Supplemental Experimental Procedures

Mice and cell culture

All animal studies were approved by the Institutional Animal Care and Use Committee at Stony Brook University. Wild-type mice (Jackson Laboratory) were maintained on the 129/Sv background. Oct4-GFP mice (Jackson Laboratory) Nestin-GFP transgenic mice were maintained on a mixed 129/B6 genetic background. MEFs were derived from E12.5-E13.5 embryos using standard procedures. Mouse Oct4-GFP ES cells were kindly provided by Antonella Galli (Columbia University). Primary MEFs were cultured in DMEM/10%FBS on 0.1% gelatinized tissue culture plates. ES and iPS cells were cultured inKSR medium consisting of F12 DME supplemented with 10% KSR (Invitrogen), 1 x nonessential amino acids, 1 x Glutamine, 1 x Pen/Strep, 0.1 μM β-mercaptoethanol, and 1000 μ/ml LIF (Millipore).

For teratoma assays, 1 x10^6 iPS cells were resuspended in Matrigel (Invitrogen) and injected subcutaneously into CD1 athymic nude mice (Harlan). When tumors reached 1 cm^3 in size (6 to 8 weeks), they were harvested and processed histologically. For blastocyst injection, 4-week-old C57BL/6N (Taconic) female mice were superovulated by administration of 5 IU of pregnant mare serum gonadotropin and 5 IU of human chorionic gonadotrophin, and then mated with C57Bl/6N males. All blastocysts were collected at day 3.5 after detection of vaginal plugs and flushed in M2 medium (Millipore). The blastocysts were washed in M2 medium and cultured in KSOM+AA medium (Millipore) at 37^0C, 5% CO2 incubator. Five days before injection, the iPS cells were thawed and passaged once. Ten to fifteen iPS cells were injected into each blastocyst for generation of chimeric mice.

Flow cytometry

Cells were harvested using Accutase (Sigma), stained with anti-Sca1, Thy1.2, CD34, CD133, EpCAM, CDH1, SSEA1, NGFR and c-Kit antibodies (BD Pharmingen or eBioscience), and

analyzed using FACS Calibur (Becton-Dickinson) and CellQuest software. Hematopoietic cells (Mac1- and CD45-positive) were excluded from the sorting gates.

Expression analyses

Total cellular RNA was extracted using the RNeasy Mini Kit (Quiagen). One-color hybridizations of cRNAs (2 technical replicas) were performed against the mouse Ilumina MouseRef-8 Expression BeadChip 25K microarray according to the manufacturer's instructions by Empire Genomics LLC (Buffalo NY). Microarray signals were processed with GenomeStudio Gene Expression Module (GSGX) Version 1.6.0. Data were background corrected and quantile normalized. Differentially expressed genes were defined as > 2-fold change. For scatter plot analyses, Pearson correlation coefficients between two samples were calculated using normalized signals. Scatter plots were generated using log2 transformed gene expression values. The heatmaps were generated by calculating ratios of expression in each sample versus control. The log2 values were then supplied to the heatmap function of the R statistical package. Module expression analysis was conducted as describe in (Kim et al., 2010). Average gene expression values (log2) of all genes were set as baseline 0. The gene expression values (log2) of each module relative to the overall average were represented as mean ± SEM. Definition of each module is as follows: the Core module is composed of genes co-occupied by at least seven factors among nine factors shown in the Core cluster (Smad1, Stat3, Klf4, Oct4, Nanog, Sox2, Nac1, Zfp281, and Dax1) (Kim et al., 2010); the ESC-like module (Wong et al., 2008); the Myc module comprises targets of Myc, Max, N-Myc, Dmap1, E2F1, E2F4, and Zfx (Kim et al., 2010); the PRC module is composed of targets of PRC cluster proteins, Phc1, Rnf2, Eed and Suz12 (Kim et al., 2010). We confirmed reproducibility of our microarray by quantitative real-time PCR, using RNA samples that had been independently isolated from newly infected MEFs. RT-PCR was performed using total RNAs and gene-specific primers. The sequences of primers are listed in Supplemental Table 4.

Western blotting

Western analyses were performed on a routine basis. Representative results are shown. Briefly, cells were lysed in buffer containing 10 mM Tris (pH 7.4), 150 mM NaCl, 10% glycerol, 1% Triton, 1 mM EDTA, 0.1% SDS and protease inhibitor cocktail (Roche). Aliquots of whole-cell lysates (50–80 μg of protein) were separated on SDS–acrylamide gels and blotted onto BA85 nitrocellulose membranes (Schleicher & Schuell). They were then incubated with antibodies specific for Oct4 (ab19857), Klf4 (ab34814, both from Abcam); Sox2 (AB5603, Millipore), c-Kit (DI3A2, Cell Signaling), CDH1 (610404, BD Pharmingen), EpCam (E144, Novus), and Erk1,2 (05-157, Upstate).

Supplemental Figures

Suppl. Figure 1.

Genomic PCR to detect exogenous transgene integration in OS-generated (**A**) or OK-generated (**B**) iPS clones. Mouse ES cells were used as negative controls.

Suppl. Figure 2.

