

Virus-free induction of pluripotency and subsequent excision of reprogramming factors

Keisuke Kaji¹, Katherine Norrby¹, Agnieszka Paca¹, Maria Mileikovsky², Paria Mohseni^{2,3} & Knut Woltjen²

Reprogramming of somatic cells to pluripotency, thereby creating induced pluripotent stem (iPS) cells, promises to transform regenerative medicine. Most instances of direct reprogramming have been achieved by forced expression of defined factors using multiple viral vectors^{1–7}. However, such iPS cells contain a large number of viral vector integrations^{1,8}, any one of which could cause unpredictable genetic dysfunction. Whereas *c-Myc* is dispensable for reprogramming^{9,10}, complete elimination of the other exogenous factors is also desired because ectopic expression of either Oct4 (also known as Pou5f1) or Klf4 can induce dysplasia^{11,12}. Two transient transfection-reprogramming methods have been published to address this issue^{13,14}. However, the efficiency of both approaches is extremely low, and neither has been applied successfully to human cells so far. Here we show that non-viral transfection of a single multiprotein expression vector, which comprises the coding sequences of *c-Myc*, *Klf4*, *Oct4* and *Sox2* linked with 2A peptides, can reprogram both mouse and human fibroblasts. Moreover, the transgene can be removed once reprogramming has been achieved. iPS cells produced with this non-viral vector show robust expression of pluripotency markers, indicating a reprogrammed state confirmed functionally by *in vitro* differentiation assays and formation of adult chimaeric mice. When the single-vector reprogramming system was combined with a *piggyBac* transposon^{15,16}, we succeeded in establishing reprogrammed human cell lines from embryonic fibroblasts with robust expression of pluripotency markers. This system minimizes genome modification in iPS cells and enables complete elimination of exogenous reprogramming factors, efficiently providing iPS cells that are applicable to regenerative medicine, drug screening and the establishment of disease models.

Efficient multiprotein expression has been reported in a variety of cell types, including human embryonic stem (ES) cells, using the 2A peptide sequence of foot and mouth disease virus (F2A) or 2A-like sequences from other viruses^{17,18}. Recently, this multiprotein expression strategy has also been applied for reprogramming with transient transfection¹⁴ and a doxycycline (dox)-inducible lentiviral vector^{19,20}. Here we have taken advantage of the strategy to generate virus-free, factor-removable iPS cells using a single plasmid with a 2A-peptide-linked reprogramming cassette, *c-Myc-Klf4-Oct4-Sox2* (MKOS)-*IRES-mOrange*, flanked by *loxP* sites, pCAG2LMKOSimO (Supplementary Fig. 1). Initially we investigated whether the 2A-peptide-mediated multiprotein expression could achieve robust expression of *c-Myc*, *Klf4*, *Oct4* and *Sox2* when transcribed from the ubiquitously expressed synthetic CAG enhancer/promoter²¹. When a vector pCAGMKOSiE, which carries the same 2A-peptide-linked MKOS reprogramming cassette, was transfected into HEK293 cells, expression of *Klf4*, *Oct4* and *Sox2* could be detected by immunoblotting (Supplementary Fig. 2a). Whereas high expression of

endogenous *c-Myc* in HEK293 cells precluded clear identification of exogenous *c-Myc*, a phosphorylated form at Thr58 that was subjected to subsequent ubiquitination²² was enriched in the transfectants, indicating that excess *c-Myc* was degraded (Supplementary Fig. 2a, b). Appropriate nuclear localization of exogenous Oct4 and Sox2 was also observed in the transfected HEK293 cells (Supplementary Fig. 2c).

When the vector pCAG2LMKOSimO was introduced into mouse embryonic fibroblasts (MEFs), some mOrange-positive cells converted to an ES-cell-like morphology at day 5–6, and by day 9 colonies containing alkaline-phosphatase-positive cells appeared (data not shown). Moreover, morphologically ES-cell-like colonies picked between days 20 and 30 succeeded to grow maintaining an ES-cell-like morphology on gelatin (Supplementary Fig. 3a). We then went on to estimate the reprogramming efficiency using Nanog reactivation as a marker of reprogramming^{3,4}. MEFs from TNG mice, which have a green fluorescent protein (GFP) reporter inserted at the Nanog start codon²³, and MEFs from wild-type 129 mice were transfected with the pCAG2LMKOSimO plasmid and cultured on either irradiated MEFs or gelatin. The number of transiently transfected mOrange-positive cells was measured by flow cytometry at day 2. The number of reprogrammed colonies judged by GFP positivity (TNG MEFs) or anti-Nanog immunofluorescence (129 MEFs) (Supplementary Fig. 3b) was scored on day 28 (Table 1). By comparing stable transfection efficiency with nucleofection (3.6% of transiently transfected cells, see Supplementary Fig. 3c for details) and the number of reprogrammed colonies, we calculated overall reprogramming efficiency to be an average of 2.5% (Supplementary Table 1). Although the estimation method is different from that used in viral reprogramming systems (efficiency of ~0.1%^{2,3,7}), this relatively high efficiency may depend on several factors in this non-viral method, including expression of the four reprogramming factors from a single transcript and use of the CAG enhancer/promoter, which may be less prone to silencing.

