

A Rate-limited Step Guided by Different Regulation of Chromatin Modifiers Should be Undergone after Initiation Phase in Cellular Reprogramming

Rong HU^{1, 2}, Xian-Hua DAI^{1, a, *}, Zhi-Ming 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

^aemail: issdxh@mail.sysu.edu.cn

*Corresponding author

Keywords: IPS, Reprogramming, Gene Expression, Histone Modification.

Abstract. At early stage and late stage of cellular reprogramming the transgenes arouse two waves of transcription and epigenetic abruptly, while it is relatively quiet in intermediate stage. In this paper we analyzed the gene expression and chromatin modification data at different time point of mouse embryonic fibroblasts (MEFs) reprogramming at the single-cell resolution to assay how the cells progressing to reprogramming limit the transition rate after initiation phase. By comparing the gene expression profiling of refractory cells and progressing cells at day 3 and day 6 after factor induction, we observe over four times more differential expressed genes in refractory cells at this stage. These differentially expressed genes obtain more H3k27me3 to become bivalent, which resulting in aberrated downregulation of genes related to cell cycle, cell adhesion and development. Chromatin modifiers are found to be different regulated in these two cell populations. These observations reveal that the refractory cells responded the transcription signals earlier than the progressing cells and then fail to limit the transition rate in the intermediate phase. Differential stage-specific regulation of chromatin modifiers guide the cells into different way for reprogramming by controlling the raising of bivalent sites.

Introduction

Induced pluripotent stem cells (iPSCs) have been generated from somatic cell types by enforced expression of transcription factors such as Oct. 4, Sox2, Klf4, and c-Myc (OSKM) [1]. This finding raises the possibility of creating patient-specific stem cells for regenerative medicine [2]. However, reprogramming is a complicated and inefficient process, only 0.0001% to 29% of somatic cells can be reprogramming to iPSCs [1, 3]. To characterize the molecular mechanisms of reprogramming, many groups focus on transcriptional and epigenetic changes in cell populations at different time points after factor induction [4-11]. These studies revealed three phases of reprogramming: early phase, intermediate phase and late phase. The initiation phase is marked by a mesenchymal-to-epithelial transition (MET) [5] and is described to be 'stochastic'; The late maturation and stabilization phases have been studied at single cell level and is reported to be 'deterministic' or more 'hierarchical'[12]. OSKM transgenes is 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 reveal two distinct waves

of major gene activity: the first wave occurred between days 0 and 3, and the second wave started after day 9, which is towards the end of the process (day 12). It suggests that reprogramming is a multi-step process that follows a series of molecular events. The overexpression of transgenes at late stage is harmful to cells' reprogramming.

The intermediate phase is a bottleneck before transiting to stable iPSCs. It is probably a rate-limiting step, but what defines it and how it works is still not clear [13]. Here we compare gene expression of cells progressing to reprogramming and refractory to reprogramming at first 6 days after OSKM factors induction, to find the subset of over-changed genes individually in refractory cells. We also check the chromatin modifications data of the over-changed genes and some chromatin modifiers' change in different stage of reprogramming. The analysis results show the progressing cells' transition rate can be kept in gradually increasing, while the refractory cells obtained more bivalent sites and over-rate changed at day 6 resulting from the different regulation of chromatin modifiers at distinct stage.

Results

Expression Analysis Demonstrates the Refractory Cells are Less Rate-limited

We profile Genome-wide gene expression of refractory cells and the progressing cells at the different time points during the MEF reprogramming. The result shows that there is a peak of different expressed genes at day 3 in both of the progressing cells (SSEA1+) and the refractory cells (Thy+), but Thy+ cells fail to arouse the second peak which occur at about day 12. We notice that the Thy+ cells received a sub-peak at day 6 after the first 3 days' transcription change peak, whereas in the SSEA1+ cells this stage is more quiet (Fig. 1). It is reasonable to consider it is likely these genes' over-changing in early stage that contributed to the failure of Thy1+ cells to reprogram. We compare the differentially expressed genes (DEGs, FC>2, t test $p < 0.05$) of these two cell populations at day 3 and day 6. As a result, 3342 and 2120 DEGs are obtained at day 3 respectively. The DEGs of SSEA1+ cells are much more than that of Thy+ cells at day 3; however, at day 6 Thy+ cells activate much more genes, especially downregulated genes which were over four times of the SSEA1+ cells.