A, B. Heat maps from microarray analyses showing induction of endogenous transcription factors in Sca1-SP, DN, Thy1-SP and DP MEFs transduced with OS- (**A**) or OK-expressing retroviruses (**B**). Values are given relative to their own uninfected controls. Expression in corresponding uninfected Sca1-SP and DN (summarily called 'Thy1-') *versus* Thy1-SP and DP (summarily called 'Thy1+') cells are shown as controls.

C. Heat maps of shared transcription factor genes in OS- and OK-transduced Thy1-negative (Sca1-SP and DN) and Thy1-positive (Thy1-SP or DP) cells.

Suppl. Figure 3.

Average gene expression values of the ESC-like, Core pluripotency, Polycomb repressive complex (PRC) and Myc transcriptional modules in control uninfected Thy1-negative (Sca1-SP and DN) and Thy1-positive (Thy1-SP and DP) cells as well as in Sca1-SP, DN and Thy1-SP subpopulations transduced with OS or OK retroviruses. Microarray analyses were performed 4 days post infection with the indicated retroviruses. Data are represented as mean +/- SEM.

Suppl. Figure 4.

The Thy1+ fractions (composed of Thy1+ single-positive, Thy1-SP; and Thy1+Sca1+, double positive, DP cells) represent the majority of MEFs (>70%), while Thy1-negative fractions (composed of Sca1+ single-positive, Sca1-SP; and Thy1-Sca1-, double negative, DN cells) represent a relative minority (**A**). When sorted cells were maintained under standard culture conditions for 7 days, Sca1-SP MEFs gradually lost their phenotype and converted to DN and Thy1-SP cells (**A**). Likewise, DP cells largely lost the expression of Sca1, becoming Thy1-SP cells, and the same drift was seen in DN cells (**A**). In contrast, the Thy1-SP fraction was stable in retaining the corresponding phenotype over extended periods of time (**A**). Thus, DP, DN and Sca1-SP cells spontaneously gravitate towards acquisition of Thy1 single positivity. Retroviral expression of Klf4, alone or in combination with Oct4 and Sox2 rendered all fractions positive for Sca1, thereby reducing the proportion of DN and Thy1-SP cells (**B**). On the other hand, retroviral expression of Oct4 and Sox2 either alone or in combination with each other, produced no significant changes (**4B** and data not shown). Thus, the sorted MEF fractions acquire different phenotypic features depending on whether they were infected with OK versus OS viruses or other factor combinations.

Suppl. Figure 5.

A. FACS analysis of Nestin-GFP+ MEFs. Nestin-GFP+ cells (left) were gated (R1) and analyzed for Thy1 expression (right). The majority of Nestin-GFP+ cells are Thy1-negative (81%).

B. FACS analysis of NGFR+ MEFs. Nestin-GFP+NGFR+ cells (left) were gated (R1) and analyzed for Thy1 expression (right). The majority of NGFR+ cells are Thy1-negative (79%).

Suppl. Figure 6.

A. Western blot analysis of GFP expression in FACS-sorted Nestin-GFP-negative MEFs. Cells were analyzed 4 days post-infection with empty control retroviruses (Mock) or OS-, OK-, or OSK-expressing viruses. Note that expression of GFP (i.e., Nestin) is significantly reduced upon conversion of cells into iPS cells (OS-generated iPS colony is shown). ES cells derived from Oct4-GFP transgenic mice are shown for control. MAPK is a control for equal loading.

B. Induction of GFP expression in FACS-sorted Nestin-GFP-negative MEFs transduced with retroviruses as in (**A**). Cells were analyzed by FACS on days 4, 7, 11, and 14 post-infection.

Suppl. Figure 7.

A. Quantitative real-time RT-PCR analysis of gene expression levels in Nestin-GFP negative (yellow group) and Nestin-GFP bright MEFs (blue group). Cells were analyzed before and 1 or 11 days after transduction with OS, OK or OSK viruses. Data were normalized to the corresponding HPRT value to obtain relative changes in the indicated mRNAs.