Stable cell lines were established by picking colonies derived from transfection of 129 MEFs (8 out of 12 morphologically ES-cell-like colonies picked) and TNG MEFs (5 out of 9 GFP-positive colonies picked) with pCAG2LMKOSimO. We examined the expression of the reprogramming factors in eight cell lines, imO1–imO8. All cell lines showed robust reactivation of endogenous *Oct4* and *Sox2*, comprising most of total *Oct4* and *Sox2* transcripts detected by common primers for endogenous and exogenous cDNAs (Fig. 1a). Endogenous *c-Myc* expression, which was higher in MEFs than in ES cells, became similar to levels in ES cells in all cell lines, whereas there was no large change in endogenous *Klf4* expression levels (Fig. 1a). Total *c-Myc* and *Klf4* expression were high relative to expression in ES cells, but total *Oct4* and *Sox2* expression were not, although the exogenous transcript encodes all four genes. This observation could be explained by the fact

¹MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, University of Edinburgh, Edinburgh EH9 3JQ, UK. ²Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario M5G 1X5, Canada. ³Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

Table 1 | Nucleofection conditions and number of Nanog–GFP/Nanog-positive colonies

Nucleofection condition				Two days after nucleofection			Number of Nanog–GFP/Nanog-positive colonies at day 28	
Mice	Experiment number	Passage number	DNA (μ g)	Cells per well	mOrange-positive (%)	Positive cells per well	On γ MEFs	On gelatin
TNG	1	4	2	100,000	8.2	8,200	–	6
	2	3	2	35,000	13.3	4,655	–	3.5*
129	1	3	2	30,000	21.0	6,300	–	3.3†
	2	4	2	20,000	18.6	3,720	6	2
	3	5	5	50,000	10.4	5,200	8	1
	4	5	10	50,000	12.0	6,000	7	6

2×10^6 MEFs from TNG or 129 mice (passage number is shown) were transfected using the indicated amount of linearized DNA in each experiment. mOrange-positive cell number at day 2 was estimated from the collected cell number from one well and the percentage of mOrange-positive cells. The number of Nanog–GFP colonies from TNG MEFs or Nanog-positive colonies (by immunofluorescence) from 129 MEFs at day 28 on irradiated MEFs (γ MEFs) or on gelatin is shown.

–, no data; *, average of two wells; †, average of three wells.

that the expression level of *c-Myc* and *Klf4* is lower than that of *Oct4* and *Sox2* in ES cells ($c-Myc < Klf4 < Oct4 \approx Sox2$, quantified by real-time polymerase chain reaction (PCR), data not shown). The total amount of the four reprogramming factor proteins was also comparable to that in ES cells in all cell lines examined (Supplementary Fig. 4a). In conjunction with *Oct4* and *Sox2*, 13 additional pluripotent markers, which are highly enriched in both ES cells and iPS cells^{1,4,24}, were reactivated in all cell lines screened (Fig. 1b). Bisulphite genomic sequencing analyses revealed that the promoter region of *Oct4* and *Nanog* is highly unmethylated in the iPS cells (Supplementary Fig. 4c). These results indicate that efficient reprogramming had occurred due to the non-viral single vector.

The number of vector-integration sites was analysed by Southern blotting in cell lines imO1–imO8, as well as in five TNG iPS cell lines, TNGimO1–TNGimO5 (strategy shown in Fig. 2a). Of the 13 cell lines, imO7 and TNGimO5 were found to have a single-vector integration in one site, indicating that a stable single-copy insertion of the non-viral expression vector can achieve direct reprogramming, effectively minimizing genome modification in iPS cells (Fig. 2b). We did not find reprogrammed cell lines without vector integration. The

integration sites of imO1, imO3 and imO7, as well as the tandem repeats in imO1 and imO3, were identified by inverse PCR and confirmed by genomic PCR (Fig. 2c and Supplementary Fig. 5). All integration sites of the individual cell lines were different, supporting the notion that reprogramming does not depend on vector integration at specific loci⁸. Expression of the genes at the integration sites was detectable in MEFs, ES cells and all examined iPS cell lines, without significant influence imposed by vector integration (Supplementary Fig. 5d). Adult cells (footpad fibroblasts) were also successfully reprogrammed by means of definite single-copy vector integration (Supplementary Fig. 6).

