SUPPLEMENTARY INFORMATION

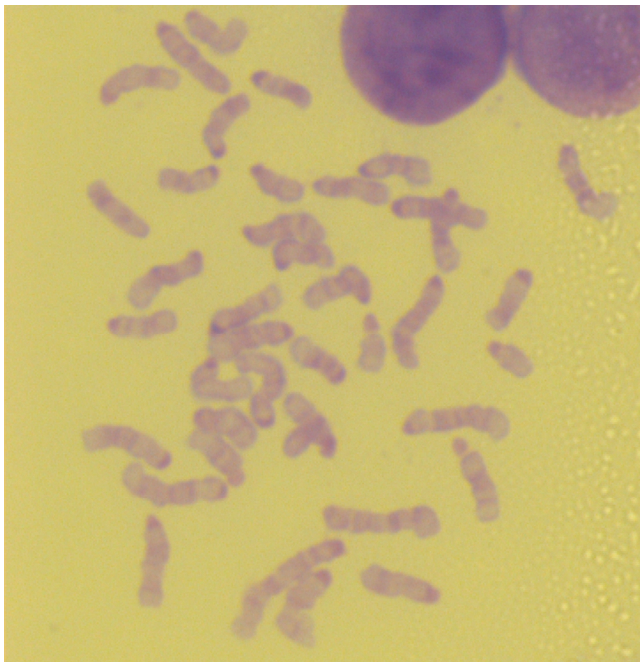


Figure S1: Karyotype analysis of iPS cell line

One hundred individual chromosomal spreads were counted for each of the iPS samples. Shown here is an IP14D-1 spread at passage 15, predominantly diploid. More than 75% of the cells showed normal mouse karyotype of 40 chromosomes.

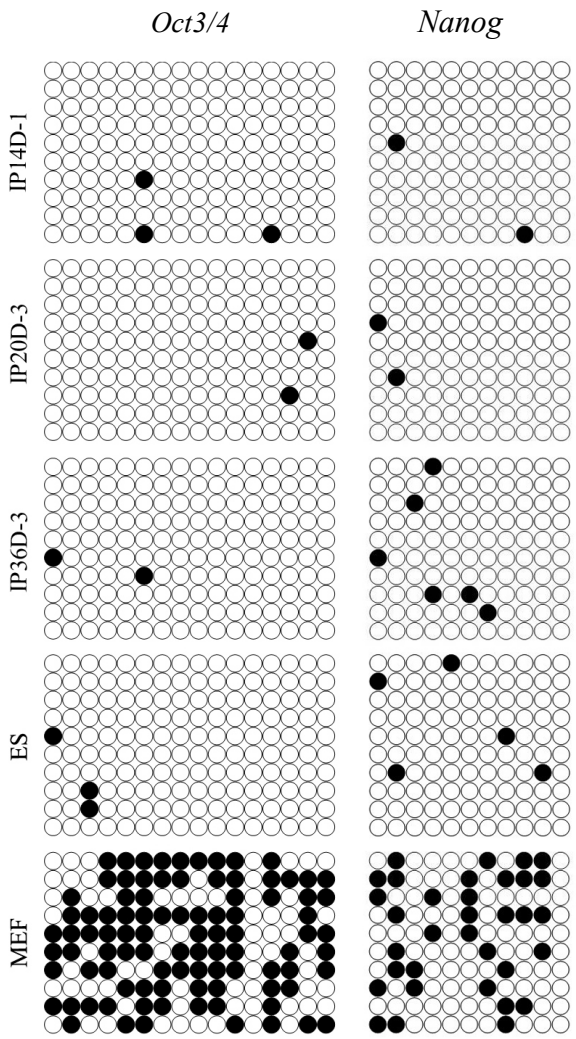


Figure S2. Methylation analysis of *Oct4* and *Nanog* promoter regions.

