

# Lentiviral Vector Design and Imaging Approaches to Visualize the Early Stages of Cellular Reprogramming

Eva Warlich<sup>1,2</sup>, Johannes Kuehle<sup>1,2</sup>, Tobias Cantz<sup>2-4</sup>, Martijn H Brugman<sup>1</sup>, Tobias Maetzig<sup>1,2</sup>, Melanie Galla<sup>1,2</sup>, Adam A Filipczyk<sup>5</sup>, Stephan Halle<sup>6</sup>, Hannes Klump<sup>1,7</sup>, Hans R Schöler<sup>2,4</sup>, Christopher Baum<sup>1,2</sup>, Timm Schroeder<sup>5</sup> and Axel Schambach<sup>1,2,8</sup>

<sup>1</sup>Department of Experimental Hematology, Hannover Medical School, Hannover, Germany; <sup>2</sup>Cluster of Excellence REBIRTH, Hannover, Germany; <sup>3</sup>Junior Research Group Stem Cell Biology, Hannover Medical School, Hannover, Germany; <sup>4</sup>Department of Cell and Developmental Biology, Max-Planck-Institute for Molecular Biomedicine, Münster, Germany; <sup>5</sup>Institute of Stem Cell Research, Helmholtz Center Munich, Neuherberg, Germany; <sup>6</sup>Institute of Immunology, Hannover Medical School, Hannover, Germany; <sup>7</sup>Institute for Transfusion Medicine, University Hospital Essen, Essen, Germany; <sup>8</sup>Integrated Research and Treatment Center Transplantation, IFB-Tx, Hannover Medical School, Hannover, Germany

Induced pluripotent stem cells (iPSCs) can be derived from somatic cells by gene transfer of reprogramming transcription factors. Expression levels of these factors strongly influence the overall efficacy to form iPSC colonies, but additional contribution of stochastic cell-intrinsic factors has been proposed. Here, we present engineered color-coded lentiviral vectors in which codon-optimized reprogramming factors are co-expressed by a strong retroviral promoter that is rapidly silenced in iPSC, and imaged the conversion of fibroblasts to iPSC. We combined fluorescence microscopy with long-term single cell tracking, and used live-cell imaging to analyze the emergence and composition of early iPSC clusters. Applying our engineered lentiviral vectors, we demonstrate that vector silencing typically occurs prior to or simultaneously with the induction of an Oct4-EGFP pluripotency marker. Around 7 days post-transduction (pt), a subfraction of cells in clonal colonies expressed Oct4-EGFP and rapidly expanded. Cell tracking of single cell-derived iPSC colonies supported the concept that stochastic epigenetic changes are necessary for reprogramming. We also found that iPSC colonies may emerge as a genetic mosaic originating from different clusters. Improved vector design with continuous cell tracking thus creates a powerful system to explore the subtle dynamics of biological processes such as early reprogramming events.

Received 23 June 2010; accepted 27 December 2010; published online 1 February 2011. doi:10.1038/mt.2010.314

## INTRODUCTION

Differentiated somatic cells can be converted into induced pluripotent stem cells (iPSC) with properties similar to embryonic stem (ES) cells by expressing a defined set of reprogramming

factors (RFs). After the first proof-of-principle was obtained and a set of four transcription factors (Oct4, Klf4, Sox2, c-Myc) was identified in the pioneering study, several reports have reproduced these findings with murine and human cells and started to investigate the underlying mechanisms.<sup>1-7</sup> While the initial studies used  $\gamma$ -retroviral vectors for delivery of the RF, promising delivery technologies such as nonviral piggyBac transposon-based vectors, episomal vectors, tetracycline-regulated systems, stabilized mRNAs,<sup>8</sup> and protein delivery via protein transduction domains have been developed.<sup>9-13</sup> Aiming for highly efficient generation of iPSCs, polycistronic lentiviral vectors were used that co-express all RF from one construct to ensure that all RF are available in each transduced cell.<sup>14,15</sup>

Important variables guiding reprogramming factor expression in a lentiviral context are located on the transcriptional (*i.e.*, promoter choice) and post-transcriptional level. Addressing the latter, codon-optimization of the RF (to more favored human tRNA usage) might increase their basal expression levels<sup>16,17</sup> due to improved mRNA stability and translation. Likewise, the insertion of a woodchuck hepatitis virus post-transcriptional regulatory element (wPRE) also enhances post-transcriptional transgene expression (mRNA stability, export, and translation).<sup>18,19</sup> However, both potentially beneficial avenues have not been systematically addressed so far.

Important components of the signalling pathways involved in reprogramming have been elucidated<sup>6,20</sup> and small molecules<sup>21,22</sup> or alternative cell sources<sup>23</sup> were introduced to enhance this process or circumvent the need for some of the RF.<sup>6,20,21,24</sup> Still, the mechanism underlying reprogramming is poorly understood and it remains enigmatic why only a minor fraction of cells expressing the RF is capable to fully convert into iPSC and where this fraction derives from. The “stochastic” model,<sup>25</sup> in which most of the differentiated cells have the potential to become iPSC presupposing epigenetic chromatin remodelling, has only recently

**Correspondence:** Timm Schroeder (for time-lapse microscopy and tracking), Institute of Stem Cell Research, Helmholtz Center Munich, Ingolstaedter Landstraße 1, D-85764 Neuherberg, Germany. E-mail: [tim.schroeder@helmholtz-muenchen.de](mailto:tim.schroeder@helmholtz-muenchen.de) or Axel Schambach (for the remaining work), Department of Experimental Hematology, Hannover Medical School, Carl Neuberg Str. 1, D-30625 Hannover, Germany. E-mail: [schambach.axel@mh-hannover.de](mailto:schambach.axel@mh-hannover.de)

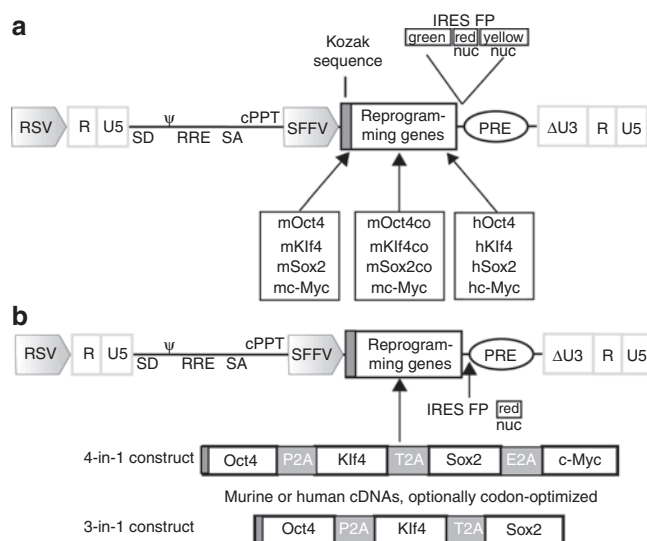
been substantiated by Hanna *et al.* using a conditional transgenic model.<sup>26</sup> This and other groups recently investigated morphological and molecular changes associated with reprogramming,<sup>27,28</sup> and monitored these processes every few hours or days albeit without directly marking RF expression.

