEAA, 1x 1-glutamine, 1x-P/S, 50 ug/ml of ascorbic acid, and 4 ng/ml bFGF for 4-days post-splitting. Medium was changed to HES media on day 10 of reprogramming, and medium changes were performed every 2-3 days until day 28.

Initial iPSC expansion and characterization
To establish iPSC lines, iPSC clones were handpicked about 1-month post-infection based on cell morphology and size. Clones were subsequently passaged and expanded in HES medium for a minimum of 20 passages to erase residual epigenetic memory[2-3] Characterization of the clones included pluripotent ESC markers and qPCR analysis: DMNT3B, ABCG2, REX1, OCT4, SOX2, NANOG, cMYC, KLF4 gene expression. Flow cytometry was utilized to identify pluripotency antigens: SSEA3/4, TRA-1-81 and TRA-1-60.

RNA isolation and real-time quantitative PCR analysis
Total RNA and miRNA were isolated from cells using the miRNeasy kit (Qiagen, Inc), with the optional column RNAse-free DNase treatment, according to the manufacturer's instructions. The cDNA was produced using random hexamers with Superscript III Reverse Transcriptase (Invitrogen), from 200 ng to 1 ug of starting RNA. Human gDNA was used for the qPCR standard curve with a 10-fold gDNA diution series, ranging from 0.1 to 10 ng per reaction to evaluate the efficiency of the PCR and calculate copy number of each gene relative to the housekeeping gene Cyclophilin. The expression level was expressed as number of molecules of RNA for each indicated gene per number of molecules of cyclophilin, following the protocol previously published.[1] Real-time quantitative PCR reactions were set up in

triplicate with the SYBR Green QPCR Master Mix (Roche, Indianapolis, IN, http://www.roche.com) and run on a Light Cycler 480II qPCR System (Roche, Indianapolis, IN, http://www.roche.com). Primer sequences are: CYCLOPHILIN: F' GAA GAG TGC GAT CAA GAA CCC ATG AC, R' GTC TCT CCT CCT TCT CCT CCT ATC TTT ACT T; DNMT3B: F' TAC AGA CGT GTG CAG TTG TAG GCA, R' GTG CAG ACT CCA GCC CTT GTA TTT; REX1: AAA GCA TCT CCT CAT TCA TGG T, R' TGG GCT TTC AGG TTA TTT GAC T; ABCG2: TCA GGA GAC CAC ATT TCA TCT AGC CC, R' CAG GGC ACC CAC TGA CAA ACT AAA; NANOG: F' CCT GAA GAC GTG TGA AGA TGA G, R' GCT GAT TAG GCT CCA ACC ATA C; OCT4: AAC CTG GAG TTT GTG CCA GGG TTT, R' TGA ACT TCA CCT TCC CTC CAA CCA, SOX2: AGA AGA GGA GAG AGA AAG AAA GGG AGA GA, R' GAG AGA GGC AAA CTG GAA TCA GGA TCA AA; KLF4: GAG GGA AGA CCA GAA TTC CCT TGA, R' AGA ACC AAG ACT CAC CAA GCA CCA; cMYC; F' ATG CAA CCT CAC AAC CTT GGC TGA, R' GCC CAA AGT CCA ATT TGA GGC AGT.

Southern blot analysis

Genomic DNA (10 ug, for hSTEMCCA vector Southern blot) was digested with BamHI, separated on a 0.8% agarose gel and blotted onto Hybond-N+ membrane (Amersham Biosciences). Membrane was probed using WRPE element as previously published[16], and labeling was performed using ^{32}P-α-dCTP with the High Prime Random Labeling (Roche, Indianapolis, IN) following the manufacturer's instructions.

Teratoma formation

For teratoma formation, iPSCs were plated on a Matrigel-coated dish, and a minimum of 10^6 cells were injected subcutaneously into the neck of a Fox Chase SCID Beige mouse. Mice were sacrificed 6-9 weeks later, teratomas isolated and processed for histological analysis. The images were captured by the Zeiss Axioskip2 microscope (Munich, Germany) (Original magnification x20-40).

Cre-mediated hSTEMCCA Excision

Transgene removal was performed as previously published[1]. Briefly, iPSCs were plated on a matrigel-coated dish containing puromycin resistant irradiated MEFs. Cells were transfected with 2 ug/well of pHAGE2-Cre-IRES-PuroR plasmid DNA using Hela Monster transfection reagent (Mirus, Madison, Wi) according to manufacturer's instructions. Medium was changed to puromycin (1.2 ug/ml) selection medium 24 hours post-transfection for 48 hours. The re-emergence of iPSCs was observed between 7-10 days post-puromycin selection, and a total of 12 colonies were picked when colonies reached between 50-100 cells per cluster. Genomic DNA was collected using DNeasy Blood & Tissue kit (Qiagen, Valencia, CA) from each subclone after 2-3 weeks in culture and screened for hSTEMCCA transgene using the following primers and conditions: OCT4-Tg-F: 5'-GGT GCG CCA GTA AAG CAG ACA TTA AAC 3';KLF4-Tg-R: 5'-CAG ACG CGA ACG TGG AGA AAG A-3' and GAPDH-F: 5'-GTG GAC CTG ACC TGC CGT CT-3'; GAPDH-R: 5'- GGA GGA GTG GGT GTC GCT GT-3'; 95°C for 10 minutes; followed by 30-cycles of 95°C for 30 seconds, and 60°C for 1 minute.

