Different regulation of chromatin modification guide cellular reprogramming into a rate-limited step after initiation phase

**Background.** In the early and late stages of cell reprogramming to induced pluripotent stem cells (iPSCs) ectopic expression of Oct4, Sox2, Klf4 and Myc (OSKM) aroused two peaks of transcriptional and epigenetic change respectively. However, it was relatively quiet in the intermediate stage. In this paper our aim is to gain insight into the molecular events that occur after the initiation phase of pluripotency induction. **Methods.** GSE42379 containing 28 gene expression profiles and GSE42477 containing 10 genome binding/occupancy profiles of mouse embryonic fibroblasts (MEF) were downloaded from GEO. These datasets included untreated MEFs, OSKM-induced MEFs progressing and refractory to reprogram at 3, 6, 9, 12 day post-transduction, and iPS cell lines. Differentially expressed genes (DEGs) were identified between different cell lines. The Chip-seq peaks and putative target gene were obtained from GeneProf website. Gene ontology analysis was performed on the Ensemble website. **Result.** Compared with the progressing cells, the refractory cells obtained more than two times DEGs at 6 day post-transduction, in particular, down-regulated DEGs related to cell cycle, cell adhesion and development were over 4 times of that in progressing cells. The expression of the DEGs which could only be detected in refractory cells at 6 day were traced back to day 3, and we found the expression of the up-regulated DEGs at 3 day were higher in refractory cells, whereas the expression of the down-regulated DEGs at 3 day were lower in refractory cells. The analysis of histone modification in genome-wide and in DEGs showed that during the reprogramming process the increase of bivalent sites in genome were mainly attributed to gaining of H3K27me3 and losing of H3K4me3. Different regulation of H3K27me3 in DEGs was the key to regulate the expression differently in progressing and refractory cells. The expression of chromatin modifiers in the two cell populations were checked and found to be differential regulated at different time point during
reprogramming process. **Conclusion.** Genes related to immune response, cell adhesion, DNA replication and cell cycle in the refractory cells responded to the induction factor earlier than in the progressing cells, which led to excessive conversion rate. We supposed that after initiation phase cellular reprogramming required to undergo a rate-limited step guided by different regulation of chromatin modification.
Different Regulation of Chromatin Modification Guide

Cellular Reprogramming into A Rate-limited Step after Initiation Phase

Rong Hu¹,², Xianhua Dai¹*, Zhiming Dai¹, Qian Xiang¹

¹School of Information Science and Technology, Sun Yat-sen University, Guangzhou, China
²Department of Applied Mathematics, Guangdong University of Finance, Guangzhou, China

*Corresponding author:

Xianhua Dai¹
Email address: issdxh@mail.sysu.edu.cn

Rong Hu¹,²
Email address: sandarahu@163.com

Abstract

Background. In the early and late stages of cell reprogramming to induced pluripotent stem cells (iPSCs) ectopic expression of Oct4, Sox2, Klf4 and Myc (OSKM) aroused two peaks of transcriptional and epigenetic change respectively. However, it was relatively quiet in the intermediate stage. In this paper our aim is to gain insight into the molecular events that occur after the initiation phase of pluripotency induction.

Methods. GSE42379 containing 28 gene expression profiles and GSE42477 containing 10 genome binding/occupancy profiles of mouse embryonic fibroblasts (MEF) were downloaded from GEO. These datasets included untreated MEFs, OSKM-induced MEFs progressing and refractory to reprogram at 3, 6, 9, 12 day post-transduction, and iPS cell lines. Differentially expressed genes (DEGs) were identified between different cell lines. The Chip-seq peaks and putative target gene were obtained from GeneProf website. Gene ontology analysis was performed on the Ensemble website.