To remove the exogenous reprogramming factors, we performed transient *Cre* transfection and tracked the loss of fluorescent reporter expression. Many of the mOrange-negative colonies started to differentiate after *Cre* transfection (Supplementary Fig. 7a). This excision-mediated differentiation was prevented by culturing the cells in the presence of the fibroblast growth factor (Fgf) receptor inhibitor PD173074, which inhibits mouse ES cell differentiation²⁵ (Fig. 3a). The percentage of differentiated colonies after *Cre*-mediated excision differed among clones, possibly reflecting variable stability of the

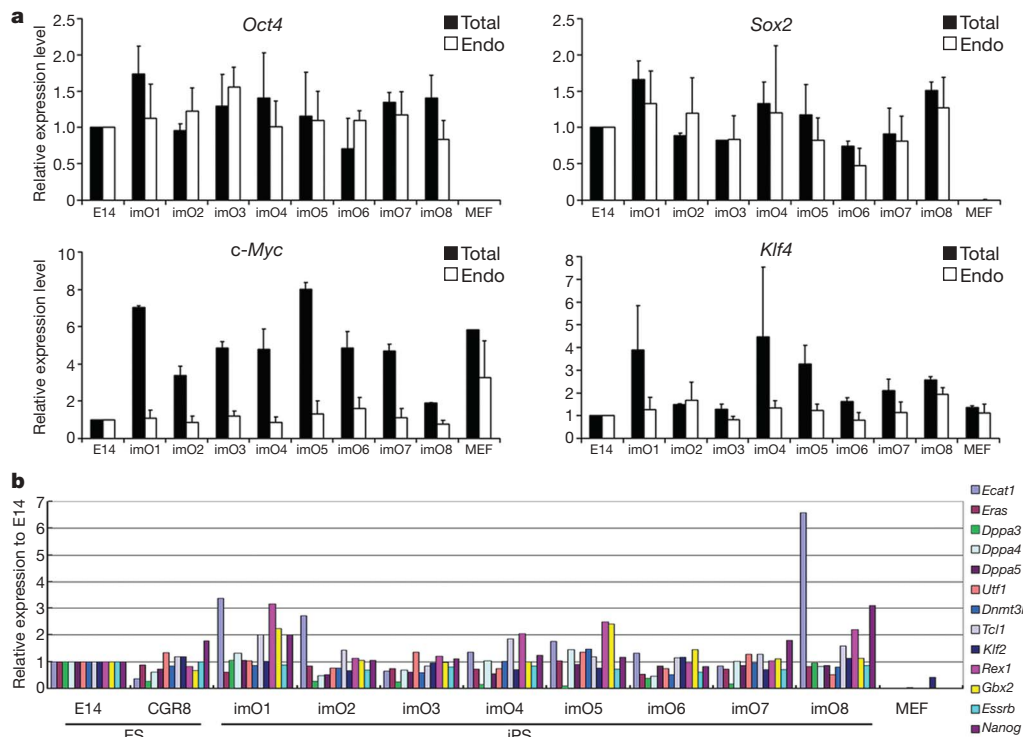


Figure 1 | Efficient reactivation of pluripotent markers in iPS cells generated by a non-viral multiprotein expression vector. **a**, Quantitative PCR for total and endogenous *c-Myc*, *Klf4*, *Oct4* and *Sox2* expression. Data are shown as relative expression to an ES cell line, E14Tg2a (E14). Error bars indicate the s.d. generated from triplicates. **b**, Quantitative PCR for

pluripotent markers. Two independent ES cell lines, E14Tg2a (E14) and CGR8, were analysed together with iPS cell lines. Data are shown as relative expression to E14Tg2a, and represent one of two independent experiments. *Ecat1* (also known as 2410004A20Rik), *Dppa5* (*Dppa5a*), *Rex1* (*Zfp42*).