Cell Sorting and Genetic analysis

IPSCs and ESCs were grown to 85-90% confluence on MEFs plated on 1:6 Matrigel (BD Biosciences). In order to limit variability across culturing conditions, all cells were grown in

identical conditions; gelatin, matrigel, bFGF, FACS sorting antibodies, TrypLE, DMEM/F12 (50:50), and KOSR lots were fixed. Cells were disassociated with accutase for 4-min to remove MEFs, followed by an additional 10 min at RT to establish a single cell suspension. Cell counts and viability were assessed for all lines (>96%, respectively). Cells were pelleted and resuspended at $5x10^6$ cells/ml in staining medium (IMDM + 5% Knockout S/R + DNase). Cells were stained with SSEA3-FITC (1:100; Biolegend) and Tra-1-81-APC (1:100; Biolegend) antibodies for 30-min on ice. Cells were washed 2x and passed through a 70 uM filter to remove clumps. Cells were sorted (FACS Aria II, Becton Dickenson), based on SSEA3-positive and Tra-1-81-positive expression (Figure 1A). The cells were resuspended in Qiazol and snap frozen for global gene array and miRNA analyses. All samples were stored in -80°C until further analysis.

Gene Expression Analysis

Three biological replicates were prepared for all iPSC and hESC-derived samples. Quantified RNA samples were hybridized to Affymetrix GeneChip Human Gene 1.0 ST arrays. The cells were analyzed for mRNA expression using Affymetrix GeneChip Human Gene 1.0 ST arrays. All gene array data obtained from Affymetrix passed quality control steps and were sent to RMA normalization and log transformation by Affymetrix Expression Console. Triplicate data were averaged and analyzed by Gene Cluster 3.0[4] for principal components assay (PCA) by excluding genes with expression levels lower than 5. PCA data were visualized with TreeView. Significance Analysis of Microarrays (SAM) was carried out to identify the top 25 up-regulated and down-regulated genes with more than 1.5 fold change with significant p value ($p < 0.05$) by comparing triplicates of CHOPWT2.1 to CHOPWT2.2 and CHOPWT2.3. Gene list and fold changes are presented in Supplemental Table 1.

SNP Array Analysis

DNA was isolated from each fibroblast and iPSC line using DNeasy Blood and Tissue Kit (Qiagen, Inc), and genotyped on the Illumina BeadChip at the Center for Applied Genomics, Children's Hospital of Philadelphia. The CHOPWT1 and CHOPWT2 fibroblast and iPSC lines were genotyped on the HumanHap610-Quad arrays with 620,901 markers, whereas the BMC fibroblast and iPSC lines were genotyped on HumanOmni1-Quad arrays with 1,140,419 markers. CNV analysis was performed using the PennCNV software[5] based on the signal intensity (Log R Ratio and B Allele Frequency) data exported from Illumina GenomeStudio. The default parameters were used, and only large CNVs (>200kb) were retained for subsequent analysis. The signal intensity data around all CNV calls were plotted and visually inspected to ensure accuracy. We used the scan_region.pl program within PennCNV to determine if a CNV call is unique to an iPSC line, in comparison to the parental fibroblast as a reference.

iPSC differentiations into hematopoietic progenitors cells using adherent monolayer cultures

For hematopoietic differentiation, between 5 – 10 x 10^4 iPSCs were plated on matrigel (BD Biosciences) coated 6-well plates and cultured for 1-2 days in HES media supplemented with 5 ng/ml bFGF. Upon reaching ~50% confluence, differentiation was started by adding RPMI media P/S, glutamine, ascorbic acid & monothioglycerol (MTG) including BMP4 (5 ng/ml), Wnt3a (2 mM) and VEGF (50 ng/ml) for the first 2 days. Cells were maintained one additional day in RPMI media plus BMP4 (5 ng/ml), VEGF (50 ng/ml), and bFGF (20 ng/ml). On day 4,

RPMI media was replaced with StemPro34 (SP34) (Invitrogen) plus BMP4 (5 ng/ml), VEGF (50 ng/ml), and bFGF (20 ng/ml). From day 4-5, SP34 media was supplemented with VEGF (15 ng/ml) and bFGF (5 ng/ml). On day 6, SP34 media was replaced by serum free media (SFD; Cheng et. al 2012) plus VEGF (50 ng/ml), IL6 (10 ng/ml), SCF (50 ng/ml), bFGF (100 ng/ml), and FLT3 ligand (5 ng/ml). On day 7-9, media was replaced daily with SFD supplemented with VEGF (50 ng/ml), IL6 (10 ng/ml), SCF (50 ng/ml), bFGF (100 ng/ml), TPO (50 ng/ml), FLT3 ligand (5 ng/ml). Non-adherent common myeloid progenitors (CMPs) were collected (day 7-9) and used in methylcellulose or megacult colony assays (Stem Cell Technologies, Inc) according to manufacturer instructions. Images were captured by the Zeiss Axioskop2 microscope (Munich, Germany) (Original magnification x20). Megakaryocyte size was determined as previously published[6]. The CMPs were also expanded for 4-5 days in megakaryocyte liquid culture conditions: SFD supplemented with TPO (50 ng/ml), SCF (50 ng/ml), and IL3 (10 ng/ml).