Result. Compared with the progressing cells, the refractory cells obtained more than two times DEGs at 6 day post-transduction, in particular, down-regulated DEGs related to cell cycle, cell adhesion and development were over 4 times of that in progressing cells. The expression of the DEGs which could only be detected in refractory cells at 6 day were traced back to day 3, and we found the expression of the up-regulated DEGs at 3 day were higher in refractory cells, whereas the expression of the down-regulated DEGs at 3 day were lower in refractory cells. The analysis of histone modification in genome-wide and in DEGs showed that during the reprogramming process the increase of bivalent sites in genome were mainly attributed to gaining of H3K27me3 and losing of H3K4me3. Different regulation of H3K27me3 in DEGs was the key to regulate the expression differently in progressing and refractory cells. The expression of chromatin modifiers in the two cell populations...
were checked and found to be differential regulated at different time point during reprogramming process.

**Conclusion.** Genes related to immune response, cell adhesion, DNA replication and cell cycle in the refractory cells responded to the induction factor earlier than in the progressing cells, which led to excessive conversion rate. We supposed that after initiation phase cellular reprogramming required to undergo a rate-limited step guided by different regulation of chromatin modification.

**Introduction**

Induced pluripotent stem cells (iPSCs) had been generated from various somatic cell types by ectopic expression of transcription factors such as Oct4 (also known Pou5f1), Sox2, Klf4, and Myc (OSKM) (Takahashi, et al., 2006, 2007). These findings raised the possibility of creating patient-specific stem cells for regenerative medicine (Nishikawa, et al., 2008). However, cellular reprogramming was a complicated and inefficient process, only 0.0001% to 29% of somatic cells could be reprogrammed to iPSCs (Takahashi, et al., 2006, 2007, Yu, et al., 2007, Huangfu, et al., 2008). Better understanding of molecular mechanisms underlying the reprogramming would help to define optimal strategies for reprogramming and answer fundamental questions regarding how cellular identity was maintained and transformed.

To characterize the molecular mechanisms of reprogramming, many groups focused on transcriptional and epigenetic changes in different cell populations at different time points after factor induction (Mikkelsen, et al., 2008, Li, et al., 2010, Samavarchi-Tehrani, et al., 2010, Golipour, et al., 2012, Polo, et al., 2012, Hansson, et al., 2012, Zhang, et al., 2012, Sancho-Martinez, et al., 2013). These studies revealed three phases of reprogramming: early (or initiation) phase, intermediate phase and late (maturation and stabilization) phase. The initiation phase was marked by a mesenchymal-to-epithelial transition (MET) (Li, et al., 2010, Mah, et al., 2011) and was described to be ‘stochastic’; The late maturation and stabilization phases had been studied at single cell level and was reported to be ‘deterministic’ or more ‘hierarchical’ (Buganim, et al., 2012).

OSKM transgenes was required at the early and intermediate stage of reprogramming process, but should be removed for the transition from the maturation to the stabilization phase. Genome-wide analysis of the intermediate cell populations revealed two distinct waves of major gene activity: the first wave occurred in 0-3 day post-transduction, and the second wave towards the end of the process started after 9 day. These suggested that reprogramming was a multi-step process followed by a series of molecular events. The over expression of transgenes at late stage is harmful to cellular reprogramming.

The intermediate phase was considered to be a bottleneck before converting to stable iPSCs. It was probably a rate-limiting step, but what defined it and how it worked was still not clear (Buganim, et al., 2013). In this study we compared the gene expression of MEFs, intermediate cells progressing and refractory to reprogramming at 3, 6, 9, 12 day post-transduction and iPSCs, to find the difference of expression regulation in refractory and progressing cells. We then checked the chromatin
modifications change in genome and in the differentially expressed genes (DEGs) at intermediate stage of reprogramming. Expression change of some chromatin modifiers in different stage of reprogramming had been also checked. The results of analysis showed the conversion rate of progressing cells could be maintained in a gradual growth, while the refractory cells obtained more bivalent sites and the expressions of DEGs changed too much at intermediate stage which was likely to result in the failure of arousing the second wave of transcription and epigenetic alteration at the late stage of reprogramming process and finally caused the cells’ being trapped in the partially reprogrammed state. The observation of expression pattern of chromatin modifiers showed that the different regulation of chromatin modifiers at distinct stage might be the key to guide the cells into different way for reprogramming by triggering the correct time series of molecular events.

