Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution

Citation: Stelzer, Yonatan et al. "Tracing Dynamic Changes of DNA Methylation at Single-Cell Resolution." Cell 163.1 (2015): 218–229.

As Published: http://dx.doi.org/10.1016/j.cell.2015.08.046

Publisher: Elsevier

Persistent URL: http://hdl.handle.net/1721.1/108812

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

HHS Public Access

Author manuscript

Cell. Author manuscript; available in PMC 2016 September 24.

Published in final edited form as:

Cell. 2015 September 24; 163(1): 218–229. doi:10.1016/j.cell.2015.08.046.

Tracing dynamic changes of DNA methylation at single cell resolution

Yonatan Stelzer[1,3], **Chikdu Shakti Shivalila**[1,2,3], **Frank Soldner**[1], **Styliani Markoulaki**[1], and **Rudolf Jaenisch**[1,2,4]

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA

SUMMARY

Mammalian DNA methylation plays an essential role in development. To date, only snapshots of different mouse and human cell types have been generated, providing a static view on DNA methylation. To enable monitoring of methylation status as it changes over time, we establish a **R**eporter of **G**enomic **M**ethylation (RGM) that relies on a minimal imprinted gene promoter driving a fluorescent protein. We show that insertion of RGM proximal to promoter-associated CpG islands reports the gain and loss of DNA methylation. We further utilized RGM to report endogenous methylation dynamics of non-coding regulatory elements, such as the pluripotency-specific super enhancers of *Sox2* and *miR290*. Loci-specific DNA methylation changes, and its correlation with transcription was visualized during cell state transition following differentiation of mouse embryonic stem cells and during reprogramming of somatic cells to pluripotency. RGM will allow the investigation of dynamic methylation changes during development and disease at single cell resolution.

INTRODUCTION

DNA methylation is recognized as a principal contributor to the stability and regulation of gene expression in development and maintenance of cellular identity (Bird, 2002; Cedar and Bergman, 2012; Jaenisch and Bird, 2003; Reik et al., 2001). Changes in DNA methylation are dynamic and it is still largely unknown how they dictate spatial and temporal gene expression programs (Smith and Meissner, 2013). Recent advancements in sequencing technologies enabled the establishment of methylation maps for multiple cell types in both human (Roadmap Epigenomics et al., 2015; Schultz et al., 2015; Smith et al., 2014; Ziller et

[4]Correspondence should be addressed to R.J., jaenisch@wi.mit.edu, Phone: 617-258-5186.

[3]These authors contributed equally to this work

Author Contributions

Y.S. and R.J conceived the idea for this project. Y.S. and C.S.S. designed and conducted experiments and interpreted the data. F.S. assisted with targeting of mESCs, and S.M. conducted blastocyst injections. Y.S. and R.J. wrote the manuscript with input from all other authors.

al., 2013) and mouse (Hon et al., 2013), thus providing a framework for identifying key lineage-specific regulators (Rivera and Ren, 2013). DNA methylation is a dynamic process and current methods are only bulk and provide a static “snapshot” view of the methylation state during cell state transitions. The difficulty in translating real-time epigenetic changes into a traceable readout, is, to date, a limiting factor in our ability to follow the dynamics of DNA methylation. Therefore a key challenge in the field is to generate tools that allow tracing changes in DNA methylation over time.

Here we set out to generate a DNA methylation reporter system that is capable of visualizing genomic methylation states at single cell resolution. The design of the reporter was based on two premises: (i) previous observations suggesting that CpG sites can serve as cis-acting signals, affecting the methylation state of adjacent CpGs (Brandeis et al., 1994; Mummaneni et al., 1995; Turker, 2002); (ii) a methylation-sensitive promoter that, when introduced in proximity to a CpG region of choice, may be utilized to report on methylation changes of the adjacent sequences. Thus, a key issue in establishing a DNA methylation reporter was identifying a methylation-sensitive promoter, which is not independently regulated by the DNA methylation machinery, but can be affected by exogenous methylation changes. Constitutively active genes usually contain hypomethylated high density CpG islands (CGIs) in their promoter regions and are not regulated by DNA methylation (Deaton and Bird, 2011) whereas gene promoters associated with low density CGI are activated and repressed in a tissue-specific manner. Because methylation of both classes of promoters is either not regulated by the DNA methylation machinery in all tissues or regulated in a tissue-dependent manner, these promoters cannot be utilized as DNA methylation reporters. In contrast, imprinted gene promoters exhibit inherent sensitivity to DNA methylation of adjacent genomic regions resulting in transcriptional activation or silencing. This mechanism has been established for a subgroup of germline-derived differentially methylated regions (DMRs) that affect in cis the methylation state of secondary regulatory promoter elements, which in turn control imprinted gene activity. Importantly, following their establishment, promoter-associated imprinted DMRs are not regulated by the DNA methylation machinery in a tissue-specific manner (Ferguson-Smith, 2011). We hypothesized that these intrinsic characteristics of imprinted gene promoters make them attractive candidates for methylation sensors. Perhaps one of the best-studied example is the Prader-Willi Angelman region, in which an imprinted DMR resides at the small nuclear ribonucleoprotein polypeptide N (*Snrpn*) gene promoter region controlling its parent-of-origin monoallelic expression (Buiting et al., 1995; Kantor et al., 2004). Furthermore, *Snrpn* is expressed in most of the tissues and thus serves as an attractive candidate to generate a DNA methylation reporter.

Changes in DNA methylation occur mostly at non-CGIs, some of which are associated with tissue-specific gene promoters (Jones, 2012). Nevertheless, a growing body of evidence suggests that the bulk of tissue-specific changes in DNA methylation is associated with noncoding sequences (Irizarry et al., 2009) such as distal regulatory elements, which include enhancers and transcription factor binding sites (Hon et al., 2013; Stadler et al., 2011; Ziller et al., 2013). Recent reports identified super-enhancers (SE) as clusters of TF and mediatorbinding sites associated with bona-fide enhancer chromatin marks to control the

expression of key cell identity genes (Dowen et al., 2014; Hnisz et al., 2013; Whyte et al., 2013). Global genomic comparisons of tissue-specific DNA methylation and transcription factor (TF) chromatin immunoprecipitation sequencing (ChIP-seq) data correlated the chromatin with the methylation state (Xie et al., 2013). Thus, many tissue-specific enhancers are hypomethylated in tissues where the target genes are expressed, but are hypermethylated in tissues where the target genes are silent (Hon et al., 2013).

In this paper we establish a **R**eporter of **G**enomic **M**ethylation (RGM) that enables the visualization of changes in DNA methylation in live cells. We show that a minimal Snprn promoter can report on the DNA methylation state of endogenous gene promoters. We also generated reporter cell lines for the pluripotency-specific *miR290* and *Sox2* SEs and demonstrate that RGM can be used to capture dynamic DNA methylation changes in distal non-coding regulatory regions. An attractive aspect of RGM is its utility to visualize DNA methylation changes in development and disease at single cell resolution in the same sample.

RESULTS

A methylation-sensitive reporter system based on a minimal imprinted promoter

To establish a methylation reporter, we generated a minimal *Snrpn* promoter that includes the conserved elements between human and mouse and contains the endogenous imprinted DMR region (Figure S1A). The minimal promoter region driving GFP was cloned into a sleeping beauty transposon vector (Ivics et al., 1997) to facilitate stable integration into the genome. Recent studies have demonstrated that different CGI vectors, when stably inserted into mouse embryonic stem cells (mESCs), adopt a methylation pattern that corresponds to the *in vivo* methylation pattern of the respective endogenous sequence (Sabag et al., 2014). To test whether DNA methylation can propagate into the *Snrpn* promoter region *in vivo*, we designed an experimental system in which the CGI regions of *Gapdh* and *Dazl* were cloned upstream of our reporter (Figure 1A). The promoter of *Gapdh* encompasses a hypomethylated CGI consistent with constitutive expression in all tissues. In contrast, the *Dazl* promoter-associated CGI is hypermethylated in all tissues excluding the germ cells (Hackett et al., 2013). Given the different expression and methylation patterns of both genes, upon stable integration of the two reporter vectors into mESCs the *Gapdh* CGI is expected to maintain its hypomethylated state, while the *Dazl* CGI would be subjected to *de novo* methylation (Sabag et al., 2014). Figure 1B show that more than 95% of cells carrying the *Gapdh* reporter expressed GFP. In contrast, more than 30% of cells carrying the *Dazl* reporter were GFP negative, corresponding to reporter silencing. The effect of the *Dazl* reporter becomes more robust upon continued passage, with more than 80% of the cells silencing their reporter within 4 weeks (Figure 1B).

To assess the DNA methylation levels of the *Gapdh* and *Dazl* reporters following introduction into mESCs, we sorted *Gapdh* GFP positive and *Dazl* GFP negative cell populations (Figure 1C). The GFP expression state was stable upon continuous culture and passaging of the two sorted cell populations for over 7 weeks (Figure 1C). DNA was extracted from both *Gapdh* GFP positive and *Dazl* GFP negative cells and subjected to bisulfite conversion and PCR sequencing. Figure 1D shows that *Gapdh* GFP positive cells

maintained the hypomethylated state at both *Gapdh* CGI and the *Snrpn* promoter regions, whereas *Dazl* GFP negative cells became highly *de novo* methylated at the *Dazl* CGI region and its corresponding downstream *Snrpn* promoter (Figure 1E). These results are consistent with the hypothesis that DNA methylation can be propagated from the CGI into the *Snrpn* promoter region resulting in repression of transcriptional activity.