We check the common and individual DEGs at day 6 of SSEA1+, Thy+ and Thy- cells. We find that the most binding sites of the common DEGs are targeted by Sox2 in ESCs, while the individual DEGs of Thy+ cells are mostly targeted by Oct4 in ESCs. This indicates Oct4 plays more important role in determining the direction of reprogramming at initiation phase.

The temporal profiles of the Thy+ individual DEGs quantified at day 0, day 3 and day 6 is partitioned into four categories with distinct expression patterns. Some genes are gradually upregulated (cluster 1, Fig. 2). Genes in this cluster are related to extracellular space/matrix, plasma membrane, and immune response processes (e.g., Igfbp4, Col11a1, Igdcc4, Rab3d, Fcgrt, and Colec12). Genes in the second category are downregulate during day 0 to day 3, while are upregulate in the next 3 days (cluster 2). Most of genes in this category are involved in cell adhesion and cell-cell contacts (e.g., Cdh11, Itga9, Fgf9, and Pkd1). The third category is transiently up-regulated at the former stage and down-regulated later (cluster 3). Genes in this category include DNA replication, cell division processes and DNA binding (e.g., Rpa2, Cdc45l, Prim2, Ccnb1, Lig1, Aspm, Uhrf1, and Hmga1). The fourth category (cluster 4) includes gradually downregulated genes relating to cell cycle (e.g., Ccna2, Prc1, Spag5, and Chek1).

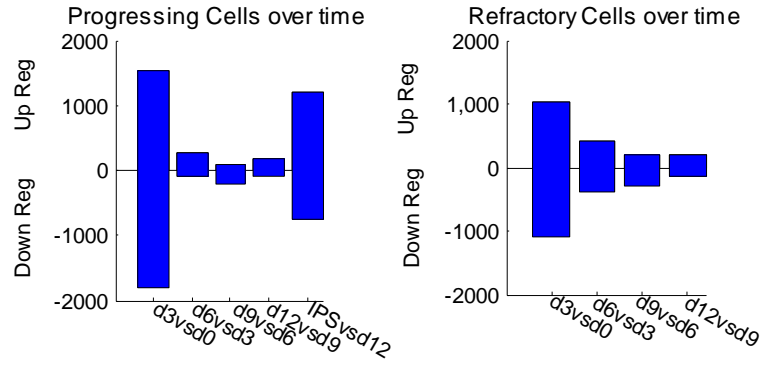


Figure 1. Number of DEGs during reprogramming process

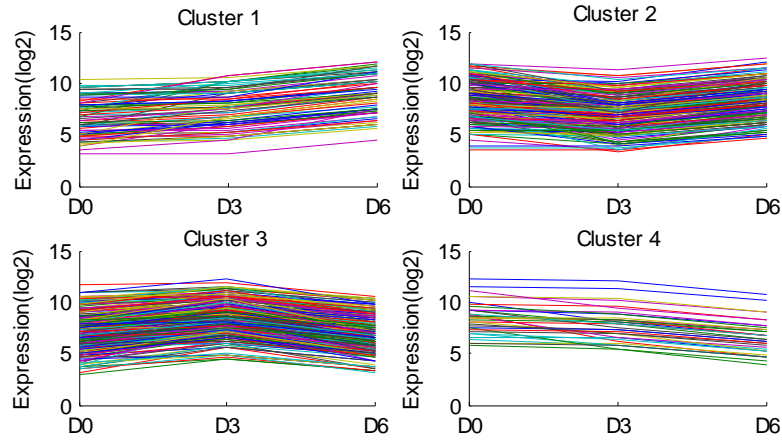


Figure 2. Expression Patterns of 4 Clusters of individual Thy+ DEGs

To further understand the molecular reasons for the inability of Thy+ limiting the transition rate, we trace expression of the four categories genes back to day 3, and find that the genes that are upregulated at day 6 are more activated in Thy+ cells than in SSEA1+ cells at day 3, while the downregulated genes are more repressed in Thy+ cells at day3. These observations reveal that Thy+ cells responded the transcription signals earlier than the SSEA1 cells and then fail to limit the transition rate in the intermediate phase (Fig. 3).

