Article

Global epigenetic changes during somatic cell reprogramming to iPS cells

Anna Mattout[†], Alva Biran[†], and Eran Meshorer*

Department of Genetics, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
* Correspondence to: Eran Meshorer, E-mail: meshorer@cc.huji.ac.il

Embryonic stem cells (ESCs) exhibit unique chromatin features, including a permissive transcriptional program and an open, decondensed chromatin state. Induced pluripotent stem cells (iPSCs), which are very similar to ESCs, hold great promise for therapy and basic research. However, the mechanisms by which reprogramming occurs and the chromatin organization that underlies the reprogramming process are largely unknown. Here we characterize and compare the epigenetic landscapes of partially and fully reprogrammed iPSCs to mouse embryonic fibroblasts (MEFs) and ESCs, which serves as a standard for pluripotency. Using immunofluorescence and biochemical fractionations, we analyzed the levels and distribution of a battery of histone modifications (H3ac, H4ac, H4K5ac, H3K9ac, H3K27ac, H3K4me3, H3K36me2, H3K9me3, H3K27me3, and γH2AX), as well as HP1α and lamin A. We find that fully reprogrammed iPSCs are epigenetically identical to ESCs, and that partially reprogrammed iPSCs are closer to MEFs. Intriguingly, combining both time-course reprogramming experiments and data from the partially reprogrammed iPSCs, we find that heterochromatin reorganization precedes Nanog expression and active histone marking. Together, these data delineate the global epigenetic state of iPSCs in conjunction with their pluripotent state, and demonstrate that heterochromatin precedes euchromatin in reorganization during reprogramming.

Keywords: chromatin, pluripotency, reprogramming, iPSCs, ESCs, histone modifications, epigenetics

Introduction

Pluripotent embryonic stem cells (ESCs) are self-renewing cells derived from the mammalian blastocyst inner cell mass, and have the potential to differentiate to all the cell types of the embryo (Evans and Kaufman, 1981; Thomson et al., 1998). A method that directly induces pluripotent stem cells from somatic cells by the exogenous expression of central transcription factors (Oct4, Sox2, Klf4, and c-Myc) of the pluripotency network has been established recently (Takahashi and Yamanaka, 2006; Yu et al., 2007), which opened novel potential paths to the use of induced pluripotent stem cells (iPSCs) (Yamanaka, 2007; Jaenisch and Young, 2008; Wu et al., 2009).

iPSCs were found to be very similar to ESCs, with ESC-like colony morphology and expression of pluripotency markers (Jaenisch and Young, 2008; Stadtfeld and Hochedlinger, 2010). During the reprogramming process, the transcriptional network changes gradually from the somatic to the pluripotent state (Mikkelsen et al., 2008). Fully reprogrammed iPSCs express endogenous pluripotent markers, form teratomas, which contain cells of the three germ layers, and can contribute to the formation of 'all-iPSC mice' by tetraploid complementation

(Zhao et al., 2009). An intermediate, stable state between mouse embryonic fibroblasts (MEFs) and iPSCs was also identified. Such intermediates that have undergone incomplete reprogramming were referred to as 'partial iPSCs' or PiPS (Mikkelsen et al., 2008; Plath and Lowry, 2011).

iPSCs were originally described as undistinguishable from ESCs, but accumulating evidence demonstrates some epigenetic changes (Hanna et al., 2010; Stadtfeld and Hochedlinger, 2010). However, of note is that the general variation among ESC and iPSC lines seems to be in the same order as between ESCs and iPSCs (Bock et al., 2011).

Direct reprogramming of somatic cells to iPSCs provides a favorable system to study the epigenetic features that are prerequisite for pluripotency. The reprogramming process has been shown to be accompanied by DNA hypomethylation and by some epigenetic changes mostly on pluripotency-related or developmentally regulated gene promoters such as the typical ESC bivalent marking of H3K4me3/H3K27me3 of lineage-specific genes (Bernstein et al., 2006; Maherali et al., 2007; Mikkelsen et al., 2008; Koche et al., 2011).

Nuclear morphological transformations following reprogramming indicate that the changes that somatic cells must accommodate during the reprogramming process cannot be limited to specific gene expression programs and chromatin marks on pluripotent genes. We therefore reasoned that changes in iPSCs are manifested in a more global way towards an 'open' and

dynamic chromatin state that enables the plasticity of these cells akin to mouse ESCs (Meshorer et al., 2006). The global chromatin architecture in pluripotent cells has a more dispersed chromatin conformation characterized by a paucity of compacted areas compared with differentiated cells (Efroni et al., 2008; Ahmed et al., 2010; Fussner et al., 2011; Gaspar-Maia et al., 2011). This observation is further supported by a significant increase in H3K9me3-positive heterochromatin foci during mouse ESC differentiation (Meshorer et al., 2006). Moreover, the open chromatin feature of pluripotent cells is not solely an outcome of the paucity of heterochromatin but also an abundance of euchromatic marks (pan-H3ac, pan-H4ac, H3K36me2, H3K9ac, and H3K4me3) (Meshorer et al., 2006; Efroni et al., 2008).

Here we studied the global changes in the distribution and levels of histone modifications during somatic cell reprogramming to the pluripotent state in murine cells. We characterized five cell lines: R1, ESCs; 1D4, fully reprogrammed and previously characterized iPSCs (Maherali et al., 2007); Rr5, a mixed population of fully and partially reprogrammed iPSCs; Rr2, a partially reprogrammed iPSC line; and primary MEFs, from which the iPSC lines were derived. To avoid confusion, we refer to 'fully reprogrammed' cells as Nanog-positive throughout the paper. We show that most of the euchromatin/active marks, including H3ac, H3K9ac, H3K27ac, H4ac, H4K5ac, H3K4me3, and H3K36me2, are higher in the fully pluripotent state, correlating with Nanog expression levels. In contrast, in partial iPSCs, in which Nanog is not yet present, euchromatin-related histone modification levels are similar to those in MEFs. We were able to correlate the epigenetic and the pluripotent states by means of immunofluorescence (IF), which allows capturing data for single cells. This is an important advantage when measuring iPSC populations that are very often heterogeneous to some extent. Using IF, we also show that the heterochromatin protein HP1α and the heterochromatin-associated histone modification, H3K9me3, both redistribute during reprogramming from distinct foci toward a more diffused pattern. Interestingly, these changes were also observed in the partially reprogrammed lines. Finally, using time-course experiments, we find that heterochromatin remodeling begins relatively early, by day 6 of reprogramming, prior to the changes observed in active histone modification levels. This study is the first to thoroughly characterize global chromatin features during the transition from a differentiated to a fully reprogrammed state. We show the conversion of histone modifications from a somatic to a pluripotent state, and demonstrate that changes in heterochromatin distribution precede Nanog expression, while the global open chromatin epigenetic landscape occurs simultaneously with Nanog expression.