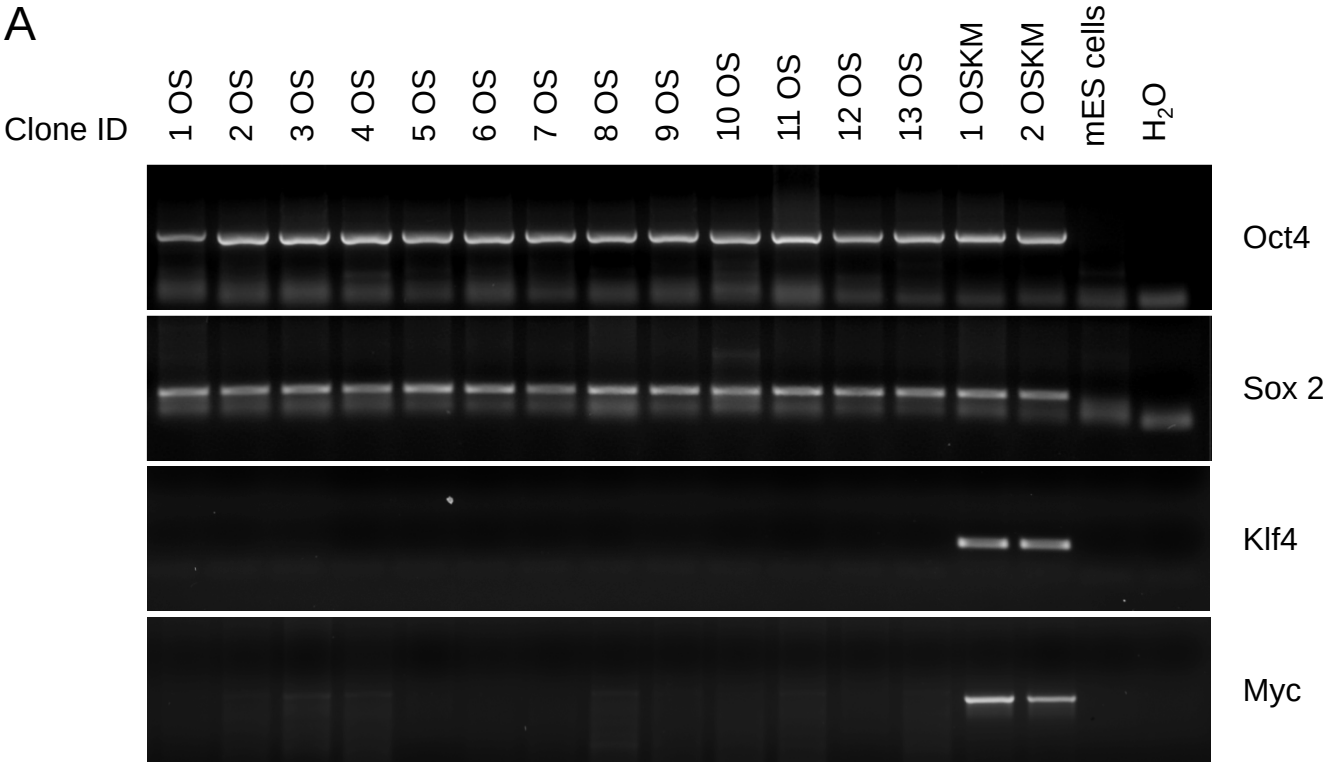
B. Quantitative real-time RT-PCR analysis of NGFR expression levels in total unsorted Nestin-GFP MEFs and sorted populations of Nestin-GFP+NGFR+ and Nestin-GFP-NGFR- MEFs. Cells were analyzed before and 1 or 11 days after transduction with OS, OK or OSK viruses. Data were normalized to the corresponding HPRT value to obtain relative expression changes.

Suppl. Figure 8.

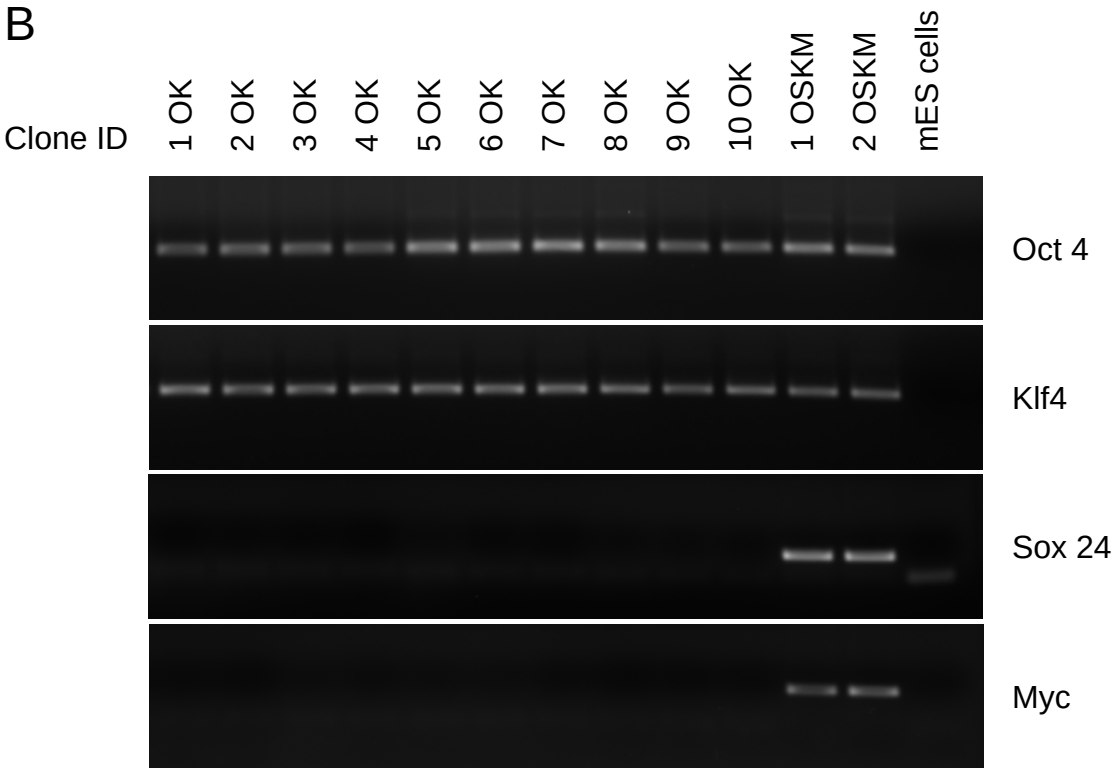
Schematic of reprogramming stages: disruption of cellular homeostasis (A), cell switch to transient states characterized by concurrent expression of mixed lineage markers (B), activation

of pluripotency networks (C) and the conversion into IPS cells (D). The OSK and OS combinations of factors are more efficient than OK in activation of pluripotency networks, while the OSK and OK combinations are more efficient and faster than OS in inducing ES cell-specific morphological changes.