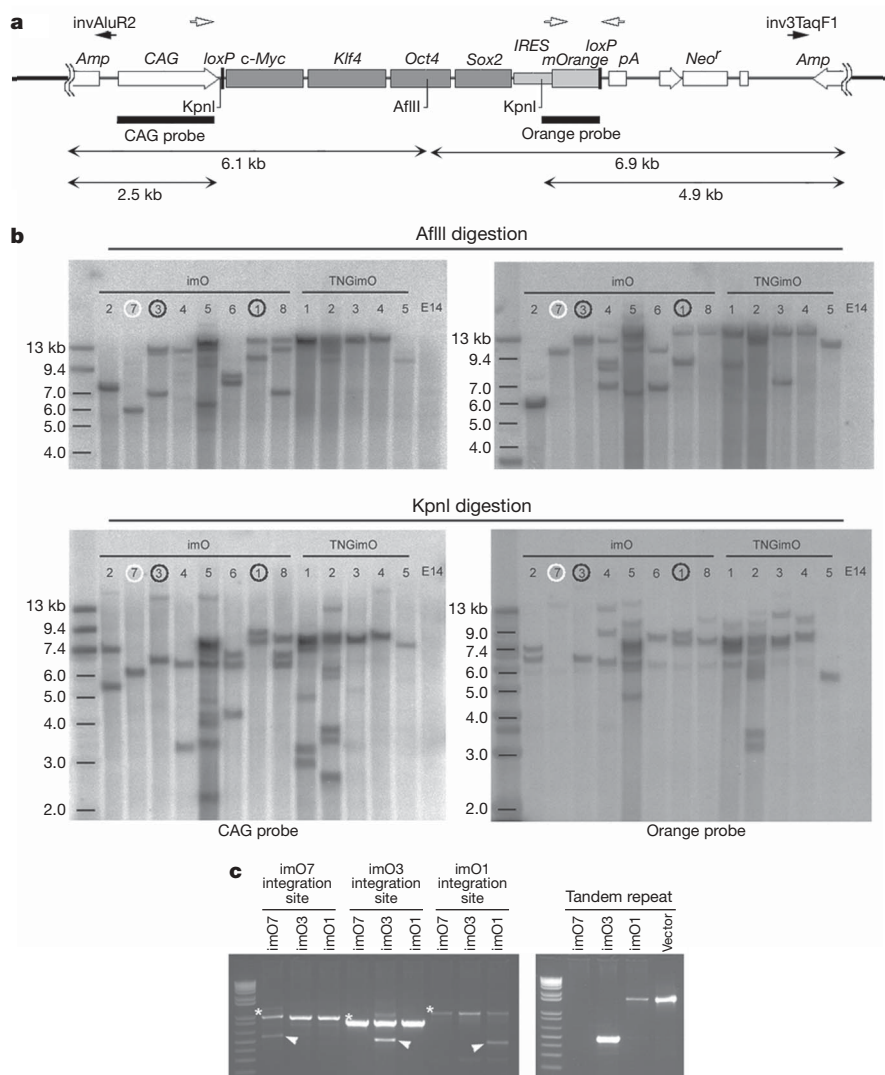


Figure 2 | Non-viral iPS cells with a single-vector integration site.

a, Schematic diagram of restriction enzyme sites, AflIII and KpnI, in pCAG2LMKOSimO and two probes, CAG probe and Orange probe, used for Southern blotting (bars). Black arrows (*invAluR2*, *inv3TaqF1*) indicate the position of primers used to detect tandem repeat integration in **c**. *pA*, *Amp* and *Neo^r* represent polyA signal, ampicillin-resistant gene and neomycin-resistance gene, respectively. **b**, Southern blotting analysis for the AflIII- (top) and KpnI- (bottom) digested genome of imO1–imO8 and TNGimO1–TNGimO5 using the CAG probe (left panels) and Orange probe

(right panels). **c**, Validation of the integration site and tandem repeat integration. A single integration site of imO7 (white circle in **b**) and a single integration site with tandem integration (same orientation) of imO1 and imO3 (black circles in **b**) was identified by inverse PCR (data not shown) and validated by genomic PCR. Asterisks, band from wild-type allele. Arrowheads, integration-site-specific band. Detail of the integration site is shown in Supplementary Fig. 5. The different band size of tandem repeats is caused by vector degradation accompanied by random integration.

reprogrammed state (Fig. 3b). All tested reprogramming-cassette-excised clones, which were generated from imO2, imO3, imO7 and TNGimO5 and expanded in the presence of PD173074, could subsequently maintain an undifferentiated morphology for at least five passages after removal of PD173074 (data not shown). The presence of PD173074 may allow cells to adapt upon rapid loss of exogenous reprogramming factors at single colony density. Whereas Cre-excised cell lines from imO7 and TNGimO5 could be maintained over 20 passages while maintaining an undifferentiated morphology, cell lines derived from imO2 and imO3 were prone to generate more differentiated cells at later passages in the absence of PD173074 (data not shown). The endogenous gene expression of *c-Myc*, *Klf4*, *Sox2* and *Oct4* was maintained in the Cre-excised cell lines imO2Ec3 derived from imO2 with additional constitutive enhanced GFP (eGFP) expression (imO2E1), imO3c8 and imO7c8, which are derived from imO3 and imO7, respectively (Supplementary Fig. 8a). Expression of other pluripotency genes was also sustained (Supplementary Fig. 8b), indicating that our single-vector system has enabled complete

elimination of exogenous genes without disturbing maintenance of the iPS cell state.