Results

Generation and characterization of partially and fully reprogrammed iPSCs

To characterize chromatin features before and after reprogramming of MEFs into iPSCs, we compared primary MEFs, a mixed population of fully and partially reprogrammed iPSCs (Rr5), partially reprogrammed iPSCs (Rr2), mouse iPSCs (1D4) and ESCs (R1). The iPSC line 1D4 is a well-characterized, fully reprogrammed line, which can generate chimeric mice (Maherali et al., 2007). The characterization of the iPSC state was determined by several criteria. First, the morphology of the induced colonies was examined. Fully reprogrammed colonies (in Rr5 cells, for example) appeared compact and round, similar to ESCs, while partially reprogrammed colonies appeared dispersed and amorphic (Supplementary Figure S1A and B). Another morphological criterion is the size of the nucleus, which significantly shrinks during reprogramming, reaching a size similar to that of R1 ESC nucleus. The average nuclear size of MEFs is at least twice that of the cells that undergo reprogramming (Supplementary Figure S1C). Along with morphological changes, the proliferation rate of all the fully and partially reprogrammed iPSCs increases, reaching a rate similar to that of R1 ESCs (data not shown). The most unequivocal feature which distinguishes the partially from the fully reprogrammed iPSCs is Nanog expression. Since Nanog is not a part of the reprogramming cocktail, it serves as an excellent indicator for pluripotency. Using IF, we find that apart from R1 ESCs, only the fully reprogrammed iPSCs (1D4 and some Rr5 cells) express Nanog (Supplementary Figure S1C). Therefore, in each of the IF experiments, we always double-stained the cells with Nanog. This allowed us to follow the epigenetic reprogramming at a single cell level, and to correlate the epigenetic state accurately with the pluripotent state of the cells. Fluorescence intensity of Nanog in MEFs was used as background level, and only signal above this threshold was considered positive (Supplementary Figure S1C).

We further examined the presence of pluripotent transcripts using RT−PCR. Endogenous Oct4 was present in R1 ESCs, 1D4 iPSCs, Rr5 iPSCs but not in Rr2 partial iPSCs or MEFs (Supplementary Figure S1D). E-Ras, a tumor-related gene that is considered to play a role in the long-term maintenance of pluripotency (Takahashi and Yamanaka, 2006), was present in all pluripotent cells but also in the Rr2 partial iPSCs, demonstrating their partial reprogramming (Supplementary Figure S1D). Although the endogenous Oct4 transcript is not present in the Rr2 partial iPSCs (Supplementary Figure S1D), IF revealed a strong Oct4 staining, demonstrating that the exogenous Oct4 was not silenced (Supplementary Figure S1G). This, again, indicates the incomplete reprogramming of the Rr2 partial iPSC line, since complete reprogramming follows retroviral silencing (Maherali et al., 2007; Okita et al., 2007). We also examined the expression level of lamin A, which is a nuclear envelope protein found in differentiated cells but not in ESCs (Constantinescu et al., 2006). As expected, we found that lamin A is absent from the nuclear envelope in fully reprogrammed iPSCs and ESCs whereas in the partially reprogrammed Rr2 cells, lamin A is present (Supplementary Figure S1H). Thus, lamin A may serve as a viable marker for complete reprogramming.

Finally, we performed alkaline phosphatase (AP) staining and teratoma formation assays. The Rr5 iPSCs were properly stained with AP, and also gave rise to the three expected germ layers in teratoma assays (Supplementary Figure S1E and F), suggesting that at least some of the transplanted cells were fully pluripotent. Rr2 partial iPSCs generated small tumors that are comprised of

undifferentiated cells with no apparent differentiation (data not shown), suggesting that they do not contain fully reprogrammed cells.

Global and specific histone acetylation levels are increased in a uniform manner in ESCs and fully reprogrammed iPSCs

The iPSCs are very similar, if not identical to ESCs in terms of morphology, proliferation rate, gene expression pattern, differentiation potential, and a few epigenetic marks that have been tested to date (Jaenisch and Young, 2008; Stadtfeld and Hochedlinger, 2010). To extend this further at a global scale, we wished to characterize the epigenetic landscape of iPSCs by analyzing the presence and level of 10 histone modifications, and to compare the global epigenetic landscape of ESCs, iPSCs, partial iPSCs and MEFs, from which the iPSCs/partial iPSCs have been generated.

We first analyzed the global acetylation level of histone H3 by double IF with Nanog. In all experiments, we tested R1 ESCs, 1D4 iPSCs, Rr5 iPSCs, Rr2 iPSCs, and MEFs (Figure 1A). Nanog staining distinguishes the 'true' pluripotent cells (ESCs and fully reprogrammed iPSCs) from the cells not yet completely reprogrammed and from MEFs (Figure 1A). The partial iPSCs and the MEFs exhibit a low Nanog fluorescent signal, which represents the background

for this antibody. DAPI staining allows visualizing the DNA content of all cells, and to easily detect the presence of MEFs around the ESC/iPSC colonies. For ESC and iPSC cell culture, a feeder layer of MEFs is commonly added in order to maintain the undifferentiated state. In our IF experiments, we used dividing MEFs for the feeder layer. This provided us with an additional internal control. Another important control is the use of the heterogeneous population of the Rr5 iPSCs, in which some of the cells are fully reprogrammed (Nanog-positive) and some cells are only partially reprogrammed (Nanog-negative) (Figure 1A and B). This cell line allowed us to compare histone modification levels within the same image field, of two cell populations that differ in their pluripotent state, but with otherwise identical features (i.e. morphology, size, proliferation rate, nuclear volume).

Quantification of the fluorescence intensity for each nucleus in its optimal confocal plane showed that the global H3 acetylation (H3ac) level is about 2-fold higher in ESC colonies compared with MEFs (Figure 1A and B, examples of MEFs in the same image field as ESCs or iPSCs are indicated by asterisks). To validate the fluorescence intensity, we show a western blot with similar results (Figure 1C). These observations are in agreement with previous
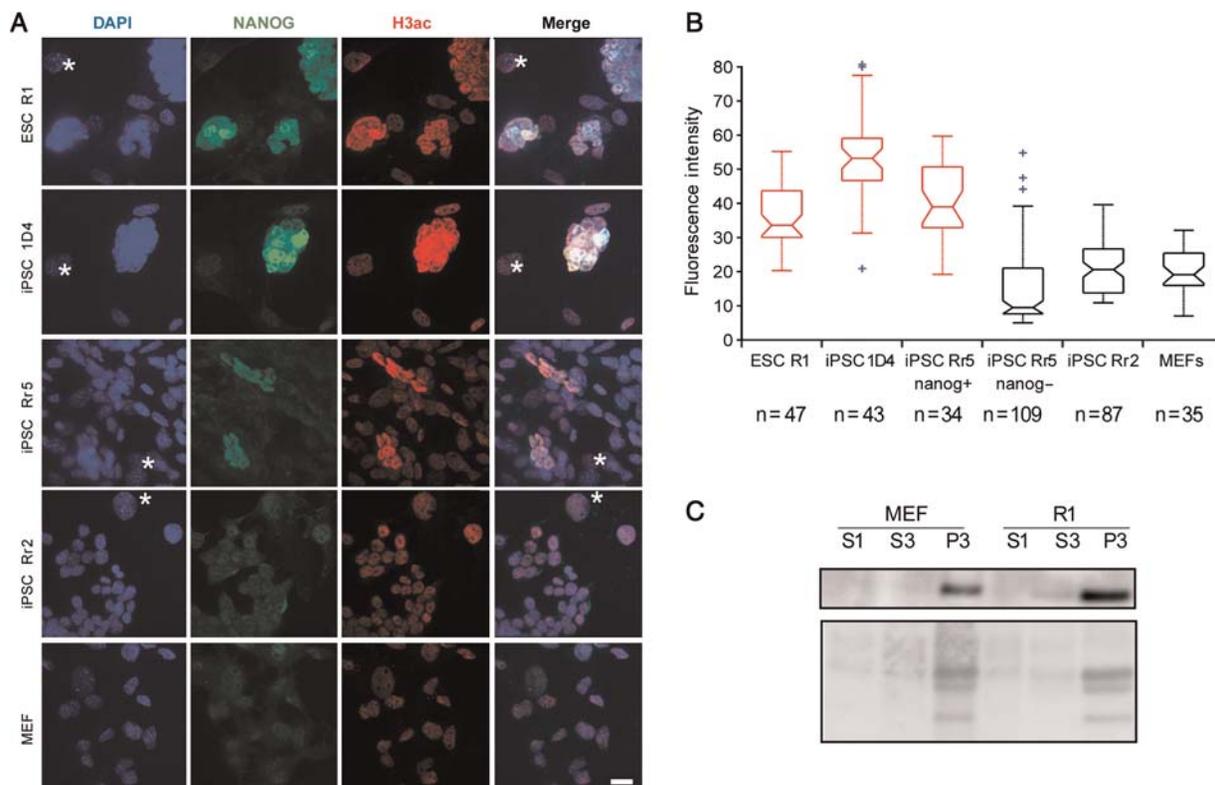
Figure 1 Fully reprogrammed iPSCs have a similar level of global histone acetylation as ESCs. (**A**) Maximum projection (3D reconstructed image) of all the tested cell types we used in this study, immunostained for Nanog (green) and histone H3 pan-acetylated (H3ac, red), and counterstained with DAPI (blue). MEFs used as feeder layer are indicated by asterisks. Bars = 15 μm. (**B**) The fluorescence intensity of H3ac is represented as boxplots for each cell type we tested. Boxes correspond to center quartiles, the bar marks the median and the whiskers indicate 1.5× of interquartiles. Fluorescence intensity was measured in confocal sections (2D) where each nucleus was at its optimal focal plane and clearly distinguishable from surrounding nuclei. The number of measured nuclei is indicated under each cell type (n). (**C**) Western blot for H3ac (top panel) in MEFs and R1 ESCs, fractionated to cytoplasmic fraction (S1), nucleoplasmic (nuclear chromatin-unbound) fraction (S3) and chromatin-bound fraction (P3). Protein staining with PonceauS (bottom panel) in the histone range of the blot was used as a loading control.