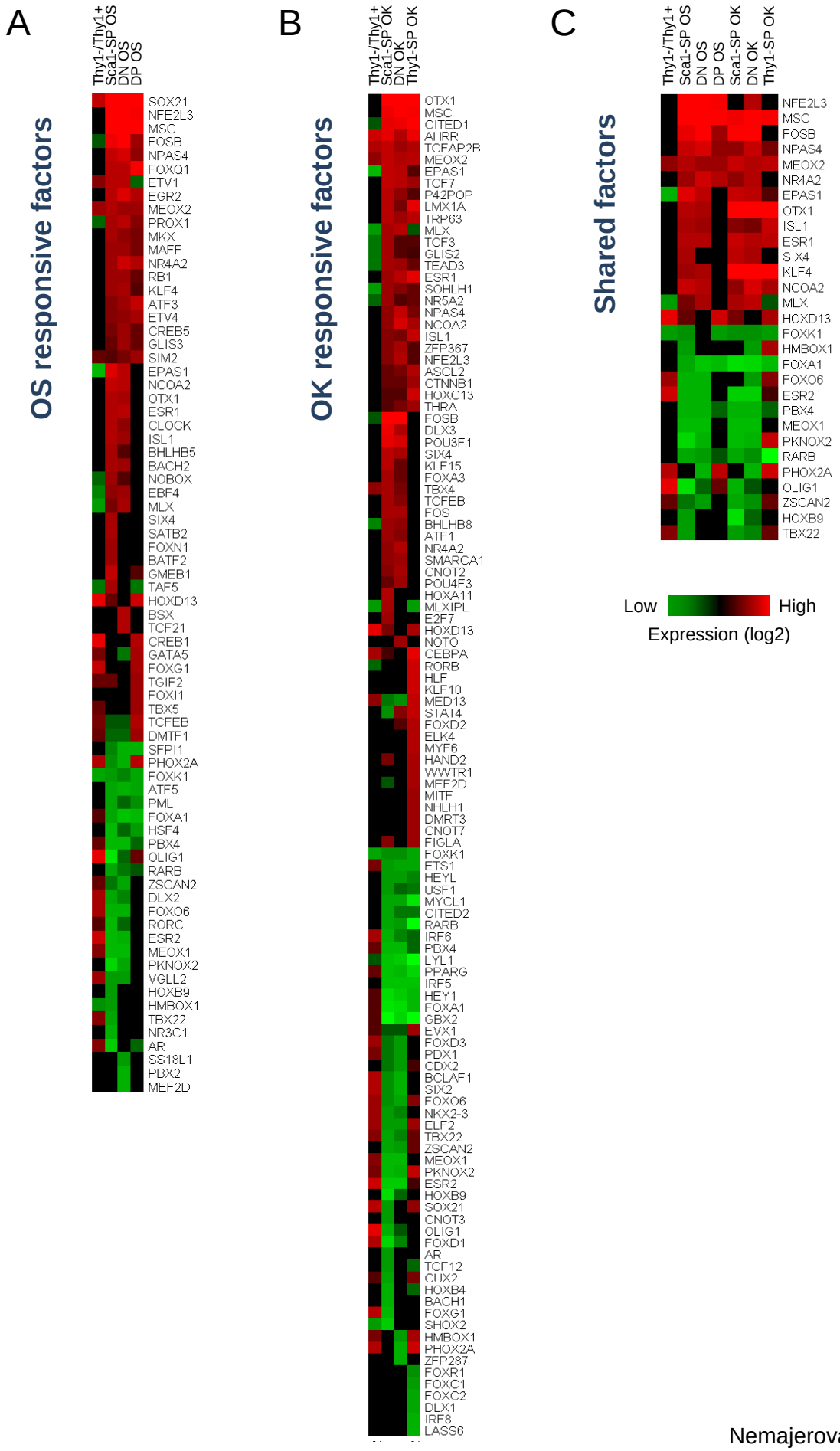
A



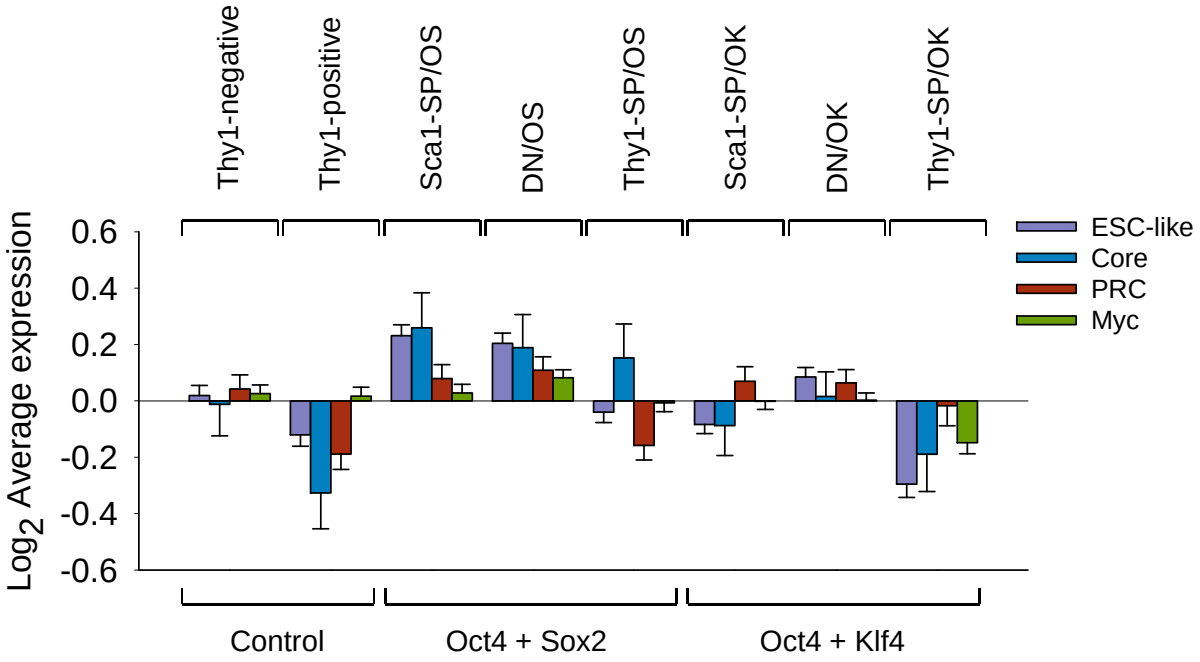
B



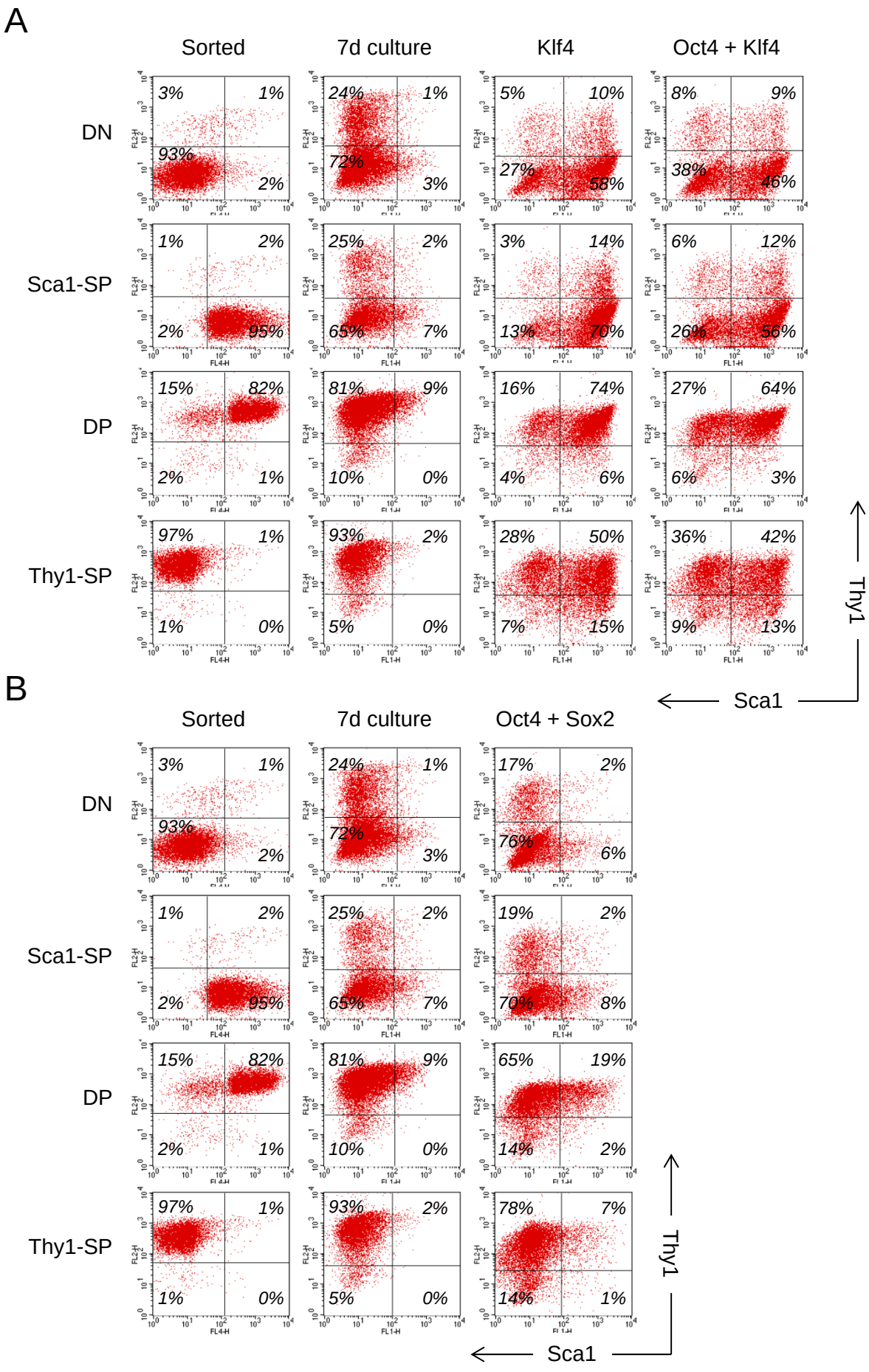
Nemajerova et al., Suppl. Figure 1



Nemajerova et al., Suppl. Figure 2

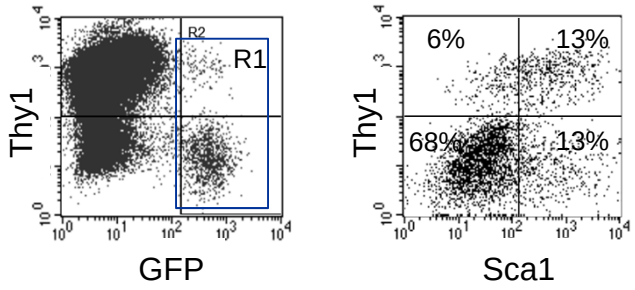


Nemajerova et al., Suppl. Figure 3

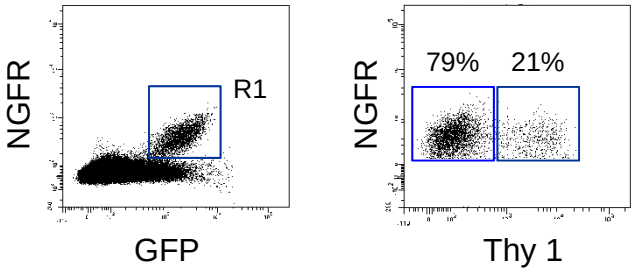


Nemajerova et al., Suppl. Figure 4

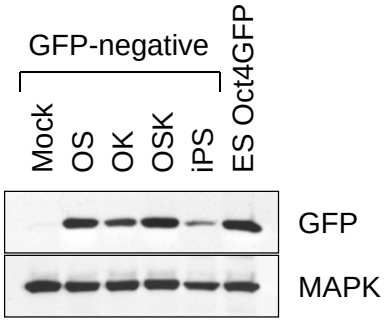
A



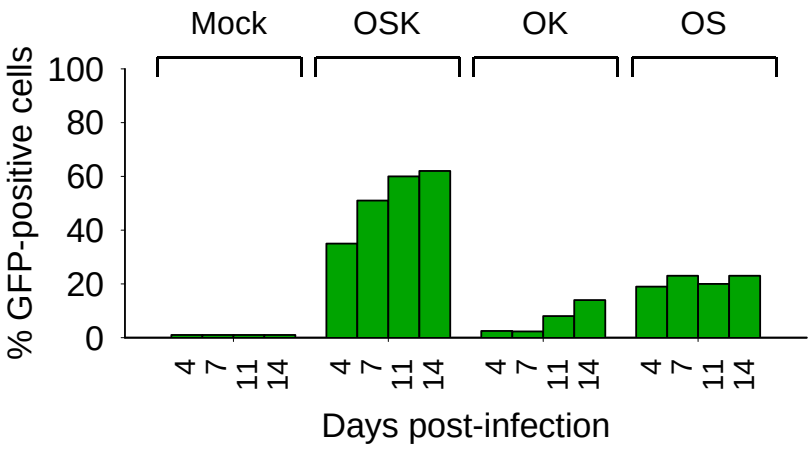
B



A

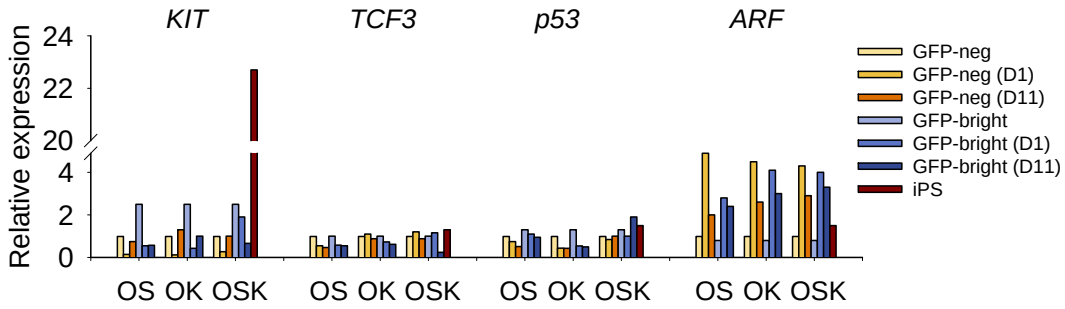


B

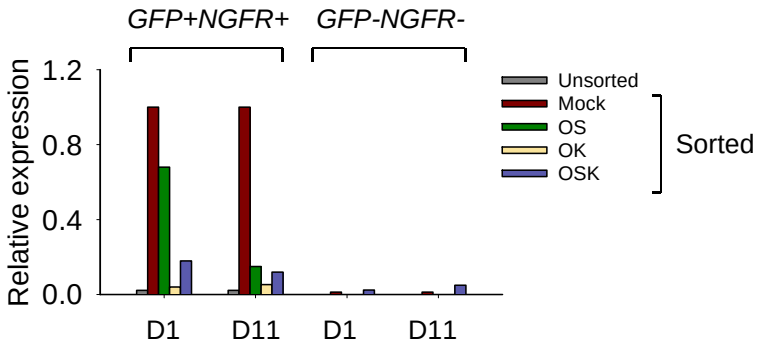


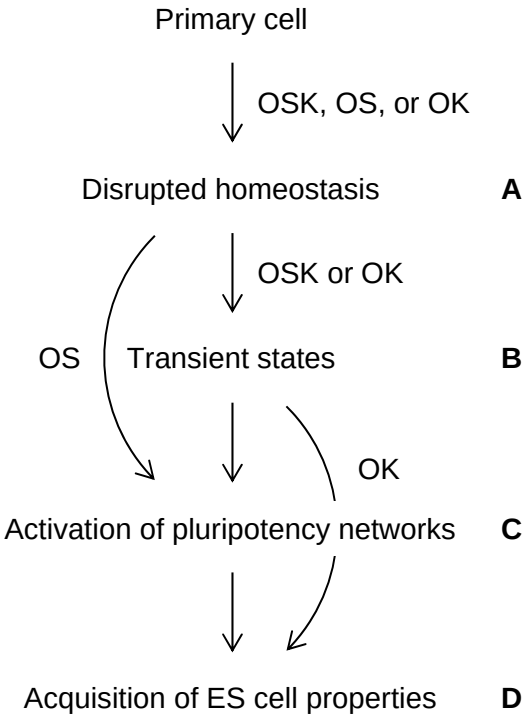
Nemajerova et al., Suppl. Figure 6

A


KIT
TCF3
p53
ARF
Relative expression
24
22
20
4
2
0
OS OK OSK
OS OK OSK
OS OK OSK
OS OK OSK
GFP-neg
GFP-neg (D1)
GFP-neg (D11)
GFP-bright
GFP-bright (D1)