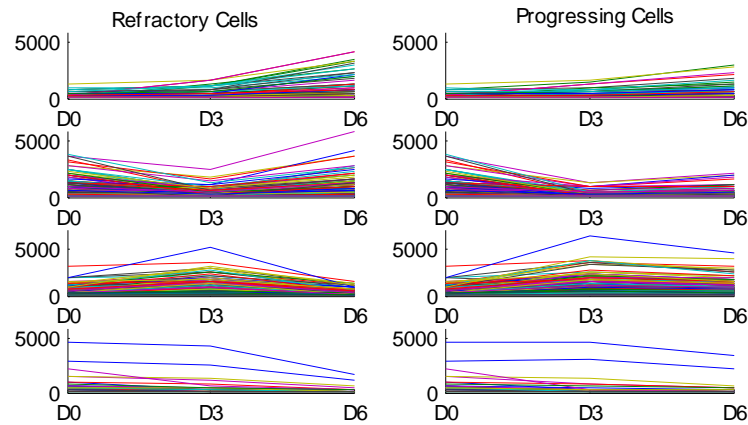


Figure 3. Expressions of Refractory Cells and Progressing Cells from Day 0 to Day 6

Progressing Cells Limited Transition Rate by Confining the Gaining of H3K4me2 to Control the Raising of Bivalent Loci

To dissect the differential transcription changes in progressing and refractory cells, we also check the epigenetic modification of these two cell populations at day 3. Active and repressive histone methylation marks (histone H3 lysine 4 and lysine 27 trimethylation H3K4me3/H3K27me3, abbreviated as K4/K27) are analyzed in genome-wide and in the DEGs. Statistical data shows absolute majority of K4 target loci and minority of K27 (81% vs. 30%) in genome-wide genes of MEF and in these loci 27% are bivalent. After 3 days' induction the proportion of genes enriched for K4 and K27 become to 82% and 41%, the bivalent loci goes to 38% in SSEA1+ cells. This reveals that at the initiation phase the increasing of bivalent loci mainly because of obtaining of K27.

Generally, the dynamics of K4/K27 show increase of repressive mark and decrease of active mark (Table 1). In SSEA1+ DEGs K27 increase relatively 25%, K4 decrease relatively 26%, while in DEGs of Thy+ cells the proportion goes to 16% for K27 and 56% for K4. However, we notice that the proportion of gaining K4 mark is higher than losing K4 mark in SSEA1+ DEGs, while in Thy+ DEGs we can get contrary result. The bivalent loci increase 18% in SSEA1+ DEGs and 13% in Thy+ DEGs (Table 2).

Thus, We can supposed that at the early stage of reprogramming the epigenetic change can be described as increase of repressive mark and decrease of active mark, but the rate of gaining H3K4 is confined in progressing cells in order to control the raising of bivalent loci.

Table 1. Proportion of K4 and K27 Changed in the Different Cell Populations

Cell Population	Increased K27	Decreased K27	Increased K4	Decreased K4
SSEA1+	0.4632	0.212	0.329	0.597
Thy+	0.322	0.162	0.179	0.740

Table 2. Proportion of Increased bivalent site in Different Cell Populations

Cell Population	Gaining K27	Losing K27	Gaining K4	Losing K4	Increased bivalent sites
SSEA1+	0.260	0.087	0.074	0.035	0.177
Thy+	0.211	0.079	0.0049	0.0246	0.133

Differential Stage-specific Regulation of Chromatin Modifiers Guide the Direction of Reprogramming

Since the state of histone modification are indicators of the fate of reprogramming cells, it is reasonable to inspect the change of chromatin modifiers during the reprogramming. In this paper several histone modifiers which is related to H3K4me3 and H3K27me3 are chosen to check how the histone modifiers are targeted to genes with an altered expression that is crucial to the conversion process.

It is reported that the WD repeat protein 5 (WDR5) can interacts with OCT4 on pluripotency gene promoters, and this maintains global and localized H3K4me3 distribution [14]. The H3K27me3 demethylase enzyme UTX physically interacts with OSK to remove H3K27me3 from early pluripotency genes [15]. BMI1, RING1, EZH2, EED and SUZ12 are involved in maintaining the transcriptional repressive state of genes [16]. SMARCA4 and SMARCC1 enhance reprogramming by establishing a euchromatic chromatin state and enhancing binding of reprogramming factors to key reprogramming gene promoters [17].

