Molecular Human Reproduction, Vol.20, No.6 pp. 538–549, 2014
Advanced Access publication on February 5, 2014 doi:10.1093/molehr/gau012

molecular human reproduction
MHR | basic science of reproductive medicine

ORIGINAL RESEARCH

ONSL and OSKM cocktails act synergistically in reprogramming human somatic cells into induced pluripotent stem cells

Laura Jung[1,†], Philippe Tropel[1,10,†], Yohann Moal[2], Marius Teletin[1,3], Eric Jeandidier[4], Régis Gayon[2], Christian Himmelspach[4], Fiona Bello[1], Cécile André[1], Adeline Tosch[1], Ahmed Mansouri[5,6,7], Catherine Bruant-Rodier[8], Pascale Bouillé[2], and Stéphane Viville[1,9,*]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de Santé et de Recherche Médicale (INSERM) U964, Centre National de Recherche Scientifique (CNRS) UMR1704, Université de Strasbourg, Illkirch 67404, France [2]Vectalys SAS CanalBiotech II, 3 rue des Satellites, Toulouse 31400, France [3]Service de Biologie de la Reproduction, Centre Hospitalier Universitaire, Strasbourg F-67000, France [4]Service de Génétique, Centre Hospitalier de Mulhouse 20, rue du Dr Laennec, Mulhouse Cedex 68070, France [5]Department of Molecular Cell Biology, Molecular Cell Differentiation, Max-Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Goettingen, Germany [6]Department of Clinical Neurophysiology, University of Goettingen, Goettingen, Germany [7]Genome and Stem Cell Center, GENKOK, Erciyes University, Kayseri, Turkey [8]Service de Chirurgie Plastique Reconstructrice et Esthétique, Chirurgie B – Hôpital Civil, Strasbourg F-67000, France [9]Centre Hospitalier Universitaire, Strasbourg F-67000, France [10]Present Address: Aix Marseille Université, INSERM, GMGF UMR_F 910, 13385 Marseille, France

* Correspondence address. E-mail: viville@igbmc.fr

Submitted on June 17, 2013; resubmitted on December 13, 2013; accepted on January 9, 2014

ABSTRACT: The advent of human induced pluripotent stem cells (hiPSC) is revolutionizing many research fields including cell-replacement therapy, drug screening, physiopathology of specific diseases and more basic research such as embryonic development or diseases modeling. Despite the large number of reports on reprogramming methods, techniques in use remain globally inefficient. We present here a new optimized approach to improve this efficiency. After having tested different monocistronic vectors with poor results, we adopted a polycistronic cassette encoding Thomson's cocktail OCT4, NANOG, SOX2 and LIN28 (ONSL) separated by 2A peptides. This cassette was tested in various vector backbones, based on lentivirus or retrovirus under a LTR or EF1 alpha promoter. This allowed us to show that ONSL-carrier retrovectors reprogrammed adult fibroblast cells with a much higher efficiency (up to 0.6%) than any other tested. We then compared the reprogramming efficiencies of two different polycistronic genes, ONSL and *OCT4*, *SOX2*, *KLF4* and *cMYC* (OSKM) placed in the same retrovector backbone. Interestingly, in this context ONSL gene reprograms more efficiently than OSKM but OSKM reprograms faster suggesting that the two cocktails may reprogram through distinct pathways. By equally mixing RV-LTR-ONSL and RV-LTR-OSKM, we indeed observed a remarkable synergy, yielding a reprogramming efficiency of >2%. We present here a drastic improvement of the reprogramming efficiency, which opens doors to the development of automated and high throughput strategies of hiPSC production. Furthermore, non-integrative reprogramming protocols (i.e. mRNA) may take advantage of this synergy to boost their efficiency.

Key words: human iPSC cells / reprogramming / retrovirus / lentivirus / OSKM, ONSL cocktails

Introduction

The possibility to reprogram differentiated somatic cells into pluripotent stem cells, so-called induced pluripotent stem cells (iPSC), is currently revolutionizing experimental approaches in many research avenues such as development (Zhu and Huangfu, 2013) and cancer researches (Bernhardt *et al.*, 2012), cell biology (Ramalho-Santos, 2009), reproductive biology (Woods and Tilly, 2012), human pathology modeling (Onder and Daley, 2012), drug discovery assays (Bellin *et al.*, 2012) and human cell therapy strategies (Daley and Scadden, 2008). All of these areas previously suffered from the lack of pertinent human models. Recently, iPSC technology has opened new fields of research that were hardly accessible

† Equal first author.

with embryonic stem cells (ESC), mainly because of poor availability of human embryos and of ethical dilemmas. iPSC relieve most of these obstacles. Accordingly, interest in iPSC was illustrated by the 2012 Nobel Prize being awarded to Professor Shinya Yamanaka, only 6 years after his inaugural publication (Takahashi and Yamanaka, 2006).

In their initial work, Takahashi and Yamanaka identified a minimal set of factors, i.e. OCT4 (also known as Pou5f1), SOX2, KLF4 and cMYC (also known as MYC) (OSKM) that were able to successfully reprogram mouse embryonic and adult fibroblasts into embryonic stem (ES)-like cells. Mouse iPSCs share mESC cells markers and properties: they express OCT4, NANOG, SSEA1 and alkaline phosphatase (AP), and can differentiate *in vitro* (embryoid bodies) or *in vivo* (teratomas), giving rise to tissues derived from the three embryonic germ layers. Mouse iPSC can also participate in chimera formation after injection into host blastocysts and colonize the germinal lineage, allowing germline transmission of iPSC genetic traits (Takahashi and Yamanaka, 2006; Okita *et al*., 2007). In 2007, two groups independently reported human fibroblasts reprogramming using either the same original cocktail OSKM or a new set of reprogramming factors including OCT4, NANOG, SOX2 and LIN28 (ONSL cocktail) (Takahashi *et al*., 2007; Yu *et al*., 2007). Human iPSCs share all human ESC characteristics, i.e. they express pluripotent markers such as OCT4, NANOG, SOX2, TRA-1-60, TRA-1-81 and AP and are able to form embryoid bodies and teratomas where all three germ layers derivatives can be found, fulfilling gold standard definition of human pluripotent cells (Maherali and Hochedlinger, 2008).

The initial methods for reprogramming factors delivery were based on integrative virus-derived vectors like retrovectors (Takahashi *et al*., 2007) and lentivectors (Yu *et al*., 2007). In addition to a low efficiency, a major drawback of such methods is that they permanently integrate into the host genome. The latter event could potentially disrupt important sequences like open reading frame or regulatory regions, resulting in insertional mutations (Daley and Scadden, 2008; Robinton and Daley, 2012). In addition, the permanent presence of the transgenes in genomic DNA could also cause residual reprogramming factor expression either in iPSC or in their progeny. Both types of artifacts may affect iPSC quality and differentiation potential (Yu *et al*., 2007) or even result in tumorigenesis (Okita *et al*., 2007), precluding any usage in cell therapy protocols, one of the major challenges of the iPSC field.

In order to prevent such mutagenesis effects, many efforts have been made to improve the general efficiency of the reprogramming process and to develop non-integrative methods. One of the first improvements was the construction of polycistronic genes, allowing the efficient synthesis of all reprogramming factors from one mRNA (Carey *et al*., 2009; Sommer *et al*., 2009). Such synthetic genes are based on the use of internal ribosome entry sequences and/or of 2A peptides (de Felipe, 2002; Szymczak and Vignali, 2005). Currently, most of the reprogramming technologies are based on such a polycistronic strategy, which has clearly brought a simplification, but not dramatic improvement, in the procedure efficiency.

Alternative technologies avoiding genomic integration issues have been developed (Gonzalez *et al*., 2011; Robinton and Daley, 2012): (i) excisable vector using either Cre-loxP system (Somers *et al*., 2010) or piggyBac transposon-based vectors (Kaji *et al*., 2009; Woltjen *et al*., 2009); (ii) non-integrative viruses, such as adenovirus (Stadtfeld *et al*., 2008; Zhou and Freed, 2009) or Sendaï virus (Fusaki *et al*., 2009); (iii) DNA delivery, such as serial plasmid vectors transfection (Si-Tayeb *et al*., 2010), auto-replicative episome vectors (Yu *et al*., 2009), minicircle DNA (Jia *et al*., 2010); (iv) mRNA (Warren *et al*., 2010; Yakubov *et al*., 2010; Mandal and Rossi, 2013) or miRNA (Anokye-Danso *et al*., 2011; Miyoshi *et al*., 2011) transfections; and (v) protein delivery (Kim *et al*., 2009; Zhou *et al*., 2009).

So far, none of those techniques brings together all expected features of an 'ideal' technique (Robinton and Daley, 2012). Therefore, we considered that there was still work to be accomplished in setting up an efficient non-integrative reprogramming process. We reasoned, as others (Yu *et al*., 2009), that the first hurdle to be passed concerns the efficiency and that it will be worthwhile to transpose an efficient integrative system to a non-integrative system. In addition, this will open the possibility to develop automation of the process. We therefore decided to first concentrate our efforts on the improvement of integrative protocols.

To this aim, we compared the reprogramming efficiency of two polycistronic cocktails ONSL (Yu *et al*., 2007) and OSKM (Takahashi *et al*., 2007) placed in the same vector backbones, based on lentivirus (LV-EF1) or retrovirus (RV-LTR). We showed that the best option was a vector derived from a retrovirus backbone driving the polycistronic cassette under its own internal LTR promoter, and that the ONSL polycistronic cassette always gave a much better efficiency compared with the OSKM cassette. When mixed, ONSL and OSKM polycistronic genes synergized the reprogramming process, allowing it to reach an efficiency of over 2% of the treated cells.