Methods

Data acquisition

The expression and histone modification data of single cell from progressing cells and refractory cells at different time points were obtained from the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) through accession numbers GSE42379 and GSE42477.

Gene expression analysis and unsupervised clustering

The differentially expressed genes (DEGs) were identified for different comparisons by a t-test with P-value <0.05 and fold change (FC) >2. Statistical associations including mean calculations, standard deviation calculations, Pearson correlations, t-tests were calculated using Matlab (The Mathworks, Inc.).

Transcription factor activity analysis

We identified three ChIP-Chip data sets from the published literature for Oct4, Sox2, Klf4, c-Myc, and Nanog (Kim, et al., 2008, Marson, et al., 2008, Sridharan, et al., 2009). Of our differentially expressed mRNA targets, we identified those that were bound by each of these factors in at least two published experiments.

ChIP-seq analysis

ChIP-seq coverage was detected by centering a -10kb and +10kb window on the transcriptional start site for each gene. The Chip-seq peaks and putative target gene were obtained from GeneProf website. (https://www.geneprof.org/GeneProf/browse_exp.jsp).

Gene ontology

Gene ontology analysis was performed on the Ensemble website. (http://asia.ensembl.org/biomart/martview/4ffc325a8de7f3fa7a6e94835f2b237d).

Results

Expression analysis demonstrated that the refractory cells were less rate-limited
To analyze the key factors to block cells reprogramming to pluripotent state, we profiled genome-wide gene expression of refractory cells and the progressing cells at the different time points during the MEF reprogramming.

Firstly, the differentially expressed genes (DEGs) of iPSC and MEF were selected by mavolcanoplot function in Matlab. The results of setting the fold change (FC) > 2, p-value < 0.05 showed that there were 3349 up-regulated genes and 3150 down-regulated genes between iPSCs and MEFs, which confirmed that the gene expression profiles of these two cell populations were very different (Figure 1).

Then we compared the difference between the expression of the intermediate cell populations and iPSCs respectively at different time points during the reprogramming. Scatter plot of the gene expression in iPSC and in progressing cells (also called SSEA1+ cells) and refractory cells (also called Thy+ cells) were shown in Figure 2. We observed that at 3 day post-transduction the differences between SSEA1+ cells and iPSCs, and differences between Thy+ cells and iPSCs were both greatly reduced, but at 6 day the expression profile of SSEA1+ cells was similar to that at 3 day, however it seemed to be back to MEF state in Thy+ cells. Since then, the great difference of gene expression of Thy+ cells and iPSCs had been kept to the end of the process, while the difference between SSEA1+ cells and iPSCs was slowly shrinking.

To make clear what hindered the reprogramming of Thy+ cells at this stage, we compared the expression profile of SSEA1+ cells and Thy+ cells in the reprogramming process to find out the differentially expressed genes between the two cell populations at various time periods. The result showed that there was a peak of DEGs at 3 day both in SSEA1+ and Thy+ cells, but Thy+ cells failed to arouse the second wave which occurred at about 12 day. We noticed that a sub-peak of DEGs could be detected in Thy+ cells at day 6, whereas the SSEA1+ cells were quieter in this stage (Figure 3). We supposed that these over changed genes might be the key to the failure of Thy+ cells to reprogram. By comparing the DEGs (FC>2, t test p < 0.05) of these two cell populations at 3 day and 6 day post-transduction, 3342 and 2120 DEGs were obtained at 3 day, and 373 and 807 DEGs were obtained at 6 day, respectively (data was shown in Table 1). Noticed that the SSEA1+ cells got much more DEGs than Thy+ cells at 3 day, however, Thy+ cells got much more DEGs at 6 day, in particular, down-regulated DEGs related to cell cycle, cell adhesion and development were over 4 times of that in progressing cells.