In the present study, we developed fluorescence-coded lentiviral vectors that initially trigger high-level expression of the RF and are rapidly silenced in reprogramming cells. Transducing murine embryonic fibroblasts (MEFs) of an established and well-characterized Oct4-EGFP reporter mouse strain (OG2),<sup>29–32</sup> we used continuous single-cell tracking<sup>33</sup> with short intervals (minutes)<sup>34,35</sup> to film the “birth” of pluripotent cells in cell clusters expressing and silencing the RF. Kinetic analyses and cell tracking provided supporting evidence for the stochastic emergence of reprogrammed cells. Our data also showed that early clonal colonies containing reprogrammed cells are frequently contaminated with cells that fail to undergo full reprogramming, and that iPSC colonies are often invaded by cells derived from surrounding clusters.

## RESULTS

### Design of lentiviral vectors promoting efficient onset of reprogramming gene expression and fast epigenetic silencing in pluripotent cells

Our developed modular lentiviral vector system encodes murine or human versions of the canonical RF (Oct4, Klf4, Sox2, c-Myc). We aimed for a functional, easily interchangeable design allowing efficient coexpression of RF and a fluorescent marker (preferably dTomato or a nuclear-localized derivative for cell tracking in condensed cell clusters) on the same mRNA to monitor RF expression. We constructed either 1-factor vectors (expressing just one RF) or combinatorial (3-in-1 or 4-in-1) constructs coexpressing Oct4, Klf4, Sox2, and optionally c-Myc via different 2A-proteinase sites<sup>36</sup> (Figure 1), that have previously been shown to mediate almost complete separation of recombinant proteins. In addition, we introduced a number of modifications to the expression cassette to improve RF and fluorescent marker expression, as conversion to pluripotency requires robust RF expression and a clear fluorescence signal is a prerequisite for imaging/cell tracking studies. Therefore, we modified mRNA processing by equipping the fully sequenced complementary DNAs (cDNAs) with a Kozak consensus element for efficient initiation of translation and added a well-characterized post-transcriptional regulatory element derived from the woodchuck hepatitis virus (wPRE).<sup>18,19</sup> Furthermore, we tested different promoters (PGK: phosphoglycerokinase; UCOE: ubiquitous chromatin opening element; SFFV: spleen focus-forming virus U3 promoter) and introduced synthetic cDNAs for the RF in which codon usage was optimized for expression in human and murine cells.<sup>16,17</sup> In addition, by this approach eight putative splice donor sequences were removed from the 4-in-1 vector. Codon-optimization generally resulted in enhanced expression levels of the individual murine and human RF and also increased the expression from the multicistronic 3-in-1 and 4-in-1 vectors (Figure 2a–d and Supplementary Figure S1). Most of our expression comparisons were based on dTomato fluorescence intensity measured by flow cytometry. This primarily reflects general increases in mRNA stability by codon-optimization of the RF.



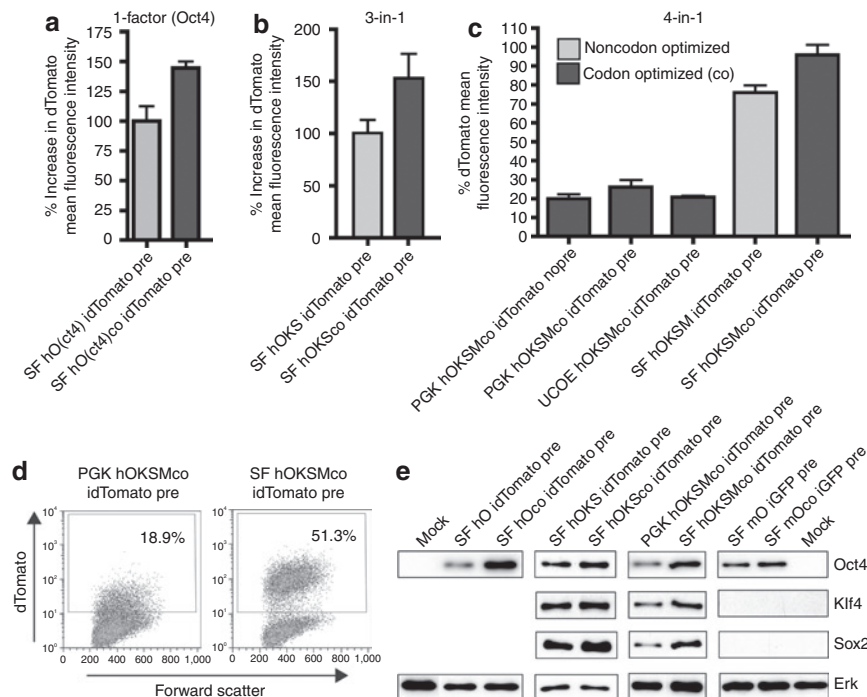
**Figure 1** Design of 1-factor and combinatorial multicistronic reprogramming vectors. **(a)** Modular configuration of the self-inactivating (SIN) vector backbones for expression of the murine (m) or human (h) reprogramming factor (RF) Oct4, Klf4, Sox2, and c-Myc, optionally codon-optimized (co).  $\Delta$  marks the SIN configuration with partially deleted U3 of the 3' long terminal repeat. cPPT, central polypurine tract; FP, fluorescent protein (green: GFP; red: dTomato; yellow: Venus); IRES, internal ribosomal entry site; nuc, nuclear membrane-localized derivative; PRE, post-transcriptional regulatory element; RRE, rev-responsive element; RSV, Rous sarcoma virus U3 promoter; SA, splice acceptor; SD, splice donor; SFFV, spleen focus-forming virus promoter;  $\psi$ , packaging signal. **(b)** Multicistronic all-in-one SIN vectors expressing either 4 or 3 RF via 2A self-cleavage sites. cDNA, complementary DNA; E2A, equine rhinitis A virus 2A; P2A, porcine teschovirus 2A; T2A, thesava asigna virus 2A.

Immunoblot analysis of cells transduced with 1-factor and multicistronic (3-in-1 and 4-in-1) vectors showed clearly improved RF expression (Figure 2e).

We and others had previously demonstrated that addition of the wPRE element clearly increased the expression of transgenic EGFP (up to eightfold) and virus titer.<sup>19,37</sup> As effects of the wPRE are often context dependent, we evaluated the inclusion of wPRE in the multicistronic PGK driven 4-in-1 vector. Here, increased dTomato expression (located on the same mRNA as the RFs) was visible (Figure 2c), and the titer improved more than fivefold when the wPRE was present (data not shown).

Of special importance for the present study is our choice of a retroviral promoter (SFFV), which mediates efficient expression in fibroblasts and other somatic cell types<sup>38</sup> but is rapidly silenced in cells undergoing epigenetic remodelling,<sup>39–41</sup> as indicated in differentiating ES cells (Supplementary Figure S2) as well as reprogramming cells (see below). Furthermore, SFFV mediated an at least fourfold higher transgene expression as compared to the PGK and UCOE promoter elements (see below; Figure 2c–e and Supplementary Figure S1). The SFFV promoter thus leads to an easily separable population of RF expressing cells, as required for imaging and cell tracking (Figure 2d and Supplementary Figure S1).