Materials and Methods

Monocistronic and polycistronic lentiviral plasmids

Three plasmids were used in order to produce recombinant lentiviral particles. A first plasmid pLVgagpol (p8.74) provides a nucleic acid encoding wild-type endogenous viral gag and pol genes lacking vif, vpr, vpu and nef genes, but including rev. A second plasmid pVSVG (pMDG) provides a nucleic acid encoding the vesicular stomatitis virus envelope glycoprotein (VSV-G). A third pLV-EF1-Gene of interest (GOI) self-inactivating plasmid (U3 region of the LTR was deleted as shown in Supplementary data, Fig. S1), produced in an endofree method, provides the woodchuck hepatitis post-transcriptional regulatory element (WPRE) and the cPPT/CTS sequence under control of the human elongation factor 1 alpha promoter (EF1alpha). According to the experiments the GOIs refer to polycistronic genes *OCT4*-P2A-*NANOG*-F2A-*SOX2*-T2A-*LIN28* (ONSL) or to the monocistronic genes *OCT4*, *SOX2*, *NANOG* and *LIN28*.

The polycistronic cassette: ONSL sequence is given in Supplementary data, Fig. S2. Construction of plasmid vector containing polycistronic ONSL under control of the EF1alpha promoter was generated after MluI/PmeI cloning from a lentiviral vectors backbone (pLV-EF1-MCS) produced by Vectalys as described in Supplementary data, Fig. S3. All constructs were generated using unique restriction sites after amplification by PCR (Supplementary data, Table S1). Constructions of plasmid vectors containing monocistronic genes *OCT4, SOX2, NANOG* and *LIN28* under control of the EF1alpha promoter, were generated according to the same strategy (for PCR primers see Supplementary data, Table S1).

Monocistronic and polycistronic retroviral plasmids

Three plasmids were used in order to produce recombinant retroviral particles. A first plasmid pRVgagpol (pMNgag-pol) provides a nucleic acid encoding viral gag and pol genes. A second plasmid pVSVG (pMDG) provides a nucleic acid encoding the vesicular stomatitis virus envelope glycoprotein

(VSV-G). A third pRV-Gene of interest (GOI) was produced in an endofree method, under control of MoMLV wild-type endogenous LTR (Supplementary data, Fig. S4) or under control of the human elongation factor 1 alpha promoter (EF1alpha). MoMLV wild-type endogenous LTR was deleted into plasmid carrying the EF1alpha promoter as shown in Supplementary data, Fig. S5. According to the experiments the GOIs refer to polycistronic genes ONSL or *OCT4*-P2A-*SOX2*-T2A-*KLF4*-E2A-*cMYC* (OSKM) or to the monocistronic genes *OCT4*, *SOX2*, *NANOG*, *LIN28*, *KLF4* or *cMYC*.

The polycistronic cassette: OSKM was produced by GeneArt (Regensburg, Germany) according to Carey *et al*. (2009) with a blunt 5′ extremity and a BamH1 3′ extremity. OSKM coding sequence is given in Supplementary data, Fig. S6.

Constructions of plasmid vectors containing polycistronic cassette ONSL or OSKM under the control of MoMLV wild-type endogenous LTR were generated after BamH1/NaeI cloning from a retroviral vector backbone (pRV-MoMLV LTR-MCS) produced by Vectalys as described in Supplementary data, Fig. S4. Constructions of plasmid vectors containing monocistronic genes *OCT4*, *NANOG, SOX2*, *LIN28, KLF4* or *cMYC* under control of MoMLV wild-type endogenous LTR were generated according to the same strategy using MfeI/NaeI restriction sites.

Constructions of plasmid vectors containing polycistronic cassette ONSL or OSKM under control of the EF1alpha promoter were generated after BamH1/MluI or BamH1/NaeI cloning, respectively, from a retroviral vectors backbone (pRVsin-EF1-MCS) produced by Vectalys as described in Supplementary data, Fig. S7. All constructs were generated using unique restriction sites after amplification by PCR (primers are given in Supplementary data, Table S1).

Production of monocistronic and polycistronic retrovirus and lentivirus

All viral vectors were designed and produced by Vectalys (Toulouse, France). Viral vectors were produced in the human embryonic kidney HEK293T cell line. HEK293T cells were used to seed a 10-layer CellSTACK (6320 cm^2, Corning, Tewksbury, MA, USA) and were transfected 2 days later, in fresh DMEM without FCS supplemented with 1% penicillin/streptomycin and 1% ultraglutamine (PAA Laboratories GmbH, Pasching, Austria). Cells were simultaneously transfected with three plasmids: pVSV-G, pGagPol, pGOI. The supernatant was discarded 24 h after transfection, and was replaced with fresh non-supplemented DMEM. The vectors were harvested twice during the 24 h after transfection. The harvested vectors were clarified by centrifugation for 5 min at 3000*g*, followed by microfiltration through a sterile filter unit with 0.45 μm pores (Stericup, EMD Millipore Corporation, Billerica, MA, USA). The crude vector preparation was concentrated and purified by tangential flow ultrafiltration. The supernatant was then diafiltered against DMEM. Once the diafiltration was complete, the retentate was recovered and further concentrated by ultrafiltration.

Functional particle quantification by qPCR

HCT116 cells were used to seed 96-well plates at a density of 12 500 cells per well, in 250 μl of DMEM supplemented with 10% FCS, 1% penicillin/streptomycin and 1% ultraglutamine (complete medium). Five serial dilutions were performed 24 h later, with complete medium, for each vector sample and an rLV-EF1-GFP internal standard. The cells were transduced with these serial dilutions in the presence of 8 μg/ml Polybrene (Sigma-Aldrich Co., St. Louis, MO, USA). For each sample series, one well of non-transduced cells was included as a control. Four days after transduction, the cells were released by trypsin treatment and harvested by centrifugation and each cell pellet was resuspended in 250 μl of PBS. The titer was calculated by determining the number of integrated genome per milliliter (IG/ml) by qPCR. The final titer was then converted in transducing units per milliliter (TU/ml) using an internal standard whose titer was previously determined by FACS. The titer in TU/ml is then used to determine the right volume of each viral vector required to achieve the targeted multiplicity of infection (MOI) according to the following equation:

$$\text{Viral vectors volume required (}\mu\text{l)} = \frac{\text{Number of cells seeded}}{\text{Viral vectors titer (TU/ml)}} \times \text{M.O.I} \times 1000$$

In this equation, the MOI is defined as the ratio of viral particles to target cells. As an example, an MOI of 5 represents 5 times more viral particles than target cells.

Cell culture and viral infection

All cultures were performed at 37°C, under 5% CO_2 atmosphere. Primary dermal fibroblasts were established and maintained on gelatin-coated dishes in DMEM 1 g/l glucose Glutamax 1 × supplemented with Antibiotic Antimycotic 1 × and 0.1 mM non-essential amino acids (all from Invitrogen, Life Technologies, Carlsbad, CA, USA) and 10% fetal bovine serum (FBS; Sigma-Aldrich Co.).

Primary dermal fibroblast cell line, called CPRE2, was established from a skin biopsy obtained after an abdominoplasty of a 38 years old woman suffering from obesity, after the informed consent of the patient. The tissue was sequentially treated with trypsin 0.25% for 30 min and then overnight with collagenase/dispase in basic DMEM medium. Epidermis was mechanically removed and isolated dermal fragments were seeded on gelatin-coated dishes. Dermal fibroblasts were then selected upon cultivation. Passage 5 cells were preferably used for reprogramming procedure. The project was submitted to and approved by the local hospital ethical committee, *Comité de Protection de le Personne* of Strasbourg university Hospital.

On Day 0, a constant number (10^5) of dermal fibroblasts were seeded per 35 mm dish coated with 0.1% gelatin. On Day 1, cell infection was performed into a final volume of 1 ml of fibroblast medium containing 8 μg/ml of polybrene (Sigma-Aldrich Co.). On Day 2, medium was replaced with fresh medium. At Day 3, infected cells were transferred onto a 100-mm dish containing 10^6 feeder cells (passage 3 mitomycin-C treated mouse embryonic fibroblasts). From Day 6 to 9, fibroblast medium was progressively switched to human induced pluripotent stem cell (hiPSC) medium (KO-DMEM, 20% KOSR, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1 × Penicillin-Streptomycin, 0.1 mM β-mercaptoethanol; all from Life Technologies, Carlsbad, CA, USA) supplemented with 10 ng/ml of bFGF (R&D Systems, Minneapolis, MN, USA). hiPSC clones were individually picked at Week 4 and expanded on matrigel coated 35 mm-dish (BD Biosciences, San Jose, CA, USA) in mTeSR1 medium (Stemcell Technologies, Vancouver, BC, Canada) for further characterization.

All polycistronic reprogramming experiments were performed in triplicate, namely the infection of three cell dishes, each containing 10^5 cells. Two infected cell dishes were maintained until the end of the reprogramming protocol for hiPSC detection, and reprogramming efficiency calculation. The third infected cell dish was used to purify either mRNA or proteins 3 days post-infection. The monocistronic reprogramming strategy was performed in duplicate. One infected cell dish was maintained until the end of the reprogramming protocol and hiPSC clones were counted. The second infected cell dish was used to purify mRNA and proteins 3 days post-infection.

