**Abstract**

**Introduction**

It has been more than 6 years since the four Yamanaka transcription factors—Oct4, Klf4, Sox2 and c-Myc—boosted research on reprogramming somatic cells into embryonic-like cells. However, methods for increasing the reprogramming efficiency and identifying the reprogramming-resistant cells remain unknown. To summarize the updated knowledge and the progress of the molecular mechanism of reprogramming, this review has introduced and evaluated the work done by Polo et al., who established a method to define the intermediate state of pre-reprogramming and the two waves in which reprogramming is triggered. This article provides a comprehensive and critical overview of the study conducted by Polo et al. by discussing the first wave triggered by c-Myc/Klf4 and the second one regulated by Oct4/Sox2/Klf4. First, we focused on the methods used for the definition of the two waves followed by a discussion of how to distinguish the reprogrammable and reprogramming-resistant cells from mouse embryonic fibroblasts with flow cytometry. It also placed significant emphasis on the necessity of the Yamanaka factors in the observed waves. Finally, the article asked whether or not the reprogramming model used by Polo et al. could be adapted to human cell reprogramming. Topics ranging from how to comprehensively understand the pre- and post-reprogramming processes to future complementary work in the field are also discussed in this article.

**Conclusion**

This review highly evaluates the work of Polo et al., which provided evidence of two waves during the post-reprogramming process. Moreover, we raise some critical questions which would be more helpful to understand the molecular mechanism of reprogramming.

**Introduction**

The four famous Yamanaka transcription factors Oct4, Klf4, Sox2 and c-Myc (OKSM) are critical to the formation of induced pluripotent stem cells (iPSCs)\(^1\),\(^2\), but the mechanism of reprogramming and the intermediate state during the reprogramming remain unclear. To uncover the black box of reprogramming, intriguingly, Polo et al. found two-wave process during iPS cell formation\(^3\). In this study, by gene-targeting doxycycline-inducible reprogramming system\(^4\), they generated iPSCs from mouse embryonic fibroblast cells (MEFs) with OKSM overexpression. The results revealed two waves from pre-reprogrammed fibroblast cells to post-reprogrammed iPSCs. The first wave was triggered by c-Myc/Klf4 and the second was driven by Oct4/Sox2/Klf4. Somatic cells that failed to be reprogrammed could not start the second wave, even though they could activate the first wave. In addition, the researchers found that after the first wave, some bivalent genes gradually developed, but DNA methylation happens only after the second wave, during which a stable pluripotent state is acquired. The conclusions from this study provide us with detailed insights into the characteristics and molecular mechanisms of inherent molecular events in reprogramming. Since Yamanaka generated the iPSCs, we have entered a post-reprogramming era. However, the precise regulation of reprogramming factors and why the OKSMs can reprogram MEFs have remained unknown. Polo et al. have first defined the intermediate state of reprogramming into separate time points. By starting with defining the reprogramming process by FACSorted Thy1-negative, SSEA1-positive and Oct4-GFP-positive cells, they demonstrated the different cell-surface markers in the bulk cells. Combined with this observation, the gene-expression assay also showed the reprogramming and embryonic stem cell (ESC)-related gene level in different points. Based on these observations, they found two waves before fully being reprogrammed: days 0–3 is pre-reprogramming wave, day 9 is the post-reprogramming wave and days 3–9 is the intermediate state. This report is critical for improving the efficiency of reprogramming and its future therapeutic applications, but the results are preliminary and some important limitations must be addressed. The aim of this review was to discuss cell reprogramming in the post-Yamanaka era.

**Discussion**

**Question about the definition of the time point of two waves**

In this paper, by using a large amount of data, these authors tried to define the two waves by days 0–3 and days
3–9. However, their effort remains unpromising. For example, Li and Smavarchi-Tehrani reported that mesenchymal epithelial transition (MET) took place in the reprogramming state5,6. This paper defined days 0–3 as the first wave. The authors tried to confirm that they had found the intermediate state in which MET was triggered, but MET was committed from days 0 to 3 and the reverse epithelial mesenchymal transition (EMT) was found from days 3 to 6. This is inconsistent with Li and Smavarchi-Tehrani’s reports in which METs were detected from days 5 to 8 and even longer; which is beyond the first wave.