reports (Meshorer et al., 2006; Bian et al., 2009; Krejčí et al., 2009). However, a novel observation is that the fully reprogrammed 1D4 iPSCs and the Nanog[+] colonies of the Rr5 iPSCs also contain elevated levels of H3ac, similarly to ESCs, and interestingly, the partially reprogrammed cells (Nanog[−] Rr5 cells and Rr2 partial iPSCs) have lower H3ac signal, similar to MEFs (Figure 1A). This established that only the fully reprogrammed iPSCs share the same level of H3ac with ESCs, and that this high level of H3ac correlates with Nanog levels in the analyzed cells.

Similarly, we found that the level of H4ac is ∼2.5-fold higher in ESCs and in fully reprogrammed iPSCs than in partially reprogrammed iPSCs and MEFs (Supplementary Figure S2A). Because of space consideration, in this IF figure and all the subsequent ones, although we always analyzed all five cell lines depicted earlier, we only present the R1 ESCs and the Rr5 iPSCs but provide the representation of the quantified fluorescence as boxplots for each cell type, for each histone modification tested.

Histone acetylation is generally considered to correlate with transcriptional activity (Zhou et al., 2011), and many acetylations

on specific lysine residues are known to correlate with active chromatin (Creyghton et al., 2010; Rada-Iglesias et al., 2010). We thus tested the status of several specific histone acetylation marks, including lysine 5 on histone H4 (H4K5ac), H3K9ac and H3K27ac. As with H3ac and H4ac, ESCs and fully reprogrammed iPSCs displayed higher levels of these modifications than MEFs or partially reprogrammed iPSCs (Supplementary Figure S2B–D).

These results show that the tested modifications (H3ac, H4ac, H4K5ac, H3K9ac, and H3K27ac) are more abundant in ESCs than in differentiated cells; that histone acetylation patterns are globally identical in ESCs and iPSCs, and that partial reprogramming is not sufficient to reach ESC-like histone acetylation levels.

Active methylation marks are similarly abundant in ESCs and fully reprogrammed iPSCs

Next, we wondered whether histone methylation marks are also equivalent in iPSCs and ESCs. Histone methylations are associated with both active and repressive chromatin. H3K4me3 is the active binomial of the known bivalent mark (H3K4me3/ H3K27me3) in mouse and human, which was previously found to be slightly elevated in ESCs (Efroni et al., 2008); and
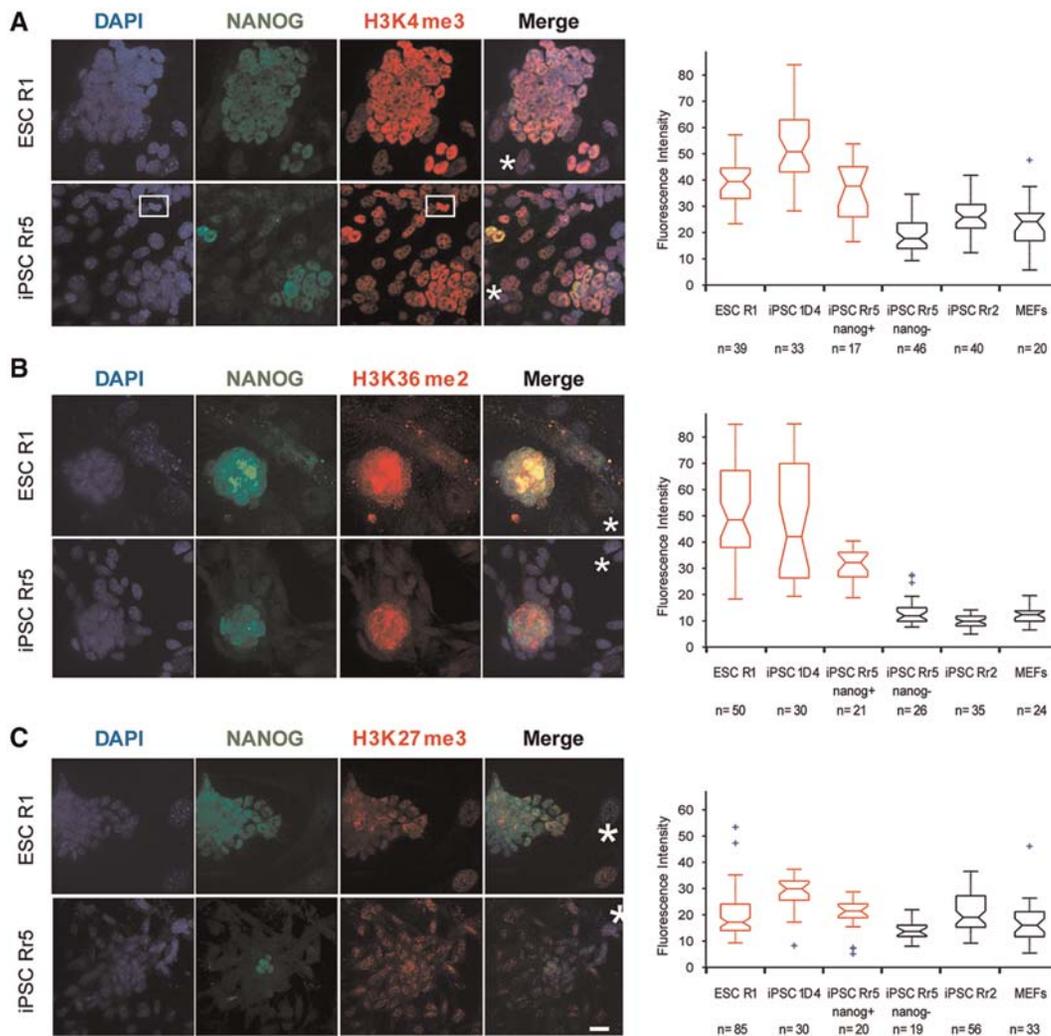
Figure 2 Fully reprogrammed iPSCs and ESCs have similar levels of euchromatic histone methylations. (**A**–**C**) 3D reconstructed images of R1 ESCs and Rr5 iPSCs immunostained for Nanog (green) and the histone modification (red): H3K4me3 (**A**), H3K36me2 (**B**), and H3K27me3 (**C**), and counterstained with DAPI (blue). MEFs used as feeder layer are indicated by asterisks. Bars = 15 μm.