We next examined the differentiation ability of the cell lines before and after reprogramming-cassette excision. Cre-excised cell lines imO3c8 and imO7c8 downregulated the pluripotency markers *Oct4*, *Nanog* and *Rex1* and upregulated markers of all three germ layers in embryoid bodies, thus behaving in the same way as ES cells²⁶ (Supplementary Fig. 9a). Furthermore, β -tubulin-positive neurons were generated efficiently in a monolayer neural differentiation protocol²⁷ (Supplementary Fig. 9b). Differentiation of the factor-retained parental cell lines imO3 and imO7 was less efficient (Supplementary Fig. 9a, b). Both imO7 and imO7c8 produced teratocarcinomas when transplanted to the kidney capsule, although imO7 tumours contained more undifferentiated cells than did imO7c8 tumours (Fig. 3c). Finally, we injected the reprogramming-vector-free iPS cells into C57BL/6 blastocysts (results summarized in Supplementary Table 2). To identify chimaeric embryos, two cell lines with constitutive eGFP expression, imO3Ec5 and imO7Ec3, were used, and high-contribution

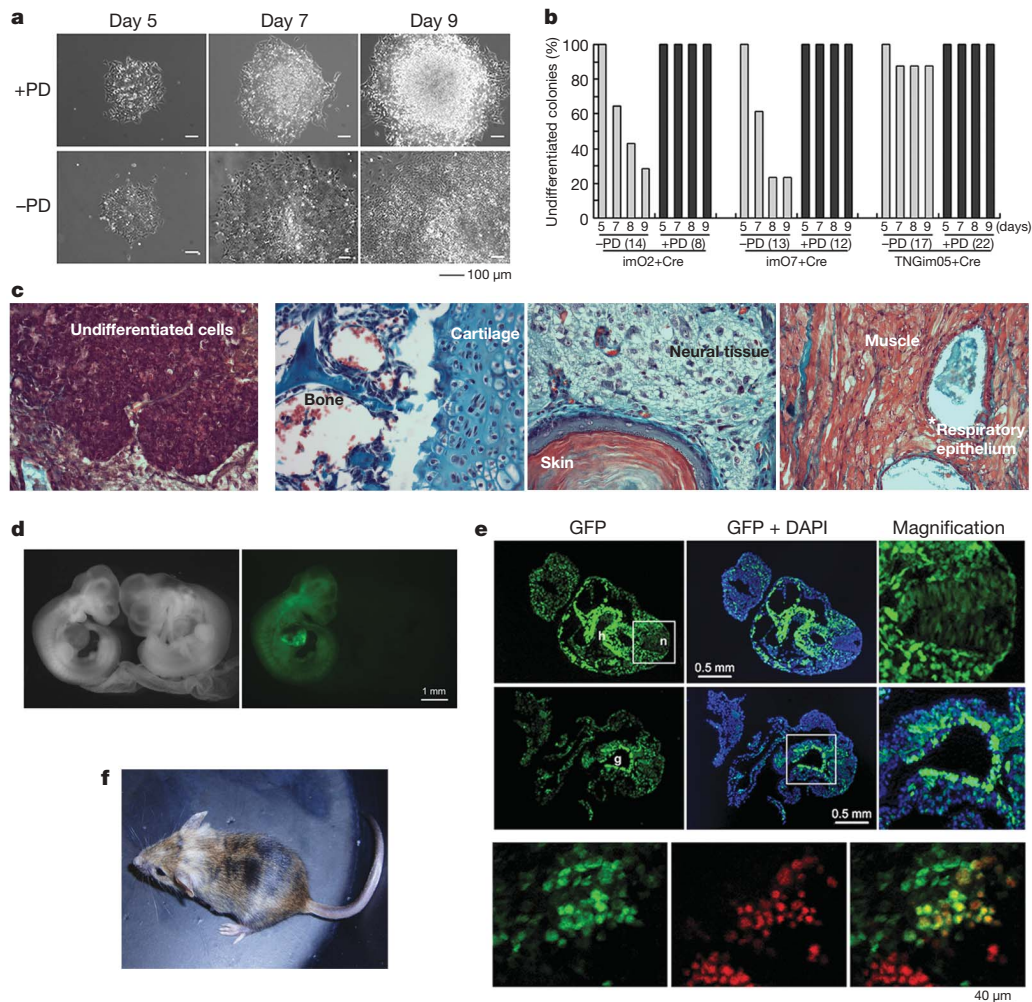


Figure 3 | Reprogramming cassette excision and pluripotency of the non-viral iPS cells. **a**, Many of the undifferentiated colonies at 5 days post *Cre* transfection differentiated by day 9 in the absence of the fibroblast growth factor receptor inhibitor PD173074 (bottom). **b**, Percentage of reprogramming-cassette-free undifferentiated colonies in the absence ($-PD$) and presence ($+PD$) of PD173074. Numbers of monitored reprogramming-cassette-excised colonies are indicated in parentheses. Experiments were performed in three cell lines, imO2, imO7 and TNGimO5. **c**, Undifferentiated cells in imO7 teratoma (left) and various tissues in

chimaeric embryos were generated at high frequency from both cell lines (Fig. 3d and Supplementary Table 2). Contribution to all three germ layers—ectoderm, mesoderm and endoderm—was observed, in addition to Oct4-positive germ cells in the 12.5 days post coitum (d.p.c.) genital ridge (Fig. 3e). Three *Cre*-excised iPS cell lines, imO1c5, imO7c8 and TNGimO3cC5, gave rise to live chimaeras, indicating that iPS cells derived with the non-viral multiprotein expression vector are genuinely pluripotent (Fig. 3f).