FACS data showed that fully reprogrammed cells are Thy1 negative and SSEA1 positive. They also demonstrate that some SSEA1-positive cells (13.2%) are not iPSCs, which are Oct4-GFP negative. According to this paper, day 3 is a critical breakpoint for understanding reprogramming. However, data did not show the FACS pattern for Thy1 and SSEA1 double-positive cells on day 3. This raised a question of why the Thy1 was decreasing and SSEA1 was increasing from days 0 to 3 and 6.

**Question about the definition of reprogrammable and reprogramming refractory cells by Thy1 as cell-surface marker**

Since Thy1 is a lineage-associated gene in both mouse and human EFCs7,8, it is unclear whether Thy1-negative cells, which are 12.1% mixed with Thy1-positive cells on day 0, might be the source of contamination cells for Thy1-negative cells on days 3–12. These authors must consider whether the following intermediate reprogrammed cells are derived from Thy1-negative cells. As a result, whether the Thy1-negative cells on days 3–12 are derived from Thy1-expressing cells on day 0 is still unknown. To resolve the doubt, a better experiment strategy would be to use Thy1-positive cells, excluding Thy1-negative cells as starting cells (control). Moreover, according to the data of Polo et al., overexpression of OSKM will lead Thy1-positive cells to form iPSCs; thus, Thy1 is not so critical in distinguishing reprogrammable cells. In other words, repeated transduction could increase reprogramming efficiency, regardless of Thy1 cells being positive.

Thy1 gene-expression changes in mouse fibroblast reprogramming have been discussed previously9. However, this paper did not evaluate that study. In the older study, c-Myc, independent of OKS, is thoroughly discussed, too. Moreover, Thy1 repression was also originally described. To an extent, there are similar points with this paper. The older paper should be taken as a reference, and its results should be compared with those found in this report.

**Question about the four factors used in the study: Are they necessary in the two waves?**

In c-Myc targets3, which genes are c-Myc, Oct4/Sox2 and Klf4 targets? According to Vierbuchen and Wernig’s10 report, there is a pioneering transcription factor first binding to the chromatin, MyoD, in the cardiomyocyte lineage reprogramming. We note that ‘the MyoD enhancer also contains an Oct4-binding site’10. We hypothesize that increasing iPSC-forming efficiency may require more Oct4 copies, while c-Myc is not as important as Oct4.

Yamanaka reported that the transcription factor c-Myc is not necessary in both mouse and human reprogramming, but c-Myc helped IPS form faster11. In Yamanaka’s report, which compared three OKS factors, more Nanog-negative colonies were found in c-Myc/IPS, which is consistent with the conclusion in this paper: Some SSEA1-positive cells are Oct4-GFP negative, indicating they are IPS background cells. Moreover, according to this report, c-Myc accounted for cell activation, and Thy1-negative cells, which are more easily transduced by c-Myc, are activated and proliferated during reprogramming. In conclusion, c-Myc is an important but unnecessary transcription factor.

In this experiment, these authors analysed relative gene-expression change in both Thy1-positive cells and programmed cells, but they did not list the endogenous and exogenous OKSM gene expression, respectively, although they claimed ‘endogenous Oct4 and Sox2 transcripts were detectable’. Because doxycycline was administered along with the cell culture from days 0 to 12, no evidence compared the differences between endogenous and exogenous Oct4 and Sox2. In this paper, Polo et al. put Oct4 and Sox2 together for targeting gene analyses, but Oct4 and Sox2 played different roles during the reprogramming12,13, separately. We do not know the single gene targets. Consequently, if the author can present the independent factor analysis, it will be informative and useful for understanding the mechanism of reprogramming. It would be more convincing to clearly identify the intermediate state when OKSM gene expression changes, if the endogenous gene expression changes in the Thy1 and iPSC.