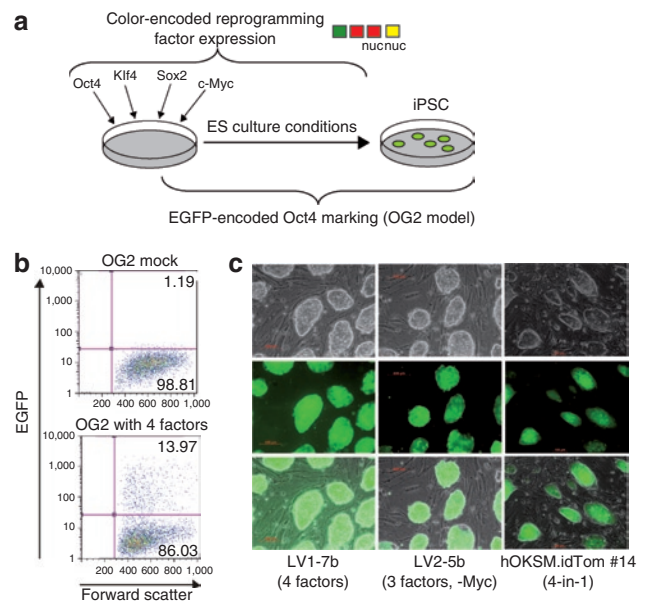
Taken together, the combination of transcriptional (promoter choice) and post-transcriptional modifications generated an improved vector backbone with relatively high RF expression and clearly visible fluorescent marker expression.



**Figure 2** Increased reprogramming factor (RF) expression by vector modifications. Effects of codon-optimization of RF on dTomato (located on the same mRNA) expression demonstrated by changes in fluorescence intensity. Changes are illustrated regarding the human (a) Oct4 1-factor vector, (b) combinatorial 3(factors)-in-1, and (c) 4(factors)-in-1 vectors. For the latter 4-in-1 reprogramming cassette also influences of different promoters (PGK, phosphoglycerokinase; UCOE, ubiquitous chromatin opening element; SF, spleen focus) and the woodchuck hepatitis virus post-transcriptional regulatory element (wPRE) are shown (PRE/wPRE). (d) Flow cytometry analysis comparing dTomato expression from codon-optimized 4-in-1 reprogramming vectors mediated by the PGK and SFFV promoter. (e) Influences of codon-optimization and promoter choice on protein levels of the RF.

### Evidence for efficient reprogramming by the described vector system

Using these modified vectors, we first validated our approach for iPSC generation. **Supplementary Figure S3** shows the correct processing of reprogramming factor mRNAs and proteins. When transducing OG2 MEFs, we relied on EGFP expression as a sensitive indicator for the activation of a pluripotency-associated transcriptional network (see **Figure 3a** for overview of experimental system), as validated earlier.<sup>29–32</sup> This reporter was also shown to correlate with SSEA-1 and Oct4 expression in iPS lines (**Supplementary Figure S4a,b**). Based on the detection of EGFP<sup>+</sup> cells, reprogramming in presence of valproic acid was triggered with high efficiency (>10% Oct4-EGFP<sup>+</sup> cells at day 11 post-transduction (pt); **Figure 3b**). We picked single colonies and demonstrated ES cell-like morphology and marker expression in representative iPSC colonies (**Figure 3c** and **Supplementary Figure S4**) obtained with a cocktail of RF or alternatively a “4-in-1” construct. Nanog and Oct4 promoters were widely unmethylated, as typically observed in ES cells (**Supplementary Figure S5**). iPSC lines were competent to form teratoma with differentiation to the three germ layers after injection into flanks of immunodeficient mice (**Supplementary Figure S6**). Furthermore, the developed vectors efficiently converted other cell types such as adult fibroblasts, bone-marrow-derived mesenchymal cells or blood cells into permanent iPSC lines with ES-like morphology and pluripotency marker expression (T. Cantz, E. Warlich, J. Kuehle, and A. Schambach, unpublished results).



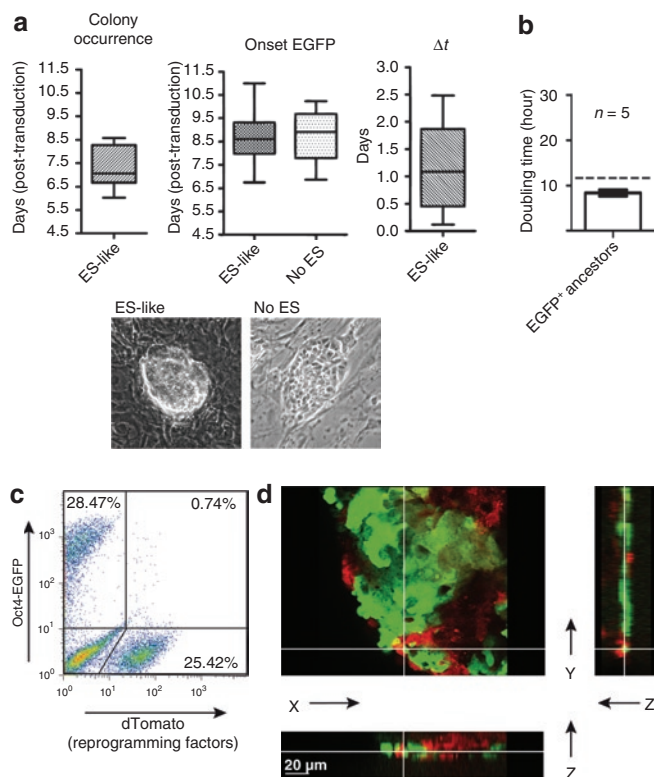
**Figure 3** Efficient reprogramming by the novel vector system. (a) Schematic illustration of reprogramming by virus-encoded transduction of Oct4-EGFP (OG2) transgenic murine embryonic fibroblasts (MEFs). (b) Representative flow cytometry analysis of Oct4-EGFP expression in OG2 MEFs 11 days post-transduction with all 4 factors (lower) and their untransduced counterparts (upper). (c) Fluorescence microscopy of selected OG2-derived induced pluripotent stem cell (iPSC) clones demonstrating embryonic stem cell morphology and Oct4-EGFP activation. Upper brightfield, middle EGFP fluorescence, lower overlay.

## Kinetic analysis of early reprogramming

Next, we applied our lentiviral reprogramming vectors to monitor the kinetics of successful reprogramming. After transduction of OG2 MEFs red fluorescence indicated RF expression, while green fluorescence in combination with careful morphological analyses indicated the emergence of iPSC. EGFP half-life time had previously been determined to be in the range of 20 hours with maturation occurring within 30 minutes and our data indicate that the half-life of dTomato is likely to be in the range of hours. We co-transduced OG2 MEFs with all four single-factor reprogramming vectors (the Sox2 transgene linked to an IRES-dTomato cassette), monitoring the time window of day 4.5 to day 11.5 pt. While red fluorescent transduced cells displayed significant variance in the formation of cell clusters with a converted morphology, we identified—with some variance—the emergence of colonies containing cells with a characteristic ES-like phenotype around day 7 pt. The onset of EGFP expression occurred as little as 3 hours until >60 hours after the emergence of these colonies (Figure 4a). These putative pluripotent EGFP<sup>+</sup> cells appeared within colonies of transduced dTomato<sup>+</sup> cells and rapidly expanded. While tracking of individual cells expressing EGFP is difficult in the condensed structure of late colonies, the investigation of doubling times in the limited number of “early-onset” EGFP<sup>+</sup> cells suggested proliferation rates around 10 hours (Figure 4b) similar to the previously determined proliferation rates of ES cells.<sup>42,43</sup> In line with this, time-lapse observation suggested the absence of frequent apoptotic events in RF transduced cells. However, an additional conversion of further cells within the colony from dTomato to EGFP expression might also contribute to the rapid expansion of EGFP<sup>+</sup> cells. Despite this expansion the majority of iPSC-bearing colonies also contained cells that failed to be reprogrammed within the observation period and often continued to express the exogenous RF (Supplementary Videos S1a,b and S2).