RGM is a reporter for *in vivo* demethylation

The experiments described above showed that RGM reports on *de novo* methylation imposed *in vivo* on the unmethylated *Dazl* CGI donor test sequence. Conversely, we were interested to assess whether a methylated and silent donor *Snrpn* promoter can be reactivated by means of demethylation acquired *in vivo*. For this we used the CpG methyltransferase M.SssI to *in vitro* methylate both *Gapdh* and *Dazl* reporter constructs. Treatment of the plasmids with M.SssI enzyme followed by bisulfite conversion, PCR amplification and sequencing, confirmed the complete hypermethylation of both the CGI and *Snrpn* promoter regions (Figures 2A, S1B and S1C). ESCs were transfected with either *Gapdh* or *Dazl* reporter and selected for cells carrying stably integrated vectors. Following one week of culture we identified robust activation of GFP in virtually all cells carrying the integrated *Gapdh* reporter, whereas cells carrying the *Dazl* reporter remained GFP negative (Figures 2B, C, D). To assess the DNA methylation state of the *Gapdh* and *Dazl* CGI and the respective downstream *Snrpn* promoter regions, DNA was extracted from the two cell lines, subjected to bisulfite conversion, PCR amplification and sequencing. Figure 2E demonstrates that, consistent with high GFP expression, the *Gapdh* CGI and its downstream *Snrpn* promoter had become fully demethylated. In contrast, the *Dazl* CGI and its downstream *Snrpn* promoter sequences maintained the hypermethylated state in agreement with complete repression of the GFP signal (Figure 2F). Thus, our data support the hypothesis that a *Snrpn* promoter can report on *in vivo* demethylation of the CGI in its proximity.

Dnmt1, 3a and 3b mediate methylation and reporter activity

We used ESCs deficient for the DNA methyltransferases *Dnmt1, Dnmt3a* and *Dnmt3b* to gain mechanistic insights into demethylation and *de novo* methylation imposed on the *Snrpn* promoter in transfected ESCs. Figure 2G shows that introduction of an *in vitro* methylated *Dazl Snrpn* vector into Dnmt1 mutant cells resulted in about 80% GFP positive cells by passage five, in contrast to no GFP positive cells when inserted into wt cells. In agreement with the role of Dnmt1 as being the maintenance DNA methyltransferase (Li et al., 1992), bisulfite sequencing analysis on the sorted GFP positive cells confirmed that reactivation of the methylated *Dazl* reporter occurred by passive demethylation (Figure 2H). To clarify the mechanism of *de novo* methylation, we introduced an unmethylated version of both vectors into mESCs deficient for both *de novo* DNA methyltransferases *Dnmt3a* and *Dnmt3b* (Pawlak and Jaenisch, 2011). Figure 2I shows that the vast majority of cells carrying the *Dazl* or the *Gapdh* reporters were positive for GFP unlike *Dazl* reporter expression in control V6.5 cells (Figure 2I), which is consistent with *Dnmt3a/b* mediating *de novo* methylation and reporter silencing.

Recent studies have shown that culturing mESCs in 2i medium (inhibitors of MEK and GSK3), and leukemia inhibitory factor (LIF) results in downregulation of *Dnmt3a* and *Dnmt3b*, consequently leading to global hypomethylation (Lee et al., 2014). To assess whether these culture conditions affect reporter activity, we transfected the unmethylated *Gapdh* and *Dazl* reporters into wt mESCs cultured in 2i and LIF. Figure 2I shows that the great majority of the stably transfected cells were GFP positive, consistent with 2i-mediated downregulation of the *Dnmt3a* and *3b*.

RGM can report on methylation associated with endogenous gene promoters

To test whether the *Snrpn* promoter could also report on DNA methylation levels associated with endogenous gene promoters, we utilized CRISP/Cas-mediated gene editing to target the endogenous CGI's located at the promoter regions of *Gapdh* and *Dazl* (Figures 3A, S2A and S2B). Figure 3B shows 35/36 *Dazl* targeted clones were GFP negative indicating robust silencing of the *Dazl* reporter whereas 20/21 *Gapdh* targeted clones were GFP positive (Figure 3B). FACS analysis of correctly targeted clones confirmed that *Gapdh* reporter cells were all GFP positive with the CGI and *Snrpn* promoter unmethylated (Figure 3C,D) in contrast to *Dazl* GFP negative clones with the corresponding sequences methylated (Figure 3E,F). Our results demonstrate that *Snrpn* reporter activity reports on the methylation state of its surrounding sequences and does not alter their methylation state. Furthermore, the endogenous targeting results suggested that the partial repression of the *Dazl* reporter (Figure 1B), observed at early passages of the transgene experiment, may be due to multiple genome integration and position effects.

RGM can report on methylation of pluripotency specific super-enhancers

Methylation of super enhancers (SEs) has been shown to change during differentiation. We tested whether RGM would report on the active and hypomethylated state of the pluripotencyspecific SEs associated with the *miR290* and *Sox2* genes in mESCs and their methylated and inactive state in somatic cells (Figure 4A and Figure S3A). In contrast to the CGIs located at gene promoters (*Gapdh* and *Dazl*), the SE regions of both *Sox2* and *miR290* represents lowdensity CpG sequences. Utilizing CRISP/Cas mediated gene editing, we inserted a *Snrpn* tdTomato reporter into the endogenous *miR290* and *Sox2* enhancer (Figure 4B and Figure S3B, respectively). As recipient cells, we used the previously established *Oct4, Sox2, Klf4* and *c-Myc* (OSKM) polycistronic dox-inducible secondary reprogrammable mESCs (Carey et al., 2011), which also carried a GFP reporter knocked into the endogenous *Nanog* locus. Correct integration of the vector was validated by PCR and Southern analysis (Figure S3C). Figure 4C shows that both targeted ESC lines (*miR290* #21 and *Sox2* #2) expressed tdTomato as well as *Nanog*-GFP. To assess whether the tdTomato expression correlated with hypomethylation of the inserted RGM, DNA extracted from the bulk mESCs population was bisulfite converted, amplified by PCR and sequenced with the PCR amplification including both the SE CpG region and the downstream *Snrpn* promoter. As predicted from the methylation maps (Figure 4A and Figure S3A), both endogenous *miR290* and *Sox2* CpG regions were mostly hypomethylated (Figure 4D). Importantly, the *Snrpn* promoter was also hypomethylated consistent with reporter expression. Of note, a few highly methylated alleles were detected (Figure 4D), possibly reflecting an inherent variation in the bulk population due to the presence of cells that carry

an inactive reporter. To test this possibility, we analyzed the *Sox2* SE region in the untargeted parental cell, which identified the presence of both methylated and unmethylated alleles at the same frequency as the targeted reporter cell line (Figure S3D). We conclude that RGM can report on the methylation state of distal genomic regulatory regions.

Dynamic *de novo* DNA methylation during differentiation

To monitor real-time changes in genomic DNA methylation during *in vitro* differentiation, mESCs carrying the tdTomato reporters reflecting DNA methylation levels at the SE regions, were exposed to Retinoic Acid (RA), which induces a rapid exit from pluripotency, and cellular differentiation (Rhinn and Dolle, 2012). The presence of the *Nanog*-GFP reporter allowed monitoring exit from pluripotency by loss of GFP expression. Sorted double positive ($tdTomato^+$ / GFP^+) *miR290* and *Sox2* cells were plated on feeder-free gelatin coated plates, treated with 0.25uM RA the following day (Figure 5A) and analyzed at different times after addition of RA (Figure 5A and B). As expected, undifferentiated cells were double positive ($tdTomato^+$ / GFP^+). However, upon induction of differentiation a gradual reduction in the fraction of double positive cells was observed with most disappearing over the time course of 7 days, resulting in a largely double negative cell population (Figures 5B and 5C). This is in contrast to control *Gapdh* reporter cells which, as expected, appeared completely GFP positive following 7 days of RA differentiation (Figure S4A). tdTomato and *Nanog*-GFP positive cells disappeared with different kinetics: while singly tdTomato positive cells ($tdTomato^+$ / GFP^-) appeared after 2 days, only a few single *Nanog*-GFP positive cells ($tdTomato^-$ / GFP^+) were detected during differentiation (Figure 5B and 5C) suggesting that *Nanog* was silenced prior to methylation and silencing of the *miR290* and *Sox2* SEs.

To confirm that loss of the tdTomato signal correlated with accumulation of *de novo* methylation in both SE regions, we sorted the main populations at different time points during RA differentiation (Figure 5C). DNA was extracted from the different cell populations and subjected to bisulfite sequencing, thus allowing a comprehensive analysis of the methylation state in both the endogenous *miR290* and *Sox2* SE and their respective *Snrpn* promoter regions (Figures 5D,E, S4B, C). In contrast to the bulk population of mESCs (Figure 4D), the sorted double positive cells did not harbor completely methylated alleles, consistent with the notion that methylated alleles in the bulk population represent intrinsic variation. The methylation of both *miR290* and *Sox2* in single positive cells ($tdTomato^+$ / GFP^-) was low, consistent with tdTomato expression. The overall increased *de novo* methylation in the single positive cells, compared with the double positive cells, may suggest that DNA methylation mediated silencing was already initiated in this intermediate cell population. Notably, our analysis identified completely methylated genomes in the *Sox2* single positive ($tdTomato^+$ / GFP^-) cell population (Figure 5E). This suggest that during rapid changes of *de novo* methylation the half-life of the fluorescent protein (FP) may lead to an over-estimation of cells that are still hypomethylated during cell state transitions. Finally, in agreement with the silencing of tdTomato expression, the double negative cells ($tdTomato^-$ / GFP^-) exhibited robust hypermethylation on both endogenous SE regions and their respective *Snrpn* promoters (Figures 5D, E, S4B, C). To test whether the targeted reporter allele correlated with the methylation levels of the untargeted allele (wt), we

analyzed the wt allele in *Sox2* reporter cells at different time points during differentiation. Figure S4D shows that similar to the reporter allele, the wt allele exhibited low levels of methylation in the sorted double positive cells, and high levels of methylation following seven days of differentiation. We conclude that RGM allows dynamic monitoring *de novo* methylation events that are imposed on genomic sequences upon exiting from pluripotency. Our data suggest that the differentiation of ESCs induces silencing of *Nanog* prior to *de novo* methylation of the two *miR290* and *Sox2* SEs.