**Can the conclusion in this paper be applied to human iPSCs?**

Evidence indicates that Thy1 is highly expressed in undifferentiated human ESCs14–16, while in this paper, Thy1 is a critical marker to distinguish reprogramming refractory cells and fully reprogrammed iPSCs, which are more similar to human ESCs. Other reports showed that two factors, Oct4 and Klf412 or Oct4 and Sox217, can reprogram fibroblast cells, raising a question about the necessity of c-Myc. This experiment was based on mouse iPSCs, but in human somatic cell reprogramming, Yu et al. used Oct4, Sox2, Nanog and Lin2818. Moreover, in this report, data
showed that mmet-let-7i was down-regulated, while its target Lin28 started increasing as late as day 9, in which the second wave started. In Yu’s report, instead of c-Myc and KIf4, Lin28/Nanog plus Oct4/Sox2 succeeded in reprogramming iPSCs. If the conclusion drawn by Polo can be addressed with human iPSCs, it raises a question about Yu’s model: Besides Oct4/Sox2 and Lin28, which factor triggered the first wave? As a result, this result cannot be applied to human iPSCs.

There are some unclear descriptions in the figure interpretation. For example, in the text, the authors mentioned that 90–95% Thy1+/Thy1– cells failed while 5–10% cells experienced reprogramming, but no data supported this result. Moreover, in the relative gene-expression figures, the Y-axis lost the scale of relative gene expression, so whether the gene expressions changed insignificantly is unknown. As a result, it is necessary to confirm the absolute value of OKSM expression in MEFCs. Perhaps the MEFCs have experienced c-Myc and KIf4 expression previously, so they are amenable to KM over-expression, while OS are quite low in the starting cells.

In this paper, by way of ChIP assay and network component analysis, these authors demonstrated c-Myc, Oct4/Sox2 and KIf4 targeted gene expressions. However, we did not see the targeted gene lists in the text. Whether stemness genes such as Nanog, Oct4 and CXCR4 are highly expressed or whether they are c-Myc targeted remained unknown. Further work should be done with transduction of four factors with different orders to answer these questions.

Polo et al. opened an era for understanding reprogramming thoroughly and the reason for low reprogramming efficiency. This study contributed highly to the reprogramming research because it is the first attempt to separate the pre-reprogramming state into several steps and provided us a guide to analyse the molecular mechanism of reprogramming. Even though this paper reported a novel idea about the intermediate state of reprogramming, the matter still needs to be more thoroughly discussed. According to the questions identified above, the definition of the time point in which the so-called waves were triggered merits further discussion. For future studies, more references should be considered, cells under more precise time point should be undertaken and the endogenous gene must be analysed carefully. But their findings that OKSM are indispensable in the two waves should not be interpreted as indicating in other reprogramming experiments which did not use the four factors, e.g., reprogramming iPSCs without c-Myc or with Lin28 and Nanog instead of Oct4/Sox2 in human reprogramming, or reprogramming somatic cell with only one factor. Despite some limitations, this paper is innovative because it explains the molecular mechanisms of mouse somatic cell reprogramming. The authors give us with evidence to distinguish the reprogramming-resistant cells from the pool cells. The paper provides us with a landmark for finding critical pre-reprogramming steps so that it will be easier to generate iPSCs in the near future.

Conclusion
To thoroughly understand the mechanism of iPSC reprogramming, based on the outstanding work done by Polo et al., more questions should be answered.

Abbreviations list
EMT, epithelial mesenchymal transition; ESC, embryonic stem cell; iPSC, induced pluripotent stem cell; MEFC, mouse embryonic fibroblast cell; MET, mesenchymal epithelial transition; OKSM, Oct4, KIf4, Sox2 and c-Myc

Acknowledgements
This work was supported by Grants-in-Aid (18591421, 20591531 and 23591872) for scientific research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan.

References

For citation purposes: Li B, Zheng YW. Uncover the black box of reprogramming in the post-Yamanaka era. OA Biotechnology 2013 Jan 31;2(1):5.
Review