The results obtained from live-cell imaging were confirmed using a “4-in-1” vector in which we linked codon-optimized cDNAs of Oct4, Klf4, Sox2, and wild-type c-Myc, devoid of untranslated sequences, by different 2A self-cleavage sites coupled with the dTomato transgene (via an IRES site) (Figure 1b). With this vector, analysis of iPSC containing colonies by flow cytometry (Figure 4c and Supplementary Figure S7a) revealed a downregulation of SFFV-driven dTomato expression in emerging EGFP<sup>+</sup> cells. Individual cells coexpressing EGFP and dTomato were hardly observed (<2% double positive cells) at early time points and were almost completely absent 11 days pt. Considering the necessity for dTomato degradation or dilution in proliferating cells and the requirement for EGFP accumulation, these data indicate that vector silencing typically occurred even before or at least simultaneously with the induction of the Oct4-EGFP allele.

Apoptome 3D-microscopy-based reconstruction showed that colonies were composed of both, EGFP<sup>+</sup> potential iPSC and dTomato<sup>+</sup> cells still expressing the RF, whereas cells coexpressing both markers were rare (Figure 4d and Supplementary Video S3). In summary, these studies underlined that we developed a system in which the onset and silencing of reprogramming factor expression can be monitored with high sensitivity, and which shows a remarkable frequency of reprogramming cassette silencing in

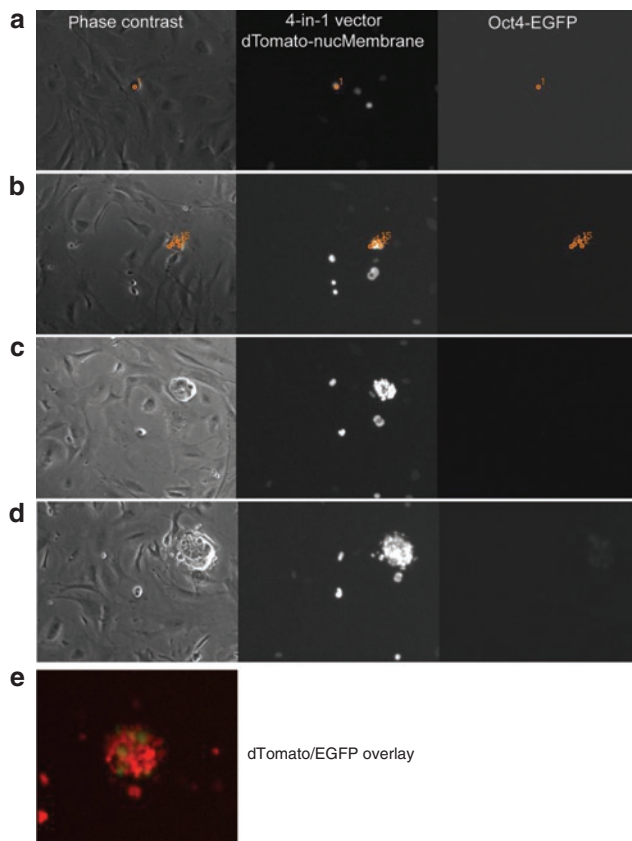


**Figure 4** Typical events in the early stages of reprogramming. **(a)** Colony-wise analysis of time-lapse imaging recorded over a period of 7 days (days 4.5–11.5) following co-transduction of OG2 murine embryonic fibroblasts with all four reprogramming factors. Analyzed are the emergence of embryonic stem (ES) cell-like morphology (ES-like,  $n = 21$ , left plot) and onset of Oct4-EGFP in these colonies (middle and right plot,  $\Delta t = t(\text{onset EGFP}) - t(\text{colony occurrence})$ ). EGFP also appeared in many proliferative clusters that did not demonstrate all criteria for ES cell-like morphology (no ES,  $n = 33$ , central plot). Boxes extend from 25th to 75th percentile with line at median and whiskers spanning minimum to maximum. **(b)** Analysis of doubling times in early EGFP<sup>+</sup> ancestors of reprogramming factor transduced cells by time-lapse imaging. Hatched line indicates doubling times of ES cells as previously described.<sup>42</sup> **(c)** Expression levels of fluorescent markers in induced pluripotent stem cell (iPSC)-like cells as analyzed by flow cytometry 15 days post-transduction (pt) using a 4-in-1 vector linked to internal ribosomal entry site (IRES)-dTomato. A homogeneous cell population resembling the size of iPSC was gated. **(d)** 3D-microscopy of an iPSC-like colony (induced using a 4-in-1 vector linked to IRES-dTomato) on day 13 pt. The view from above is shown with z sections in the x and y planes (depicted on the right and below). White lines indicate where the sections have been made.

reprogramming cells especially when using a cassette for combinatorial expression of RF.

## Heterogeneity of iPSC colony composition and of induction of reprogramming

When undertaking a careful analysis of reprogramming colony composition in the first days after transduction with RF, we observed frequent heterogeneity of Oct4-EGFP expression in iPSC colonies. EGFP not only occurred with some variance after emergence of ES-like colonies, but also rather appeared at one or more distinct spots within a colony than being homogeneously distributed (Supplementary Videos S1a,b, S2, S5b, and S6). Interestingly, EGFP also occurred in a number of colonies that



**Figure 5** Exemplary sequence of induced pluripotent stem cell (iPSC)-like colony development and associated marker expression. A single fibroblast (1) (**a**) transduced with a 4-in-1 reprogramming vector (linked to nuclear-localized dTomato), (**b**) gives rise to progeny (serially numbered and tracked in orange, also see tracking tree in **Supplementary Figure S8**), (**c**) that adopts an iPSC-like morphology. (**d**) Consecutively, an Oct4-EGFP reporter switches on in a subset of cells. Left phase contrast; middle nuclear-localized dTomato; right Oct4-EGFP signal. (**e**) An overlay is shown of the dTomato-nucMembrane signal and the Oct4-EGFP of the same colony as in **d**.

displayed a transformed phenotype rather than matching all morphological criteria of ES cells (round, sharp border, shiny, compact structure) (**Figure 4a**). While early after colony formation, cells expressed at least one of the fluorescence markers, occasionally some cells within a colony neither expressed the RF nor the Oct4-EGFP reporter at later time points (**Supplementary Figure S7b**).

To further ease the traceability of individual cells within compacted iPSC colonies, we employed nuclear Venus or dTomato (**Supplementary Video S4**, also see text below). Due to the high resolution of the time-lapse imaging we were able to document that many colonies did not arise from single-transduced cells, but rather incorporated neighboring transduced cells while proliferating. Also, confluence of neighboring colonies was not only often observed (**Supplementary Video S1a**) and most pronounced at high reprogramming efficiencies but also with remarkable frequency at relatively low plating densities. As a consequence, “colonies” were in part nonclonal, likely containing cells with different vector integration sites, and at different reprogramming levels. Both findings, colony aggregation and EGFP pattern, reflect the potentially genetic and epigenetic heterogeneity of iPSC colonies

and underline the power of high-frequency imaging even for compacted and fast-dividing cells such as iPSC.