GFP-bright (D11)
iPS

B


GFP+NGFR+
GFP-NGFR-
Relative expression
1.2
0.8
0.4
0.0
D1
D11
D1
D11
Unsorted
Mock
OS
OK
OSK
Sorted


Primary cell
OSK, OS, or OK
Disrupted homeostasis
A
OSK or OK
OS
Transient states
B
OK
Activation of pluripotency networks
C
Acquisition of ES cell properties
D

Supplemental Table 1.
Sca1-SP MEFs show increased expression of genes involved in embryonic morphogenesis, tissue morphogenesis and regulation of cell development.

Term	Count	%	P Value
GO:0048598~embryonic morphogenesis	8	6.1	0.016666071
GO:0048729~tissue morphogenesis	8	6.1	0.001862836
GO:0060284~regulation of cell development	8	6.1	1.69E-04
GO:0048736~appendage development	5	3.8	0.011506968
GO:0035107~appendage morphogenesis	5	3.8	0.010242797
GO:0060348~bone development	5	3.8	0.011181983
GO:0008544~epidermis development	4	3.0	0.064658192
GO:0001654~eye development	5	3.8	0.028622431
GO:0007507~heart development	7	5.3	0.006052495
GO:0003007~heart morphogenesis	5	3.8	0.002142245
GO:0001822~kidney development	4	3.0	0.044328062
GO:0035108~limb morphogenesis	5	3.8	0.010242
GO:0060173~limb development	5	3.8	0.011506968
GO:0014031~mesenchymal cell development	3	2.3	0.046788171
GO:0048762~mesenchymal cell differentiation	3	2.3	0.050424149
GO:0030182~neuron differentiation	10	7.6	0.002676603
GO:0045664~regulation of neuron differentiation	7	5.3	1.04E-04
GO:0007423~sensory organ development	8	6.1	0.002867616
GO:0001501~skeletal system development	9	6.8	0.001192193
GO:0014706~striated muscle tissue development	4	3.0	0.067142518
GO:0001655~urogenital system development	6	4.5	0.004445775
GO:0051216~cartilage development	6	4.5	2.69E-04
GO:0006928~cell motion	8	6.1	0.018575252
GO:0007155~cell adhesion	12	9.1	0.002739769
GO:0030574~collagen catabolic process	4	3.0	2.44E-04
GO:0030198~extracellular matrix organization	5	3.8	0.00653851
GO:0007167~receptor protein signaling pathway	7	5.3	0.015413984
GO:0042981~regulation of apoptosis	9	6.8	0.050942203
GO:0001558~regulation of cell growth	4	3.0	0.030321804
GO:0040008~regulation of growth	6	4.5	0.040578565
GO:0051252~regulation of RNA metabolic process	17	12.9	0.077021932
GO:0006355~regulation of transcription, DNA-dependent	17	12.9	0.068985832
GO:0006357~regulation of transcription from RNA polymerase II promoter	10	7.6	0.037471912

Supplemental Table 2.
DN MEFs are enriched for myofibroblast-committed progenitors.