Flow Cytyometry

Cells from differentiations were analyzed for the following cell surface antigens: CD31-PE-Cy7, KDR-APC, CD41-PE, CD235-APC, CD43-FITC, CD45-PB, CD42-PE, and CD18-APC as previously published[7] and data not shown. The cells were acquired with a FACS canto II flow cytometer (Becton Dickenson). Analysis was performed using FlowJo software (Tree Star Inc.).

Mills et al. Suppl. Table 1.

RefseqRNA	Gene Symbol	Fold Change
NM_016588	NRN1	7.26
NM_003155	STC1	6.31
NM_175063.4	C19orf63	6.12
NM_001042422.1	SLC16A3	6.01
NM_007000	UPK1A	5.91
NM_198581.2	ZC3H6	5.89
NM_199286	DPPA3	5.79
NM_003529	HIST1H3A	5.46
NM_001828	CLC	5.30
BC005107	C21orf105	5.20
NM_005840	SPRY3	5.04
NM_001172	BNIP3	4.53
NM_020358	TRIM49	4.51
NM_001678	ATP1B2	4.48
NC_000011.9	LOC692247	4.28
NM_002522	NPTX1	4.13
NM_138492.4	PRELID2	4.02
NC_000017.10	SNORA38B	3.75
NM_014367.3	FAM162A	3.64
NM_021623.1	PLEKHA2	3.57
NM_144670.2	A2ML1	3.55
NM_007261	CD300A	3.32
NM_020358	LOC729384	3.32
NM_001039	SCNN1G	3.30
NM_001159296.1	KLF8	3.26
NM_020225.1	STOX2	-2.18
NM_018837	SULF2	-2.18
NM_001838	CCR7	-2.19
NM_000610	CD44	-2.22
NM_022068.2	FAM38B2	-2.24
NM_002442	MSI1	-2.24
NM_021977	SLC22A3	-2.25
NM_002167.4	ID3	-2.25
NM_017780.2	CHD7	-2.26
NM_001048251.1	CMTM3	-2.27
NM_000965	RARB	-2.28
NM_001082967.1	FAM19A5	-2.28
NM_030753	WNT3	-2.28
NM_153377	LRIG3	-2.31
NM_014420	DKK4	-2.32
NM_001025194.1	CES1	-2.35
NM_001428	HTR1A	-2.40
NM_000930	PLAT	-2.41
NM_006537	USP3	-2.41
NM_003877	SOCS2	-2.47
NM_006183	NTS	-2.48
NM_014344.2	FJX1	-2.55
NM_016591	GCNT4	-2.57
NM_014899	RHOBTB3	-2.57
NM_178826.3	ANO4	-2.94