Genomic DNA from iPS cell lines (IP20D-3, IP36D-3, IP14D-1) at passage 10 as well as from MEFs and ES cells were isolated and bisulfite treated. *Oct4* and *Nanog* promoter regions were amplified with nested primers (Supplementary Table S2). Ten randomly selected clones were sequenced and analyzed. Open and filled circles represents unmethylated and methylated CpG dinucleotides, respectively. The three iPS cell lines are very different in methylation patterns compared to the parental Oct4-GFP MEFs, but are very similar to those from normal ES cells, reflecting the epigenetic remodeling that occurred in concert with the reprogramming events.

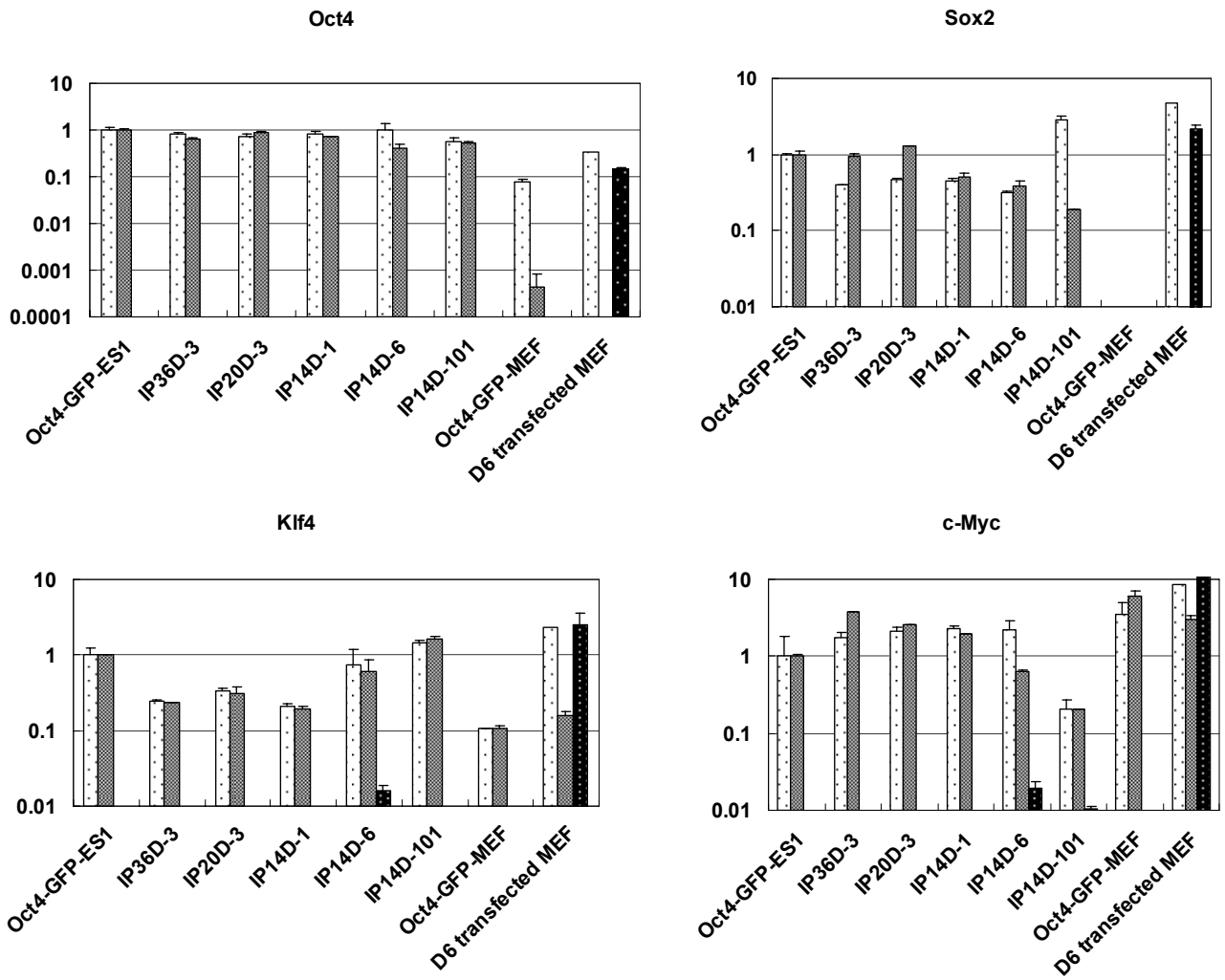
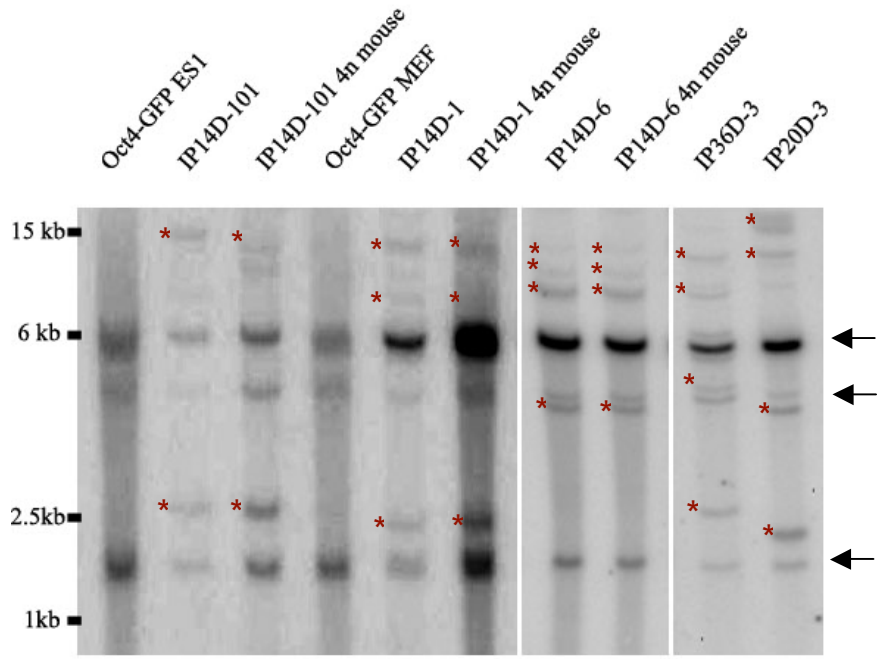
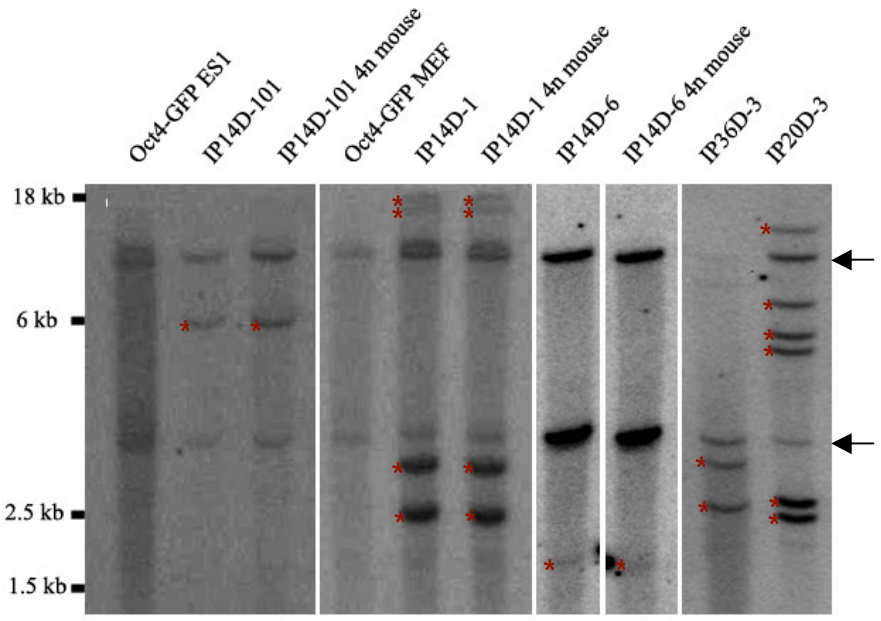


