



Published in final edited form as:

*Cell Stem Cell*. 2010 December 3; 7(6): 649–650. doi:10.1016/j.stem.2010.11.018.

## Noncoding RNA landmarks of pluripotency and reprogramming

Zhengqing Ouyang, Grace Zheng, and Howard Y. Chang

Howard Hughes Medical Institute and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305.

### Abstract

Noncoding RNAs have emerged as important determinants of pluripotency and reprogramming. In this issue, Kosik and colleagues (Neveu et al., 2010) now provide a detailed map of microRNA expression patterns to infer the biological states of embryonic and induced pluripotent stem cells.

---

What are the molecular determinants of pluripotency? This question lies at the heart of current debates about the equivalence of embryonic stem cells (ESCs) and induced pluripotent cells (iPSCs) (reviewed by Loh and Lim, 2010). While the developmental potential of mouse iPSCs can be evaluated by their ability to contribute to or generate entire embryos, such experiments are not possible for human iPSCs envisioned for use in regenerative medicine applications. Thus, it is necessary to uncover molecular markers that correspond with the biological properties of hiPSCs in order to gauge their degree of pluripotency. Prior studies have examined genome-wide patterns of chromatin state, mRNA expression, and on a more limited level, microRNA (miRNA) expression (Loh and Lim, 2010). Neveu et al. now provide a new perspective based on detailed analysis of miRNAs in ESCs, iPSCs, and various differentiated and cancer cells (Neveu et al., 2010).

The authors measured the expression of 330 miRNAs using a highly quantitative real time PCR approach in a diverse set of 49 samples. These samples included authentic ESCs, differentiated cells, cancer cells, as well as iPSCs generated with several combinations of reprogramming factors and methods of factor delivery. Unbiased clustering divided the samples into four distinct categories: differentiated cells, cancer cells, and two subsets of pluripotent cells. Furthermore, the use of multiple supervised classification methods generated a so-called ‘miRMap’ which could be used in a predictive fashion with additional data sets and revealed that certain pluripotent cells share an overlapping signature with cancer cells.

In order to establish the miRMap, the authors looked for similarities and differences in the miRNA expression profiles between the individual samples. The main sources of variation in the patterns of miRNA expression can be visualized as two-dimensional maps using principal components analysis (PCA), in which each dimension represents a group of coordinately regulated miRNAs. Principal components are a mathematical abstraction. In brief, given a matrix of numbers (e.g. 330 miRNAs  $\times$  49 samples), the first principal component (PC) is the set of miRNAs that accounts for most of the variation among the samples. The second PC is the next set of miRNAs that account for most of the remaining variation, in a manner that is independent of the first PC. Additional PCs account for further

---

© 2010 Il Press. All rights reserved.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

variation. In this way, PCA can mathematically identify the main sources of variation in a complex matrix. Neveu et al. nicely show that the first three PCs appear to have meaningful biological associations. Using this analysis, pluripotent cells are clearly separated from all lineage-committed cells, based on the pluripotency-related miRNAs that make up the 2<sup>nd</sup> PC. Intriguingly, PC3 divides the human pluripotent cells into two classes. The first class (Class 1) consists of most hESCs and some virus-derived iPSCs, while the second (Class 2) contains episome-, protein-derived iPSCs, some virus-derived iPSCs, and the H9 hESC. Surprisingly, PC3 also separates differentiated cells from cancer cells.

The distinction between the four categories defined by the PCA was optimized using supervised learning methods, which resulted in the miRMap, a 2D classification system (Neveu et al., 2010). The classifier that distinguishes Class 1 and Class 2 pluripotent cells is driven by the expression of a dozen of miRNAs, which are absent in the latter. The authors applied miRMap to several published microarray datasets, and found that individual lines could be prospectively segregated into the two classes of human pluripotent cells. Importantly, the miRMap can also be used to visualize the gene expression trajectory of cells during directed differentiation or reprogramming. Differentiation and reprogramming do not proceed by direct transitions from one pattern, or quadrant of the miRMap to another, but instead proceed via a circuitous route through an intermediate stage that exhibits the Class 2 pattern, which is characteristic of cancer cells and some iPSCs. These findings raise the intriguing possibility that reprogramming cell fate requires the sequential interplay of proliferation and cell fate commitment, which needs further investigation (Singh and Dalton, 2009).

To look more closely at the differences between the two classes of pluripotent cells revealed by miRMap, the authors performed Gene Set Enrichment Analysis on published microarray data from hESC and iPSCs. Not only is the expression of a significant portion of genes in the p53 pathway dysregulated in Class 2 pluripotent cells, the expression of p53 mRNA is also lower, suggesting that p53 is important in controlling states of pluripotency. Indeed, several means of p53 inactivation, including enforced expression of miR-92 and miR-141 that target p53, shifted the miRNA profile of the disrupted cells from Class 1 to Class 2. Despite their capacity to induce this observed shift miR-92 and miR-141 are not prominent classifiers within miRMap. However, p53 bears predicted target sites for other miRNAs, and so it is likely that alternative miRNAs or factors regulate p53 expression in the two subtypes of pluripotent cells. Why persistent differences in p53 pathway activity are observed between the two iPSC classes remains unclear. Nevertheless, these experiments suggest an important link between the status of the p53 network and a given miRNA profile during reprogramming, and raise the question of what players in the network are connected to miRNAs expressed by hESCs and iPSCs.