H3K27me3 is a classic mark of repressive chromatin in many cell types and is the repressive binomial of the bivalent mark (Azuara et al., 2006; Bernstein et al., 2006). H3K36me2 is found to correlate with 'on' genes (Rao et al., 2005) and to be more abundant in ESCs compared with NPCs but to a lesser extent (Efroni et al., 2008). When we tested these modifications in our cells, we found that the global levels of H3K4me3 and H3K36me2 are significantly higher in ESCs and iPSCs than in MEFs or partial iPSCs (Figure 2A and B). We also observed, somewhat surprisingly, that the global levels of H3K27me3 are very similar in ESCs and MEFs (Figure 2C).

Some pluripotent heterochromatin features are present in partially reprogrammed iPSCs

HP1α is important for formation and maintenance of heterochromatin (Hediger and Gasser, 2006; Zeng et al., 2010). In differentiated somatic cells, HP1α is distributed in distinct nuclear foci, and these foci are known to usually co-localize with constitutive heterochromatic regions of the genome (Wreggett et al., 1994). In ESCs, HP1α is distributed in fewer, larger foci with less defined boundaries than in differentiated cells (Aoto et al., 2006; Meshorer et al., 2006). To examine whether this unique distribution of HP1α is also characteristic of iPSCs, we co-immnunostained our cells with Nanog and HP1α (Figure 3A). The number of HP1α foci per nucleus was found to be lower in all pluripotent cell types we analyzed than in MEFs (Figure 3A and B). Interestingly, the partial iPSC cell line Rr2 has a similar number of HP1α foci per nucleus as pluripotent cells, but the Nanog⁻ Rr5 partial iPSCs possess an intermediate number of HP1α foci between that of pluripotent cells and the MEFs (Figure 3B). These observations suggest that the two partial iPSCs, Rr2 and Nanog⁻ Rr5 were blocked at different stable stages during the reprogramming process. Based on HP1α organization, it is likely that the Nanog⁻ Rr5 cells were blocked at an earlier stage during the process than Rr2.

An additional characteristic of HP1α staining in pluripotent cells that we noticed is its significant staining in the nucleoplasm, in addition to foci (Figure 3A). Since the distribution of HP1α is not uniform, in contrast to the previous modifications that we have assessed, we measured the nucleoplasmic signal by line profile analyses (Figure 3C). Line profiles allow us to determine the intensity and width of the foci, in addition to the nucleoplasmic background intensity (Figure 3C). The nucleoplasmic signal (in 30–60 cells analyzed) in ESCs, 1D4 iPSCs, Nanog⁺ Rr5 iPSCs and Rr2 partial iPSCs is 3–4-fold higher than in MEFs, but only 1.4-fold higher in Nanog⁻ Rr5 iPSCs compared with MEFs. These results argue for a decrease in the number of HP1α foci and an increase in nucleoplasmic background signal during reprogramming, even in cells that are only partially reprogrammed, reflecting reorganization of heterochromatin during the reprogramming process.

To ask whether heterochromatin reorganization is a more general phenomenon during reprogramming, we analyzed an additional heterochromatin mark, H3K9me3. The number of H3K9me3 foci per nucleus was previously shown to be significantly lower in ESCs than in differentiated cells (Aoto et al., 2006; Meshorer et al., 2006). Consistent with this and with the HP1α foci distribution in iPSCs (Figure 3), the number of H3K9me3 foci also dropped in cells that underwent
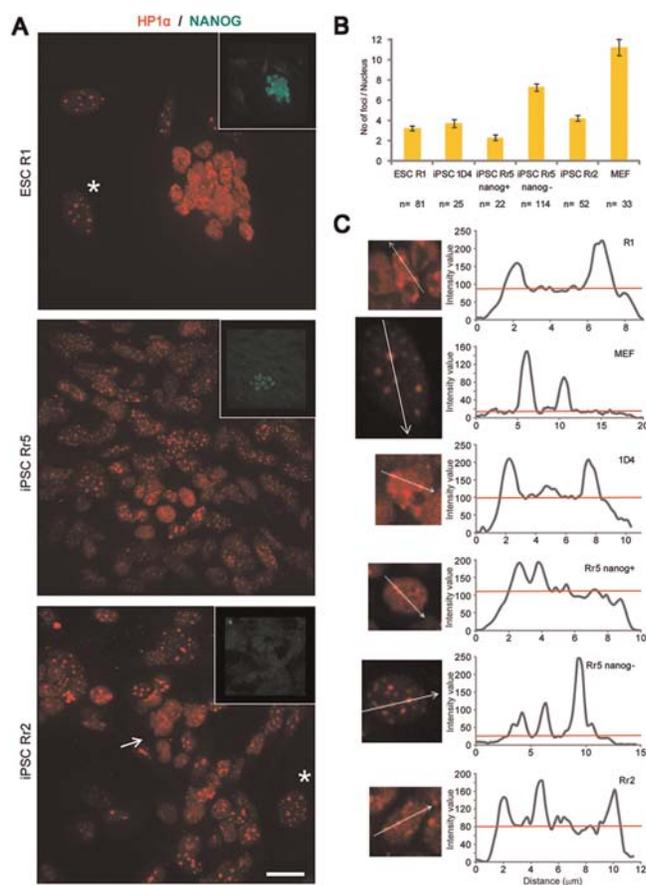
Figure 3 HP1α rearrangements in partially and fully reprogrammed iPSCs. (**A**) 3D reconstructed images of R1 ESCs, Rr5 iPSCs and Rr2 partial iPSCs immunostained for HP1α (red) and Nanog (green insets). MEFs used as feeder layer are indicated by asterisks. Bars = 15 μm. (**B**) Average foci number per nucleus is displayed for each cell type analyzed. Data were mean ± SEM. (**C**) Representative line scan plots of R1 ESCs, MEFs, 1D4 iPSCs, Nanog⁺ Rr5 iPSCs, Nanog⁻ Rr5 iPSCs, and Rr2 partial iPSCs. The scanned nucleus is shown on the left and the corresponding line plot on the right. Peaks represent foci and the red line indicates the nucleoplasmic background intensity level.

reprogramming compared with MEFs (Supplementary Figure S3A and B). Our IF results also suggest that the nucleoplasmic background of H3K9me3 staining is higher in ESCs and iPSCs than in differentiated cells (Supplementary Figure S3A and C). We found that the average nucleoplasmic signal (in 30–60 cells analyzed) in ESCs, 1D4 iPSCs, Nanog⁺ Rr5 iPSCs is almost 2-fold higher than in MEFs, and in Rr2 partial iPSCs, only 1.5-fold higher than in MEFs.