To test whether *in vivo* differentiation resulted in silencing of the tdTomato reporter in both *miR290* and *Sox2* SE regions, we analyzed 13.5 dpi chimeric embryos. As control, we injected ESCs harboring the *Gapdh* CGI reporter driving a GFP sequence, which had also been infected with lentiviruses resulting in constitutive expression of tdTomato. The robust expression of GFP in the *Gapdh* control embryos, demonstrated the widespread expression signature of the *Snrpn* promoter throughout mouse tissues (Figure 6A). Unlike the *Gapdh* control, both *miR290* and *Sox2* embryos were completely negative for both GFP and tdTomato, demonstrating robust repression of *Nanog* and the *Snrpn* promoter during *in vivo* differentiation (Figure 6A)

DNA demethylation during cellular reprogramming

Reprogramming of somatic cells to iPS cells involves demethylation and activation of the pluripotency SEs *Sox2* and *miR290* (see Figures 4A and S3A). We investigated whether RGM could be used to capture demethylation events that are gradually acquired during cellular reprogramming. For this we used secondary Dox-inducible reprogrammable mouse embryonic fibroblasts (MEF)s isolated from 13.5 dpi chimeric embryos that had been injected at the blastocyst stage with the OSKM DOX inducible ESCs (Carey et al., 2011) carrying *Nanog*-GFP and the tdTomato reporter reflecting DNA methylation levels at the *Sox2* or *miR290* SE alleles (see Figure 6B). Culture of these MEFs in DOX induces the reprogramming factors while *Nanog*-GFP activation allows monitoring the course of reprogramming in the bulk somatic cell population (Buganim et al., 2012). As expected, MEFs isolated from 13.5 dpi embryos were negative for both GFP and tdTomato expression, as measured by fluorescent microscopy and FACS analysis (Figures 6C and S5A). Importantly, consistent with tdTomato repression, both endogenous *miR290* and *Sox2* SE regions as well as their corresponding downstream *Snrpn* promoter regions were hypermethylated (Figure 6D). Further analysis of the wt allele in *Sox2* MEF showed high correlation with the targeted reporter allele, demonstrating robust repression of the SE region *in vivo* (Figure S5B).

To test whether reprogramming-induced demethylation can be visualized by RGM, we treated the secondary MEFs with serum and LIF medium supplemented with 2 ug/ml doxycycline (Dox). Both *miR290* and *Sox2* MEFs were successfully reprogrammed, resulting in double positive cells ($tdTomato^+$ / GFP^+, data not shown). It was recently shown that a combination of three chemicals, TGF-β antagonist ALK5 inhibitor II; GSK3b antagonist CHIR99021 and Ascorbic Acid, an enzymatic cofactor (from here on referred to as 3C), results in more efficient and synchronous reprogramming (Vidal et al., 2014). To achieve more synchronized and efficient reprogramming, both *miR290* and *Sox2* MEFs were

subjected to 3C culture conditions and the dynamics of reporter activation was monitored by flow cytometry. While the first expression of tdTomato$^+$ and GFP$^+$ cells emerged at day 16 (Figure 6E), reporter activation of both *miR290* and *Sox2* occurred with different kinetics. Figure 6E shows accumulation of *miR290* reporter cells that activated both GFP and tdTomato (tdTomato$^+$ / GFP$^+$) over time. A small population of single positive GFP cells appeared in late stages of reprogramming consistent with a stochastic sequence of events in the reprogramming of the *miR290* SE region. Unlike *miR290* reporter cells, however, *Sox2* cells showed a more robust and defined dynamics of activation of both reporters. By day 16 a population of single positive GFP cells (tdTomato$^-$ / GFP$^+$) had accumulated, which gradually shifted to become double positive (tdTomato$^+$ / GFP$^+$) over time (Figures 6E and S5C). To test whether the single positive GFP cells give rise to double positive cells, we sorted the single positive GFP cells and replated them on feeders using Dox independent culture conditions. Consistent with the repression of the tdTomato signal, bisulfite sequencing confirmed that the single positive GFP cells exhibit high levels of methylation in the SE region, as well as in the downstream *Snrpn* promoter region (Figure S5D). Upon further culture tdTomato positive cells appeared demonstrating that single positive GFP cells give rise to double positive cells (Figure S5E).

Our results suggest that reprogramming of both *miR290* and *Sox2* SE regions are late events, with the *Sox2* SE region being reprogrammed subsequently to the activation of endogenous *Nanog*. *miR290* and *Sox2* double positive (tdTomato$^+$ / GFP$^+$) cells invariably proceed to a Dox independent iPS cell state (Figure 6F). To assess the methylation state of the *Sox2* and *miR290* SEs, we performed bisulfite sequencing on DNA extracted from sorted double positive (tdTomato$^+$ / GFP$^+$) iPS cells. As shown in Figure 6G, both *miR290* and *Sox2* SE regions, and their corresponding downstream *Snrpn* promoters were demethylated. These results confirmed that RGM can visualize demethylation of regulatory genomic regions during reprogramming with single cell resolution.

Discussion

In this work we have generated a DNA methylation reporter (RGM) that allows imaging of DNA methylation with single cell resolution. The design of the reporter system took advantage of the intrinsic characteristics of imprinted gene promoters, for which the transcriptional activity reflects the DNA methylation state of adjacent sequences. Importantly, imprinted promoters are neutral to developmental or tissue specific DNA methylation changes, with their activity strictly dependent on the methylation state of the adjacent regulatory elements. This is in contrast to CGI sequences such as *Gapdh* or tissue-specific elements such as the *Dazl* promoter associated sequences, which become demethylated or *de novo* methylated, respectively, when inserted into the genome of ESCs (Brandeis et al., 1994; Sabag et al., 2014). This indicates that methylation of these elements as opposed to imprinted promoters is sequence – dependent and subject to trans-acting signals and cell state-dependent regulation.

The RGM reporter system described here is based on the *Snrpn* minimal promoter that is not subjected to methylation changes by itself, and therefore GFP expression is solely dependent on the methylation state of surrounding sequences. Consistent with this premise, ES cells

appeared GFP positive when stably transfected with the methylated or unmethylated *Gapdh*/*Snrpn*-GFP vector, but were GFP negative when transfected with the methylated or unmethylated *Dazl*/*Snrpn*-GFP reporter. This indicates that the *Snrpn* promoter region can be used as a faithful sensor for regional methylation changes of adjacent sequences.

To investigate whether RGM can report on the methylation state of endogenous loci we targeted CGIs located at *Gapdh* and *Dazl* promoter regions, resulting in differential methylation and activity of the *Snrpn* reporter. Thus, the *Snrpn* promoter effectively reflects local methylation patterns without affecting the endogenous epigenetic state. As most of the tissue-specific DNA methylation changes occur in low-density CpG regulatory regions, we asked whether RGM could report on the methylation state of non-coding low-density CpG regions. We chose two pluripotency-specific SEs that are associated with the *miR290* and *Sox2* genes and are known to be active and unmethylated in ESCs but become methylated and inactive upon cellular differentiation. CRISPR/Cas mediated insertion of the *Snrpn*-tdTomato reporter into ESCs resulted in tdTomato positive clones but tdTomato expression was silenced in mid-gestation chimeric embryos, which reflects the demethylation state of the SEs in pluripotent cells and their *de novo* methylation upon induction of differentiation. Conversely, MEFs isolated from chimeric embryos were tdTomato negative with both elements highly methylated. Upon conversion of the MEFs into iPSCs, however, the cells became tdTomato positive reflecting demethylation of the SEs during reprogramming to pluripotency. Our results establish that RGM reporter activity mirrors the changes of DNA methylation imposed on endogenous CGI and low-density CpG genomic elements during development, upon cellular differentiation and during reprogramming. Extensive epigenomic analyses of multiple tissues and cell types in both human and mice, suggest that embryonic development and cell-type specification are associated with massive epigenomic remodeling at discrete enhancers (Hon et al., 2013; Roadmap Epigenomics et al., 2015; Schultz et al., 2015; Ziller et al., 2013). It will thus be of interest to test whether RGM can be utilized to report on the DNA methylation state associated with more discrete regulatory regions. Implementing the methylation reporter to tissue-specific DMRs holds the promise to further elucidate the link between DNA methylation and other epigenetic mechanisms, with cell fate regulation.