We checked the common and individual DEGs at 6 day of SSEA1+ and Thy+ cells, and found that the most binding sites of the common DEGs were the target site of Sox2 in ESCs, while the individual DEGs of Thy+ cells were mostly the target site of Oct4 in ESCs. This indicated that the regulation of Oct4 maybe play more important role in determining the direction of reprogramming at initiation phase, and Sox2 was the key regulation factor of the late stage.
Table 1. The Count of DEGs in Three Cell Populations

<table>
<thead>
<tr>
<th>Contrastive group</th>
<th>Cell Population</th>
<th>The number of DEGs</th>
<th>The number of up-regulated DEGs</th>
<th>The number of down-regulated DEGs</th>
</tr>
</thead>
<tbody>
<tr>
<td>D3 vs D0</td>
<td>Ssea1+</td>
<td>3342</td>
<td>1528</td>
<td>1814</td>
</tr>
<tr>
<td></td>
<td>Thy+</td>
<td>2120</td>
<td>1028</td>
<td>1092</td>
</tr>
<tr>
<td>D6 vs D3</td>
<td>Ssea1+</td>
<td>373</td>
<td>279</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Thy+</td>
<td>801</td>
<td>403</td>
<td>398</td>
</tr>
<tr>
<td>D9 vs D6</td>
<td>Ssea1+</td>
<td>284</td>
<td>87</td>
<td>197</td>
</tr>
<tr>
<td></td>
<td>Thy+</td>
<td>496</td>
<td>205</td>
<td>291</td>
</tr>
<tr>
<td>D12 vs D9</td>
<td>Ssea1+</td>
<td>271</td>
<td>194</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>Thy+</td>
<td>333</td>
<td>204</td>
<td>129</td>
</tr>
<tr>
<td>iPS vs D12</td>
<td>Ssea1+</td>
<td>2490</td>
<td>1876</td>
<td>1614</td>
</tr>
</tbody>
</table>

The expression profile of the Thy+ DEGs at MEF, 3 day and 6 day post-transduction was partitioned into four categories by distinct regulation pattern. Some genes were gradually up-regulated (cluster 1, Figure 4). Genes in this cluster were related to extracellular space/matrix, plasma membrane, and immune response processes (e.g., Igfbp4, Col11a1, Igdec4, Rab3d, Fcgrt, Colec12). Genes in the second category were down-regulated during 0 day to 3 day, and up-regulated in the next three days (cluster 2). Most of genes in this category were involved in cell adhesion and cell-cell contacts (e.g., Cdh11, Itga9, Fgf9, Pkd1). The third category was transiently up-regulated at the former stage and down-regulated later (cluster 3). Genes in this category included DNA replication, cell division processes and DNA binding genes (e.g., Rpa2, Cdc45l, Prim2, Ccnb1, Lig1, Aspm, Uhrf1, Hmga1). The fourth category (cluster 4) contained the gradually down-regulated genes associated with cell cycle (e.g., Ccnb2, Prc1, Spag5, Chek1).

To further detect the expression change of the DEGS in Thy+ cell, the expressions of the DEGS in four categories were traced back to 3 day. We observed that the expression of the up-regulated DEGs at 3 day were higher in refractory cells, whereas the expression of the down-regulated DEGs at 3 day were lower in refractory cells. These observations revealed that the refractory cells responded the transduction signals earlier than the progressing cells. This reflected the refractory cells’ incapability of controlling the conversion rate in the intermediate phase (Figure 5). Therefore, we supposed that after initiation phase cellular reprogramming required to undergo a rate-limited step.