We next verified these findings biochemically by dividing cells into three fractions: cytoplasm, nucleoplasm (nuclear chromatin-unbound), and chromatin-bound (Mendez and Stillman, 2000). We found that in MEFs more than 90% of H3K9me3 is in the chromatin-bound fraction and ~10% in the nucleoplasmic fraction, whereas in ESCs, the nucleoplasmic fraction of H3K9me3 increased to more than 20% (Supplementary Figure S3D).

On the whole, the reduction in heterochromatin foci and the increase in the diffuse nucleoplasmic signal of HP1α and

H3K9me3 suggest that major heterochromatin rearrangements occur during reprogramming. Interestingly, such changes that reflect the heterochromatin state in ESCs are also apparent in partially reprogrammed iPSCs, in contrast to the open chromatin marks that we have tested that were present only in fully reprogrammed iPSCs. These observations imply that heterochromatin reorganization occurs earlier during the reprogramming process than the marking of chromatin with active modifications.

Active histone marks coincide with Nanog during reprogramming, whereas heterochromatin rearrangements precede Nanog expression

Our data so far concentrated on the relative changes of epigenetic marks in established cell lines, relative to one another. However, to clearly determine the timing of the observed changes during the reprogramming process, time-course experiments are required.

To this end, we reprogrammed primary MEFs directly on glass bottom plates, and fixed the cells at day 0, 6, 7, 8, 9, and 12 of the reprogramming. Cells were then subjected to co-immunostaining for Nanog together with H3K27ac, H3K4me3, H3K9ac, H3K9me3, or HP1α. Nanog staining began to appear at day 7 (Figure 4A), coinciding with the active marks tested (Figure 4A–C). Both Nanog expression and the levels of the examined histone modifications increased simultaneously, starting from day 7–8. At day 12, iPSC colonies were morphologically clear and Nanog expression was high, as well as the level of H3K27ac (Figure 4A) as H3K4me3 (Figure 4B) and H3K9ac (Figure 4C). In the case of H3K9me3 (Figure 4D and F) and HP1α (Figure 4E and G), we found that small forming colonies at day 6–7, which are still Nanog-negative, possessed significantly less H3K9me3 and HP1α foci than the surrounding cells that did not begin to reprogram. In those colonies, HP1α also displayed a significantly higher nucleoplasmic signal (Figure 4G). At later time points during reprogramming, the nucleoplasmic signals of H3K9me3 and HP1α continued to increase (Figure 4D and E).

These time-course experiments allowed us to conclude that heterochromatin rearrangement is the first chromatin-related change that we can detect by light microscopy in the tested marks during reprogramming. Heterochromatin changes can thus be considered as an early event during reprogramming, at the same time when the first small iPSC colonies appear, before Nanog expression.
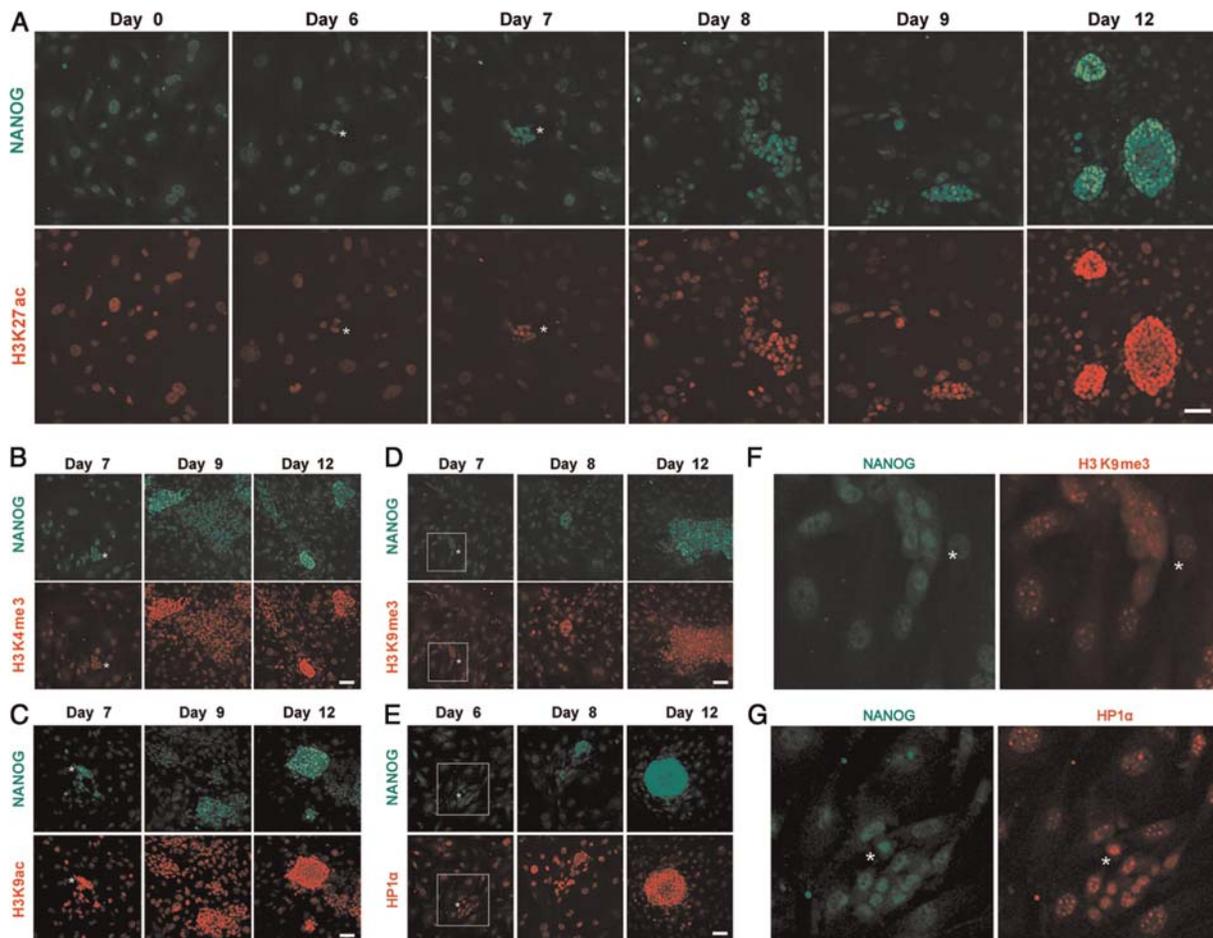
Figure 4 Heterochromatin rearrangements precede Nanog expression during reprogramming. (**A–E**) Time-course reprogramming experiments in day 0, 6, 7, 8, 9, and 12. Cells were fixed at the indicated day and immunostained with Nanog (green, top) and the indicated chromatin mark (red, bottom). Enlarged images of emerging iPSC colonies that do not yet express Nanog (green, with levels similar to background) are shown in **F** (H3K9me3) and **G** (HP1α). MEFs used as feeder layer are indicated by asterisks. Bars = 15 μm.