AP staining

To detect AP activity, dishes were rinsed twice with PBS and then fixed with cold methanol for 2 min and dried. hiPSC colonies were stained with BCIP/NBT (Sigma-Aldrich Co.) for 20 min according to manufacturer instructions.

Karyotype analysis

For karyotype analysis, hiPSC cells were first treated with colchicine (Sigma-Aldrich Co.) for 4 h. After trypsination, cells were shocked with hypotonic KCl 0.075 M solution for 20 min at 37°C. Then, cells were fixed in methanol:acetic acid solution (3:1) and conserved at −20°C until analysis. Conventional cytogenetic has been performed applying RHG banding on metaphase chromosomes. Chromosome examination was carried out on Axio Plan 2 microscope (Zeiss, Ulm, Germany) equipped with the Ikaros 2 software (Metasystems, Althusheim, Germany). For each clones, 30 metaphases were analyzed except for one (RV ONSL+OSKM 2) for which only 25 metaphases have been observed. Karyotypes were described accordingly to the ISCN 2013 (International Standing Committee on Human Cytogenetic Nomenclature *et al*., 2013) at a resolution of 300 bps.

FACS analysis of pluripotent markers

Markers surface expression was analyzed using anti-Tra-I-60, anti-Tra-I-81 and anti-SSEA4 antibodies (EMD Millipore Corporation). Mouse IgG3 and IgM isotypes were provided by Abcam, and FITC labeled anti-mouse polyvalent Ig-G, A, M secondary antibodies by Sigma. Flow cytometry analysis of surface markers was performed on accutase resuspended cells in PBS BSA 1% Sodium azide 0.1%. Both primary and secondary antibody incubation were carried out at room temperature for 30 min. Analysis was performed on a FACS Calibur flow cytometer (BDIS, San Jose, CA, USA) using the CellQuest acquisition and analysis software (BDIS). Final data and graphs were analyzed and prepared in the FlowJo software (Tree Star, Inc., Ashland, OR, USA).

Quantitative RT–PCR

Total RNA was prepared with RNeasy Mini Kit (Qiagen) with on-column DNase I digestion. For each sample, 1 μg total RNA was reverse-transcribed by random priming using Superscript II (Invitrogen, Life Technologies). qPCR reaction was carried on using SYBR® green JumpStart™TaqReadyMix™ (Sigma-Aldrich Co.) and LightCycler 480 (Roche, Basel, Switzerland). The efficiency and specificity of each primer pair was checked using a cDNA standard curve. All samples were normalized to endogenous *GAPDH* expression. To check the expression of the transgenes in hiPSC clones, cells at Day 3 post-infection were used as positive control. Overlapping primers, i.e. forward primer in the first gene and reverse primer in the second gene, were designed to discriminate polycistronic gene expression from the endogene counterpart. Primers sequences are listed in Supplementary data, Table S1.

Quantification of integrated copy number

Total DNA was extracted using the NucleoSpin Tissue (Macherey-Nagel, Dueren, Germany) according to supplier recommendations. The transgene copy number of lentiviral or retroviral vectors was determined by quantitative PCR analysis. Real-time PCR was carried out with 150 ng of total DNA diluted in RNase-Free water, SYBR GreenER (Invitrogen, Life Technologies), specific primers binding to the WPRE sequence and albumin (housekeeping gene) in a final volume of 20 μl, using Step One Real-Time PCR System (Applied Biosystems, Life Technologies, Carlsbad, CA, USA). The DNA copy numbers were calculated by referring the Ct values for each sample to a standard plasmid curve. All qPCRs were performed in duplicate and normalized by gDNA sample coming from monoclonal cell line (HCT116-GFP) containing one copy number of GFP.

DNA methylation of *OCT4* promoter

Genomic DNA was bisulfite treated using the EpiTect bisulfite kit (Qiagen). 0.5 to 1 μg of gDNA was treated and 50 ng of treated DNA was used for each PCR reaction using previously published primers and conditions (Primer couples 3 and 9 for analysis of *OCT4* promoter and enhancer (Freberg *et al*., 2007)). PCR products were purified using a PCR purification Kit (Qiagen), cloned into a Topo TA vector (Invitrogen, Life Technologies) and sequenced (ATGC biotech, Konstanz, Germany).

Embryoid body formation and differentiation markers analysis

For embryoid body formation, hiPSC were dissociated with dispase solution (1 mg/ml, Stemcell Technologies, Vancouver, BC, Canada), resuspended in 1 ml of Aggrewell medium (Stemcell Technologies, Vancouver, BC, Canada) containing 2 μM Y27632 (Stemgent, Cambridge, MA, USA), centrifuged in Aggrewell plates for 3 min at 80*g* and further incubated at 37°C for 24 h. The next day, embryoid bodies were retrieved and transferred in one well of a 6-well low binding plate (Nunc, Waltham, MA, USA) in 3 ml of Aggrewell medium. The following days, medium was progressively switched to KO-DMEM, 20% FBS, 2 mM L-glutamine, 0.1 mM Non-Essential Amino Acids, 1× penicillin–streptomycin (all from Invitrogen, Life Technologies). After 30 days, embryoid bodies were collected and total RNA and cDNA were prepared as described in previous paragraph. The expression of 90 validated genes associated with stem cell pluripotency and differentiation to all three germ layers were analyzed using the TaqMan Human Stem Cell Pluripotency Array (Applied Biosystems, Life Technologies) according to manufacturer instructions. Briefly, cDNA samples from hiPSC and corresponding embryoid bodies (collected at the same passage) were loaded on the same fluidic card according to manufacturer instructions. The fluidic card was run on an Applied Biosystems 79000HT machine with the SDS2.4 software (Applied Biosystems, Life Technologies). Raw data (.sds) were exported on the SDS RQ manager software (Applied Biosystems, Life Technologies) for analysis. For analysis settings, Ct was fixed manually with a threshold at 0.2. Analyzed results data were exported to .txt format for further analysis in Data Assist software (Applied Biosystems, Life Technologies). Maximum allowable Ct value was fixed at 35 and endogenous control genes were selected according to manufacturer instructions (18S-Hs99999901_s1, ACTB-Hs99999903_m1, CTNNB1-Hs00170025_m1, EEF1A1-Hs00742749_s1, GAPD-Hs99999905_m1, RAF1-Hs00234119_m1) for ΔCt calculation. Data were graphically represented through heat map using Pearson's correlation as distance measurement, average linking as clustering method and assay centric as map type.

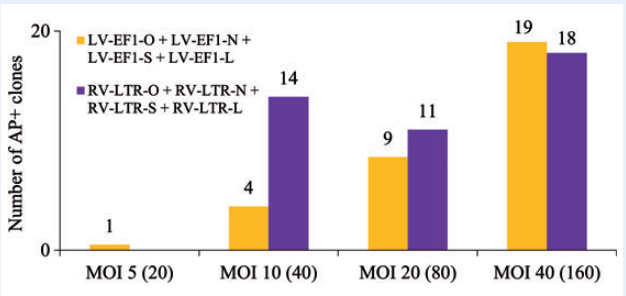

Figure 1 Monocistronic gene-carrying retro- and lentiviruses reprogram poorly dermal fibroblasts. Number of alkaline phosphatase positive clones stained 30 days post-infection of 10^5 CPRE2 (one dish counted). For each condition, four monocistronic viruses were used at the same MOI, i.e. MOI 5, MOI 10; MOI 20 or MOI 40. The total MOI corresponds to the sum of the four monocistronic viruses MOIs (indicated in brackets), i.e. MOI 20, MOI 40, MOI 80, MOI 160. LV-EF1-O: lentivirus carrying *OCT4* gene; LV-EF1-N: lentivirus carrying *NANOG* gene; LV-EF1-S: lentivirus carrying *SOX2* gene; LV-EF1-L: lentivirus carrying *LIN28* gene; RV-LTR-O: retrovirus carrying *OCT4* gene; RV-LTR-N: retrovirus carrying *NANOG* gene; RV-LTR-S: retrovirus carrying *SOX2* gene; RV-LTR-L: retrovirus carrying *LIN28* gene.