Time-lapse imaging allowed tracking of iPSC colony generation from a single cell (**Figure 5**, **Supplementary Figure S8** and **Videos S5a,b and S6**) based on the software tool (TTT). Previous studies described a discrepancy between the number of genetically modified cells and successfully reprogrammed cells, suggestive of stochastic gate-keeping events. To investigate the clonal derivation of the subset of successfully reprogrammed cells in detail, we initiated continuous cell tracking reflecting RF and pluripotency marker expression under real-time conditions from very early time points (from day 4.5 until day 11.5 pt). Since a single-transduced cell gives rise to genetically identical progeny, the influence of secondary events can be investigated. In **Figure 5** (sequential picture panel) and **Supplementary Video S5a,b** we documented such a case, in which only a fraction of the genetically identical progeny from a single-transduced (red) cell switched on Oct4-EGFP. To illustrate the clonality during iPSC emergence, single colony forming cells and their progeny were tracked with TTT (**Supplementary Video S5a and Figure S8**). We confirmed these data in a second setting, in which we used a 4-in-1 cassette linked to the newly developed nuclear-localized dTomato to ease single cell tracking and monitored a wider time window (days 1–13 pt) (**Supplementary Video S6**).

Taken together, the combination of our vector system with high-frequency live-cell imaging allowed us to capture the conversion from RF expression to the induction of pluripotency markers, and to document the substantial heterogeneity within a clonal colony of genetically identical cells undergoing reprogramming.

## DISCUSSION

We developed and validated lentiviral vectors which express the RF from codon-optimized coexpression cassettes under control of a retroviral promoter with high activity in somatic cells but rapid silencing in pluripotent cells. Of note, we used a vector backbone with a primer binding site not sensitive to TRIM28-mediated silencing,<sup>44</sup> because this mechanism is also potentially active in some somatic cells<sup>38</sup> which would hamper reprogramming. Employing codon-optimized transgene sequences not only increases vector expression<sup>16,17</sup> but also facilitates the discrimination between endogenous and exogenous RF (with identical protein sequence) based on PCR. Our system could thus, *e.g.*, be used to determine copy numbers of integrates or to investigate the expression, silencing and reactivation of the factors even in absence of fluorescent markers.

As functional validation, we used the described vector system in the current study to visualize the early stages of reprogramming by monitoring the emergence of iPSC within cell colonies expressing and silencing the canonical RF. In MEFs with a transgenic Oct4 indicator system, we analyzed the subsequent dynamic changes by serial microscopy, including Apotome 3D reconstruction and time-lapse fluorescence microscopy, the latter enabled by an advanced live-cell imaging tool called TTT. This program has previously been validated to provide reliable kinetic analyses of cell fate dynamics over several days of observation. In these studies, TTT was successfully used for demonstrating the continuous single-cell imaging of blood generation from hemogenic

endothelium<sup>34</sup> and the instructive character of extrinsic cytokines to direct hematopoietic progenitor lineage choice.<sup>35</sup> Thus, this was an ideal sensitive experimental system to monitor the early stages of reprogramming with high temporal resolution to confirm and expand recently published data from other groups:

(i) RF expression was efficiently downregulated simultaneously or even before the expression of endogenous Oct4 pronounced especially at late time points after transduction (>12d) and when a 4-in-1 reprogramming vector was used (**Supplementary Video S3** and **Figure 4c,d**). (ii) Confluence of neighboring unrelated cells or colonies frequently resulted in genetic mosaics (**Supplementary Video S1a**), especially with increasing reprogramming efficiency. (iii) We could track the apparently random birth of iPSC-like cells in multiple locations of single cell-derived colonies expressing RF, surrounded by a majority of transduced cells that failed to express the Oct4-EGFP reprogramming marker (**Supplementary Videos S5a,b** and **S6**).

Sadelain and colleagues<sup>45</sup> have generated human iPSCs derived from bicistronic lentiviral vectors coexpressing each single reprogramming factor with a fluorescent protein. In line with another study which suggested that the epigenetic extinction of the transgene cassette is mechanistically advantageous to acquire pluripotency,<sup>46</sup> Sadelain *et al.* suggested that “pronounced (vector) silencing is a hallmark of successful reprogramming” and that it “follows the acquisition of pluripotent cell markers”.<sup>45</sup> Due to the 20–120-fold higher temporal resolution of our approach our data suggest that vector silencing may rather be a prerequisite for the switch to the pluripotent state as prolonged ectopic RF expression may interfere with induction or maintenance of a delicately balanced pluripotency transcriptional network needed for pluripotent cell homeostasis.<sup>47</sup>

By analyzing the average output of populations of reprogrammed cells, first evidence for the epigenetic stochasticity of reprogramming was obtained.<sup>26</sup> Here, we provide supporting evidence for stochastic reprogramming by analyzing not only anonymous populations, but also individual cells in live-cell imaging with high temporal resolution. These findings argue in support of Yamanaka’s hypothesis of a stochastic model for iPSC generation,<sup>25</sup> and our experimental system may not only be useful to address this important point but also to examine the role of small molecules and genetic determinants of the reprogramming process.

In a recent study, serial photographic imaging allowed the distinction of hiPSC from partially reprogrammed cells.<sup>28</sup> Here, vector expression and pluripotency markers monitored every 2–3 days revealed that proviral silencing and pluripotency marker expression indicate the fully reprogrammed state. The authors also reported the conversion of partially to fully reprogrammed cells in a subset of cells within colonies. However, imaging in 2-day intervals does not rule out the possibility of a nonclonal situation, as our present investigations suggest the frequent genetic heterogeneity of iPSC colonies. Recently, Smith and colleagues used fluorescently labelled fibroblasts and analyzed their fate during the reprogramming process using live-cell imaging.<sup>27</sup> However, the fluorescent proteins (actin fusion proteins) marked the entire cells but did not monitor the reprogramming process itself. Thus, it was still difficult to track single converting cells. In our study, the small time intervals (every 12 minutes to 2 hours) allowed the

tracking of single-cell fate demonstrating that early colonies often reflect mosaics or patchworks formed on the basis of an epigenetic variability within transduced cells in conjunction with the potential confluence of neighboring transduced cells and colonies. Detailed studies aimed at the characterization of molecular events supporting the early stages of reprogramming should thus use advanced methods of cell purification to address the correct cell source, *e.g.*, using antibody-based methods or the EOS pluripotency indicator.<sup>48</sup>

In summary, improved vector design with continuous live single-cell observation at high temporal resolution creates a powerful system to assess the subtle kinetics and morphology during biological processes such as the early stages of reprogramming. The experimental system described herein could be useful to further explore reprogramming events as well as to screen pluripotency markers, reprogramming factor candidates, roadblock inhibitors, supporting micro-environments, and other novel tools for reprogramming.

## MATERIALS AND METHODS

**Ethics statement.** All animal work has been performed strictly according to institutional guidelines and national regulations and was approved by the state office for protection of nature, environment, and consumers (LANUV) of North-Rhine Westphalia, Germany.