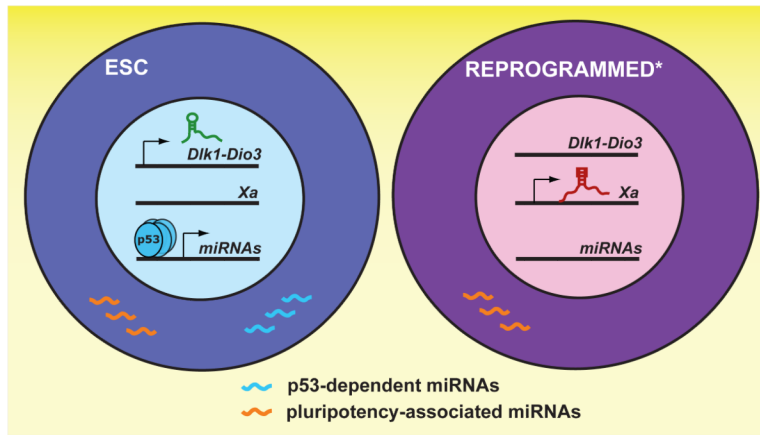
These new findings add to the emerging theme that noncoding RNAs serve as key determinants and barriers of epigenetic reprogramming. Many long noncoding RNAs are involved in the regulation of chromatin states, specifically in marking chromosomal regions in an allele-specific or cell-specific fashion (Lee, 2009). As such, long ncRNAs can be important regulators of imprinting, dosage compensation, and lineage determination—events that need to be properly reset to return to an ESC-like epigenetic state. Indeed, Hochedlinger and colleagues found that aberrant silencing of an imprinted locus, including long ncRNAs and other genes, poses a barrier for high quality iPSC generation, but re-expression of the locus rescues iPSC pluripotency (Stadtfeld et al., 2010). Likewise, reprogramming of somatic nuclei by nuclear transfer into eggs is substantially improved by ensuring the proper regulation of Xist, the long noncoding RNA that inactivates one of two X chromosomes in female cells (Inoue et al., 2010). Now we can add p53 pathway miRNAs as another important source of variation between iPSC and ESCs (Neveu et al., 2010). The

large number of miRNAs that modify p53 function may reflect the degree to which this genome guardian pathway is activated during epigenetic reprogramming—which is known to limit reprogramming efficiency and may be selected against (reviewed in Deng and Xu, 2009). Recent evidence of lineage memory in iPSCs suggests that other ncRNAs involved in lineage commitment or positional identity may also need to be reset to ensure successful reprogramming (Kim et al., 2010)(Polo et al., 2010).

The findings of Neveu et al. also illuminate the ongoing debate regarding the proposed equivalence of ESCs and iPSCs (Loh and Lim, 2010). If noncoding RNAs, and specifically miRNAs, constitute a large part of the functional difference between ESCs and some iPSCs, such differences may be easily missed in studies of chromatin state or mRNA profiling. As deep sequencing technologies continue their rapid ascent towards greater coverage and affordability, the panel biomarkers that define ESC and iPSC states will likely grow. As Neveu et al. demonstrated, such markers may also provide potential inroads into uncovering mechanisms that regulate pluripotency and reprogramming.

## References

- Deng W, Xu Y. *Trends Genet.* 2009; 25:425–427. [PubMed: 19801173]
- Inoue K, Kohda T, Sugimoto M, Sado T, Ogonuki N, Matoba S, Shiura H, Ikeda R, Mochida K, Fujii T, et al. *Science.* 2010; 330:496–499. [PubMed: 20847234]
- Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LI, et al. *Nature.* 2010; 467:285–290. [PubMed: 20644535]
- Lee JT. *Genes Dev.* 2009; 23:1831–1842. [PubMed: 19684108]
- Loh KM, Lim B. *Cell Stem Cell.* 2010; 7:137–139. [PubMed: 20682438]
- Neveu P, Kye MJ, Qi S, Buchholz DE, Clegg DO, Sahin M, Park I-H, Kim K-S, Daley GQ, Kornblum HI, et al. *Cell Stem Cell* this issue. 2010
- Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, Tan KY, Apostolou E, Stadtfeld M, Li Y, Shioda T, et al. *Nat Biotechnol.* 2010; 28:848–855. [PubMed: 20644536]
- Singh AM, Dalton S. *Cell Stem Cell.* 2009; 5:141–149. [PubMed: 19664987]
- Stadtfeld M, Apostolou E, Akutsu H, Fukuda A, Follett P, Natesan S, Kono T, Shioda T, Hochedlinger K. *Nature.* 2010; 465:175–181. [PubMed: 20418860]



**Figure 1. Noncoding RNA landmarks distinguish most ESCs from imperfectly reprogrammed pluripotent cells**

Imperfectly reprogrammed cells (indicated by \*) inappropriately silence the *Dlk1-Dio3* locus, express *Xist* from the active X chromosome (*Xa*), and lose expression of p53-dependent miRNAs despite the expression of other genes associated with pluripotency.