To address the reprogramming ability of the non-viral single-vector system in human cells, we enhanced stable transfection efficiencies using a *piggyBac* (PB) transposon gene-delivery system¹⁵. Co-transfection of two PB transposons carrying a dox-inducible *MKOS-IRES-geo* cassette and a constitutively active *CAG-rtTA* transactivator construct¹⁶ was applied to human embryonic fibroblasts (Supplementary Fig. 10a). In this design, on genomic integration, the two transposons allow dox-inducible activation of *MKOS* expression in wild-type cells. We observed iPS-like colony formation 14 days post transfection (d.p.t.) when the cells were maintained in human ES cell culture conditions supplemented with dox. In total, 15 colonies were picked from 14–25 d.p.t. from 4 wells of 6-well plates, initially containing either

imO7c8 teratoma (the other three panels). **d**, imO3Ec5-derived chimaeric embryos (green) at 10.5 d.p.c. **e**, Transversal sections at 9.5 d.p.c. (upper six panels) and genital ridge at 12.5 d.p.c. (bottom panels). imO3Ec5 contributed to ectoderm (neural tissue, n; the boxed regions are magnified on the right), mesoderm (heart; h), endoderm (gut, g; the boxed regions are magnified on the right) and germ cells (red staining with anti-Oct4 antibody). DAPI, 4,6-diamidino-2-phenylindole. **f**, An adult imO7c8-derived chimaeric mouse.

3.2×10^4 or 6.4×10^4 fibroblasts per well. When dox was withdrawn on 32 d.p.t., 3 clones successfully propagated as stable cell lines, whereas the exogenous factor expression was uninduced as indicated by negative staining for lacZ activity (Supplementary Fig. 10b). All clonal lines displayed human ES cell morphology (Supplementary Fig. 10b) and were positive for alkaline phosphatase (not shown). Robust expression of endogenous pluripotency markers SSEA4, NANOG, TRA-160 and TRA-181 was confirmed in all three cell lines (Supplementary Fig. 10c). These results demonstrate that the non-viral single-vector system can reprogram human fibroblasts, and strongly indicate that it can also be applied to the production of exogenous-factor-free, non-viral human iPS cells.

We demonstrate that a single non-viral vector with 2A-peptide-linked reprogramming factors can achieve reprogramming efficiently and that the exogenous reprogramming factors can be completely removed from the iPS cells using subsequent *Cre* transfection. Absolute avoidance of unpredictable exogenous reprogramming factor reactivation is important not only for clinical applications but also for drug screening, because some small molecules affect epigenetic genome modification²⁸, which could cause unexpected reactivation

and unreliable screening results. Cre recombinase boasts the most efficient *in vivo* and *in vitro* recombination system currently known, and a part of the vector backbone remaining in the integration site following Cre-mediated factor deletion would be a tolerable remnant for *in vitro* iPS applications. We also demonstrated the single-vector reprogramming system combined with a PB transposon delivery system for human cell reprogramming. PB transposons are completely removable from their integration site without any residual change in the original DNA sequence^{15,16}. This PB-based single-vector reprogramming system will enable the generation of non-genetically modified human iPS cells as shown in the mouse¹⁶, which is ideal for regenerative medicine.

METHODS SUMMARY

A vector, pCAG2LMKOSimO, which has *c-Myc*, *Klf4*, *Oct4* and *Sox2* coding regions linked with 2A peptide sequences driven by CAG enhancer/promoter^{18,21}, was constructed as described in the Methods. The vectors were introduced into MEFs using Nucleofector II (Amaxa) and cells were cultured in ES cell culture conditions for up to four weeks. Colonies showing ES-cell-like morphology were picked and cultured on either irradiated MEFs (γ MEFs) or gelatin after trypsinization. Gene expression was analysed by quantitative PCR. Protein expression was analysed by immunoblotting. Integration number of the vector and the integration site in the established cell lines were analysed by Southern blotting and inverse PCR, respectively. The reprogramming cassette was excised by Cre transient transfection in the presence or absence of a fibroblast growth factor receptor inhibitor PD173074 (100 ng ml⁻¹), and pluripotency of the cell lines was examined *in vitro* (embryoid body formation, neural differentiation) and *in vivo* (teratoma formation, blastocyst injection).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 December 2008; accepted 12 February 2009.

Published online 1 March 2009.

- Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
- Wernig, M. *et al.* *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318–324 (2007).
- Maherali, N. *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1**, 55–70 (2007).
- Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
- Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
- Park, I. H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
- Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
- Aoi, T. *et al.* Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* **321**, 699–702 (2008).
- Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Biotechnol.* **26**, 101–106 (2008).
- Wernig, M., Meissner, A., Cassidy, J. P. & Jaenisch, R. *c-Myc* is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* **2**, 10–12 (2008).
- Hochedlinger, K., Yamada, Y., Beard, C. & Jaenisch, R. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* **121**, 465–477 (2005).
- Foster, K. W. *et al.* Induction of KLF4 in basal keratinocytes blocks the proliferation–differentiation switch and initiates squamous epithelial dysplasia. *Oncogene* **24**, 1491–1500 (2005).
- Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. & Hochedlinger, K. Induced pluripotent stem cells generated without viral integration. *Science* **322**, 954–959 (2008).
- Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. Generation of mouse induced pluripotent stem cells without viral vectors. *Science* **322**, 949–953 (2008).
- Wang, W. *et al.* Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl Acad. Sci. USA* **105**, 9290–9295 (2008).
- Woltjen, K. *et al.* piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* doi:10.1038/nature07863 (this issue).
- Hasegawa, K., Cowan, A. B., Nakatsuji, N. & Suemori, H. Efficient multicistronic expression of a transgene in human embryonic stem cells. *Stem Cells* **25**, 1707–1712 (2007).
- Szymczak, A. L. *et al.* Correction of multi-gene deficiency *in vivo* using a single ‘self-cleaving’ 2A peptide-based retroviral vector. *Nature Biotechnol.* **22**, 589–594 (2004).
- Sommer, C. A. *et al.* iPS cell generation using a single lentiviral stem cell cassette. *Stem Cells* doi:10.1634/stemcells.2008-1075 (2008).
- Carey, B. W. *et al.* Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc. Natl Acad. Sci. USA* **106**, 157–162 (2009).
- Niwa, H., Yamamura, K. & Miyazaki, J. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**, 193–199 (1991).
- Yeh, E. *et al.* A signalling pathway controlling *c-Myc* degradation that impacts oncogenic transformation of human cells. *Nature Cell Biol.* **6**, 308–318 (2004).
- Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).
- Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).
- Kunath, T. *et al.* FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**, 2895–2902 (2007).
- Kaji, K. *et al.* The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nature Cell Biol.* **8**, 285–292 (2006).
- Ying, Q. L., Stavridis, M., Griffiths, D., Li, M. & Smith, A. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nature Biotechnol.* **21**, 183–186 (2003).
- Brueckner, B., Kuck, D. & Lyko, F. DNA methyltransferase inhibitors for cancer therapy. *Cancer J.* **13**, 17–22 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Tsakiridis for advice on using the 2A peptide sequence and A. Nagy for providing data from his laboratory on generating human reprogrammed cell lines by combining the PB transposon system and the 2A-sequence-joined reprogramming factors (or MKOS single vector reprogramming system) as well as for his comments on the manuscript. We also thank I. Chambers for providing TNG MEFs and for discussion and comments on the manuscript, V. Wilson for advice on teratoma analysis, T. Kunath, S. Lowell, C. Blackburn and K. Vintersten for discussions and comments on the manuscript, and B. Hendrich for permission to start preliminary experiments of this work in his laboratory. We also thank J. Ure, L. Robertson, R. McLay and R. Wilkie for technical assistance, and Biomed unit staff for mouse husbandry. K.K. is the recipient of MRC career development fellowship in stem cell research. A.P. is the recipient of a BBSRC CASE PhD studentship. M.M., P.M. and K.W. were supported by grants from the Canadian Stem Cell Network and Juvenile Diabetes Research Foundation.

Author Contributions K.K. conceived the study, designed and executed the experiments, interpreted data and wrote the manuscript. K.N. performed immunoblotting and karyotype checking, and assisted with manuscript preparation. A.P. performed real-time PCR and *in vitro* differentiation experiments. M.M. generated human reprogrammed cells. P.M. performed immunostaining for human reprogrammed cells. K.W. constructed the PB/MKOS system, assisted human cell reprogramming experiments and analysis, and helped to prepare the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.K. (keisuke.kaji@ed.ac.uk).