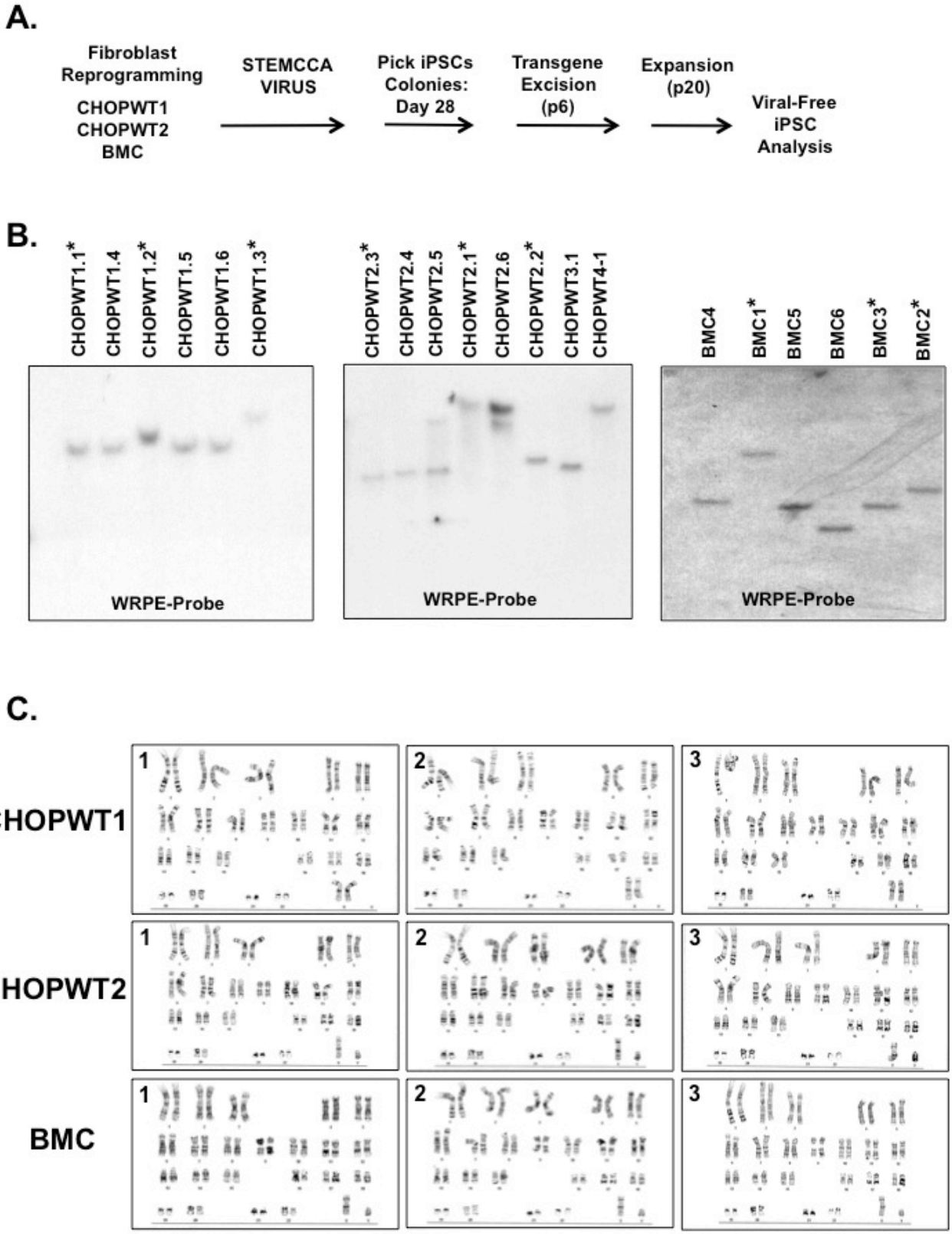
Table S1. Genes differentially expressed in WT2.1 versus WT2.2 and WT2.3 with a 1.5-fold or greater difference in expression.