Reprogramming of somatic cells into iPSCs involves extensive resetting of the epigenome (Buganim et al., 2013; Hanna et al., 2010), and coinciding with this notion, recent studies identified key role for epigenetic modifiers during this process (Mansour et al., 2012; Rais et al., 2013; Soufi et al., 2012). However, the exact kinetics of these epigenetic changes during the reprogramming process are difficult to define because of cell heterogeneity and the stochastic nature of the reprogramming process. Here we followed the methylation changes of two SEs associated with *Sox2* and *miR290*, demonstrating that demethylation of both regions are late events in the reprogramming process. Simultaneous activation of endogenous *Nanog* and *miR290* SE demethylation is consistent with *Nanog* directly regulating the expression of *miR290* cluster during reprogramming to iPS cells (Gingold et al., 2014). The gradual activation of the *Sox2* tdTomato reporter followed expression of endogenous *Nanog*, consistent with demethylation of *Sox2* SE being a late event in the process (Buganim et al., 2012). Systematic deletions of *Sox2* upstream SE region was

recently shown to dramatically affect *Sox2* expression in ESCs (Li et al., 2014; Zhou et al., 2014). Thus, the *Sox2* SE methylation reporter cells provide a rigorous experimental system to investigate how DNA methylation changes at distal regulatory region influence the expression of downstream target genes.

Changes in DNA methylation during development, lineage commitment and disease are dynamic and studies of epigenetic changes are hampered by two experimental constraints that limit mechanistic studies of methylation and gene regulation. (i) Current methodology provides only a static "snapshot" view of the methylation state during cell state transitions and (ii) that current methylation analyses require the examination of multiple cells precluding assessment of epigenetic changes in single cells. Given the overwhelming evidence of cell-cell heterogeneity in embryos, cultured cells or disease states such as cancer (Junker and van Oudenaarden, 2014), this is a serious limitation for a mechanistic understanding of the epigenetic state and gene expression during these complex processes. For example, monitoring the course of differentiation in both *miR290* and *Sox2* reporter cells confirmed the co-existence of cell populations that harbor distinct epigenetic states. In contrast, commonly used bulk methodologies would not allow isolating and distinguishing the different cell populations. Thus, sorting and isolating different cell types according to their methylation states, can be achieved only by using readout for methylation state at single-cell resolution. The RGM reporter system overcomes some of the limitations of conventional methylation analyses by providing real time visualization of DNA methylation at single cell resolution. As with any fluorescent protein-based reporter system, the accuracy to trace real-time changes depends on the half-life of the respective FP. Because the current version of the methylation reporter does not use a destabilized FP, silencing of the reporter after de novo methylation-induced repression of the *Snrpn* promoter is likely delayed. To generate a reporter that more rapidly reports on DNA methylation changes would require the use of a destabilized FP. Targeting additional loci in future studies will allow to further elucidate other possible limitations of the RGM reporter system, such as inhibition of the *Snrpn* transcriptional activity by chromatin conformation.

As RGM allows measuring dynamics of DNA methylation at single-cell resolution, it provides a framework for understanding epigenetic changes during cell state transition in heterogeneous cell populations. For example, replacing the fluorescent-based reporter system with Cre-Lox, will enable the generation of epigenetic lineage tracing maps. Furthermore, utilizing RGM together with conventional gene expression reporters may offer detailed insights into the interplay between epigenetic cues and the execution of tissue-specific gene expression programs. The use of fluorescent reporters as readout for locus-specific methylation changes may also provide an effective screening platform for the isolation of small molecule compounds that affect the methylation state of specific genomic regions.

Experimental Procedures

mESCs Cell Culture

V6.5 mouse embryonic stem cells (mESCs) were cultured on irradiated mouse embryonic fibroblasts (MEFs) with standard ESCs medium: (500 ml) DMEM supplemented with 10%

FBS (Hyclone), 10 ug recombinant leukemia inhibitory factor (LIF), 0.1 mM beta-mercaptoethanol (Sigma-Aldrich), penicillin/streptomycin, 1mM L-glutamine and 1% nonessential amino acids (all from Invitrogen). For experiments in 2i culture conditions, mESCs were cultured on gelatin-coated plates with N2B27 + 2i + LIF medium containing: (500 ml), 240 ml DMEM/F12 (Invitrogen; 11320), 240 ml Neurobasal media (Invitrogen; 21103), 5 ml N2 supplement (Invitrogen; 17502048), 10 ml B27 supplement (Invitrogen; 17504044), 10 ug recombinant LIF, 0.1 mM beta-mercaptoethanol (Sigma-Aldrich), penicillin/streptomycin, 1mM L-glutamine and 1% nonessential amino acids (all from Invitrogen), 50 ug/ml BSA (Sigma), PD0325901 (Stemgent, 1 uM), CHIR99021 (Stemgent, 3 uM).

Reporter Cell lines

To generate stably integrated *Gapdh* and *Dazl* transgene reporter cell lines, either *Gapdh*- or *Dazl*- modified PiggyBac transposon (see Extended Experimental Procedures), and a helper plasmid expressing transposase, were transfected into mESCs cells using Xfect mESC Transfection Reagent (Clontech), according to the provider's protocol. Stably integrated reporter cells were selected with puromycin (2mg/ml) for four days.

To generate *Dazl*, *Gapdh*, *miR290* and *Sox2* SE reporter cell lines, targeting vectors and CRISPR/Cas9 were transfected into mESCs using Xfect mESC Transfection Reagent (Clontech), according to the provider's protocol. 48 hours following transfection, cells were FACS sorted for GFP or tdTomato expression (respectively), and plated on MEF feeder plates. Single colonies were further analyzed for proper and single integration by southern blot and PCR analysis.

Flow Cytometry

To assess the proportion of GFP and tdTomato in the established reporter cell lines, a single cell suspension was filtered, and assessed on the LSR II SORP, LSRFortessa SORP or FACSCanto II.

Retinoic acid-induced differentiation

mESCs carrying the reporter for both *miR290* and *Sox2* SE region, were sorted for double positive GFP and tdTomato expression, and plated on gelatin coated plates in ES cell medium (+LIF). The next day, cells were washed with PBS and resuspended in basal N2B27 medium (2i medium without LIF, Insulin and the two inhibitors), supplemented with 0.25 uM RA. Medium was replaced every other day.

Blastocyst Injections for the Generation of Chimeras and secondary MEFs

Blastocyst injections were performed using (C57Bl/6xDBA) B6D2F2 host embryos. In brief, B6D2F1 females were hormone primed by an i.p. injection of PMS (Pregnant Mare Serum Gonadotropin, EMD Millipore) followed 46h later by an injection of hCG (human Chorionic Gonadrotropin, VWR). Embryos were harvested at the morula stage and cultured in a CO2 incubator overnight. On the day of the injection, groups of embryos were placed in drops of M2 medium and using a 16 um diameter injection pipet (Origio, Inc.) approximately 10 cells were injected into the blastocoel cavity of each embryo using a Piezo

micromanipulator (Prime Tech, Ltd). About 20 blastocysts were subsequently transferred to each recipient female; the day of injection was considered as 2.5 dpc. Fetuses were collected at 13.5 dpc for the extraction of embryonic fibroblasts as described before (Buganim et al., 2012).

Southern Blots

10–15 ug of genomic DNA was digested with appropriate restriction enzymes overnight. Subsequently, Genomic DNA was separated on a 0.7% agarose gel, transferred to a nylon membrane (Amersham) and hybridized with 32P random primer (Stratagene) labeled probes.

Reprogramming to iPSCs

MEFs isolated from *miR290* and *Sox2* fetuses, were plated at density of 50,000 cells per 6-well in gelatin coated plates with standard MEF medium (mESCs media without LIF). The following day MEF medium was replaced with mESCs medium containing 2mg/ml doxycycline (Sigma). Alternatively, cells were grown in mESCs medium containing 2mg/ml doxycycline and a combination 3 compounds: TGF-β antagonist ALK5 inhibitor II; GSK3b antagonist CHIR99021 and Ascorbic Acid, as described before (Vidal et al., 2014). Medium was replaced every other day during the course of reprogramming.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Thorold W. Theunissen, Patti Wisniewski and Colin Zollo for FACS analyses and cell sorting, Denes Hnisz for providing ChIP-seq tracks, Kibibi Ganz for mouse injections, Huijing Yu for help in cloning, and Stefan Semrau for help with the RA differentiation and comments on the manuscript. This study was supported by NIH grant HD 045022. Y.S. is supported by a Human Frontier Postdoctoral Fellowship and. R.J. is co-founder of Fate Therapeutics and an adviser to Stemgent.

References

Bird A. DNA methylation patterns and epigenetic memory. Genes & development. 2002; 16:6–21. [PubMed: 11782440]

Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H. Sp1 elements protect a CpG island from de novo methylation. Nature. 1994; 371:435–438. [PubMed: 8090226]

Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL, van Oudenaarden A, Jaenisch R. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell. 2012; 150:1209–1222. [PubMed: 22980981]

Buganim Y, Faddah DA, Jaenisch R. Mechanisms and models of somatic cell reprogramming. Nature reviews Genetics. 2013; 14:427–439.

Buiting K, Saitoh S, Gross S, Dittrich B, Schwartz S, Nicholls RD, Horsthemke B. Inherited microdeletions in the Angelman and Prader-Willi syndromes define an imprinting centre on human chromosome 15. Nature genetics. 1995; 9:395–400. [PubMed: 7795645]

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. [PubMed: 22136932]

Cedar H, Bergman Y. Programming of DNA methylation patterns. Annual review of biochemistry. 2012; 81:97–117.

Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes & development. 2011; 25:1010–1022. [PubMed: 21576262]

Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, Weintraub AS, Schuijers J, Lee TI, Zhao K, et al. Control of cell identity genes occurs in insulated neighborhoods in Mammalian chromosomes. Cell. 2014; 159:374–387. [PubMed: 25303531]

Ferguson-Smith AC. Genomic imprinting: the emergence of an epigenetic paradigm. Nature reviews Genetics. 2011; 12:565–575.

Gingold JA, Fidalgo M, Guallar D, Lau Z, Sun Z, Zhou H, Faiola F, Huang X, Lee DF, Waghray A, et al. A genome-wide RNAi screen identifies opposing functions of Snai1 and Snai2 on the *Nanog* dependency in reprogramming. Molecular cell. 2014; 56:140–152. [PubMed: 25240402]

Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, Surani MA. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science. 2013; 339:448–452. [PubMed: 23223451]

Hanna JH, Saha K, Jaenisch R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell. 2010; 143:508–525. [PubMed: 21074044]

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA. Super-enhancers in the control of cell identity and disease. Cell. 2013; 155:934–947. [PubMed: 24119843]

Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, Ren B. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. Nature genetics. 2013; 45:1198–1206. [PubMed: 23995138]

Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, et al. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nature genetics. 2009; 41:178–186. [PubMed: 19151715]

Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell. 1997; 91:501–510. [PubMed: 9390559]

Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nature genetics. 2003; 33(Suppl):245–254. [PubMed: 12610534]

Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nature reviews Genetics. 2012; 13:484–492.

Junker JP, van Oudenaarden A. Every cell is special: genome-wide studies add a new dimension to single-cell biology. Cell. 2014; 157:8–11. [PubMed: 24679522]

Kantor B, Kaufman Y, Makedonski K, Razin A, Shemer R. Establishing the epigenetic status of the Prader-Willi/Angelman imprinting center in the gametes and embryo. Human molecular genetics. 2004; 13:2767–2779. [PubMed: 15367489]

Lee HJ, Hore TA, Reik W. Reprogramming the methylome: erasing memory and creating diversity. Cell stem cell. 2014; 14:710–719. [PubMed: 24905162]

Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992; 69:915–926. [PubMed: 1606615]

Li Y, Rivera CM, Ishii H, Jin F, Selvaraj S, Lee AY, Dixon JR, Ren B. CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells. PloS one. 2014; 9:e114485. [PubMed: 25486255]

Mansour AA, Gafni O, Weinberger L, Zviran A, Ayyash M, Rais Y, Krupalnik V, Zerbib M, Amann-Zalcenstein D, Maza I, et al. The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. Nature. 2012; 488:409–413. [PubMed: 22801502]

Mummaneni P, Walker KA, Bishop PL, Turker MS. Epigenetic gene inactivation induced by a cis-acting methylation center. The Journal of biological chemistry. 1995; 270:788–792. [PubMed: 7822312]

Pawlak M, Jaenisch R. De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. Genes & development. 2011; 25:1035–1040. [PubMed: 21576263]

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. [PubMed: 24048479]

Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science. 2001; 293:1089–1093. [PubMed: 11498579]

Rhinn M, Dolle P. Retinoic acid signalling during development. Development. 2012; 139:843–858. [PubMed: 22318625]

Rivera CM, Ren B. Mapping human epigenomes. Cell. 2013; 155:39–55. [PubMed: 24074860]

Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, et al. Roadmap Epigenomics, C. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518:317–330. [PubMed: 25693563]

Sabag O, Zamir A, Keshet I, Hecht M, Ludwig G, Tabib A, Moss J, Cedar H. Establishment of methylation patterns in ES cells. Nature structural & molecular biology. 2014; 21:110–112.

Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, Rajagopal N, Nery JR, Urich MA, Chen H, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature. 2015

Smith ZD, Chan MM, Humm KC, Karnik R, Mekhoubad S, Regev A, Eggan K, Meissner A. DNA methylation dynamics of the human preimplantation embryo. Nature. 2014; 511:611–615. [PubMed: 25079558]

Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nature reviews Genetics. 2013; 14:204–220.

Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell. 2012; 151:994–1004. [PubMed: 23159369]

Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature. 2011; 480:490–495. [PubMed: 22170606]

Turker MS. Gene silencing in mammalian cells and the spread of DNA methylation. Oncogene. 2002; 21:5388–5393. [PubMed: 12154401]

Vidal SE, Amlani B, Chen T, Tsirigos A, Stadtfeld M. Combinatorial Modulation of Signaling Pathways Reveals Cell-Type-Specific Requirements for Highly Efficient and Synchronous iPSC Reprogramming. Stem cell reports. 2014; 3:574–584. [PubMed: 25358786]

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013; 153:307–319. [PubMed: 23582322]

Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. Cell. 2013; 153:1134–1148. [PubMed: 23664764]

Zhou HY, Katsman Y, Dhaliwal NK, Davidson S, Macpherson NN, Sakthidevi M, Collura F, Mitchell JA. A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. Genes & development. 2014; 28:2699–2711. [PubMed: 25512558]

Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, De Jager PL, Rosen ED, Bennett DA, Bernstein BE, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013; 500:477–481. [PubMed: 23925113]

Highlights

- A Reporter for endogenous genomic DNA Methylation (RGM) is established
- RGM can capture endogenous methylation state of promoters and non coding regions
- RGM allows tracing of methylation changes both in-vitro and in-vivo
- RGM allows monitoring dynamics at single cell resolution during cell fate changes

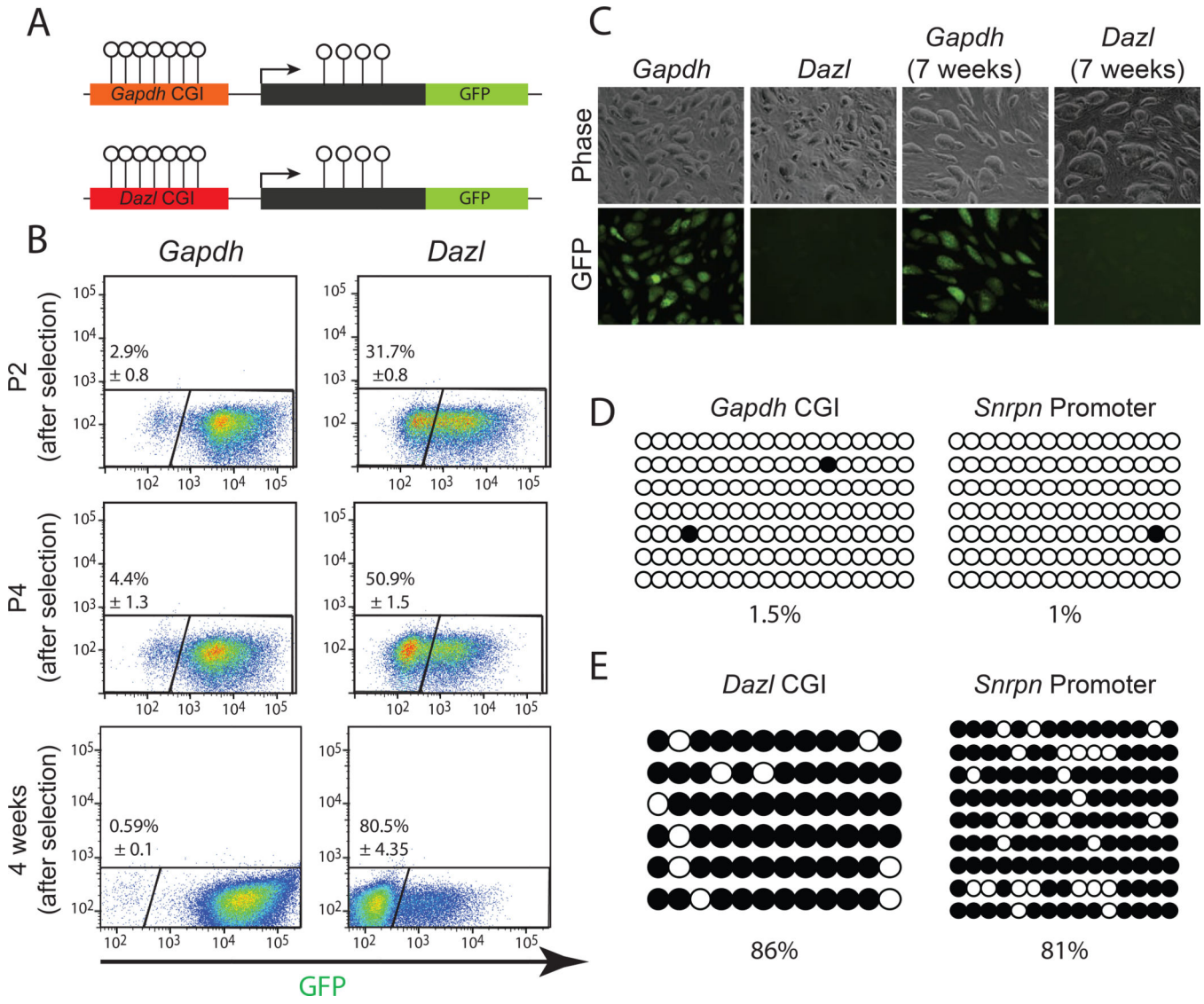

Figure 1. An active minimal *Snrpn* promoter can be repressed in *cis* by means of spreading of DNA methylation into the promoter region

(**A**) Schematic representation of the sleeping-beauty based vectors. Endogenous CpG Islands (CGI) of *Dazl* and *Gapdh* genes were cloned upstream of a minimal *Snrpn* promoter region - driving GFP. Open circle lollipops schematically represent individual unmethylated CpG.