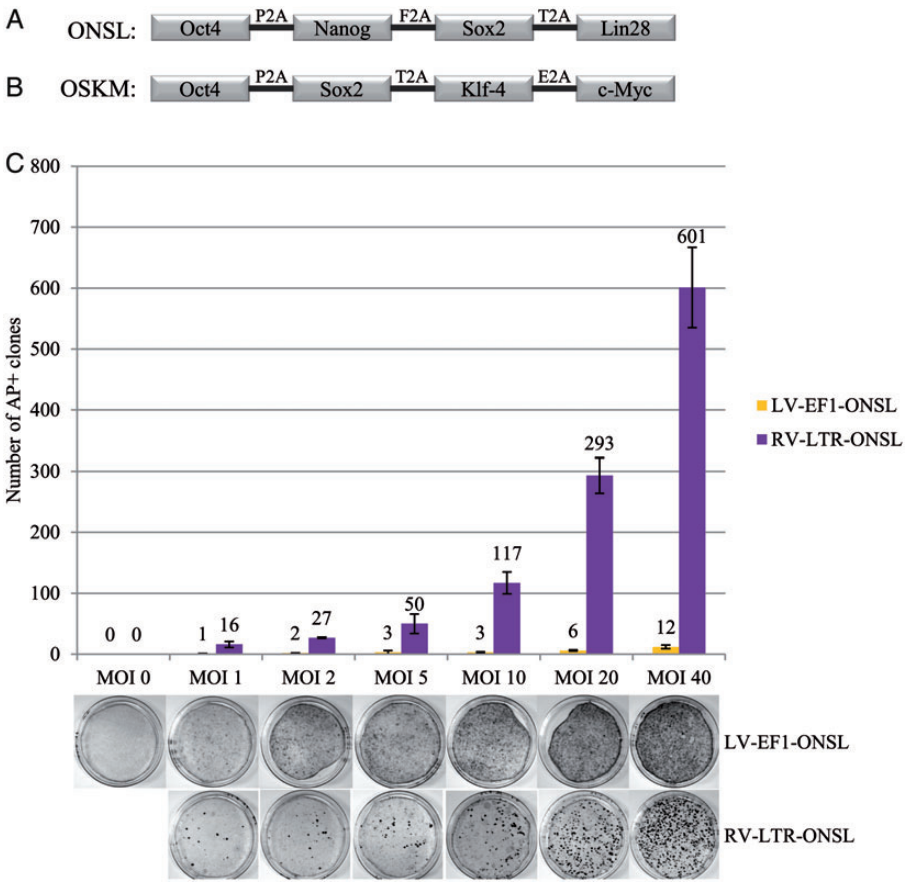

Figure 2 ONSL polycistronic gene reprograms fibroblasts with various efficiency depending on MOI and virus backbone. (**A**) ONSL cassette consisting of a single reading frame of human *OCT4*, *NANOG, SOX2*, and *LIN28* coding sequences linked by P2A peptides (porcine teschovirus-1), F2A (foot-and-mouth disease virus) and T2A (Thosea asigna virus) sequences, respectively. (**B**) OSKM cassette consisting of a single reading frame of human *OCT4, SOX2, KLF4* and *cMYC* coding sequences linked by P2A peptides (porcine teschovirus-1), T2A (Thosea asigna virus) and E2A (equine rhinitis A virus (ERAV) 2A) sequences, respectively. (**C**) Number of alkaline phosphatase positive clones stained 30 days post-infection of 10^5 CPRE2 and the corresponding stained B100 dishes. Data correspond to average from duplicate and error bars represent standard deviation.

Teratoma formation

For *in vivo* teratoma formation, cells from one Matrigel coated 60 mm-dish were collected by dispase treatment and resuspended in 75 µl of KO-DMEM, mixed with 75 µl of Matrigel (BD Biosciences, San Jose, CA, USA) and kept on ice. The hiPSC-Matrigel mixture was injected subcutaneously in 8-week-old NOD/SCID female mice (Charles River Laboratory, Wilmington, MA, USA). Two mice were injected for each hiPSC clone. After 6 weeks to three months, teratomas were dissected and fixed in formalin. Samples were embedded in paraffin and processed with hematoxylin and eosin staining at the histology laboratory of the Institute Clinique de la Souris (Illkirch, France).

Results

Monocistronic viral strategies yield poor reprogramming efficiency

In the prospect of deriving hiPSCs from adult patients, we decided to establish our own adult human primary skin fibroblasts cell line CPRE2 as a working model, instead of performing experiments on fetal model cells like BJ or IMR90. In a first attempt to establish efficient reprogramming protocols, we compared the ONSL cocktail efficiency when transfected by either four individual lentivectors, or retrovectors. We focused on this reprogramming gene set primarily because of the absence of the proto-oncogenes *KLF4* and *cMYC*. We compared lentivectors to retrovectors showing specific levels of purification and concentration adapted to primary cells and carrying the same four genes.

For both type of vectors we tested, on a constant number of cells (10^5), different multiplicities of infection (MOI) from 5 to 40 for each of the four individual vectors, corresponding to a total MOI of 20–160 (Fig. 1). Whatever the vector, we observed almost no colonies at MOI 5 (total MOI 20) and a slight increase in colony number when MOI increased up to 40 (total MOI 160). This poor efficiency (0.02%) was visually observed on a regular basis in our hiPSC derivation activity from normal (Lapillonne *et al.*, 2010) or genetic disease-carrying cells (Hick *et al.*, 2013) confirming previously published data (Takahashi *et al.*, 2007; Yu *et al.*, 2007). During these experiments, we noticed that cytotoxicity increased with MOI, which may explain why efficiency was poorly improved when increasing MOI. At this stage, vector

backbone did not influence reprogramming efficiency since colony numbers obtained with retro- and lentivectors were quite similar.

Modified polycistronic gene reprograms fibroblasts efficiently in retrovirus but not lentivirus backbone

In order to limit the MOI, we adopted a new strategy based on a polycistronic gene allowing expression of the four reprogramming factors from the same mRNA. For this purpose, we synthesized a polycistronic gene carrying *OCT4*, *NANOG*, *SOX2* and *LIN28*, in this order, separated by 2A peptide derived sequences (*OCT4*-P2A-*NANOG*-F2A-*SOX2*-T2A-*LIN28*). Encoding sequence was optimized for human codon usage (Fig. 2A, Supplementary data, Fig. S2). This new synthetic gene (later on called ONSL) was tested for reprogramming when inserted in lentivector (LV-EF1-ONSL) and retrovector (RV-LTR-ONSL). The reprogramming efficiencies of both vectors were tested by infecting CPRE2 cells with an adapted purification level and an increasing MOI, from 1 to 40, respectively.

LV-EF1-ONSL poorly reprogrammed, with a slight increase of AP^+ colonies number up to 12 at MOI 40 (Fig. 2C). In contrast, RV-LTR-ONSL showed a drastic improvement in reprogramming efficiency with up to 600 AP^+ colonies, for 10^5 infected cells, at MOI 40 (0.6% efficiency). By repeating the experiments, using the same original cells, we could show that the process is highly reproducible, with an efficiency rate going from 0.3 to 0.6% (Figs 2C and 4A). In addition to quantitative improvement, we also observed general background decrease making hiPSC colonies identification much easier (Fig. 2C).

In order to check the quality of the transduction performed by lenti- and retrovectors, we replaced the ONSL cassette by a GFP gene in the retrovector (RV-LTR-GFP) and in the lentivector (LV-EF1-GFP). Three days post-transduction, using a range of MOI from 1 to 40, CPRE2 fibroblasts were analyzed by flow cytometry (Fig. 3). At the level of the transduction rate, RV-LTR-GFP and LV-EF1-GFP performed almost equally 95 and 91%, respectively, whatever the MOI. In contrast, RV-LTR-GFP performed better than LV-EF1-GFP relating to two parameters: (i) the mean intensity is higher with RV-LTR-GFP than with LV-EF1-GFP and (ii) transduced cell populations are more homogenous with RV-LTR-GFP than with LV-EF1-GFP. Those data support the idea that retrovector leads to a globally more efficient expression of GOI than lentivectors.

Virus backbone architecture greatly influences reprogramming efficiency

Retro- and lentivectors differ in many points considering their gross structure, but the main differences focuses on security issues. For safety reasons, third-generation pSIN lentivectors carry mutated LTR impeding transcriptional activity. To express the gene of interest (GOI), heterologous regulatory sequences like EF1 alpha promoter were added. In contrast, Moloney-derived retrovectors allow LTR-driven expression of GOI. This major difference in vector architecture provides a first hypothesis to explain differences in GFP expression and in reprogramming.

To test this point, we produced RV-based vector harboring mutated LTR promoting sequences where transcriptional activity is supported by EF1alpha heterologous promoter (RV-EF1-ONSL), mimicking LV situation. When used in reprogramming experiment, RV-EF1-ONSL was 30 times less effective than 'original' RV-LTR-ONSL, producing no more AP+ colonies than LV-EF1-ONSL in other experiments (Fig. 4A). This result strongly suggests that LTR promoter presents crucial characteristics for successful reprogramming.

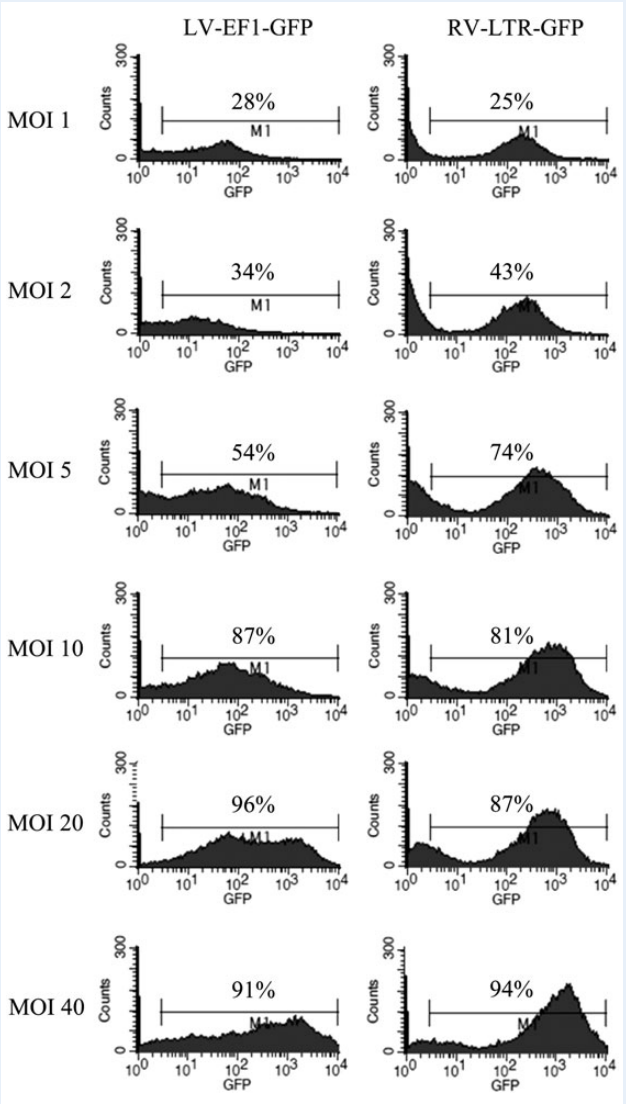

Figure 3 Comparison of LV-EF1-GFP and RV-LTR-GFP efficiency at 3 days post-infection. Diagrams correspond to FACS analysis of GFP expression for each condition of CPRE2 infection using LV-EF1-GFP or RV-LTR-GFP from MOI 1 to MOI 40. M1 corresponds to the GFP positive region, i.e. percentage of transfected cells indicated and was determined with GFP negative CPRE2. LV-EF1-GFP: lentivirus carrying GFP gene; RV-LTR-GFP: retrovirus carrying GFP gene.