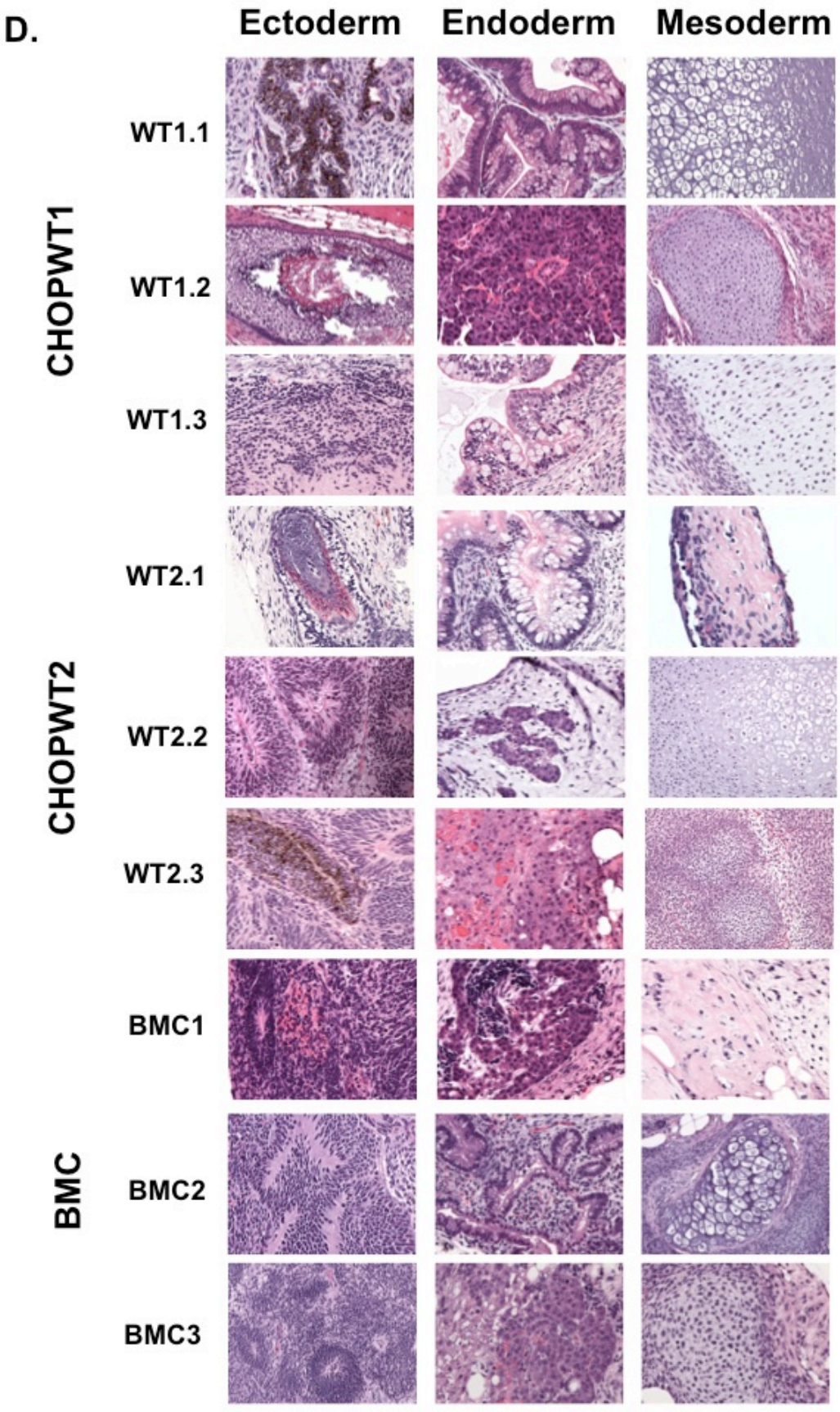
Mills et al. Supplementary Table 2

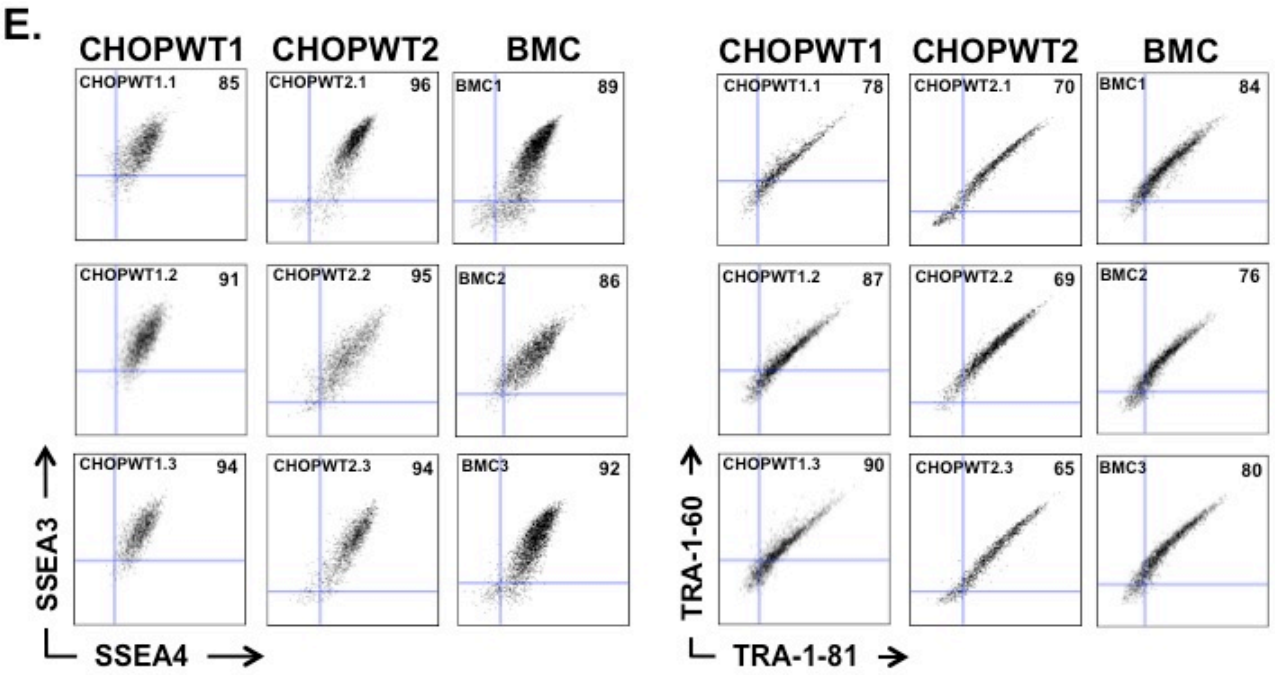
iPSC Line	Chromosome	Region	Size	Type	No. of Genes	Genes
CHOPWT1.1	hg7(q11.22)	69817650-70852210	760KB	Duplication	2	Auts2, WBSCR17 (MIR3914-1, MIR3914-2)
CHOPWT1.2	hg7(q11.22)	69817650-70852210	760KB	Duplication	2	Auts2, WBSCR17 (MIR3914-1, MIR3914-2)
CHOPWT1.3	hg5 (q34)	167069407-167222548	200KB	Duplication	1	ODZ2
	hg3 (q26.32)	175046386-176164717	1.3MB	Duplication	1	NAALADL2
CHOPWT2.1	hg3(q26.32)	176682972-176892871	200KB	Deletion	1	TBL1XR1
	hg6(q21)	111242190-112175075	1MB	Duplication	10	GTF3C6, RPF2, GSTM2P1, SLC16A10, KIAA1919, REV3L, TRAF3IP2-AS1, AF216583, TRAF3IP2, FYN
	hg16(q23.3)	82329444-82701613	372KB	Duplication	8	CDH13, HSBP1, MLYCD, OSGIN1, NECAB2, EFCBP2, SLC38A8, MBTPS1
	hg20 (q11.1-q11.21)	29311279-30545148	1.2MB	Duplication	29	FRG1B, MLLT10P1, DEFB115, DEFB116, DEFB118, DEFB119, DEFB121, DEFB121,DEFB122, DEFB123, DEFB124, BC016143, REM1, LINC00028, TRNA_Pseudo, HM13, PSIMCT-1, ID1, COX4I2, BCL2L1, TPX2, MYLK2 , FOXS1, DUSP15, TTLL9, U1, (MIR3193)
CHOPWT2.2		No Changes Detected				**None**
CHOPWT2.3	hg6(q21)	110835963-111057837	200KB	Duplication	3	DDO, SLC22A16, CDC2L6
	hg6(q21)	111442009-112506443	1MB	Duplication	11	SLC16A10, KIAA1919, REV3L, TRAF3IP2-AS1, AF216583, TRAF3IP2, FYN, WISP3, LAMA4,TUBE1, C6orf225
	hg14 (q21.3)	50311239-50656217	346KB	Duplication	7	NEMF, ARF6, AK092589, C14orf182, C14orf183, METTL21D, SOS2
BMC1		No Changes Detected				**None**
BMC2		No Changes Detected				**None**
BMC3	hg3 (q24.1)	29557812-29765655	200KB	Deletion	1	RBMS3