**Vector construction.** To construct a modular set of lentiviral reprogramming vectors, we used a 3rd generation lentiviral vector<sup>49</sup> (pRRL.PPT.PGK.EGFPpre, kindly provided by L Naldini, Milano, Italy) and equipped it with *NheI*, *AgeI* and *SalI* sites.<sup>50</sup> We introduced the retroviral SFFV U3 promoter amplified via PCR as *NheI* and *AgeI* fragment. The human and murine cDNAs encoding the RF Oct4, Sox2, Klf4, and c-Myc were engineered with a Kozak sequence and amplified as *AgeI* and *SalI* fragments via Pfu PCR. Furthermore, cDNAs for Oct4, Klf4, Sox2 were codon-optimized (Mr. Gene, Regensburg, Germany), thereby improving translation, mRNA half-life time, and removing cryptic splice and polyA sites.<sup>16,17</sup> To allow tracking of reprogramming factor expression, we introduced fluorescent markers, namely EGFP, Venus (YFP derivative, here nuclear localized), and the red fluorescent protein dTomato, via an EMCV IRES (internal ribosomal entry site) sequence in the *SalI* site downstream of the RF. We also constructed 3-in-1 or 4-in-1 lentiviral vectors harboring Oct4, Sox2, Klf4 (3-in-1) plus optionally c-Myc (4-in-1), linked by 2A self-cleavage sites (as described in **Figure 1**), and equipped them with an IRES-dTomato cassette. To ease single-cell tracking, we additionally used a nuclear-localized dTomato (dTomato-nucMembrane). For promoter activity comparison we equipped the 4-in-1 constructs alternatively with the PGK and UCOE promoters. All constructs were validated by sequencing. Cloning details are available on request. See **Supplementary Materials and Methods** for primer details.

**Cell lines, vector production, and lentiviral transduction.** Murine (SC-1) and human HT1080 fibroblast lines, and the human embryonic kidney line 293T were cultured in Dulbecco’s modified Eagle’s medium (PAA, Pasching, Austria) supplemented with 10% fetal calf serum (PAA), 2 mmol/l L-glutamine and sodium pyruvate (all PAA).

Virus production was performed as previously described.<sup>50</sup> In short,  $5 \times 10^6$  293T cells were seeded 24 hours prior to transfection in 10 cm dishes. Cells underwent transfection with 5 µg lentiviral vector, 12 µg pcDNA3.GP.4xCTE (expressing HIV-1 gag/pol), 5 µg pRSV-Rev and 1.5 µg pMD.G (encoding the vesicular stomatitis virus glycoprotein) using the calcium phosphate method in presence of 10 mmol/l HEPES and 25 µmol/l chloroquine. Supernatants were harvested at time points 24, 36, 48, and 60 hours after transfection, filtered and optionally concentrated using ultracentrifugation. Titration was performed on HT1080 and SC-1

fibroblasts 5 days after transduction as previously described.<sup>50</sup> Fluorescence intensity was measured via flow cytometry (see below).

**iPSC generation and cultivation.** MEFs were prepared from pregnant (day 13.5 dpc) C3H (Charles River, Sulzfeld, Germany) and OG2 mice,<sup>29</sup> and cultured on 0.1% gelatin-coated dishes in MEF medium (low glucose Dulbecco's modified Eagle's medium, 10% fetal calf serum, 1 mmol/l L-glutamine (all from PAA), 0.1 mmol/l nonessential amino acids (Gibco, Karlsruhe, Germany), 100 μmol/l β-mercaptoethanol (Sigma, Seelze, Germany), 100 units/ml penicillin/100 μg/ml streptomycin (PAA)). To function as feeder layers, C3H MEFs were grown to confluence and inactivated with 10 μg/ml mitomycin C (Sigma). To generate iPSC, 5–7 × 10<sup>4</sup> MEFs (passage 1–3) were seeded in MEF medium in a 6-well plate 24 hours prior to transduction. MEFs were transduced with different lentiviral vectors encoding for Oct4, Klf4, Sox2, and/or c-Myc in presence of 4 μg/ml protamine sulfate for 8–16 hours. Cells were further cultured in MEF medium until day 4 pt and thereafter in ES medium (knockout Dulbecco's modified Eagle's medium (Gibco), 15% ES-tested fetal calf serum (PAA), 1 mmol/l L-glutamine (PAA), 0.1 mmol/l nonessential amino acids (Gibco), 100 μmol/l β-mercaptoethanol (Sigma), 100 units/ml penicillin/100 μg/ml streptomycin (PAA) and 10<sup>3</sup> units/ml leukemia inhibitory factor (provided by the Max-Planck-Institute, Münster, Germany, and the Department of Technical Chemistry, University Hannover, Germany)). From day 2 to 9 media were supplemented with 2 mmol/l valproic acid (Sanofi-Aventis, Frankfurt, Germany) as previously described.<sup>24</sup> Cells were either harvested via trypsinization and transferred onto mitomycin C-treated MEF feeders 4–7 days pt or reprogrammed “feeder-free” until picking of single clones (latest 15 days pt). Upon appearance of ESC-like colonies, single colonies were picked based on morphology and, in case of OG2, EGFP expression, and expanded on MEF feeder cells. ES medium was replaced every 1–3 days and iPSC was splitted every 3–4 days.

**Time-lapse imaging and single-cell tracking.** Long-term time-lapse imaging and single-cell tracking was done as described previously.<sup>34,35</sup> In brief, OG2 MEFs were transduced with fluorescently labeled LV SIN (self-inactivating) vectors expressing hOct4, hSox2, hKlf4, and hc-Myc. First, we used four single-factor vectors in which hSox2 was linked to an IRES-dTomato cassette. Second, transduction was performed with a 3-in-1 construct with hOct4, hKlf4, hSox2 linked to IRES-dTomato plus a single-factor hc-Myc vector with IRES-Venus-nucMembrane cassette (nuclear-localized Venus). Third, for even better single-cell tracking, a 4-in-1 construct linked to an IRES-dTomato-nucMembrane sequence was used for reprogramming. In setup 1 and 2, transduced MEFs were transferred onto C3H feeder cells in a T12.5 flask (BD, Heidelberg, Germany) 4 days pt and further cultured in pregassed medium. Live-cell imaging was performed at 37 °C from day 4.5 to day 11.5. Phase contrast pictures were acquired every 12 minutes, EGFP and dTomato fluorescence every 2.5 hours. In setup 3, transduced MEFs were maintained *in situ* and imaged from day 1 to day 13. Phase contrast and dTomato fluorescence pictures were taken every 12 minutes. EGFP was monitored in 2.5 hours (days 1–4) or 25-minute intervals (days 5–13).

Time-lapse microscopy was performed using a CellObserver system (Zeiss, Jena, Germany) equipped with a ×5 phase contrast objective (Zeiss), and an AxioCamHRM camera (at 1388 × 1040 pixel resolution) with Zeiss AxioVision 4.5 software. A mercury lamp (HBO103W/2) (Osram, Augsburg, Germany) was used for fluorescence illumination. EGFP and dTomato were detected using the Zeiss filter set 46HE at 1,300 ms and the Zeiss filter set 43HE at 400 ms, respectively.