The close cooperation of active and repressive histone modification marks limit the conversion rate of progressing cells

Many studies showed that cell reprogramming process was accompanied by global epigenetic remodeling (Maherali, et al., 2007, Wernig, et al., 2007, Barrero, et al., 2010). To further dissect the differential transcription changes in progressing and refractory cells, we should also check the epigenetic modification of these two cell populations. However, we did not find the data of histone modification of refractory cells at different time points of reprogramming obtained at single cell level. Therefore,
we only compared the epigenetic modification of DEGs to speculate the correct path of epigenetic regulation.

Firstly, the active and repressive marks of histone modification (H3K4me3 and H3K27me3) were analyzed in the whole genome. We compared the intensity of H3K27me3 and H3K4me3 in MEFs and intermediate cells at 3, 9, 12 day post-transduction with iPSCs. The results were shown in figure 6. Significant changes of H3K4me3 and H3K27me3 at 0-3 day and 9-12 day could be found, whereas little change occurred at other periods, which was consist with the time of two wave of changes in gene expression. In addition, we also found the slight increase of H3K27me3 and significant decrease of H3K4me3 during the reprogramming process.

Statistical data showed majority of H3K4me2 target loci and minority of H3K27me3 (~81% vs. ~30%) in genome-wide genes of MEF, and in these loci ~27% were bivalent. After three days’ induction the proportion of genes which enriched for H3K4me2 and H3K27me3 became to ~82% and ~41%, the bivalent loci went to ~38% in progressing cells. This revealed that at the initiation phase the increasing of bivalent loci mainly became as of obtaining of H3K27me3.

Generally, the dynamic change of histone modification in the reprogramming process was characterized as the increase of repressive mark and decrease of active mark in genome wide (Table 2). Then, we checked the histone modifications of DEGs at 3 day post-transduction. In SSEA1+ DEGs H3K27me3 increased relatively ~25%, H3K4me3 decreased relatively ~26%, while in DEGs of Thy+ cells the proportion went to ~16% for H3K27me3 and ~56% for H3K4me3. However, we noticed that the proportion of genes gaining H3K4me3 mark was 3.9% higher than proportion of genes losing H3K4me3 mark in SSEA1+ DEGs, whereas in Thy+ DEGs we got the opposite result that the proportion of genes gaining H3K4me3 mark was 1.9% lower than proportion of genes losing H3K4me3 mark. In addition, the bivalent loci increased 18% in SSEA1+ DEGs and 13% in Thy+ DEGs (Table 2) during the first three days of reprogramming. It was supposed that the rate of gaining H3K4me3 was confined in progressing cells in order to control the raising of bivalent loci.

We classified the SSEA1+ DEGs into four clusters according to the same classification method with the Thy+ DEGs, and then checked the dynamic change of H3K4me3/H3K27me3 of genes in each cluster. The results showed that in both DEGs, the change of H3K4me3 was similar, but there was a significant difference between the change of H3K27me3 in SSEA1+ DEGs and Thy+ DEGs.

H3K27me3 of Thy+ DEGs in cluster 1 (gradually up-regulated) and cluster 2 (down-regulated in 0-3 day, up-regulated in 3-6 day) increased more than that of SSEA1+ DEGs, so genes in cluster 1 and cluster 2 were inhibited more and got lower expression abundance in SSEA1+ cells. That was to say, genes in cluster 1 were less up-regulated at 3 day, meanwhile, genes in cluster 2 were down-regulated more at 3 day in progressing cells. Change of H3K27me3 of genes in cluster 3 (up-regulated in 0-3 day, down-regulated in 3-6 day) could be characterized as the decrease of H3K27me3, and it decreased more in Thy+ DEGs. So these genes were inhibited less and up-regulated higher in progressing cells at 3 day. In genes of cluster 4 (gradually...
down-regulated), the difference of the change of H3K27me3 between SSEA1+ DEGs and Thy+ DEGs was not significant, therefore, expression of these genes was similar in progressing cells and the refractory cells at 3 day (Figure 7a-7d).