We find that the expression of BMI1, RING1, EZH2, EED and SUZ12 is up-regulated higher in SSEA1+ cells than in Thy+ cells at day 3, while UTX is up-regulated higher in Thy+ cells. This is consistent with the result of H3K27me3 increasing more but decreasing less in SSEA1+ cells. We also find that the expression of WDR5, SMARCA4 and SMARCC1 are all higher in SSEA1+ cells than in Thy+ cells. It is confirmed that H3K4me3 modification is more active in SSEA1+ cells.

Discussion

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

To dissect the transcript and epigenetic change between the progressing cells and the refractory cells, it is useful to compare the expression and chromatin modification data of the two cell populations. Abrupt raising of the different expressed genes at day 6 and can not be triggered again in Thy+ cells, which make us supposed that it is the unlimited transition rate result in the inability of refractory cells to reprogramming. Expression analysis allows us to define four categories of dynamically expressed genes according to the distinct stage of reprogramming. Our observation illustrated that the aberrated activation of day 6 may be caused by the change of epigenetic modification much earlier.

The analysis of active and repressive mark change of epigenetic state assure that the transition rate can be limited by regulating the epigenetic environment. Almost studies report that the cells tend to become more bivalent during the reprogramming process. It is not opposite to our result. In our observation the bivalent site in SSEA1+ cells also increased generally, but the genes that exclusively different expressed in the Thy+ cells are controlled in gaining bivalent mark. So these genes were easier to get into overexpressed.

Conclusions

Integrative analysis of gene expression and chromatin modification of the progressing cells and the refractory cells make us conclude that the transition rate of the cells refractory to reprogramming are less limited than the cells progressing to reprogramming in the initiation phase. The differentially expressed genes gained more H3k27me3 to become bivalent, which resulting in the aberrated downregulation of genes related to cell cycle, cell adhesion and development. Our studies reveal that the difference regulation of chromatin modifiers between the refractory cells and the progressing cells at the early stage would guide the cells into different way.

Acknowledgements

This research was supported by National Nature Science Fund of China, Grant: 61174163; Philosophy and Social Science Planning Project of Guangdong Province, China(GD15YYJ06); Mathematic Modeling Teaching Team Building of Innovation and Strong School Project of Guangdong University of Finance, 2014; The annual project of Guangdong University of Finance(15XJ02-12).

References

- [1] Takahashi, K., and Yamanaka, S.. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676. (2006)
- [2] Nishikawa S, Goldstein RA, Nierras CR. The promise of human induced pluripotent stem cells for research and therapy. *Nat Rev Mol Cell Biol* 9:725–729. (2008)
- [3] Yu JY, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318: 1917–1920. (2007)
- [4] Mikkelsen, T. S. et al. Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454: 49–55 (2008).
- [5] Li, R. et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 7, 51–63 (2010).
- [6] Samavarchi-Tehrani, P. et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 7, 64–77 (2010).
- [7] Golipour, A. et al. A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell* 11, 769–782 (2012).
- [8] Polo, J. M. et al. A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* 151: 1617–1632 (2012).
- [9] Hansson, J. et al. Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep.*2: 1579–1592 (2012).
- [10] Zhang, J., Nuebel, E., Daley, G. Q., Koehler, C. M. & Teitell, M. A. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. *Cell Stem Cell* 11, 589–595 (2012).
- [11] Sancho-Martinez, I. & Izpisua Belmonte, J. C. Stem cells: surf the waves of reprogramming. *Nature* 493: 310–311 (2013).
- [12] Buganim, Y. et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 150: 1209–1222 (2012).
- [13] Buganim, Y., Faddah, DA., Jaenisch, R. Mechanisms and models of somatic cell reprogramming. *Nature Review Genetics* 14: 427–439 (2013).
- [14] Ang, Y. S. et al. Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* 145: 183–197 (2011).
- [15] Mansour, A. A. et al. The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* 488: 409–413 (2012).
- [16] Onder, T. T. et al. Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 483: 598–602 (2012).
- [17] Singhal, N. et al. Chromatin-remodeling components of the BAF complex facilitate reprogramming. *Cell* 141: 943–955 (2010).

[18] Mah N, Wang Y, Liao MC, Prigione A, Jozefczuk J, Lichtner B, et al. Molecular insights into reprogramming-initiation events mediated by the OSKM gene regulatory network. PLoS One 6:e24351 (2011).

[19] Barrero MJ, Boue S, Izpisua Belmonte JC. Epigenetic mechanisms that regulate cell identity. Cell Stem Cell 7: 565–570(2010)