Figure S3. Quantitative RT-PCR to detect transgene expression

Real-time RT-PCR was used to detect total, endogenous and transgenic expression of the four factors Oct4, Sox2, c-Myc and Klf4 in Oct4-GFP ES1 (an ES line with integrated Oct4-GFP expression), IP36D3, IP20D3, IP14D1, IP14D-6, IP14D-101, Oct4-GFP MEF and a positive control, day 6 transfected MEF. The assays included primers specifically designed to discriminate endogenous (black bars) from transgenic (open bars) RNA expression, and total expression is shown with grey bars. Relative expression levels (Y-axis) of these genes in each sample were first normalized to their endogenous GAPDH level, and then displayed as their relative ratio to expression of these genes in Oct4-GFP ES1. Primer sequences are listed in Supplementary Table S3.



Oct4



c-Myc

Figure S4. Validation of genomic integration of transgenes by Southern blot analysis

Total genomic DNA was extracted from Oct4-GFP ES1, Oct4-GFP MEF, IP36D3, IP20D3, IP14D1 and the corresponding tetraploid mice. DNA digested with Bam H1 and BglII were hybridized with *Oct4* and *c-Myc* cDNA probes, respectively. Arrowheads indicate endogenous *Oct4* or *c-Myc* bands found in all the samples. Asterisk indicates extra bands in the iPS samples corresponding to viral integration into the genome. Note that the different iPS lines have different integration patterns, but the three iPS cell lines and their corresponding 4N-comp animals showed the same patterns.

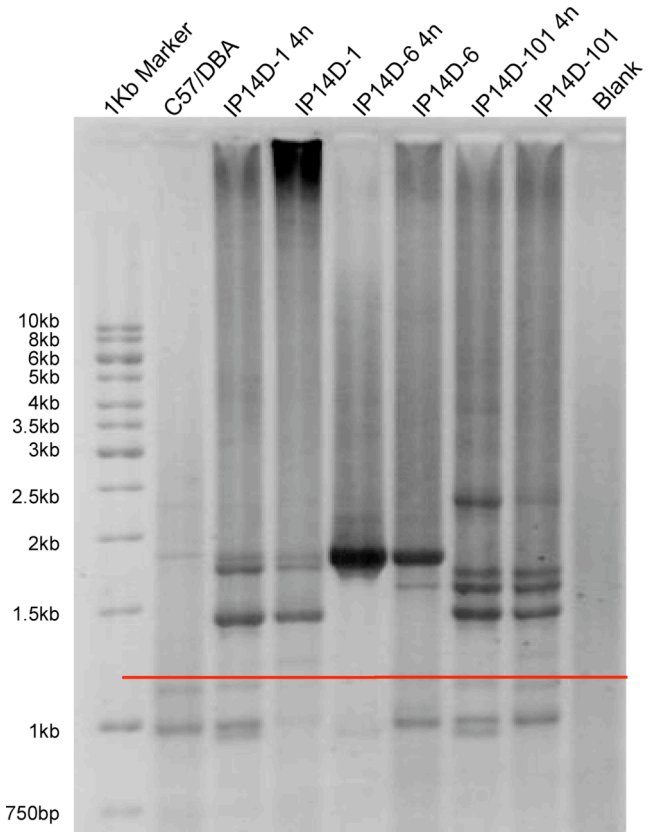


Figure S5: Identification of integration patterns in iPS cell lines and 4N-comp mice by reverse PCR.

Genomic DNA was extracted from the iPS cell lines IP14D-1, IP14D-6, IP14D-101 as well as the tails of IP14D1 4N-comp, IP14D6 4N-comp, IP14D101 4N-comp and a normal C57/DBA (B6D2F1) mice. The integration patterns of these samples were identified by reverse PCR procedures of neighboring genomic regions by nested primers specific for *Oct4*. Bands above the red line or not present in the C57/DBA control were considered positive patterns for integration. Note that the three 4N-comp mice displayed different integration patterns to each other, but similar to the iPS lines (IP14D1, P14D6 and IP14D101) they originated from.