Gene		Fold Change (DN vs. Thy1-SP)
MYLPF	fast skeletal myosin light chain 2	27.6
MYL1	myosin, light chain 1, alkali; skeletal, fast	11.6
MYL4	myosin, light chain 4, alkali; atrial, embryonic	8.3
MYL6B	myosin, light chain 6B, alkali, smooth muscle and non-muscle	5.1
MYH1	myosin, heavy chain 1, skeletal muscle, adult	2.48
MYH8	myosin, heavy chain 8, skeletal muscle, perinatal	4.24
MB	myoglobin	2.46
TNNC1	troponin C type 1	8.05
TNNC2	troponin C type 2	19.4
TNNI1	troponin I type 1	12.74
TNNI2	troponin I type 2	16.04
TNNT1	troponin T type 1	5.9
TPM2	tropomyosin 2	7.11
THY1		0.31
Sca1		1
NGFR	(CD271)	11.06
CD82	(Inducible membrane protein R2)	2.41
CD80	(T-lymphocyte activation antigen CD80)	2.37
CD68	(macrophage antigen CD68)	1
CD40	(TNF receptor superfamily member 5)	1
CD96	(T-cell surface protein tactile)	1
CD97	(heterodimeric receptor associated with inflammation)	1
CD55	(decay accelerating factor for complement)	1

Supplemental Table 3.
Sorted Sca1-SP and DN MEFs yield significantly more iPS colonies
than the corresponding Thy1-SP or DP fractions.

MEF subpopulations	p value (OSKM)*	p value (OSK)*	p value (OK)*	p value (OS)*
Sca1-SP vs. Unsorted	<0.0001	<0.0001	<0.0001	<0.0001
Sca1-SP vs. Thy1-SP	<0.0001	<0.0001	<0.0001	<0.0001
Sca1-SP vs. DP	<0.0001	<0.0001	<0.0001	<0.0001
DN vs. Unsorted	<0.0001	<0.0001	<0.0001	<0.0001
DN vs. Thy1-SP	<0.0001	<0.0001	<0.0001	<0.0001
DN vs. DP	<0.0001	<0.0001	<0.0001	0.0002
DN vs. Sca1-SP	0.17	0.97	0.039	0.42

* Cells were transduced with OSKM, OSK, OK or OS combinations of reprogramming factors.
P values were calculated using Student's t-test.

Supplemental Table 4
Primer sequences for quantitative real time RT-PCR:

Gene name	Primer sequence 5'-3'		Amplicon size (bp)
HPRT	F	GGCTATAAGTTCTTTGCTGACC	126
	R	CTCCACCAATAACTTTTATGTCC	
CDH1	F	CAGGTCTCCTCATGGCTTTGC	175
	R	CTTCCGAAAAGAAGGCTGTCC	
EPCAM	F	GCGGCTCAGAGAGACTGTG	139
	R	CCAAGCATTTAGACGCCAGTTT	
CD34	F	AAGGCTGGGTGAAGACCCTTA	157
	R	TGAATGGCCGTTTCTGGAAGT	
CD133	F	GTTGAGACTGTGCCCATGAAA	98
	R	GACGGGCTTGTCATAACAGGA	
c-KIT	F	GCCACGTCTCAGCCATCTG	90
	R	GTCGCCAGCTTCAACTATTAACT	
Oct4	F	CCATGCATTCAAACTGAGGCACCA	243
	R	AGCTATCTACTGTGTGTCCCAGTC	
Sox2	F	CGCCCAGTAGACTGCACAT	154
	R	CCCTCCCAATTCCCTTGTAT	
Klf4	F	GTGCCCCGACTAACCGTTG	185
	R	GTCGTTGAACTCCTCGGTCT	
c-Myc	F	GCTCTCCATCCTATGTTGCGG	116
	R	TCCAAGTAACTCGGTCATCATCT	
Nanog	F	TCTTCCTGGTCCCCACAGTTT	100
	R	GCAAGAATAGTTCTCGGGATGAA	
Sall4	F	CCCTGGGAACTGCGATGAAG	111
	R	TCAGAGAGACTAAAGAACTCGGC	
Tbx3	F	GAACCTACCTGTTCCCGGAAA	121
	R	CAATGCCCAATGTCTCGAAAAC	
Tcf3	F	ACGAGCTGATCCCCTTCCA	101
	R	CAGGGACGACTTGACCTCAT	
Zic3	F	TGCTGCCAGTTCAGGCTATG	83
	R	CGAGAAGGGGTTTTAGTGGTATC	
Arf	F	CTTGGTCACTGTGAGGATTCA	125
	R	CTACGTGAACGTTGCCCATCA	
p53	F	ACAGCGTGGTGGTACCTTAT	149
	R	GGTTCCCACTGGAGTCTTC	
NGFR	F	TGGGCTCAGGACTCGTGTT	189
	R	CAGGGATCTCCTCGCATTCG	

All primers are for mouse mRNA detection and have a hybridization temperature of 60°C. Oct4, Sox2 and KLF4 primers are designed to amplify only the endogenous mRNA, not the transgene.

Relative expression of the target genes was calculated using the ΔCT method described previously: Relative expression = $2^{-\Delta CT}$, where ΔCT = CT (Target gene) - CT (HPRT).