*Levels and localization of γH2AX and lamin A are similar
in ESCs and iPSCs*

γH2AX is one of the most well-established hallmarks of chromatin linked to DNA damage (Rogakou et al., 1999). We wondered whether this modification, which was found to be unexpectedly highly abundant in mouse ESCs in the absence of measurable DNA double-strand breaks (Banath et al., 2009), is similarly rich in iPSCs. Co-staining of γH2AX and Nanog revealed that the fully reprogrammed iPSCs have equivalent amount of this modification as ESCs, and that partial iPSCs exhibit similar level as MEFs (Figure 5A). This observation suggests that high levels of γH2AX mark are not only in ESCs but also in fully reprogrammed iPSCs, apparently regardless of DNA damage.

Lamin A is a nuclear membrane protein expressed in differentiated cells and is absent from ESCs (Constantinescu et al., 2006). Lamin A localizes clearly at the nuclear periphery in MEFs and in partial iPSCs, but in ESCs and in fully reprogrammed iPSCs lamin A is not expressed and does not localize at the nuclear periphery (Figure 5B). Thus, lamin A can be used as an additional marker to identify fully reprogrammed iPSCs from partially ones.

Finally, we calculated the ratio between the levels of the tested histone modifications in pluripotent versus non-pluripotent cells. This analysis showed that H3K36me2 and γH2AX have the highest ratio among the different modifications we have tested, reaching an almost 4-fold difference. H3K4me3 is estimated to ~2-fold higher in pluripotent cells, but H3K27me3 level was found to be roughly the same in pluripotent and non-pluripotent cells (Figure 6A). In addition, we calculated the correlation (R^2) between the histone modification levels and Nanog expression and found that most of the modifications correlate with Nanog expression (Supplementary Figure S4). H3K36me2 has the highest significant correlation ($R^2 = 0.77$), and consistent with our IF results, the correlation of H3K27me3 with Nanog was weak ($R^2 = 0.24$).

Ratio analyses for H3K9me3 and HP1α showed that the nucleoplasmic intensity is about 2- and 3.5-fold higher (respectively) in the pluripotent cells versus MEFs. Both marks also showed a significant decrease in the number of foci in the pluripotent state (Figure 6B).

To sum, we have analyzed the expression and distribution of a battery of histone modifications, as well as HP1α and lamin A, in ESCs, fully and partially reprogrammed iPSCs and primary MEFs. We found that pluripotent cells are enriched with γH2AX and active histone marks, which coincides with Nanog expression during reprogramming, and have a significantly larger unbound nucleoplasmic fraction of the heterochromatic HP1α and H3K9me3. Importantly, time-course experiments demonstrated that heterochromatin reorganization precedes Nanog expression and active histone marking, revealing the initial step that can be detected by light microscopy during reprogramming (Figure 6C).

Discussion

This work delineates the dynamic global changes in histone modifications and chromatin structure during the reprogramming process from somatic fibroblasts to iPSCs. Previous work
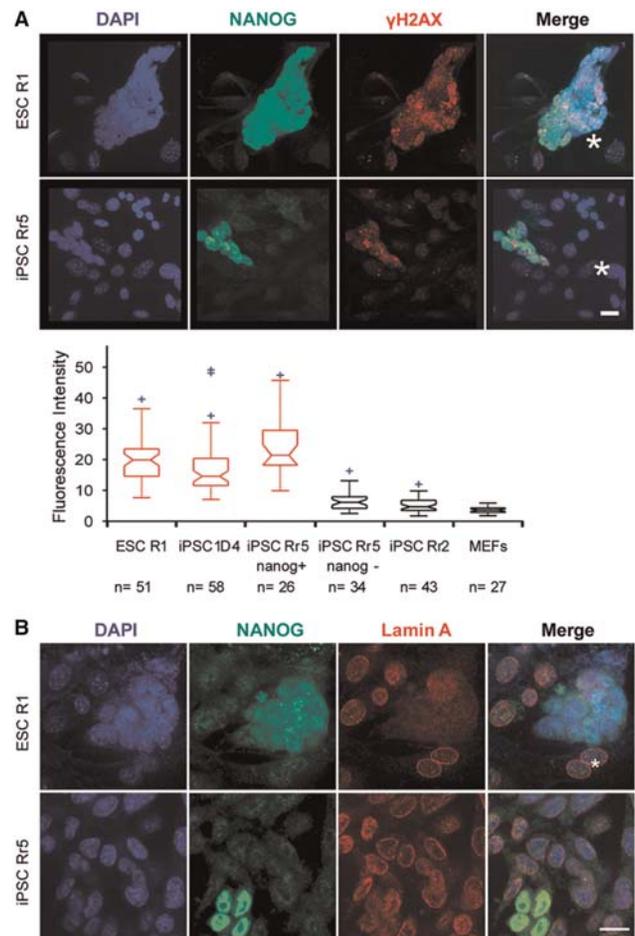
Figure 5 ESCs and iPSCs have similar high levels of γH2AX and lack lamin A expression. (**A**) 3D reconstructed images of R1 ESCs and Rr5 iPSCs immunostained for Nanog (green) and γH2AX (red), and counterstained for DAPI (blue). (**B**) Confocal sections (2D) of R1 ESCs and Rr5 iPSCs immunostained for Nanog (green) and lamin A (red), and counterstained with DAPI (blue). MEFs used as feeder layer are indicated by asterisks. Bars = 15 μm.

analyzed some molecular changes that occur during this process, revealing the kinetics of the decline in fibroblast gene expression programs and retroviral activity during the first 6–9 days (respectively) of the reprogramming process, the increase in pluripotency genes and the reversal of X-inactivation from day 10 onwards (Stadtfeld et al., 2008). More recently, a genome-wide increase in H3K4me2 was reported during the first 3 days of reprogramming in pluripotency-related or developmentally regulated gene promoters and enhancers, preceding changes in gene expression (Koche et al., 2011). Here we add the chromatin dimension to this intriguing process and demonstrate that although eventually chromatin becomes indistinguishable between iPSCs and ESCs, the transition from a somatic-like chromatin structure to an ESC-like chromatin structure occurs relatively late, at day 7 or later, but strikingly, the reorganization of heterochromatin occurs early, prior to any detectable Nanog expression, at day 6 or earlier. We further show that partially reprogrammed iPSCs, which have small nuclei similar to ESCs, show similar euchromatic histone modification patterns as MEFs, supporting the notion that
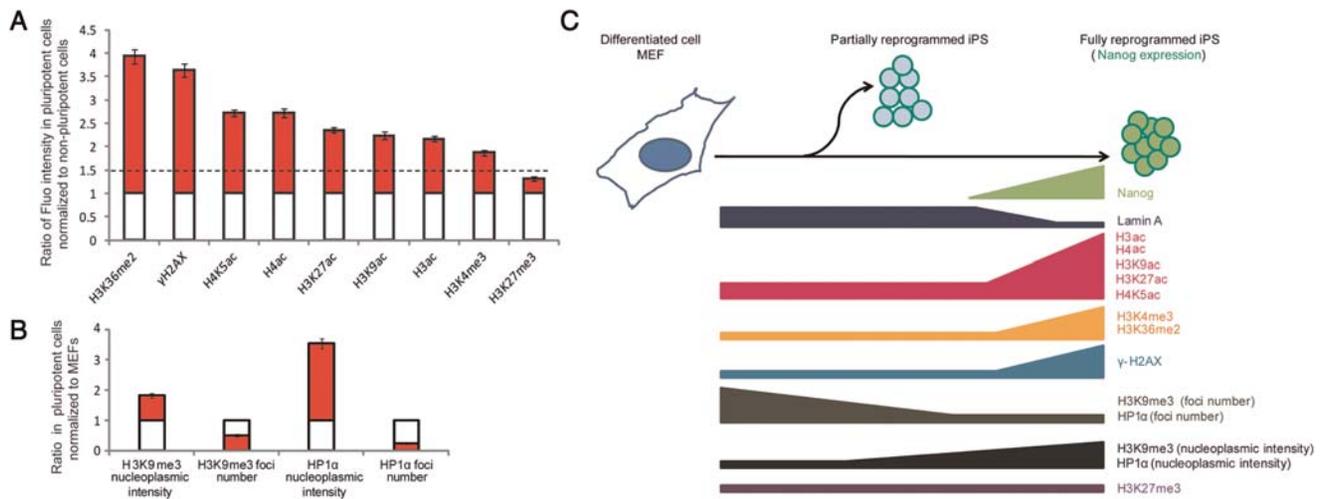
Figure 6 Summary of global epigenetic changes in iPSCs. (**A**) For each histone modification, the ratio of the average fluorescence intensity in pluripotent nuclei (R1 ESCs, 1D4 iPSCs, and Nanog$^+$ Rr5 iPSCs) versus non-pluripotent nuclei (MEFs, Nanog$^-$ Rr5 iPSCs, and Rr2 partial iPSCs) is shown in red versus the white, respectively. Data were mean $\pm$ SEM. All tested modifications, with the exception of H3K27me3, are significantly increased in pluripotent cells in comparison to non-pluripotent cells ($>$1.5-fold, $P < 0.001$ in all cases). (**B**) For the heterochromatin marks, the ratio of the average nucleoplasmic fluorescence intensity in pluripotent cells (R1 ESCs, 1D4 iPSCs, and Nanog$^+$ Rr5 iPSCs) versus MEFs (white) is shown in red. Partial iPSCs were omitted for clarity. Data were mean $\pm$ SEM. (**C**) Schematic view of the epigenetic kinetics during reprogramming. Changes represent global, and not locus-specific, trends.