ONSL and OSKM polycistronic genes differ in reprogramming efficiency and quality

Having shown the importance of regulatory sequences, we wondered how much the efficiency depends on the reprogramming cocktail nature, ONSL or OSKM. To test this, OSKM polycistronic gene was inserted in the RV vector under the control of either the endogenous LTR or the EF1alpha promoter, giving rise to RV-LTR-OSKM and RV-EF1-OSKM, respectively.

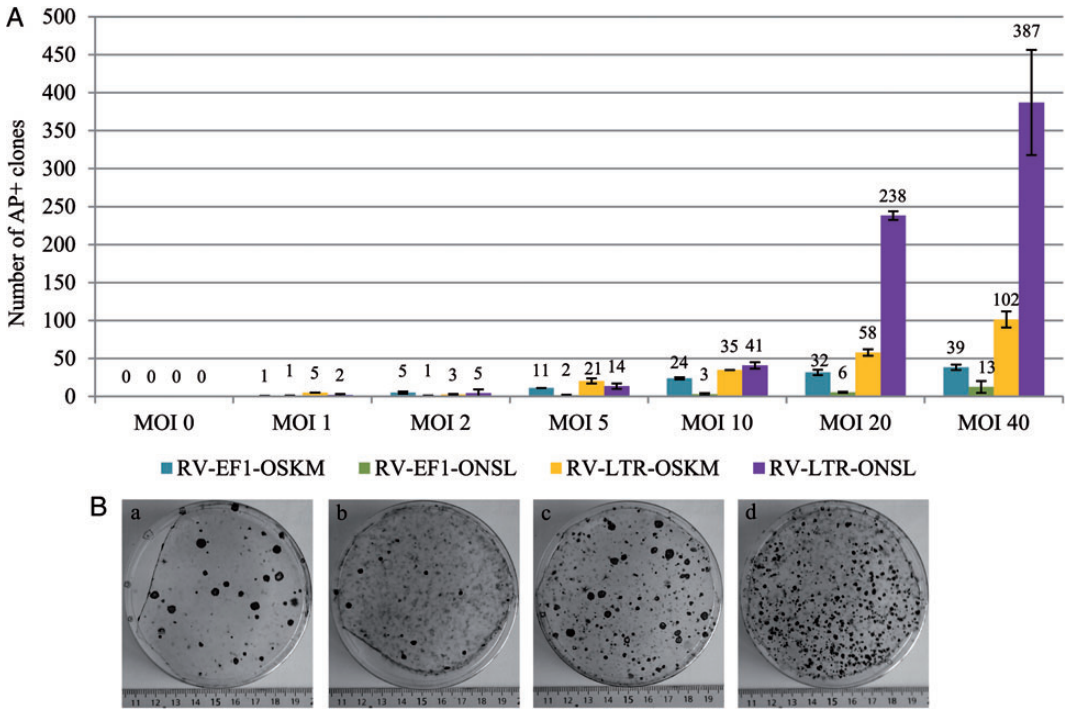

Figure 4 Comparison of ONSL and OSKM polycistronic efficiency when associated with classic (RV-LTR) or modified (RV-EF1) retrovirus backbone. (**A**) Number of alkaline phosphatase positive clones stained 30 days post-infection of 10^5 CPRE2. Data correspond to average from duplicate and error bars represent standard deviation. (**B**) Alkaline phosphatase stained B100 dishes for reprogramming with (a) RV-EF1-OSKM, (b) RV-EF1-ONSL, (c) RV-LTR-OSKM, (d) RV-LTR-ONSL all at MOI 40.

The two polycistronic genes OSKM and ONSL share two factors, i.e. *OCT4* and *SOX2*, and their general organization where reprogramming factors were separated by 2A peptides in both (Fig. 2A and B, Supplementary data, Figs S2 and S4). Transduction of CPRE2 fibroblasts with RV-LTR-OSKM yielded three to four times more AP^+ colonies than RV-EF1-OSKM (Fig. 4A), paralleling our results with RV-LTR-ONSL. Furthermore at MOI 20 or 40, RV-LTR-ONSL produced three to four times more AP^+ colonies than RV-LTR-OSKM (Fig. 4A) but RV-EF1-OSKM yielded three times more AP^+ colonies than RV-EF1-ONSL (Fig. 4A). Among vector-gene associations, RV-ONSL always harbors the highest reprogramming efficiency (Fig. 4A and B).

OSKM reprograms faster and can synergize with ONSL to improve reprogramming efficiency

When comparing ONSL and OSKM reprogramming, we observed that hiPSC colonies first appeared in RV-LTR-OSKM infected cells at Day 15, when no colony was detectable in RV-LTR-ONSL infected plates (Fig. 5A). Moreover, at Day 30, plates treated with RV-LTR-OSKM showed fewer but larger colonies than with RV-LTR-ONSL (Fig. 5B), both observations suggesting that reprogramming by the RV-LTR-OSKM occurred earlier after infection than with RV-LTR-ONSL.

These observations naturally led to the idea that they could synergize to reprogram target cells more efficiently. To challenge this hypothesis, we tested in parallel RV-LTR-OSKM and RV-LTR-ONSL in single infections at MOI 5, 10, 20 and 40 and in joint infections at MOI 2 × 5 (total 10) and 2 × 20 (total 40). When mixed, the two vectors showed a spectacular synergy leading to >1600 colonies at Day 15 and 2300 at Day 30 (MOI 2 × 20) for 10^5 transduced cells (efficiency of 2.3%; Fig. 5A and B). Confirming our previous results, at Day 15 post-infection no colony was detectable in RV-LTR-ONSL treated cells, while RV-LTR-OSKM treated cells produced up to 37 colonies per 10^5 cells at MOI 40 (Fig. 5A). At Day 30 post-infection, RV-LTR-ONSL treated cells produced up to 750 colonies per 10^5 cells at MOI 40, while, under the same conditions, RV-LTR-OSKM produced only 130 colonies.

Since the two vectors share two common genes, i.e. *OCT4* and *SOX2*, we wondered which of the non-common genes were important to synergize the reprogramming process. To test this, we sub-cloned into our RV vector *NANOG* (N), *LIN28* (L), *KLF4* (K) and *cMYC* (M) and used them to infect CPRE2 cells with the different possible combinations (ONSL+K; ONSL+M; ONSL+K+M; OSKM+N; OSKM+L; OSKM+N+L) and compared them to ONSL+OSKM. Each virus was used at MOI 5, thus combined infection of two viruses was carried out at a final MOI of 10 (2 × 5) and combination of three vectors was carried out at a final MOI of 15 (3 × 5). In order to check the effect of a dual infection, we also constructed an RV-LTR-GFP vector. Starting from RV-LTR-ONSL, we clearly observed an improvement when adding K or M or K+M, where K+M performs better than the two others, but never reached the efficiency of OSKM+ONSL combination (Fig. 6). Concerning the combinations with RV-LTR-OSKM, neither the addition of N or L or both N+L showed any improvement of the reprogramming efficiency (Fig. 6). The addition of RV-LTR-GFP had no effect (Fig. 6).