Single-cell tracking was performed using a self-written computer program (TTT) on Siemens Celsius R630 workstations with 4 GB of RAM and SUSE Linux (10.1) operating systems with KDE desktop. The tracking trees served to determine doubling times (for example see **Supplementary Figure S8**). Videos were assembled using Quick Time Player Pro Ver7.62 (Apple, Cupertino, CA). Examples of videos are displayed in **Supplementary Videos S1–S6**.

**Apotome fluorescence microscopy for analysis of iPSC colony structure.** For structural analysis OG2 MEFs were reprogrammed (using a 4-in-1 vector with IRES-dTomato cassette). Structured illumination microscopy was used to acquire optical sections of iPSC-like colonies (**Supplementary Video S3**). For this purpose single iPSC-like colonies were picked 13 days pt, transferred onto microscopical slides and embedded in Mowiol (Calbiochem, Schwalbach, Germany) for optimized optical analyses. Imaging was performed with an Axiovert 200M microscope (Zeiss) equipped with a ×40 objective (0.95 numeric aperture) and the Apotome slider (Zeiss). Imaris software (Bitplane, Zurich, Switzerland) was used for 3D visualization of raw images and colocalization analysis.

**Flow cytometry.** For monitoring fluorescent marker expression (Venus, EGFP, dTomato), cells were harvested, washed with phosphate-buffered saline, measured on a FACSCalibur or LSRII (Becton Dickinson, Heidelberg, Germany) using CellQuest, FACSDiva (Becton Dickinson) or FlowJo software (Tree Star, Ashland, OR). A gate was set on a homogeneous cell population, as determined by scatter characteristics, and 20,000 events were monitored. A marker gate was set to calculate the percentage and mean fluorescence intensity of positive cells.

**Western blot.** For protein isolation, cells were lysed in the presence of protease inhibitors (Complete Mini, Roche, Penzberg, Germany). 10–15 μg denatured proteins were separated on 12.5% sodium dodecyl sulfate-polyacrylamide gels. Proteins were transferred to a nitrocellulose membrane (Whatman, Dassel, Germany) via tank blot technique. Unspecific binding was blocked with 3% dry milk in TBS/0.05% Tween for at least 30 minutes and membranes stained with primary antibody in recommended dilution (anti-c-Myc: sc-40, anti-Oct4: sc-5279, anti-GKlf: sc-20691 and anti-Sox2: sc-17320, Santa Cruz, Heidelberg, Germany) at 4 °C overnight. After rinsing membranes were stained in appropriate peroxidase-conjugated secondary antibodies (Santa Cruz, Heidelberg, Germany) for 1 hour at room temperature, detection was performed by enhanced chemoluminescence (Thermo Scientific, Langensfeld, Germany) following manufacturer's instructions.

## SUPPLEMENTARY MATERIAL

**Figure S1.** Schematic overview of vector modifications and resulting effects.

**Figure S2.** Promoter expression characteristics in undifferentiated or differentiating ES cells.

**Figure S3.** Correct processing of reprogramming factor RNA and protein.

**Figure S4.** Analysis of ES cell marker expression in iPSC clones.

**Figure S5.** Endogenous Nanog and Oct4 promoters are highly unmethylated in generated iPSC clones.

**Figure S6.** Teratoma formation.

**Figure S7.** Relation of Oct4-EGFP reporter and reprogramming factor expression.

**Figure S8.** Exemplary single-cell tracking tree of a transduced fibroblast and its progeny.

**Video S1.** Heterogeneity of early reprogramming.

**Video S2.** Rapid expansion of Oct4-EGFP expressing iPSC-like cells.

**Video S3.** Structured illumination microscopy of an induced pluripotent stem (iPS)-like colony 15 days after transduction.

**Video S4.** Dynamics of induced pluripotent stem cell (iPSC) generation.

**Video S5.** Stochastic character of reprogramming.

**Video S6.** Stochasticity of reprogramming.

## Materials and Methods.

## ACKNOWLEDGMENTS

This work was supported by grants from Else Kröner-Fresenius-Stiftung, the German Academic Exchange Service (DAAD (0315187)), the German Ministry for Research and Education (PidNet,

IFB-Tx (01E00802), Carpod (016M0854)), the Deutsche Forschungsgemeinschaft (SPP1230, SPP1356, SFB566 and Cluster of Excellence REBIRTH (EXC 62/1)) and the European Union (Integrated Project PERSIST). We would like to thank Françoise Andre, Girmay Asgedom, Thomas Neumann (all Hannover Medical School) and Martina Bleidißel (MPI Münster, Germany) for technical assistance. In addition, we would like to thank Roger Tsien (San Diego, CA) for providing the Tomato cDNA, Michael Antoniou (London, UK) for the UCOE element and Shinya Yamanaka (Kyoto, Japan) for the original reprogramming factor cDNAs.