the transition to an ESC-like open chromatin conformation is a relatively 'late' event during reprogramming.

Partial iPSCs have been used successfully to analyze different stages of the reprogramming process (Mikkelsen et al., 2008; Plath and Lowry, 2011). In this study by using the Rr5 cell line which we generated, which includes both partial iPSCs and fully reprogrammed iPSCs, we were able to correlate chromatin features with Nanog expression levels and distinguish between Nanog-positive and Nanog-negative cells. Remarkably, many of the tested chromatin features segregated almost perfectly between the two cell populations, suggesting that chromatin structure and epigenetic signature can be used to distinguish fully reprogrammed iPSCs from cells of other various reprogramming stages. Interestingly, the histone modifications that showed the highest degree of correlation with the pluripotent state include H3K36me2, H4K5ac, and H3K4me3 (Supplementary Figure S4). H3K4me3 is strongly associated with active transcription, and recent data have shown that H3K4me2 and H3K4me3 can also modulate gene expression likely by tethering histone deacetylase (HDAC) complexes to chromatin (Pinskaya et al., 2009). H3K36me2 correlates with the 'on' or 'off' state of transcription and is established during the initial steps of gene expression (Rao et al., 2005). Interestingly, this modification has also been recently connected to the recruitment of an HDAC complex (Li et al., 2009) and to DNA repair (Fnu et al., 2011). In the case of H4K5ac, whose high level correlates strongly with Nanog expression in ESCs and iPSCs, almost no information is known about its function in transcriptional activity. In spite of that, histone acetylation in general was shown to be prevalent in undifferentiated ESCs (Krejčí et al., 2009; Mattout and Meshorer, 2010; Gaspar-Maia et al., 2011). Specifically, H3K9ac seems to correlate with pluripotency (Hezroni et al., 2011), and H3K27ac was shown to differentiate between active and poised enhancers in ESCs (Creyghton et al., 2010;

Rada-Iglesias et al., 2010). While increased histone acetylation in ESCs may support chromatin plasticity and the hyperdynamic binding of chromatin proteins to chromatin in the undifferentiated state (Meshorer et al., 2006), the elevated level of phosphorylated H2AX (γH2AX) and its relevance to pluripotency is more difficult to explain. However, it was suggested that the abundance of γH2AX in mouse ESCs is not a consequence of increased DNA damage but rather a reflection of the open chromatin conformation in these cells (Banath et al., 2009). Since human ESCs and iPSCs do not share the same elevated levels of γH2AX foci (Banath et al., 2009; Momcilovic et al., 2009, 2010), it would be important to test whether this phenomenon is species-specific or whether it reflects the differences between human and mouse ESCs. Interestingly, the only mark that did not show any significant global level differences is H3K27me3. In a recent study (Hawkins et al., 2010), the authors used ChIP-seq of various histone modifications and showed an increase in the coverage of H3K27me3 when comparing human ESCs with IMR90 human fibroblasts. This apparent discrepancy could be attributed either to species differences—mouse in our study versus human in Hawkins et al. (2010)—or to the methodology used—our approach is a global estimate that measures grossly both the nucleoplasmic and the chromatin-bound fractions, whereas ChIP-seq measures chromatin-bound fraction alone.

One of the intriguing observations regarding the transition from a somatic to a pluripotent chromatin organization is the distribution of heterochromatin-associated marks. HP1α and H3K9me3 both appeared only in discrete condensed foci in somatic cells. In pluripotent cells, the number of such foci dropped by more than half, but importantly, HP1α and H3K9me3 displayed a pronounced nucleoplasmic staining in pluripotent cells. Using chromatin versus nucleoplasmic biochemical fractionations, we were able to show that the nucleoplasmic

fraction of chromatin-associated proteins and histone marks is significantly larger in pluripotent cells, ruling out imaging-based artifacts. We also show a higher nucleoplasmic fraction of H3ac in ESCs versus MEFs, providing a potential explanation for the unbound fraction of chromatin proteins observed using fluorescence recovery after photobleaching assays in undifferentiated ESCs (Meshorer et al., 2006). It is also tempting to hypothesize that specific modified histones are found more abundantly in the chromatin-unbound fraction in pluripotent cells in order to enable rapid incorporation into chromatin during early differentiation, and therefore to regulate gene expression programs.

Intriguingly, while all euchromatin-related features (i.e. histone acetylation, H3K4 and H3K36 methylation) appeared relatively late during reprogramming, only when cells become Nanog-positive, the reduction in the number of heterochromatin foci and the appearance of nucleoplasmic 'heterochromatin' (i.e. HP1α and H3K9me3) occurred early, before Nanog expression is apparent. This suggests that physical rearrangement of heterochromatin precedes euchromatin-related epigenetic modifications during reprogramming.

Taken together, we provide the first in-depth investigation of global chromatin mark levels during reprogramming of somatic fibroblasts to iPSCs, and show that the first discernable change is the rearrangement of heterochromatin, which precedes Nanog expression.

Materials and methods

Cell culture and reprogramming

R1 mouse ESCs (passages 18–22), 1D4 mouse iPSCs (Maherali et al., 2007), and iPSCs Rr2 and Rr5 (this study) were cultured as described in Meshorer et al. (2006). For IF, cells were plated on round sterilized 12-mm coverslips in 24-well culture plates (Grenier), coated with gelatin and pre-plated with primary MEFs. The partial iPSC lines Rr2 and Rr5 were generated as previously described (Takahashi and Yamanaka, 2006). Time-course reprogramming for the time-course experiments was conducted with EF1α-STEMCCA (Sommer et al., 2009).