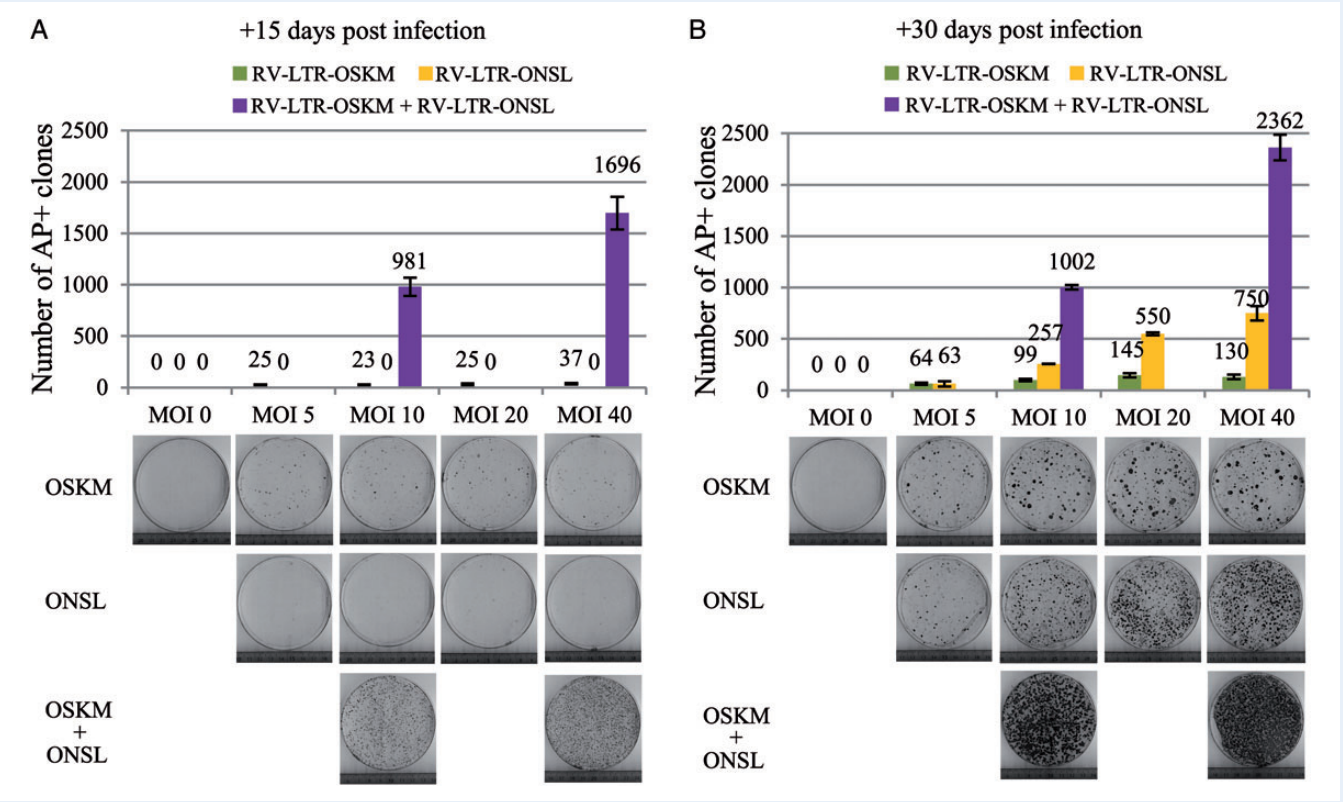

Figure 5 Comparison of RV-LTR-ONSL and RV-LTR-OSKM efficiency when used singly or mixed together at 15 days and 30 days post-infection of 10^5 CPRE2. Number of alkaline phosphatase positive clones stained at (**A**). Fifteen days post-infection (**B**). Thirty days post-infection and the corresponding stained B100 dishes. Data correspond to average from duplicate and error bars represent standard deviation.

Quality of hiPSC cells generated

For all reprogramming conditions tested during this work, clones were successfully picked-up, grown and cryopreserved. hiPSC cells were characterized according to international recommendations (Maherali and Hochedlinger, 2008), which included a morphology analysis, karyotyping, expression of pluripotent markers by RT–qPCR (*OCT4; NANOG; SOX2; LIN28*), by immunostaining and FACS analysis (SSEA4; Tra-I-60; Tra-I-81), transgene silencing, methylation status of both *OCT4* promoter and enhancer, *in vitro* differentiation with a RT–qPCR to analyze the expression of differentiation markers of the three germ layers and *in vivo* differentiation by teratomas formation followed by an histology analysis. All cell lines analyzed fully meet the conditions to define them as pluripotent (Fig. 7 and Supplementary data, Figs S8–14). The integrated copy numbers were analyzed for some of the clones. For all the cocktails tested, the results showed that the number of integration events remains low (between 1 and 2, Supplementary data, Table S2).

Discussion

Despite the vast amount of studies focusing on reprogramming methods, the major drawbacks have not been successfully bypassed (Gonzalez *et al.*, 2011). Indeed, methods remain inefficient, very costly and do not allow therapeutic usage or scale-up production and automation of the processes (Masip *et al.*, 2010; Gonzalez *et al.*, 2011). A recent paper claimed that, by Mbd3 depletion, the reprogramming efficiency of mouse fibroblast could reach 100%, but data on human adult cells remain partial and no number attesting to the efficiency was given at this stage (Rais *et al.*, 2013). We believe that effort is still needed to make a step forward to an efficient process that will then be easier to transpose to non-integrative, automatic methods.

Starting from monocistronic genes, we quickly switched to a polycistronic gene with which we tested different viral vectors. Whatever the condition, the vector ONSL cocktail as a polycistronic gene clearly performed better than monocistronic versions.

We could show that the ONSL polycistronic gene delivered by a retrovirus backbone is far more efficient than when delivered on our lentivirus backbone. Using GFP based vectors we also showed that the retrovirus vector carrying LTR promoter gives rise to a more homogenous transduced population and a higher level of GFP expression than the lentiviruses carrying an internal EF1alpha promoter. Moreover, when internal LTR is mutated and EF1alpha promoter inserted in RV backbone, the resulting vector falls in reprogramming efficiency. Others suggested a correlation between the level of expression and the efficiency of reprogramming (Hockemeyer *et al.*, 2008). Nevertheless, we cannot exclude, as recently suggested (Liu *et al.*, 2013), other parameters such as kinetics of expression driven by different promoting regions, kinetics of silencing, mRNA stability, etc. By switching from monocistronic genes to a polycistronic gene and by choosing the best viral vector, we improved the efficiency from 0.02 to 0.6%.

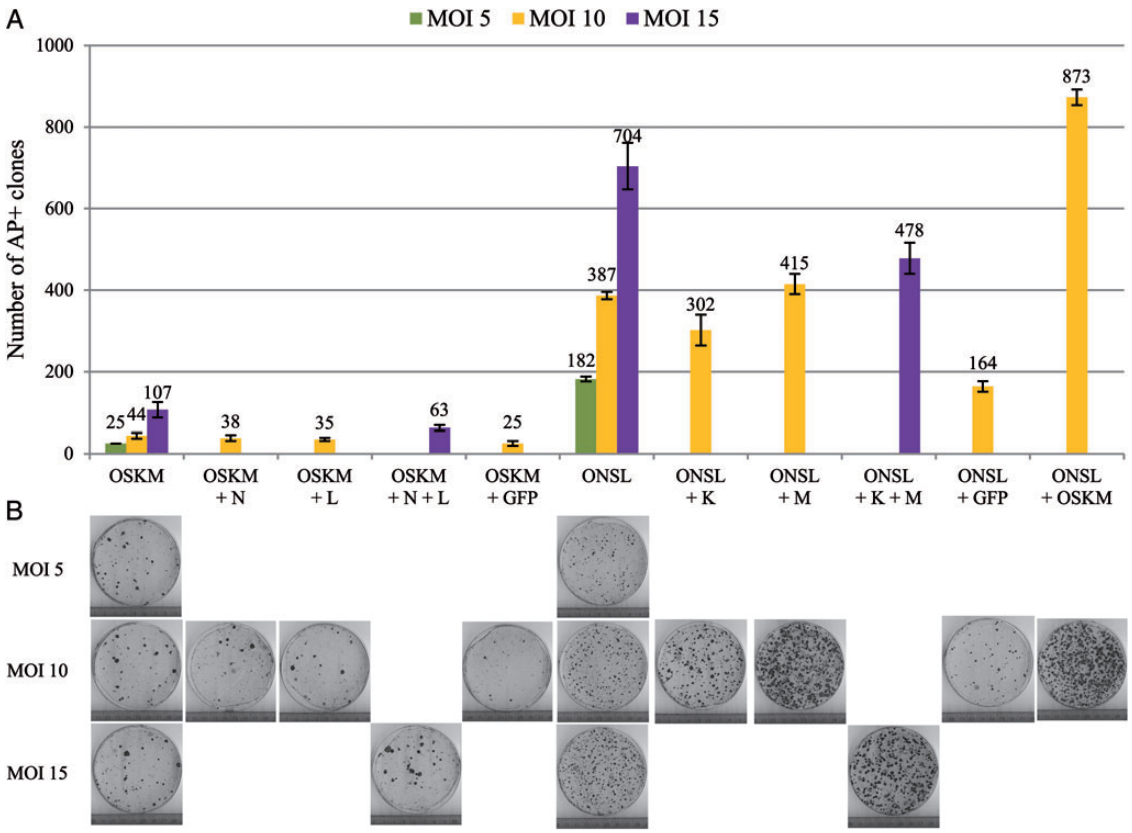

Figure 6 Comparison of reprogramming efficiency of multiple RV-LTR-ONSL and RV-LTR-OSKM combinations. Each polycistronic virus was tested in combination with one or two monocistronic viruses carrying another pluripotency factor, i.e. RV-LTR-ONSL was tested in combination with RV-LTR-K or RV-LTR-M or both RV-LTR-K + RV-LTR-M and RV-LTR-OSKM was tested in combination with RV-LTR-N or RV-LTR-L or both RV-LTR-K + RV-LTR-L. Each virus was used at MOI 5, thus combined infection of two viruses was performed at MOI 10 (2 × 5) and combination of three viruses at MOI 15 (3 × 5). To estimate the effect of the pluripotency factor addition to the starting polycistronic cocktail, we performed single infection of RV-LTR-ONSL and RV-LTR-OSKM at MOI 5, 10 and 15. To check the effect of dual infection, we used RV-LTR-GFP in combination with each polycistronic virus. The histogram represents the number of alkaline phosphatase positive clones stained 30 days post-infection of 10^5 CPRE2 and pictures are the corresponding stained B100 dishes for each condition. Data correspond to average from duplicate and error bars represent standard deviation. To increase the readability of this figure, only the symbols of the genes are indicated.