Table 2. Dynamic Change of H3K27me3 and H3K4me3 in DEGs

<table>
<thead>
<tr>
<th>Gene category</th>
<th>Proportion of genes with altered H3K27me3 (%)</th>
<th>Proportion of genes with altered H3K4me3 (%)</th>
<th>Proportion of increased bivalent sites (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSEA1+ DEGs</td>
<td>46.32/26.0 increase/gaining 21.2/8.7 decrease/losing</td>
<td>32.9/7.4 increase/gaining 59.7/3.5 decrease/losing</td>
<td>17.7</td>
</tr>
<tr>
<td>Thy+ DEGs</td>
<td>32.2/21.1 increase/gaining 16.2/7.9 decrease/losing</td>
<td>17.9/0.49 increase/gaining 74/2.46 decrease/losing</td>
<td>13.3</td>
</tr>
</tbody>
</table>

‘increase’ indicated that the intensity of H3K27me3/H3K4me3 increase. ‘gaining’ means the intensity of H3K27me3/H3K4me3 increase from zero. ‘decrease’ indicates that the intensity of H3K27me3/H3K4me3 decrease. ‘losing’ means the intensity of H3K27me3/H3K4me3 decrease to zero. Thus, we supposed that the progressing cells limited the cell type conversion rate during the reprogramming process by reshaping the chromatin modification appropriately.

**Different stage-specific regulation of chromatin modifiers guide the direction of reprogramming**

Since the dynamics of histone modification were indicators of the fate of reprogramming cells, it was reasonable to inspect the change of chromatin modifiers during the reprogramming. In this paper several histone modifiers related to H3K4me2 and H3K27me3 were chosen to check how the histone modifiers were targeted to genes with an altered expression that was crucial to the conversion process.

It was reported that the WD repeat protein 5 (WDR5) which was a core member of the mammalian Trithorax complex, could interacts with OCT4 on pluripotency gene promoters, and this maintained global and localized H3K4me3 distribution (Ang, et al., 2011). The H3K27me3 demethylase enzyme UTX physically interacted with OSK to remove the repressive mark H3K27me3 from early pluripotency genes (Mansour, et al., 2012). BMI1, RING1, EZH2, EED and SUZ12 were involved in maintaining the transcriptional repressive state of genes (Onder, et al., 2012). BRG1 (also known as SMARCA4) and BAF155 (also known as SMARCC1), two components of the BAF chromatin-remodelling complex, enhanced reprogramming by establishing a euchromatic chromatin state and enhancing binding of reprogramming factors to key reprogramming gene promoters (Singhal, et al., 2010). Over-expression of BRG1 and BAF155 induced OSKM-mediated demethylation of pluripotency genes such as Oct4, Nanog and Rex1 (also known as Zfp42) and enhances conversion to iPSCs.

We found that the direction of expression change of UTX and BMI1 in Thy+ cell was opposite to that in SSEA1+ cell in 0-3 day (Table 3). This was consistent with the
result of H3K27me3 increasing more but decreasing less in SSEA1+ cells. It indicated that Thy+ cells received a weaker inhibitory effect than SSEA1+ cells in this period. So it was likely to up-regulate the unrelated lineage genes simultaneously in Thy+ cells in initiation phase of reprogramming. We extended the expression profile of these chromatin modifiers to the 6 day and find that most of the chromatin modifiers transiently up-regulated at early stage and down-regulated at late stage in the Thy+ cells, especially WDR5 and UTX. Down-regulation of WDR5 and UTX indicated that the activation effect of epigenetic modification decreased. That meant the pluripotent genes were probably not activated in Thy+ cell during 3-6 day.