Image analysis

The fluorescence intensities of Nanog and of the selected histone modifications were measured at the optimal focal plane (2D) for each nucleus in the Z-stack using the Image processing software ImageJ (Abramoff et al., 2004). We defined the nucleus borders manually, and the intensity and the nucleus size were measured in a semi-automated way. The analysis of the heterochromatin marks HP1α and H3K9me3 also involved counting of the number of distinct foci observed in each cell. This was also measured in a semi-automatic manner with ImageJ software (analyze particles' function). The intensity of the HP1α and H3K9me3 foci and more importantly of the nucleoplasmic background was further analyzed by 30–60 line plots for each cell type and the average of the nucleoplasmic background was calculated. More Materials and methods can be found in Supplementary material.

Supplementary material

Supplementary material is available at *Journal of Molecular Cell Biology* online.

References

Abramoff, M.D., Magelhaes, P.J., and Ram, S.J. (2004). Image processing with ImageJ. Biophotonics Int. *11*, 36–42.

Ahmed, K., Dehghani, H., Rugg-Gunn, P., et al. (2010). Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. PLoS One *5*, e10531.

Aoto, T., Saitoh, N., Ichimura, T., et al. (2006). Nuclear and chromatin reorganization in the MHC-Oct3/4 locus at developmental phases of embryonic stem cell differentiation. Dev. Biol. *298*, 354–367.

Azuara, V., Perry, P., Sauer, S., et al. (2006). Chromatin signatures of pluripotent cell lines. Nat. Cell Biol. *8*, 532–538.

Banáth, J.P., Bañuelos, C.A., Klokov, D., et al. (2009). Explanation for excessive DNA single-strand breaks and endogenous repair foci in pluripotent mouse embryonic stem cells. Exp. Cell Res. *315*, 1505–1520.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Bian, Y., Alberio, R., Allegrucci, C., et al. (2009). Epigenetic marks in somatic chromatin are remodelled to resemble pluripotent nuclei by amphibian oocyte extracts. Epigenetics *4*, 194–202.

Bock, C., Kiskinis, E., Verstappen, G., et al. (2011). Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell *144*, 439–452.

Constantinescu, D., Gray, H.L., Sammak, P.J., et al. (2006). Lamin A/C expression is a marker of mouse and human embryonic stem cell differentiation. Stem Cells *24*, 177–185.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl Acad. Sci. USA *107*, 21931–21936.

Efroni, S., Duttagupta, R., Cheng, J., et al. (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell *2*, 437–447.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Fnu, S., Williamson, E.A., De Haro, L.P., et al. (2011). Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. Proc. Natl Acad. Sci. USA *108*, 540–545.

Fussner, E., Djuric, U., Strauss, M., et al. (2011). Constitutive heterochromatin reorganization during somatic cell reprogramming. EMBO J. *30*, 1778–1789.

Gaspar-Maia, A., Alajem, A., Meshorer, E., et al. (2011). Open chromatin in pluripotency and reprogramming. Nat. Rev. Mol. Cell Biol. *12*, 36–47.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell *143*, 508–525.

Hawkins, R.D., Hon, G.C., Lee, L.K., et al. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell

6, 479–491.

Hediger, F., and Gasser, S.M. (2006). Heterochromatin protein 1: don't judge the book by its cover! Curr. Opin. Genet. Dev. *16*, 143–150.

Hezroni, H., Tzchori, I., Davidhi, A., et al. (2011). H3K9 histone acetylation predicts pluripotency and reprogramming capacity of ES cells. Nucleus *2*, 300–309.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567–582.

Koche, R.P., Smith, Z.D., Adli, M., et al. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell *8*, 96–105.

Krejčí, J., Uhlířová, R., Galiová, G., et al. (2009). Genome-wide reduction in H3K9 acetylation during human embryonic stem cell differentiation. J. Cell. Physiol. *219*, 677–687.

Li, B., Jackson, J., Simon, M.D., et al. (2009). Histone H3 lysine 36 dimethylation (H3K36me2) Is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. J. Biol. Chem. *284*, 7970–7976.

Maherali, N., Sridharan, R., Xie, W., et al. (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55–70.

Mattout, A., and Meshorer, E. (2010). Chromatin plasticity and genome organization in pluripotent embryonic stem cells. Curr. Opin. Cell Biol. *22*, 334–341.

Mendez, J., and Stillman, B. (2000). Chromatin association of human origin recognition complex, Cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. Mol. Cell. Biol. *20*, 8602–8612.

Meshorer, E., Yellajoshula, D., George, E., et al. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev. Cell *10*, 105–116.

Mikkelsen, T.S., Hanna, J., Zhang, X., et al. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Momcilovic, O., Choi, S., Varum, S., et al. (2009). Ionizing radiation induces ataxia telangiectasia mutated-dependent checkpoint signaling and G(2) but not G(1) cell cycle arrest in pluripotent human embryonic stem cells. Stem Cells *27*, 1822–1835.

Momcilovic, O., Knobloch, L., Fornsaglio, J., et al. (2010). DNA damage responses in human induced pluripotent stem cells and embryonic stem cells. PLoS One *5*, e13410.

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

Pinskaya, M., Gourvennec, S., and Morillon, A. (2009). H3 lysine 4 di- and tri-methylation deposited by cryptic transcription attenuates promoter activation. EMBO J. *28*, 1697–1707.

Plath, K., and Lowry, W.E. (2011). Progress in understanding reprogramming to the induced pluripotent state. Nat. Rev. Genet. *12*, 253–265.

Rada-Iglesias, A., Bajpai, R., Swigut, T., et al. (2010). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Rao, B., Shibata, Y., Strahl, B.D., et al. (2005). Dimethylation of histone H3 at lysine 36 demarcates regulatory and nonregulatory chromatin genome-wide. Mol. Cell. Biol. *25*, 9447–9459.

Rogakou, E.P., Boon, C., Redon, C., et al. (1999). Megabase chromatin domains involved in DNA double-strand breaks *in vivo*. J. Cell Biol. *146*, 905–916.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., et al. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells *27*, 543–549.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev. *24*, 2239–2263.

Stadtfeld, M., Maherali, N., Breault, D.T., et al. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230–240.

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

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., et al. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Wreggett, K.A., Hill, F., James, P.S., et al. (1994). A mammalian homologue of *Drosophila* heterochromatin protein 1 (HP1) is a component of constitutive heterochromatin. Cytogenet. Cell Genet. *66*, 99–103.

Wu, Z., Chen, J., and Ren, J. (2009). Generation of pig induced pluripotent stem cells with a drug-inducible system. J. Mol. Cell Biol. *1*, 46–54.

Yamanaka, S. (2007). Strategies and new developments in the generation of patient-specific pluripotent stem cells. Cell Stem Cell *1*, 39–49.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

Zeng, W., Ball, A.R., Jr, and Yokomori, K. (2010). HP1: heterochromatin binding proteins working the genome. Epigenetics *5*, 287–292.

Zhao, X.-Y., Li, W., Lv, Z., et al. (2009). iPS cells produce viable mice through tetraploid complementation. Nature *461*, 86–90.

Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. Nat. Rev. Genet. *12*, 7–18.