We started our initial tests by using the ONSL cocktail, and after having gained in efficiency we wondered how the OSKM cocktail could perform. In our hands, using the same retrovirus-based vector, ONSL always performed better than OSKM. However, during the process we noted two differences: (i) OSKM cocktail reprograms faster and (ii) gives rise to granular colonies as described by others as 'partially reprogrammed colonies' (Hockemeyer *et al.*, 2008; Maherali *et al.*, 2008; Nakagawa *et al.*, 2008), which prompted us to hypothesize that the two cocktails may act differently in the way they reprogram somatic cells. Such differences in reprogramming pathways are also suggested by recent results (Liu *et al.*, 2013). Moreover, Tanabe *et al.* (2013) clearly showed that iPSC derivation requires the completion of at least two successive phases, initiation and maturation, where reprogramming factors harbor independent and complementary functions regarding the aims. All those data may explain why mixing cocktails resulted in a synergistic improvement of three to four times of the efficiency. An improvement was already observed when *NANOG* and *LIN28* genes were added to OSKM (Warren *et al.*, 2012; Mandal and Rossi, 2013; Tanabe *et al.*, 2013), but this improvement never corresponded to a synergistic effect.

Considering that both cocktails have common factors, we tried to identify factors and/or combinations crucial for synergy. None of the combinations tested reached the efficiency level of ONSL+OSKM, even when KM was added to ONSL or NL to OSKM. This is surprising since the final composition is the same.

Studies have shown that the stoichiometry may greatly influence the reprogramming efficiency (Papapetrou *et al.*, 2009; Carey *et al.*, 2011; Yamaguchi *et al.*, 2011). Indeed, it was shown that an increased level of OCT4, up to three times, improved the reprogramming process (Papapetrou *et al.*, 2009) and that too much SOX2 would be detrimental for the reprogramming efficiency (Hockemeyer *et al.*, 2008; Yamaguchi *et al.*, 2011). In our case, we do not favor a dosage effect of OCT4 and SOX2. Since, the efficiency is much higher when we used ONSL+OSKM both at MOI 20 compared with ONSL or OSKM at MOI 40. So, we rather

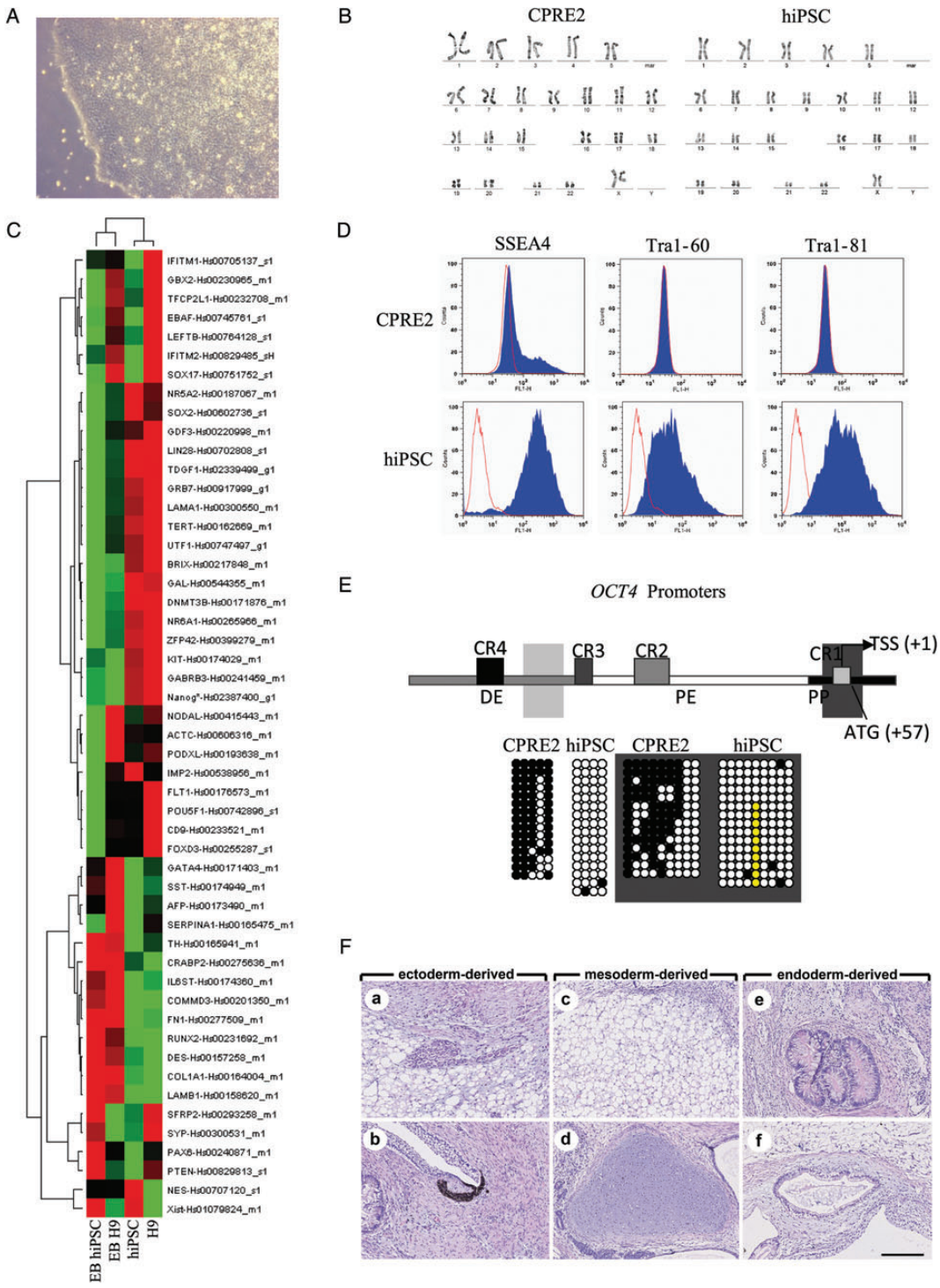

Figure 7 Pluripotency characterization of RV-ONSL1 hiPSC clone reprogrammed with RV-LTR-ONSL virus. (**A**) Bright-field image of RV-ONSL1 hiPSC clone. Scale bar represents 100 μm. (**B**) Normal karyotypes of CPRE2 fibroblasts and RV-ONSL1 hiPSC clone. (**C**) Pearson correlation analyses of global gene expression in RV-ONSL1 hiPSC clone and hESC H9 line versus corresponding embryoid bodies (EB). Red indicates increased expression compared with median levels of four samples and green corresponds to decreased expression. (**D**) Flow cytometer expression analysis of human ESC-specific cell surface markers, i.e. SSEA4, Tra1–60, Tra1–81 for CPRE2 fibroblasts and RV-ONSL1 hiPSC clone. Red line: secondary antibody control; blue solid line: antigen staining. (**E**) Analysis of the methylation status of the *OCT4* promoters (light and dark gray regions) in CPRE2 fibroblasts and hiPSC cells using bisulfite sequencing. Open circles indicate unmethylated CpG, filled circles indicate methylated CpG and yellow circles indicate incomplete data. (**F**) Hematoxylin and eosin staining of teratoma derived from RV-ONSL1 hiPSC clone. Ectoderm-derived lineage is represented by neural tissue (a) and pigmented cells (b). Mesoderm-derived tissues are represented by white adipose tissue (c) and cartilage (d). Endoderm-derived tissues are represented by glandular epithelium (e) and gut-like epithelium (f). Scale bar represents 200 μm.

favor two non-exclusive hypotheses: a direct interaction between reprogramming factors improving their reprogramming capacities and/or the activation of two different pathways that synergize the reprogramming process. Concerning the speed up of the process, it was shown in the mouse (Hanna *et al*., 2009) and in human that cMYC considerably accelerates the process (Hockemeyer *et al*., 2008; Nakagawa *et al*., 2008) most probably by speeding up the cell cycle (Hanna *et al*., 2009). In mouse, Hanna *et al.* (2009) showed that addition of NANOG to OSKM cocktail accelerates iPSC colony appearance, but in human, the overall colony number does not increase (Maherali *et al*., 2008), which would be in agreement with our results on combination experiments. Nevertheless, in mRNA-mediated iPS generation, NANOG confirms its facilitating role in human (Warren *et al*., 2012). All those data do not explain why OSKM+ONSL perform far better than single polycistronic gene supplemented with missing factors individually. Deeper insight in molecular mechanism involved requires more complex and reproducible system like inducible secondary iPSC (Hockemeyer *et al*., 2008).

Finally, reaching 2% efficiency may remain useless if integration mutagenesis increases with MOI. We also showed that it is possible to produce iPSC with only one integration event. Only one integrated copy associated with retroviral LTR promoter, which is subjected to DNA methylation and histone deacetylation (Krishnan *et al*., 2006) during the cell culture, reduces over time the residual expression of transcription factors into the genomic DNA of host cells. This improves the probability to obtain differentiated cells from iPSC cells.

Having reached such efficiency, we are now working on the possibility to scale down the number of cells to limit the cost, but also in order to automate the process. We are also working, using the combination of both polycistronic, on a non-integrative process. Even if the reprogramming mRNA transfection based methods remain time consuming, we are concentrating our efforts on this strategy, mainly because of the versatility of the system.