(**B**) Flow cytometric analysis of V6.5 mESCs cultured in serum + LIF, following stable integration of unmethylated *Gapdh* and *Dazl* reporter vectors, demonstrating robust repression of GFP signal in the *Dazl* reporter cells over time. Shown are the mean percentages of GFP negative cells ± STD of two biological replicates.

(**C**) Phase and fluorescence images of the sorted V6.5 mESCs, comprising stable integration of the *Gapdh* (left) and *Dazl* (right) vectors following prolonged culturing for 7 weeks.

(**D and E**) Bisulfite sequencing analysis of the stably transfected *Gapdh* (**D**) and *Dazl* (**E**) reporter cell lines was performed on the gene promoter-associated CGI (left) and the downstream *Snrpn* promoter region (right). Open circles represent unmethylated CpGs; Filled circles - methylated CpGs.

See also Figure S1.

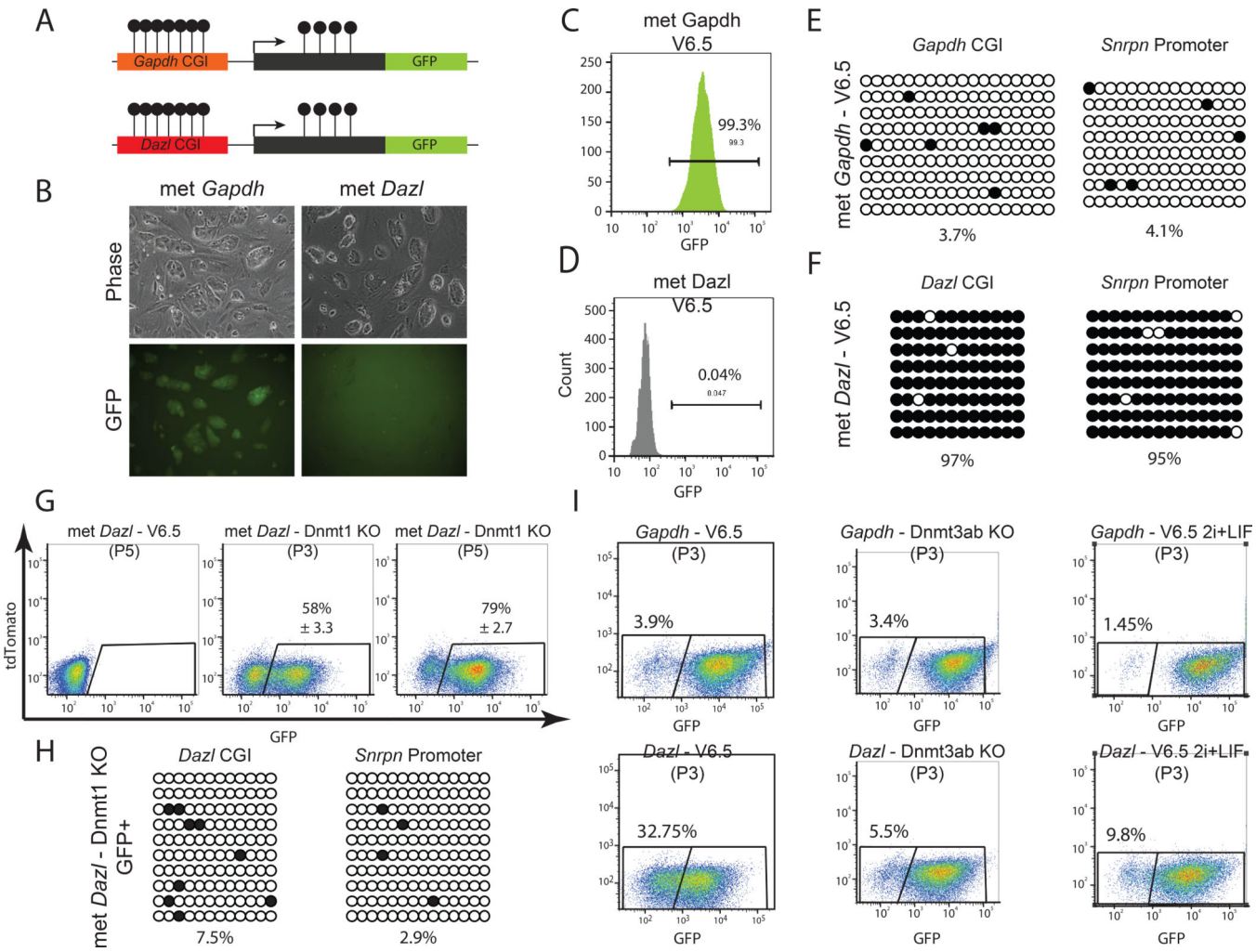

Figure 2. An *in vitro* repressed *Snrpn* promoter can be reactivated in *cis* by means of spreading of DNA demethylation into the promoter region

(**A**) Schematic representation of an *in vitro* methylated sleeping-beauty based vectors. Closed circle lollipops schematically represent individual methylated CpG.

(**B**) Phase and fluorescence images of the stably integrated V6.5 mESCs, harboring *Gapdh* (left) and *Dazl* (right) *in vitro* methylated vectors, following one week of antibiotics selection.

(**C and D**) Flow cytometric analysis of the proportion of GFP positive cells in V6.5 mESCs, stably integrated with either *Gapdh* (**C**) or *Dazl* (**D**) *in vitro* methylated vectors, following 2 weeks in culture.

(**E and F**) Bisulfite sequencing analysis of the stably transfected *Gapdh* (**E**) and *Dazl* (**F**) reporter cell lines, was performed on the gene promoter-associated CGI (left) and the downstream *Snrpn* promoter region (right).

(**G**) Flow cytometric analysis of the proportion of GFP positive cells in V6.5 mESCs and *Dnmt1* KO mESCs, stably integrated with *in vitro* methylated *Dazl* reporter vector.

(**H**) Bisulfite sequencing analysis of sorted GFP positive *Dnmt1* KO mESCs, stably integrated with *in vitro* methylated *Dazl* reporter vector.

(**I**) Flow cytometric analysis of the proportion of GFP negative cells in control V6.5 mESCs, mESCs deficient for both *Dnmt3a* and *Dnmt3b* (*Dnmt3ab* KO) and V6.5 mESCs cultured in 2i + LIF, which were stably integrated with unmethylated *Gapdh* (upper panel) and *Dazl* (lower panel) reporter vectors.
See also Figure S1.

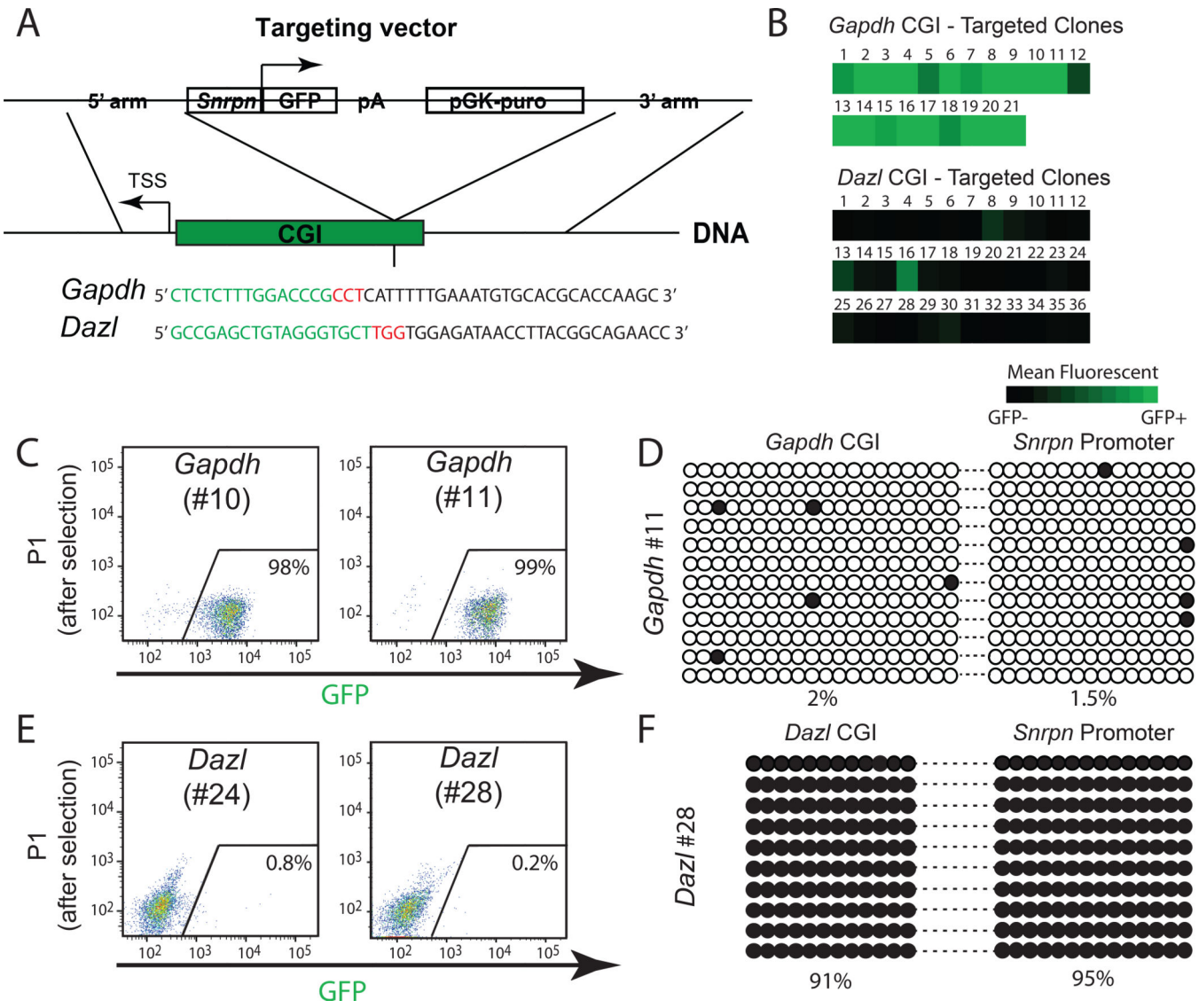

Figure 3. Generation of DNA methylation reporter cell lines for endogenous gene promoters

(**A**) CRISPR/Cas-based strategy used to integrate the DNA methylation reporter into the endogenous promoter region of *Gapdh* and *Dazl* genes. TSS - transcription start site. Green sequence - endogenous CGI region; Black sequence - targeting CRISPR; Red sequence PAM recognition site.