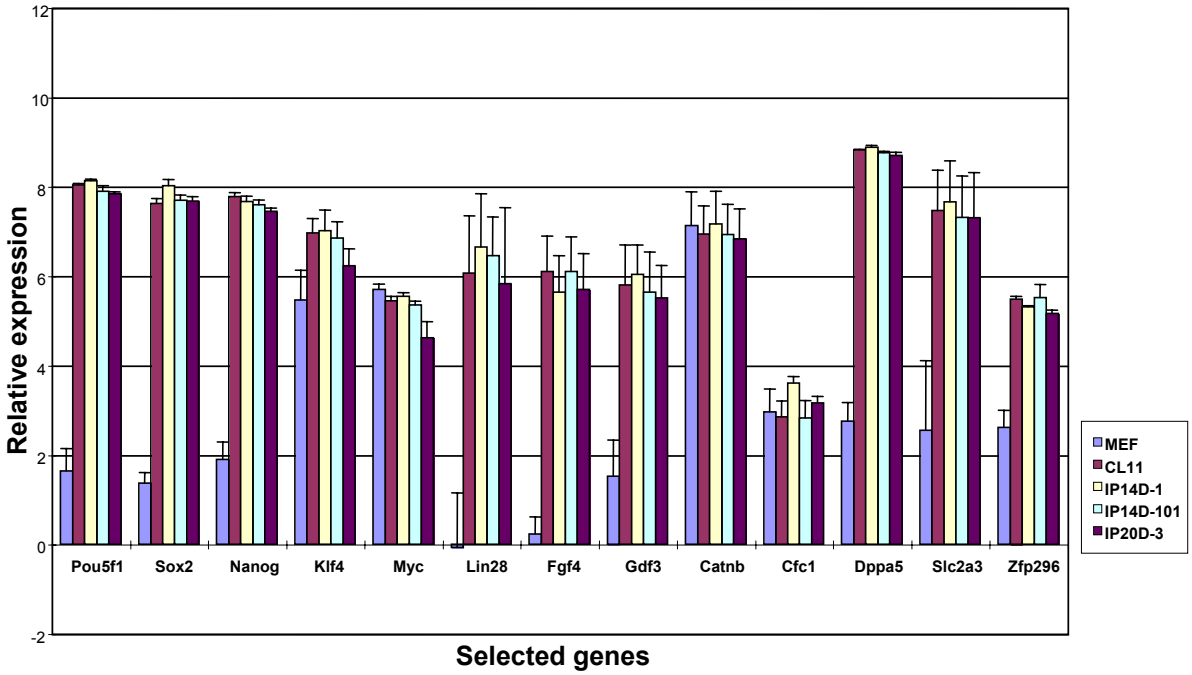


Figure S6. Expression profiles of selected pluripotent marker genes from microarray analysis.

Table S1. Summary of iPS lines derived to date.