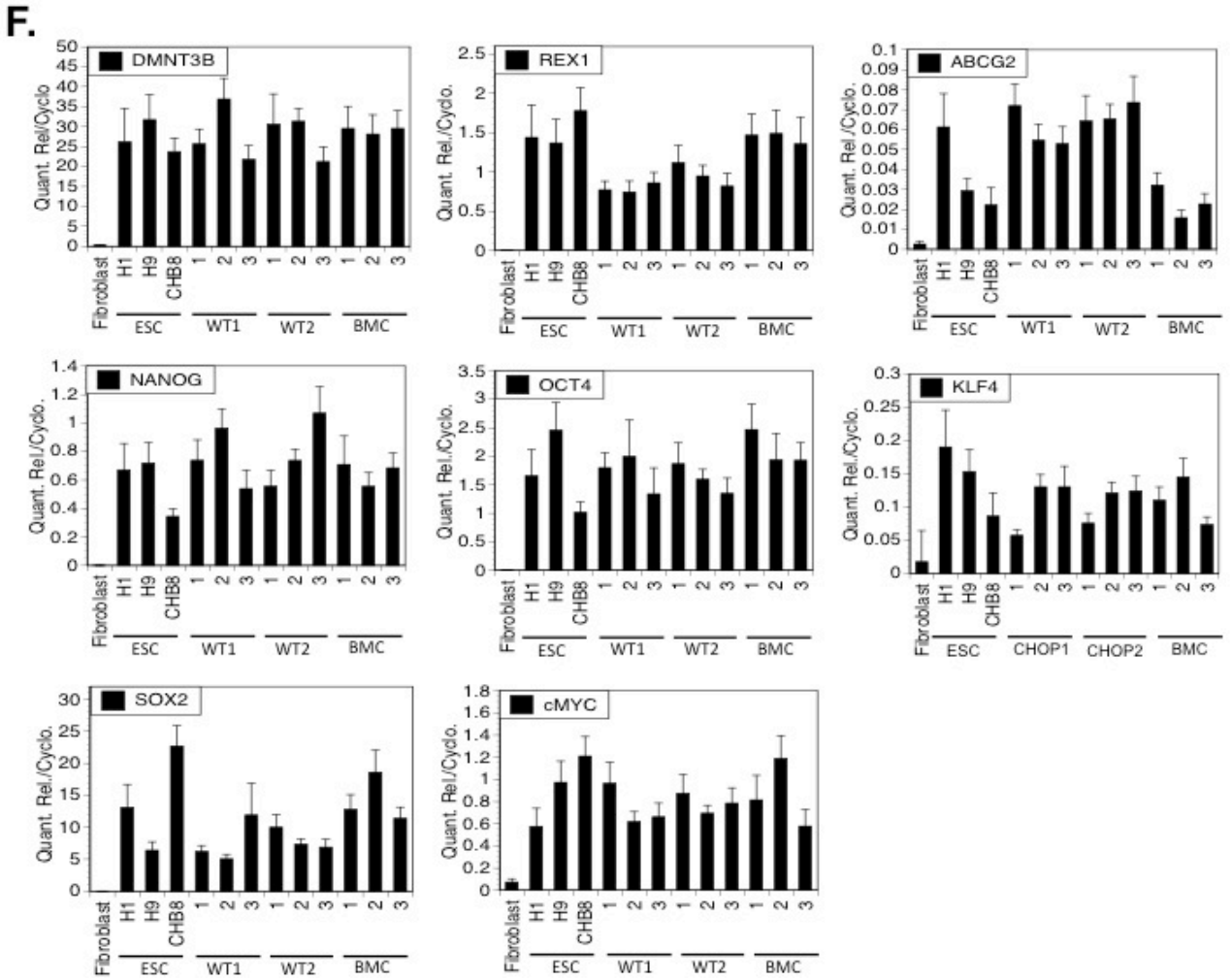
*CNVs were determined using parental fibroblast as reference to determine the number of de novo generated subchromosomal aberrations acquired during the reprogramming process.