(**B**) Flow cytometric analysis depicting the mean GFP intensity of randomly picked clones following antibiotic selection of both *Gapdh* (upper panel) and *Dazl* (lower panel) targeted V6.5 mESCs.

(**C**) Flow cytometric analysis of the proportion of GFP positive cells in two representative clones correctly targeted with the methylation reporter at the promoter region of *Gapdh*

(**D**) Bisulfite sequencing analysis was performed on mESCs harboring the DNA methylation reporter in *Gapdh* promoter region. For each cell line, the PCR amplicon (marked with dashed line) includes both the endogenous CGI (left) and the downstream integrated *Snrpn* promoter region (right).

(**E**) Flow cytometric analysis of the proportion of GFP positive cells in two representative clones correctly targeted with the methylation reporter at the promoter region of *Dazl*.
(**F**) Bisulfite sequencing analysis was performed on mESCs harboring the DNA methylation reporter in *Dazl* promoter region. For each cell line, the PCR amplicon (marked with dashed line) includes both the endogenous CGI (left) and the downstream integrated *Snrpn* promoter region (right).
See also Figure S2.

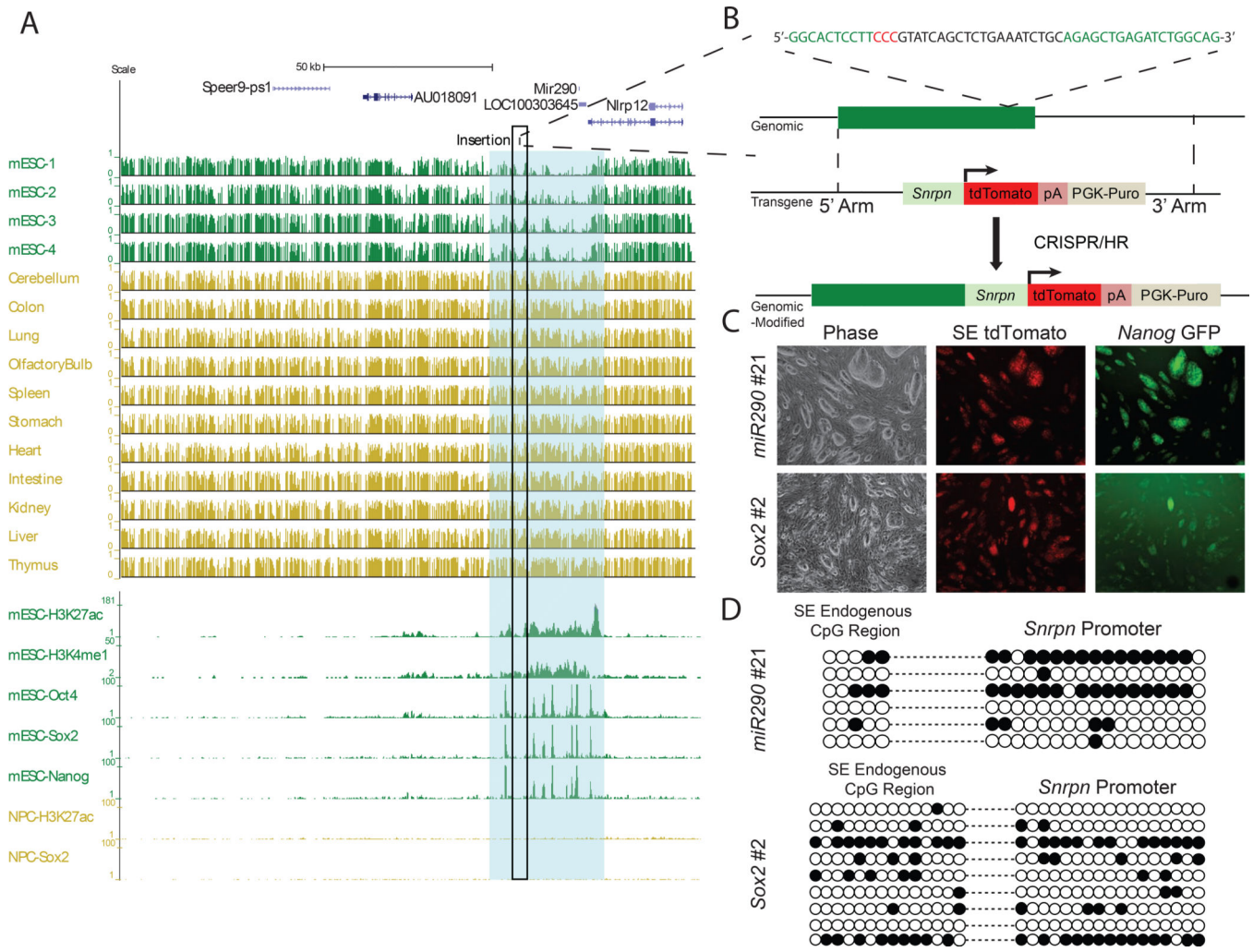

Figure 4. Generation of DNA methylation reporter cell lines for the pluripotent-specific *miR290* and *Sox2* SE regions

(**A**) Regional view depicting the DNA methylation (upper panel) and chromatin (lower panel) landscape of *miR290* upstream pluripotent-specific SE. Shown are average methylation levels and enrichment of chromatin marks in mouse undifferentiated cells (green) and in adult tissues (gold), with respect to the genomic organization of the genes. DNA methylation varies from 1-hypermethylated to 0- hypomethylated; Characteristic clusters of typical enhancer marks and binding of tissue-specific TF determine the SE region (light blue).

(**B**) CRISPR/Cas-based strategy used to integrate the DNA methylation reporter into the endogenous SE region. HR - homologous recombination. Green sequence - endogenous *miR290* CpG region; Black sequence - targeting CRISPR; Red sequence PAM recognition site.

(**C**) Phase and fluorescence images of correctly integrated DNA methylation reporter cell lines for *miR290* (upper panel) and *Sox2* (lower panel) endogenous SE regions. GFP marks endogenous expression levels of *Nanog*, whereas tdTomato reflects the endogenous DNA methylation levels at both *miR290* and *Sox2* SE regions.

(**D**) Bisulfite sequencing analysis was performed on undifferentiated mESCs harboring the DNA methylation reporter in either *miR290* SE region (upper panel) or *Sox2* SE region (lower panel). For each cell line, the PCR amplicon (marked with dashed line) includes both the endogenous CGI (left) and the downstream integrated *Snrpn* promoter region (right). See also Figure S3.