## REFERENCES

- Meissner, A, Wernig, M and Jaenisch, R (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat Biotechnol* **25**: 1177–1181.
- Takahashi, K, Tanabe, K, Ohnuki, M, Narita, M, Ichisaka, T, Tomoda, K *et al.* (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**: 861–872.
- Park, IH, Arora, N, Huo, H, Maherali, N, Ahfeldt, T, Shimamura, A *et al.* (2008). Disease-specific induced pluripotent stem cells. *Cell* **134**: 877–886.
- Yu, J, Vodyanik, MA, Smuga-Otto, K, Antosiewicz-Bourget, J, Frane, JL, Tian, S *et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**: 1917–1920.
- Maherali, N, Ahfeldt, T, Rigamonti, A, Utikal, J, Cowan, C and Hochedlinger, K (2008). A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell* **3**: 340–345.
- Hong, H, Takahashi, K, Ichisaka, T, Aoi, T, Kanagawa, O, Nakagawa, M *et al.* (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* **460**: 1132–1135.
- Raya, A, Rodríguez-Pizà, I, Guenechea, G, Vassena, R, Navarro, S, Barrero, MJ *et al.* (2009). Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* **460**: 53–59.
- Warren, L, Manos, PD, Ahfeldt, T, Loh, YH, Li, H, Lau, F *et al.* (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* **7**: 618–630.
- Okita, K, Nakagawa, M, Hyenjong, H, Ichisaka, T and Yamanaka, S (2008). Generation of mouse induced pluripotent stem cells without viral vectors. *Science* **322**: 949–953.
- Yu, J, Hu, K, Smuga-Otto, K, Tian, S, Stewart, R, Slukvin, I *et al.* (2009). Human induced pluripotent stem cells free of vector and transgene sequences. *Science* **324**: 797–801.
- Zhou, H, Wu, S, Joo, JY, Zhu, S, Han, DW, Lin, T *et al.* (2009). Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell* **4**: 381–384.
- Schambach, A, Cantz, T, Baum, C and Cathomen, T (2010). Generation and genetic modification of induced pluripotent stem cells. *Expert Opin Biol Ther* **10**: 1089–1103.
- Woltjen, K, Michael, IP, Mohseni, P, Desai, R, Mileikovsky, M, Härmäläinen, R *et al.* (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* **458**: 766–770.
- Sommer, CA, Sommer, AG, Longmire, TA, Christodoulou, C, Thomas, DD, Gostissa, M *et al.* (2010). Excision of reprogramming transgenes improves the differentiation potential of iPSC cells generated with a single excisable vector. *Stem Cells* **28**: 64–74.
- Carey, BW, Markoulaki, S, Hanna, J, Saha, K, Gao, Q, Mitalipova, M *et al.* (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc Natl Acad Sci USA* **106**: 157–162.
- Ng, YY, Baert, MR, Pike-Overzet, K, Rodijk, M, Brugman, MH, Schambach, A *et al.* (2010). Correction of B-cell development in Btk-deficient mice using lentiviral vectors with codon-optimized human BTK. *Leukemia* **24**: 1617–1630.
- Moreno-Carranza, B, Gentsch, M, Stein, S, Schambach, A, Santilli, G, Rudolf, E *et al.* (2009). Transgene optimization significantly improves SIN vector titers, gp91 phox expression and reconstitution of superoxide production in X-CGD cells. *Gene Ther* **16**: 111–118.
- Hope, T (2002). Improving the post-transcriptional aspects of lentiviral vectors. *Curr Top Microbiol Immunol* **261**: 179–189.
- Schambach, A, Bohne, J, Baum, C, Hermann, FG, Egerer, L, von Laer, D *et al.* (2006). Woodchuck hepatitis virus post-transcriptional regulatory element deleted from X protein and promoter sequences enhances retroviral vector titer and expression. *Gene Ther* **13**: 641–645.
- Ichida, JK, Blanchard, J, Lam, K, Son, EY, Chung, JE, Egli, D *et al.* (2009). A small-molecule inhibitor of TGF- $\beta$  signaling replaces Sox2 in reprogramming by inducing nanog. *Cell Stem Cell* **5**: 491–503.
- Feng, B, Ng, JH, Heng, JC and Ng, HH (2009). Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. *Cell Stem Cell* **4**: 301–312.
- Esteban, MA, Wang, T, Qin, B, Yang, J, Qin, D, Cai, J *et al.* (2010). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* **6**: 71–79.
- Kim, JB, Sebastiano, V, Wu, G, Araúzo-Bravo, MJ, Sasse, P, Gentile, L *et al.* (2009). Oct4-induced pluripotency in adult neural stem cells. *Cell* **136**: 411–419.
- Huangfu, D, Maehr, R, Guo, W, Eijkelenboom, A, Snitow, M, Chen, AE *et al.* (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol* **26**: 795–797.
- Yamanaka, S (2009). Elite and stochastic models for induced pluripotent stem cell generation. *Nature* **460**: 49–52.
- Hanna, J, Saha, K, Pando, B, van Zon, J, Lengner, CJ, Creighton, MP *et al.* (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**: 595–601.
- Smith, ZD, Nachman, I, Regev, A and Meissner, A (2010). Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat Biotechnol* **28**: 521–526.
- Chan, EM, Ratanasirintrao, S, Park, IH, Manos, PD, Loh, YH, Huo, H *et al.* (2009). Live cell imaging distinguishes bona fide human iPSC cells from partially reprogrammed cells. *Nat Biotechnol* **27**: 1033–1037.
- Szabó, PE, Hübner, K, Schöler, H and Mann, JR (2002). Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech Dev* **115**: 157–160.
- Boiani, M, Gentile, L, Gambles, VV, Cavaleri, F, Redi, CA and Schöler, HR (2005). Variable reprogramming of the pluripotent stem cell marker Oct4 in mouse clones: distinct developmental potentials in different culture environments. *Stem Cells* **23**: 1089–1104.
- Han, DW, Do, JT, Gentile, L, Stehling, M, Lee, HT and Schöler, HR (2008). Pluripotential reprogramming of the somatic genome in hybrid cells occurs with the first cell cycle. *Stem Cells* **26**: 445–454.
- Kim, JB, Zaehres, H, Wu, G, Gentile, L, Ko, K, Sebastiano, V *et al.* (2008). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* **454**: 646–650.
- Schroeder, T (2008). Imaging stem-cell-driven regeneration in mammals. *Nature* **453**: 345–351.
- Eilken, HM, Nishikawa, S and Schroeder, T (2009). Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* **457**: 896–900.
- Rieger, MA, Hoppe, PS, Smejkal, BM, Eitelhuber, AC and Schroeder, T (2009). Hematopoietic cytokines can instruct lineage choice. *Science* **325**: 217–218.
- Szymczak, AL and Vignali, DA (2005). Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther* **5**: 627–638.
- Schambach, A, Wodrich, H, Hildinger, M, Bohne, J, Kräusslich, HG and Baum, C (2000). Context dependence of different modules for posttranscriptional enhancement of gene expression from retroviral vectors. *Mol Ther* **2**: 435–445.
- Baum, C, Hegewisch-Becker, S, Eckert, HG, Stocking, C and Ostertag, W (1995). Novel retroviral vectors for efficient expression of the multidrug resistance (mdr-1) gene in early hematopoietic cells. *J Virol* **69**: 7541–7547.
- Zentilin, L, Qin, G, Tafuro, S, Dinauer, MC, Baum, C and Giacca, M (2000). Variations of retroviral vector gene expression in myeloid cells. *Gene Ther* **7**: 153–166.
- Zhang, F, Thornhill, SI, Howe, SJ, Ulaganathan, M, Schambach, A, Sinclair, J *et al.* (2007). Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells. *Blood* **110**: 1448–1457.
- Stein, S, Ott, MG, Schultze-Strasser, S, Jauch, A, Burwinkel, B, Kinner, A *et al.* (2010). Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. *Nat Med* **16**: 198–204.
- Ogawa, K, Saito, A, Matsui, H, Suzuki, H, Ohtsuka, S, Shimamoto, D *et al.* (2007). Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. *J Cell Sci* **120**(Pt 1): 55–65.
- Lee, J, Go, Y, Kang, I, Han, YM and Kim, J (2010). Oct-4 controls cell-cycle progression of embryonic stem cells. *Biochem J* **426**: 171–181.
- Wolf, D and Goff, SP (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. *Cell* **131**: 46–57.
- Papapetrou, EP, Tomishima, MJ, Chambers, SM, Mica, Y, Reed, E, Menon, J *et al.* (2009). Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. *Proc Natl Acad Sci USA* **106**: 12759–12764.
- Hotta, A and Ellis, J (2008). Retroviral vector silencing during iPSC cell induction: an epigenetic beacon that signals distinct pluripotent states. *J Cell Biochem* **105**: 940–948.
- Niwa, H, Miyazaki, J and Smith, AG (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**: 372–376.
- Hotta, A, Cheung, AY, Farra, N, Vijayaragavan, K, Séguin, CA, Draper, JS *et al.* (2009). Isolation of human iPSC cells using EOS lentiviral vectors to select for pluripotency. *Nat Methods* **6**: 370–376.
- Dull, T, Zufferey, R, Kelly, M, Mandel, RJ, Nguyen, M, Trono, D *et al.* (1998). A third-generation lentivirus vector with a conditional packaging system. *J Virol* **72**: 8463–8471.
- Schambach, A, Bohne, J, Chandra, S, Will, E, Margison, GP, Williams, DA *et al.* (2006). Equal potency of gammaretroviral and lentiviral SIN vectors for expression of O6-methylguanine-DNA methyltransferase in hematopoietic cells. *Mol Ther* **13**: 391–400.