Table 3. Expression alteration of chromatin modifiers during 0-6 day

<table>
<thead>
<tr>
<th>Gene</th>
<th>MEF</th>
<th>Thy+ cells D3</th>
<th>SSEA1+ cells D3</th>
<th>Thy+ cells D6</th>
<th>SSEA1+ cells D6</th>
<th>Thy+ cells D0-3</th>
<th>Thy+ cells D3-6</th>
<th>SSEA1+ cells D0-3</th>
<th>SSEA1+ cells D3-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utx</td>
<td>8.04</td>
<td>8.12</td>
<td>7.97</td>
<td>7.81</td>
<td>8.13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Bmi1</td>
<td>9.65</td>
<td>9.60</td>
<td>9.84</td>
<td>9.25</td>
<td>9.81</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Ring1</td>
<td>7.45</td>
<td>7.61</td>
<td>7.54</td>
<td>8.02</td>
<td>7.62</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ezh2</td>
<td>8.63</td>
<td>9.66</td>
<td>10.21</td>
<td>7.98</td>
<td>10.06</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Eed</td>
<td>9.45</td>
<td>9.64</td>
<td>9.92</td>
<td>9.17</td>
<td>9.85</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Suzl2</td>
<td>8.36</td>
<td>8.90</td>
<td>9.28</td>
<td>8.29</td>
<td>9.39</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Parp1</td>
<td>7.42</td>
<td>8.81</td>
<td>9.31</td>
<td>8.21</td>
<td>9.48</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Parp2</td>
<td>9.31</td>
<td>10.59</td>
<td>11.13</td>
<td>10.24</td>
<td>11.27</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wdr5</td>
<td>8.04</td>
<td>8.68</td>
<td>9.27</td>
<td>7.55</td>
<td>9.41</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Smarca4</td>
<td>9.89</td>
<td>10.11</td>
<td>10.22</td>
<td>9.10</td>
<td>10.12</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Smarcc1</td>
<td>8.81</td>
<td>8.99</td>
<td>9.15</td>
<td>8.79</td>
<td>9.07</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

We supposed the order of molecular events was different in Thy+ cell and SSEA1+ cell which was caused by different stage-specific regulation of chromatin modifiers. Thy+ cell tended to activate silent genes firstly, while SSEA1+ cell tended to inhibit high expressed genes at first step. In Thy+ cell a new balance between the starting and unrelated lineages might be established by up-regulating the unrelated lineage genes which blocked further reprogramming and thus cells became ‘trapped’ in the partially reprogrammed state. The observation showed that the different regulation of chromatin modifiers at distinct stage might be the key to guide the cells into different way for reprogramming by triggering the correct time series of molecular events.

Discussions

The three phase model of reprogramming was accepted by many studies. It was documented that the early phase was stochastic phase, while the late phase was a ‘deterministic’ or more ‘hierarchical’ phase. The intermediate phase was a bottleneck before transiting to stable iPSCs. It was probably a rate-limiting step, but what defines it and how it works was still not clear.

To dissect the transcript and epigenetic change between the progressing cells and

*1 represents upregulation, 0 represents downregulation.
the refractory cells, it was useful to compare the expression and chromatin modification data of the two cell populations. Aberrant increasing of the differentially expressed genes at 6 day and incapability of triggering the second wave of expression change in Thy+ cells, made us suppose that it might be the unlimited conversion rate that result in the inability of refractory cells to reprogram. Expression analysis allowed us to define four categories of differentially expressed genes according to the expression pattern in distinct stage of reprogramming. Our observations illustrated that the aberrant activation of 6 day might be caused by the change of epigenetic modification much earlier.

The analysis of active and repressive mark change of epigenetic state assured that the conversion rate could be limited by regulating the epigenetic environment. Many studies reported that the cells tended to become more bivalent during the reprogramming process. It was not opposite to our result. In our observation the bivalent site in progressing cells also increased generally, but the genes that exclusively differentially expressed in the Thy+ cells were controlled in gaining bivalent mark in progressing cells. So the expressions of these genes were easier to get into over changed in Thy+ cells.