METHODS

Plasmid construction. The sequences of all primers used for plasmid construction are listed in Supplementary Table 3. The *c-Myc* coding region plus *F2A* 5' 50 bp (*c-Myc-F2A* 5'), the *F2A* 3' 54 bp plus *Klf4* coding region plus *T2A* (*F2A* 3'-*Klf4-T2A*), the *Oct4* coding region plus *E2A* 5' 50 bp (*Oct4-E2A* 5'), and the *E2A* 3' 50 bp plus *Sox2* coding region (*E2A* 3'-*Sox2*) fragments were separately amplified by PCR using mouse ES cell complementary DNA as a template. *c-Myc-F2A-Klf4-T2A* fragment was amplified by PCR using EcoRI *Koz* 5'-*Myc* and *Klf4*-3' GSG *T2A* Xho primers after annealing and extension of *c-Myc-F2A* 5' and *F2A* 3'-*Klf4-T2A* fragment taking advantage of the complimentary *F2A* peptide region. The *Oct4-E2A-Sox2* fragment was amplified by PCR using Xho5'-*Oct4* and *Sox2*-3' EcoRI Xho primers after annealing and extension of *Oct4-E2A* 5' and *E2A* 3'-*Sox2* fragment. The individual fragments were subcloned into pTOPO-bluntII (Invitrogen), and the *Oct4-E2A-Sox2* fragment was transferred into a XhoI site located after *c-Myc-F2A-Klf4-T2A*, resulting in *c-Myc-F2A-Klf4-T2A-Oct4-E2A-Sox2* (MKOS) in pTOPO-bluntII. pCAGMKOSiE was constructed by inserting the MKOS fragment into the EcoRI site of pCIG2, which is generated by replacing CMV promoter of pIRES2-EGFP (BD Biosciences Clontech) with CAG enhancer/promoter²¹. The *IRES* fragment was amplified by NotIRES and mOrangeIRES primers using pIRES2-eGFP as a template. The *mOrange* fragment was amplified by IRESmOrange and XbaBamloxPmOrange primers using pmOrange (BD Biosciences Clontech) as a template. The *IRES-mOrange-loxP* fragment (imOloxP) was amplified using these fragments and the flanking primers, and subcloned into pTOPO-bluntII. MKOS was inserted into the EcoRI site of pcDNA3 (Invitrogen), and then annealed KpnBamloxPBamX F/R oligonucleotides was inserted before *c-Myc*, and imOloxP was inserted after *Sox2* (pCMV2LMKOSimO). pCAG2LMKOSimO was constructed by transferring a BamHI–BamHI fragment containing MKOSimO flanked by two *loxP* sites from pCMV2LMKOSimO into the pCAGDNA3 vector, which was made by replacing its CMV promoter with CAG enhancer/promoter²¹. Full sequence information of the vector and PCR conditions is available on request.

Southern blotting. Ten micrograms of genomic DNA digested with AflIII or KpnI was separated in 0.8% of agarose gel in TAE buffer, and hybridized with ³²P-labelled probes in PerfectHyb Plus Hybridization Buffer (Sigma) at 65 °C after transfer onto Hybond-XL (GE Healthcare). The expected band size from tandem integration is: same orientation, <13.0 kb, and opposite orientation, <13.8 kb or <12.2 kb in AflIII digestion; same orientation, <6.4 kb, and opposite

orientation, <9.8 kb or <5.1 kb, in KpnI digestion using either the CAG or Orange probe.

Nucleofection and reprogramming efficiency estimation. MEF Nucleofector Kit 2 (Amaxa) and program T-20 were used for nucleofection according to the manufacturer's instructions. We transfected 2×10^6 MEFs with 2–10 µg of linearized DNA, and 10^5 cells were seeded in 6-well plates on either γMEFs or gelatin in the presence of LIF. Forty-eight hours later, viable cells were collected from one gelatin well, and after counting cell number the mOrange-positive population was measured by flow cytometry (CyAn ADP, Dako). mOrange-positive cell numbers per well were estimated using the results in each experiment as shown in Table 1. The rest of the wells were kept for 4 weeks, with medium changed every 2–3 days. Nanog-GFP-positive colonies from TNG MEFs were picked and seeded on γMEFs after trypsinization. Colonies from 129 MEFs were fixed with 4% PFA and stained with anti-Nanog antibody (ab21603, Abcam) followed by Alexa-Fluor-488-conjugated anti-rabbit IgG antibody (Molecular Probes).

Reprogramming cassette excision. Cells were cultured in the presence and absence of 100 ng ml^{-1} PD173074 (Sigma) for 24 h before *Cre* transfection with Lipofectamine2000 (Invitrogen). Forty-eight hours after *Cre* transfection, cells were collected and seeded at 10^3 cells per 10 cm dish with gelatin in the presence or absence of PD173074. Undifferentiated mOrange-negative colonies were marked 5 days post *Cre* transfection, and morphology of the colonies was monitored to day 9. Colonies were picked between days 12 and 14, propagated in 96-well plates, and reprogramming cassette excision in each clone was verified by genomic PCR using CAG F, IRES mOrange and BGH R primers (Supplementary Fig. 8b, c; primer sequences are in Supplementary Table 4). Reprogramming-cassette-positive colonies by genomic PCR, which had undetectable mOrange expression by microscopy, were excluded from the total colony numbers indicated in parentheses in Fig. 3b. Differentiated colonies, which failed to grow after being picked, were considered as reprogramming-cassette-excised colonies and counted in the total colony numbers.

Quantitative PCR. Each quantitative PCR used 1% of cDNA made from 1 µg of total RNA in Light-Cycler 480 system with SYBR Green (Roche). See Supplementary Table 5 for the gene-specific primers. Data are shown as relative expression to an ES cell line, E14Tg2a, after normalization with TATA binding protein (*Tbp*) expression. Details of PCR conditions are available on request. Data represent one of two independent experiments. Error bars in *c-Myc*, *Klf4*, *Oct4* and *Sox2* PCR represent standard deviation of duplicate samples.