Table S2. DNA copy number variations (CNVs) in iPSC Lines.


A.
Fibroblast Reprogramming
CHOPWT1
CHOPWT2
BMC
STEMCCA VIRUS
Pick iPSCs Colonies: Day 28
Transgene Excision (p6)
Expansion (p20)
Viral-Free iPSC Analysis
B.
CHOPWT1.1*
CHOPWT1.4
CHOPWT1.2*
CHOPWT1.5
CHOPWT1.6
CHOPWT1.3*
WRPE-Probe
CHOPWT2.3*
CHOPWT2.4
CHOPWT2.5
CHOPWT2.1*
CHOPWT2.6
CHOPWT2.2*
CHOPWT3.1
CHOPWT4-1
WRPE-Probe
BMC4
BMC1*
BMC5
BMC6
BMC3*
BMC2*
WRPE-Probe
C.
CHOPWT1
CHOPWT2
BMC
1
2
3


D.
Ectoderm
Endoderm
Mesoderm
CHOPWT1
WT1.1
WT1.2
WT1.3
CHOPWT2
WT2.1
WT2.2
WT2.3
BMC
BMC1
BMC2
BMC3


E.
CHOPWT1
CHOPWT2
BMC
CHOPWT1
CHOPWT2
BMC
CHOPWT1.1 85
CHOPWT2.1 96
BMC1 89
CHOPWT1.1 78
CHOPWT2.1 70
BMC1 84
CHOPWT1.2 91
CHOPWT2.2 95
BMC2 86
CHOPWT1.2 87
CHOPWT2.2 69
BMC2 76
CHOPWT1.3 94
CHOPWT2.3 94
BMC3 92
CHOPWT1.3 90
CHOPWT2.3 65
BMC3 80
SSEA3
SSEA4
TRA-1-60
TRA-1-81