Supplementary data

Supplementary data are available at http://molehr.oxfordjournals.org/.

Acknowledgements

We are grateful to Dr James Turner for his critical reading of this manuscript, Valérie Rimelen, Christine Lehalle and Ornella Cammarata for their technical support. We thank the IGBMC common facilities for technical support.

Funding

This work was supported by the Centre National de la Recherche Scientifique (CNRS), Institut National de la Santé et de la Recherche Médicale (INSERM), the Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche, l'Association Française contre les Myopathies (AFM), la Fondation pour la Recherche Médicale (FRM) and Hôpitaux Universitaires de Strasbourg. This work was also supported by a collaborative project called Ship-In and funded by the Fond Unique Interministériel (FUI), the Fonds Européen de Développement Economique Régional (FEDER) and the Région Midi-Pyrénées. We also thank the two competitiveness clusters: Cancer Bio Santé and Alsace BioValley for their support in the organization of the collaborative project. We thank the Max-Planck Society and the BMBF (01GN0818) for financial support.

Conflict of interest

R.G., P.B. and Y.M. are employed by the company Vectalys. The viral vectors used in this work were produced by Vectalys. They are currently sold in Europe and will be sold in the US in 2014.

References

Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA *et al.* Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* 2011;**8**:376–388.

Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? *Nat Rev Mol Cell Biol* 2012;**13**:713–726.

Bernhardt M, Galach M, Novak D, Utikal J. Mediators of induced pluripotency and their role in cancer cells—current scientific knowledge and future perspectives. *Biotechnol J* 2012;**7**:810–821.

Carey BW, Markoulaki S, Hanna J, Saha K, Gao Q, Mitalipova M, Jaenisch R. Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc Natl Acad Sci USA* 2009;**106**:157–162.

Carey BW, Markoulaki S, Hanna JH, Faddah DA, Buganim Y, Kim J, Ganz K, Steine EJ, Cassady JP, Creyghton MP *et al.* Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell* 2011;**9**:588–598.

Daley GQ, Scadden DT. Prospects for stem cell-based therapy. *Cell* 2008;**132**:544–548.

de Felipe P. Polycistronic viral vectors. *Curr Gene Ther* 2002;**2**:355–378.

Freberg CT, Dahl JA, Timoskainen S, Collas P. Epigenetic reprogramming of OCT4 and NANOG regulatory regions by embryonal carcinoma cell extract. *Mol Biol Cell* 2007;**18**:1543–1553.

Fusaki N, Ban H, Nishiyama A, Saeki K, Hasegawa M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc Jpn Acad Ser B Phys Biol Sci* 2009;**85**:348–362.

Gonzalez F, Boue S, Izpisua Belmonte JC. Methods for making induced pluripotent stem cells: reprogramming a la carte. *Nat Rev Genet* 2011;**12**:231–242.

Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, Creyghton MP, van Oudenaarden A, Jaenisch R. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 2009;**462**:595–601.

Hick A, Wattenhofer-Donze M, Chintawar S, Tropel P, Simard JP, Vaucamps N, Gall D, Lambot L, Andre C, Reutenauer L *et al.* Neurons and cardiomyocytes derived from induced pluripotent stem cells as a model for mitochondrial defects in Friedreich's ataxia. *Dis Model Mech* 2013;**6**:608–621.

Hockemeyer D, Soldner F, Cook EG, Gao Q, Mitalipova M, Jaenisch R. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell* 2008;**3**:346–353.

International Standing Committee on Human Cytogenetic Nomenclature, Shaffer LG, McGowan-Jordan J, Schmid M. *ISCN 2013: An International System for Human Cytogenetic Nomenclature (2013)*. Basel: Karger, 2013.

Jia F, Wilson KD, Sun N, Gupta DM, Huang M, Li Z, Panetta NJ, Chen ZY, Robbins RC, Kay MA *et al.* A nonviral minicircle vector for deriving human iPS cells. *Nat Methods* 2010;**7**:197–199.

Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K. Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature* 2009;**458**:771–775.

Kim D, Kim CH, Moon JI, Chung YG, Chang MY, Han BS, Ko S, Yang E, Cha KY, Lanza R *et al*. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* 2009; **4**:472–476.

Krishnan M, Park JM, Cao F, Wang D, Paulmurugan R, Tseng JR, Gonzalgo ML, Gambhir SS, Wu JC. Effects of epigenetic modulation on reporter gene expression: implications for stem cell imaging. *FASEB J* 2006;**20**:106–108.

Lapillonne H, Kobari L, Mazurier C, Tropel P, Giarratana MC, Zanella-Cleon I, Kiger L, Wattenhofer-Donze M, Puccio H, Hebert N *et al*. Red blood cell generation from human induced pluripotent stem cells: perspectives for transfusion medicine. *Haematologica* 2010;**95**:1651–1659.

Liu X, Sun H, Qi J, Wang L, He S, Liu J, Feng C, Chen C, Li W, Guo Y *et al*. Sequential introduction of reprogramming factors reveals a time-sensitive requirement for individual factors and a sequential EMT-MET mechanism for optimal reprogramming. *Nat Cell Biol* 2013; **15**:829–838.

Maherali N, Hochedlinger K. Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2008;**3**:595–605.

Maherali N, Ahfeldt T, Rigamonti A, Utikal J, Cowan C, Hochedlinger K. A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* 2008;**3**:340–345.

Mandal PK, Rossi DJ. Reprogramming human fibroblasts to pluripotency using modified mRNA. *Nat Protoc* 2013;**8**:568–582.

Masip M, Veiga A, Izpisua Belmonte JC, Simon C. Reprogramming with defined factors: from induced pluripotency to induced transdifferentiation. *Mol Hum Reprod* 2010;**16**:856–868.

Miyoshi N, Ishii H, Nagano H, Haraguchi N, Dewi DL, Kano Y, Nishikawa S, Tanemura M, Mimori K, Tanaka F *et al*. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell* 2011;**8**:633–638.

Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 2008;**26**:101–106.

Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature* 2007;**448**:313–317.

Onder TT, Daley GQ. New lessons learned from disease modeling with induced pluripotent stem cells. *Curr Opin Genet Dev* 2012;**22**:500–508.

Papapetrou EP, Tomishima MJ, Chambers SM, Mica Y, Reed E, Menon J, Tabar V, Mo Q, Studer L, Sadelain M. Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc Natl Acad Sci USA* 2009; **106**:12759–12764.

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M *et al*. Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* 2013;**502**:65–70.

Ramalho-Santos M. iPS cells: insights into basic biology. *Cell* 2009; **138**:616–618.

Robinton DA, Daley GQ. The promise of induced pluripotent stem cells in research and therapy. *Nature* 2012;**481**:295–305.

Si-Tayeb K, Noto FK, Sepac A, Sedlic F, Bosnjak ZJ, Lough JW, Duncan SA. Generation of human induced pluripotent stem cells by simple transient transfection of plasmid DNA encoding reprogramming factors. *BMC Dev Biol* 2010;**10**:81.

Somers A, Jean JC, Sommer CA, Omari A, Ford CC, Mills JA, Ying L, Sommer AG, Jean JM, Smith BW *et al*. Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. *Stem Cells* 2010;**28**:1728–1740.

Sommer CA, Stadtfeld M, Murphy GJ, Hochedlinger K, Kotton DN, Mostoslavsky G. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* 2009;**27**:543–549.

Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K. Induced pluripotent stem cells generated without viral integration. *Science* 2008;**322**:945–949.

Szymczak AL, Vignali DA. Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther* 2005;**5**:627–638.

Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; **126**:663–676.

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007;**131**:861–872.

Tanabe K, Nakamura M, Narita M, Takahashi K, Yamanaka S. Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc Natl Acad Sci USA* 2013; **110**:12172–12179.

Warren L, Manos PD, Ahfeldt T, Loh YH, Li H, Lau F, Ebina W, Mandal PK, Smith ZD, Meissner A *et al*. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 2010;**7**:618–630.

Warren L, Ni Y, Wang J, Guo X. Feeder-free derivation of human induced pluripotent stem cells with messenger RNA. *Sci Rep* 2012;**2**:657.

Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, Hamalainen R, Cowling R, Wang W, Liu P, Gertsenstein M *et al*. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 2009; **458**:766–770.

Woods DC, Tilly JL. The next (re)generation of ovarian biology and fertility in women: is current science tomorrow's practice? *Fertil Steril* 2012; **98**:3–10.

Yakubov E, Rechavi G, Rozenblatt S, Givol D. Reprogramming of human fibroblasts to pluripotent stem cells using mRNA of four transcription factors. *Biochem Biophys Res Commun* 2010;**394**:189–193.

Yamaguchi S, Hirano K, Nagata S, Tada T. Sox2 expression effects on direct reprogramming efficiency as determined by alternative somatic cell fate. *Stem Cell Res* 2011;**6**:177–186.

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R *et al*. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 2007;**318**:1917–1920.

Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, Slukvin II, Thomson JA. Human induced pluripotent stem cells free of vector and transgene sequences. *Science* 2009;**324**:797–801.

Zhou W, Freed CR. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells* 2009; **27**:2667–2674.

Zhou H, Wu S, Joo JY, Zhu S, Han DW, Lin T, Trauger S, Bien G, Yao S, Zhu Y *et al*. Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* 2009;**4**:381–384.

Zhu Z, Huangfu D. Human pluripotent stem cells: an emerging model in developmental biology. *Development* 2013;**140**:705–717.