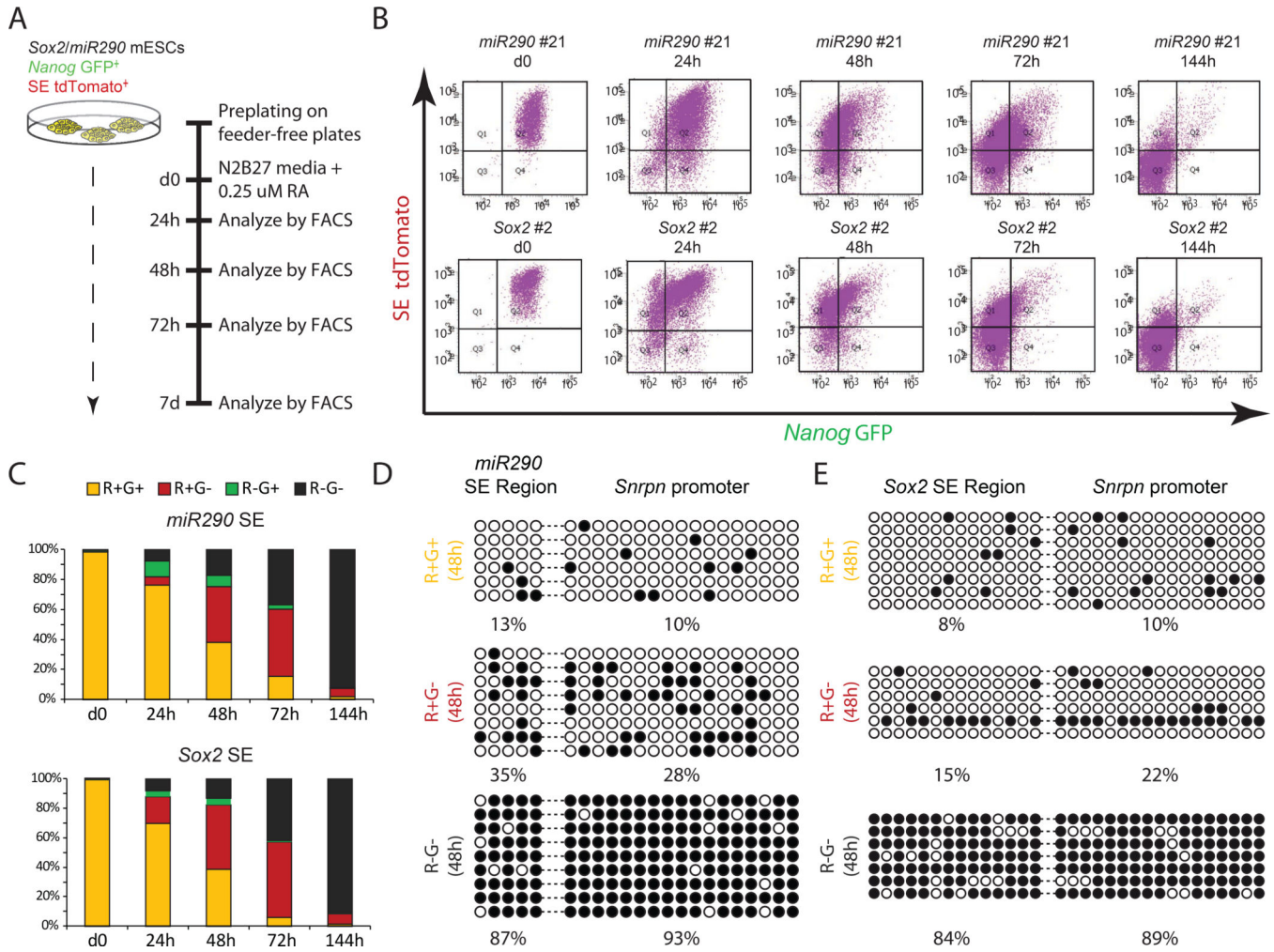

Figure 5. Dynamics of *de novo* DNA methylation of *miR290* and *Sox2* SE regions upon *in vitro* differentiation

(**A**) Schematic representation of the RA-based differentiation protocol used on *miR290* and *Sox2* reporter cell lines. GFP marks endogenous expression levels of *Nanog*, whereas tdTomato reflects the endogenous DNA methylation levels at both *miR290* and *Sox2* SE regions.

(**B**) Flow cytometric analysis of the proportion of *Nanog*-GFP positive cells (X axis) and tdTomato positive cells (Y axis) during 7 days of differentiation of *miR290* #21 (upper panel) and *Sox2* #2 (lower panel) reporter cell lines.

(**C**) Bar graph summarizing the proportion of the different cell populations during the course of 7 days RA differentiation for both *miR290* #21 (upper panel) and *Sox2* #2 (lower panel) reporter cell lines. Data represents two biological replicates. R - tdTomato ; G - GFP.

(**D and E**) Bisulfite sequencing analysis on the three main cell populations - sorted at 48 hours following initial treatment with RA. For both *miR290* #21 (**D**) and *Sox2* #2 (**E**) cell lines, the PCR amplicon (marked with dashed line) includes the endogenous CGI (left) and the downstream integrated *Snrpn* promoter region (right). R - tdTomato ; G - GFP.

See also Figure S4.

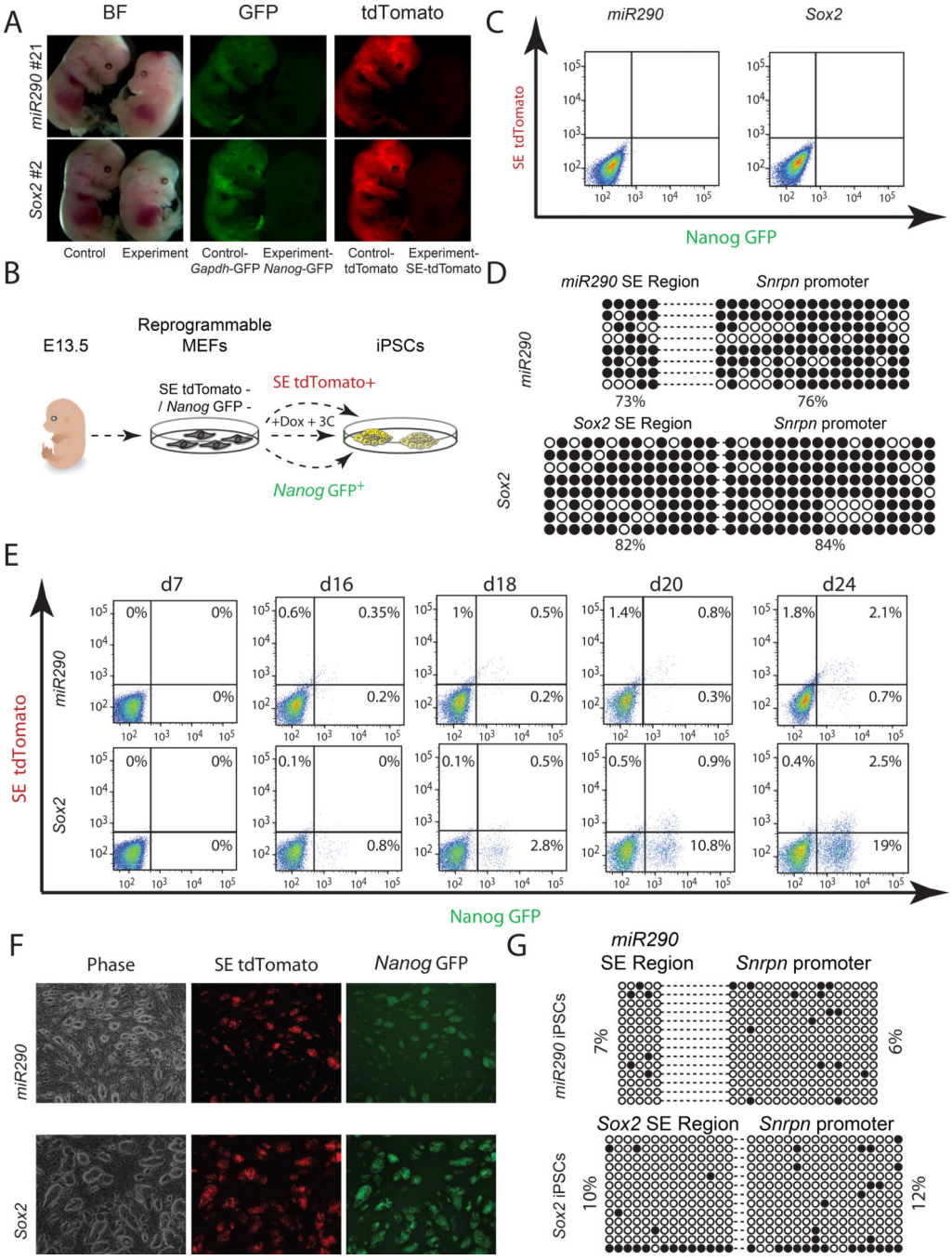

Figure 6. Dynamics of DNA demethylation of *miR290* and *Sox2* SE regions during cellular reprogramming

(**A**) *miR290* (upper panel) and *Sox2* (lower panel) reporter chimeric experimental embryos (right embryo in each panel). As controls, *Gapdh* CGI reporter mESCs driving GFP and constitutively expressing tdTomato (Control *Gapdh*-GFP and tdTomato, respectively) were injected into host blastocysts (left embryo in each panel).

(**B**) Schematic representation of the experimental procedure to monitor the dynamics of demethylation during reprogramming of *miR290* and *Sox2* reporter cell lines. GFP marks

endogenous expression levels of *Nanog*, whereas tdTomato reflects the endogenous DNA methylation levels at both *miR290* and *Sox2* SE regions.

(**C**) Flow cytometric analysis of the proportion of GFP positive cells (X axis) and tdTomato positive cells (Y axis) in P0 MEFs derived from *miR290* #21 (left) and *Sox2* #2 (right) chimeric embryos.

(**D**) Bisulfite sequencing analysis was performed on P0 MEFs derived from *miR290* #21 (upper panel) and *Sox2* #2 (lower panel) chimeras. For each cell line, the PCR amplicon (marked with dashed line) includes both the endougenous CGI (left) and the downstream integrated *Snrpn* promoter region (right).

(**E**) Analysis of the proportion of GFP positive cells (X axis) and tdTomato positive cells (Y axis) during the course of reprogramming of MEFs derived from *miR290* #21 (upper panel) and *Sox2* #2 (lower panel) chimeras. Shown are flow cytometric data from different time points following addition of dox supplemented with 3C culture condition.

(**F**) Representative images of established *miR290* and *Sox2* iPSC lines, derived from sorted double positive ($tdTomato^+$ / GFP^+) colonies.

(**G**) Bisulfite sequencing analysis was performed on P2 iPSCs derived from *miR290* #21 (upper panel) and *Sox2* #2 (lower panel) MEFs. For each cell line, the PCR amplicon (marked with dashed line) includes both the endogenous CGI (left) and the downstream integrated *Snrpn* promoter region (right).

See also Figure S5.