F.
DMNT3B
REX1
ABCG2
NANOG
OCT4
KLF4
SOX2
cMYC
Quant. Rel./Cyclo.
Fibroblast
H1
H9
CHB8
ESC
WT1
WT2
BMC
CHOP1
CHOP2

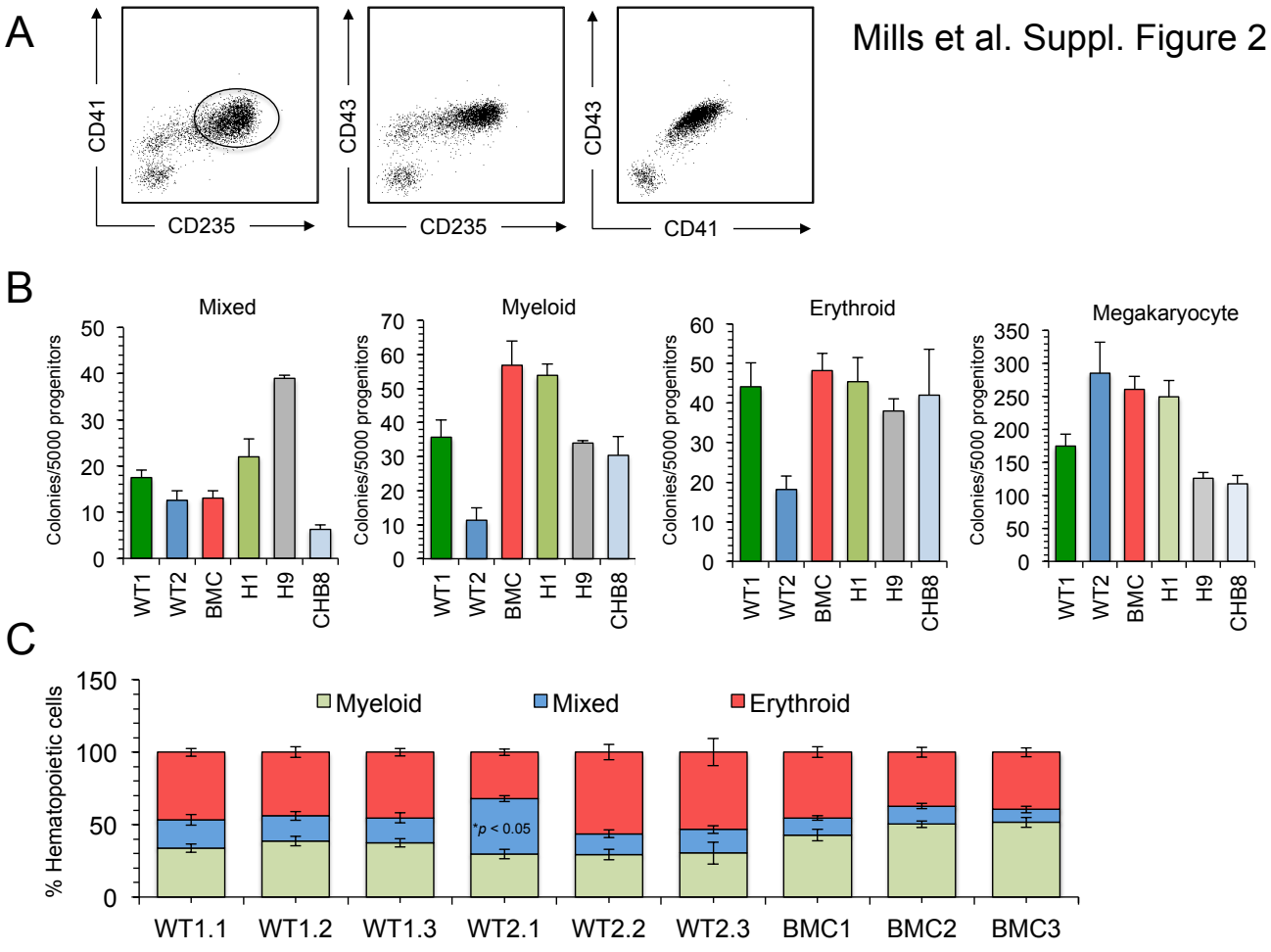

A
CD41
CD235
CD43
CD235
CD43
CD41
Mills et al. Suppl. Figure 2
B
Mixed
Myeloid
Erythroid
Megakaryocyte
Colonies/5000 progenitors
WT1
WT2
BMC
H1
H9
CHB8
C
% Hematopoietic cells
Myeloid
Mixed
Erythroid
*p < 0.05
WT1.1
WT1.2
WT1.3
WT2.1
WT2.2
WT2.3
BMC1
BMC2
BMC3

Figure S1. Characterization of fibroblast-derived iPSC clones. (A) Schema for generation of viral-free fibroblast-derived iPSC using STEMCCA cassette. (B) Southern blot analysis of single integration of STEMCCA cassette using WRPE probe. (C) Karyotype analysis of all iPSC clones showing a normal diploid G-band analysis. (D) Generation of all three germ layers in teratoma assay: ectoderm, mesoderm, and endoderm. The images were captured by the Zeiss Axioskop2 microscope (Munich, Germany) (Original magnification x20-40). (E) Expression of pluripotent surface antigen expression, SSEA3/SSEA4 and TRA-1-60/TRA-1-81, using flow cytometry. (F) Molecular profiles of iPSC lines. Real-time PCR analysis shows the expression of common pluripotency genes: DNMT3B, REX1, ABCG2, OCT4, NANOG, SOX2, cMYC, KLF4.

Figure S2. Hematopoietic potential of iPSC lines. (A) Flow cytometry analysis showing CD43+CD41+CD235+ hematopoietic progenitors cells (HPCs) released into media from adherent culture at day 7. The absolute numbers of HPCs added to colony assays were quantified using expression of CD41+CD235+CD43+ cells. (B) Methylcellulose and MegaCult assay: Colony count from HPCs as shown in part A and examined for mixed (multi-lineage colony formation), myeloid, erythrocyte, and megakaryocyte colonies. Data shown are averaged from 3 iPSC lines per individual, and 3 human ESCs were used to allow comparison across different genetic backgrounds (iPSCs: WT1: CHOPWT1.1, CHOPWT1.2 and CHOPWT2.3; WT2: CHOPWT2.1, CHPWT2.2, CHOPWT2.3; BMC: BMC1, BMC2, and BMC3; ESCs: H1, H9, and CHB8) (mean ± SEM for n=3 (iPSC); n=2 (ESC)). (C) Methylcellulose assay: Distribution of lineage-committed cells in adherent hematopoietic cultures at day 7 of differentiation for all iPSC lines, myeloid (green), mixed (blue), and erythrocyte (red) colonies (mean ± SEM for n=3; *$p < 0.05$).

References:

1. Somers A, Jean J-C, Sommer CA, et al. Generation of Transgene-Free Lung Disease-Specific Human iPS Cells Using a Single Excisable Lentiviral Stem Cell Cassette. *Stem Cells*. 2010;28:1728–1740.

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

3. Ohi Y, Qin H, Hong C, et al. Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature cell biology*. 2011;13(5):541-549.

4. De Hoon MJL, Imoto S, Nolan J, Miyano S. Open Source Clustering Software. 2004. *Bioinformatics*. 2004;20(9):1453-1454

5. Wang K, Mingyao L, Hadley D, et al. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research*. 2007;17:1665-1674.

6. Tanaka R, Koike K, Imai T, et al. Stem cell factor enhances proliferation, but not maturation of murine megakaryocytic progenitors in serum-free culture. *Blood*. 1992;80(7):1743-1749.

7. Chou ST, Byrska-Bishop M, Tober JM, et al. Trisomy 21-associated defects in human primitive hematopoiesis revealed through induced pluripotent stem cells. *Proc Natl Acad Sci U S A*. 2012;109(43):17573-17578.