Conclusions

Integrative analysis of gene expression and chromatin modification of the progressing cells and the refractory cells made us conclude that after initiation phase cellular reprogramming required to undergo a rate-limited step, but the conversion rate of the cells refractory to reprogramming were less limited than the cells progressing to reprogramming in the initiation phase. The close cooperation of active and repressive histone modification marks was the key to limit the conversion rate of progressing cells. In addition, the specific regulations of chromatin modification at different stages guided the cells into different way for reprogramming by triggering the correct time series of molecular events.

Abbreviations

OSKM: Oct4, Sox2, KIf4 and c-Myc;
ESCs: Embryonic stem cells;
iPSCs: Induced pluripotent stem cells;
MEFs: Mouse embryonic fibroblasts;
MET: Mesenchymal-to-epithelial transition;
BMP: Bone morphogenic protein;
DEGs: Differentially expressed genes;

Acknowledgments

We would like to thank the support of the National Natural Science Foundation of China (61174163), Philosophy and Social Science Planning Project of Guangdong Province, China (GD15YYJ06) and The Annual Project of Guangdong University of Finance (15XJ02-12). We also would like to thank Yanning Cai, Qianhua Zhang and
Jiayou Liu for insightful discussions and Jun Zheng for valuable comments on the manuscript.

References


reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell; 1: 55-70


Figure 1. The volcano map of DEGs between iPSC and MEF. Red dashed line indicates the location of FC=2 and p-value=0.05. Therefore, genes falling in the upper left of the plot represent the down-regulated differentially expressed genes, genes falling in the lower right of the plot represent the up-regulated differentially expressed genes.
Figure 2. Comparison of gene expression between iPSCs and cells isolated at 0, 3, 6, 9, 12 day during the reprogramming in Thy+ cells and SSEA1+ cells. The red line in the middle indicates the gene expression in the specific cells is equal to that in iPSCs. The other two red lines indicate the gene expression in the specific cells is 1/2 or two times as much as in iPSCs, respectively.
Figure 3. Number of DEGs during reprogramming process. Positive vertical axis indicates number of up-regulated genes, negative vertical axis indicates number of down-regulated genes. d0 vs d3 indicates the comparison between intermediate cells at day 0 and day 3 after induction, and so on.

Figure 4. Expression patterns of four clusters of Thy+ DEGs. Genes in cluster 1 are gradually up-regulated. Genes in cluster 2 are down-regulated during day 0-3, and up-regulated in the next three days. Genes in the cluster 3 are transiently up-regulated.
at the former three days and down-regulated later. Genes in cluster 4 are gradually
down-regulated. D0 indicates MEF, D3 and D6 indicate the intermediate cells at day 3
and day 6 after induction.

Figure 5. Expressions of four clusters Thy+ DEGs in Refractory Cells and
Progressing Cells during Day 0-6.
Figure 6. Comparison of H3K27me3 and H3K4me3 between iPSCs and cells isolated at 0, 3, 9, 12 days during reprogramming. The red line in the middle indicates the H3K27me3/H3K4me3 in the specific cells is equal to that in iPSCs. The
other two red lines indicate the H3K27me3/H3K4me3 in the specific cells is 1/2 or two times as much as in iPSCs, respectively.

Figure 7a. Change of H3K27me3 and H3K4me3 of genes in cluster 1 of Thy+ DEGs and SSEA1+ DEGs in progressing cells. The red lines on the bar indicates the mean of the increase or decrease of H3K27me3 and H3K4me3 of DEGs.

Figure 7b. Change of H3K27me3 and H3K4me3 of genes in cluster 2 of Thy+ DEGs and SSEA1+ DEGs.
Figure 7c. Change of H3K27me3 and H3K4me3 of genes in cluster 3 of Thy+ DEGs and SSEA1+ DEGs.

Figure 7d. Change of H3K27me3 and H3K4me3 of genes in cluster 4 of Thy+ DEGs and SSEA1+ DEGs.