Genetic background	Experiment number	iPS cell lines	Teratoma formation	2N chimera mice *(Number)*	% Chimerism	Germline transmission *(Number)*	4N complementation ability	
							Developed to	No. of alive 4N-comp animals
B6D2F1 (GFP$^+$)	Exp14D-1	IP14D-1	YES	YES(6)	20-95	YES (4)	D19.5	22
		IP14D-2	NT	YES			<D10.5	
	Exp14D-2	IP14D-3	NT	YES (14)	10-90		<10.5	
		IP14D-4	NT	NT			<10.5	
	Exp14D-5	IP14D-5	NT	NT			NT	
		IP14D-6	NT	YES (10)	20-90	YES (3)	D19.5	1
	Exp20D-1	IP20D-1	YES	YES (5)	5-80		D13.5	
		IP20D-2	YES	YES (1)	30		NT	
	Exp20D-2	IP20D-3	YES	YES (14)	20-80	YES (1)	D15.5	
		IP20D-4	NT	NT			NT	
		IP20D-5	NT	NT			NT	
		IP20D-6	NT	NT			NT	
		IP20D-7	NT	NT			NT	
		IP20D-8	NT	NT			NT	
		IP20D-9	NT	NT			NT	
		IP20D-10	NT	NT			NT	
		IP20D-11	NT	NT			NT	
		IP20D-12	NT	NT			NT	
		IP20D-13	NT	NT			NT	
		IP20D-14	NT	NT			NT	
	Exp20D-11	IP20D-15	YES	NT			NT	
		IP20D-16	YES	NT			NT	
		IP20D-17	YES	NT			NT	
		IP20D-18	YES	NT			NT	
		IP20D-19	YES	YES (4)	10-70	YES (1)	D13.5	
		IP20D-20	YES	NT			NT	
		IP20D-21	YES	NT			NT	
		IP20D-22	NT	NT			NT	
		IP20D-23	NT	NT			NT	
	Exp36D-2	IP36D-2	NT	NT			NT	
		IP36D-3	YES	YES (6)	5-70		D11.5	
C57x129S2 (GFP$^-$)	Exp14D-C1	IP14D-101	NT	YES(18)	10-90	YES(1)	D19.5	4
		IP14D-102	NT	NT			D13.5	
		IP14D-103	NT	NT			NT	
		IP14D-104	NT	NT			NT	
		IP14D-105	NT	NT			NT	
		IP14D-106	NT	NT			NT	

*NT = Not Tested

Table S2. PCR primer sequences for methylation studies

Gene	Accession number	Primer sequence	Nested PCR
Oct4	NM_013633	F: GAGGATTGGAGGTGTAATGGTTGTT R: CTACTAACCCATCACCCCCACCTA	1st Round
		F: CAAGCTTTGGGTTGAAATATTGGGTTTATTT R: CGGATCCCTAAAACCAAATATCCAACCATA	2nd Round
Nanog	NM_028016	F: AAGTATGGATTAATTTATTAAGGTAGTT R: AAAAAACCCACACTCATATCAATATA	1st Round
		F: AAGTATGGATTAATTTATTAAGGTAGTT R: CAACCAAATCAACCTATCTAAAAA	2nd Round

Table S3. Primer sequence for real-time RT-PCR

Target gene expression level	Name	Primer sequence
Endogenous	endo-Oct4	F: TCTTTCCACCAGGCCCCCGGCTC R: TGCGGGCGGACATGGGGAGATCC
	endo-Sox2	F: TAGAGCTAGACTCCGGGCGATGA R: TTGCCTTAAACAAGACCACGAAA
	endo-Myc	F: TGACCTAACTCGAGGAGGAGCTGGAATC R: AGTTTGAGGCAGTTAAAATTATGGCTGAAGC
	endo-Klf4	F: CCATCGGACCTACTTATCTGC R: AAAACCTCAAACCAAAACCC
Exogenous / transgene	ex-Oct4	F: CCCAGTGTGGTGGTACGGGAAATC R: AGTTGCTTTCCACTCGTGCT
	ex-Sox2	F: CCCAGTGTGGTGGTACGGGAAATC R: TCTCGGTCTCGGACAAAAGT
	ex-c-Myc	F: CCCAGTGTGGTGGTACGGGAAATC R: GCTCGCTCTGCTGTTGCTGGTGATAG
	ex-Klf4	F: CCCAGTGTGGTGGTACGGGAAATC R: GTCGTTGAACTCCTCGGTCT
Total	T-Oct4	F: GGCTTCAGACTTCGCCTCC R: AACCTGAGGTCCACAGTATGC
	T-Sox2	F: GCGGAGTGGAAACTTTTGTCC R: CGGGAAGCGTGTACTTATCCTT
	T-c-Myc	F: ATGCCCCTCAACGTGAACTTC R: CGCAACATAGGATGGAGAGCA
	T-Klf4	F: GTGCCCCGACTAACCGTTG R: GTCGTTGAACTCCTCGGTCT