SPECIAL REPORT

Dissecting reprogramming, differentiation and conversion with network biology

GQ Daley^{1,2,3}

Leukemia Supplements (2014) 3, S3–S4; doi:10.1038/leusup.2014.2

Keywords: stem cell reprogramming; CellNet; global gene regulatory networks; Context Likelihood of Relatedness Algorithm; InfoMap Algorithm; GSEA

A major preoccupation of stem cell biology is the conversion of cells from one identity to another for applications in research, drug screening and cell therapy. Numerous labs are attempting to mimic embryonic development by directing the differentiation of embryonic stem (ES) cells/induced pluripotent stem (iPS) cells using morphogens, cytokines and specialized stromal co-culture conditions. Equally vigorous efforts are being undertaken by many labs to reverse the differentiation by reprogramming—either through somatic cell nuclear transfer or, as shown by Takahashi and Yamanaka,¹ simply through the ectopic expression of C-Myc, Sox2, Oct-3/4 and KLF4—transcription factors that normally act in embryonic stem cells and can revert essentially any somatic cell to a pluripotent state. Inspired by Yamanaka's experiments but tracing its intellectual roots to Weintraub's and Lassar's classical experiments using MyoD to convert fibroblasts to muscle,² many other groups have been directly converting fibroblasts to alternative tissue types including neurons, hepatocytes or cardiomyocytes. Presently, however, we lack a quantitative tool to assess the fidelity of those cell fate conversions, and have only empiric trial-and-error strategies for manipulating lineage specifying transcription factors to effect such transitions. We have established CellNet,³ a network biology platform, to address these challenges of cellular engineering. By computational analysis of essentially all publicly available microarray expression data, we have discerned regulatory relationships among all known transcription factors and potential targets, and constructed global gene regulatory networks (GRNs). We based our approach both on direct linear and nonlinear correlations between transcription factors and their targets. The gene regulatory relationships and cell-type context were determined by applying a number of different algorithms (Context Likelihood of Relatedness, InfoMap and GSEA). The putative relationships were validated by comparison to Encode data, which has defined a subset of all transcription factor binding sites throughout the genome and thus can be used to select a threshold for defining the gene regulatory relationships. CellNet accepts as input gene array data for any type of cell fate conversion. The readout is a heat map showing the relatedness of the engineered cell to its target, a quantitative measure of GRN status of the engineered cell relative to its target tissue and a prioritized list of transcription factors that are most dysregulated. The training data for constructing these GRNs for mouse cells are the thousands of gene expression profiles for a set of 20 highly relevant cells and tissues (16 tissues

for human). In our approach, we looked at several dozen published examples of direct conversion (fibroblasts to hepatocytes, cardiomyocytes, neurons and blood), direct differentiation of ES cells along various lineages and reprogramming of somatic cells to pluripotency. A particularly efficient mode of reprogramming is to incorporate the reprogramming genes (Oct-3/4, Sox2, KLF4 and C-Myc) into the genome of fibroblasts under the influence of a doxycycline-inducible promoter. Exposure to doxycycline then converts the fibroblasts over time into iPS cells. With the CellNet algorithm one can define the kinetics of silencing of fibroblast GRNs and activation of GRNs characteristic of ES cells, and thereby shed light on the mechanisms by which cells are reprogrammed. We looked at the reprogramming data (all being early replications of the Yamanaka reprogramming to iPS cells) from at least 10 different laboratories. Nearly every one of the reprogrammed iPS lines showed near-complete GRN identity with ES cells, confirming that the reprogramming was robust. Applying CellNet to assess the fidelity by which ES cells could be differentiated into cardiomyocytes, we found on average about 60% efficacy in establishing the heart GRN, whereas attempts to directly convert fibroblasts to cardiomyocytes performed considerably less well. Similar comparisons were done for hepatocytes, neurons and blood cells, and we found that, in every case, our algorithm confirmed that the quality of the derived cell was better when a pluripotent stem cell rather than a converted fibroblast was the starting source. We also applied our algorithm to another system developed by Graf and his colleagues, in which ectopic expression of C/EBP α in a CD19+ pre-B cell converted these lymphoid cells into macrophages.⁴ Using our computational platform, we tracked the conversion of primary B cells into macrophages and demonstrated remarkable fidelity.⁵ Interestingly, however, an immortalized pre-B cell engineered to express an inducible form of C/EBP α (C10 cells) failed to extinguish the B-cell fate and despite acquiring macrophage features remained a hybrid cell. The obtained list of deregulated transcription factors included mainly early B-cell factors. When these B-cell regulators were silenced by shRNA in a follow-up experiment, we observed a robust upregulation of macrophage-specific markers in C10 cells. Indeed, our collective experience in attempting various cell fate conversions suggests that a major barrier to achieving a new cell fate is the inability to silence the old cell fate. In another system we explored the conversion of fibroblasts to induced hepatocyte-like cells by ectopic expression of the hepatocyte-associated or

¹Division of Hematology/Oncology, Children's Hospital, Harvard Medical School, Boston, MA, USA; ²Dana Farber Cancer Institute, Harvard Medical School, Boston, MA, USA and ³Howard Hughes Medical Institute, Children's Hospital, Harvard Medical School, Boston, MA, USA. Correspondence: Dr GQ Daley, Karp Family Research Building, One Blackfan Circle, Boston, MA 02115, USA.
E-mail: george.daley@childrens.harvard.edu

endoderm-specifying transcription factors Foxa1 and HNF4 in embryonic fibroblasts, as described by Sekiya and Suzuki.⁶ Using their robust method for cell fate conversion, we were able to generate cells that were essentially identical to theirs. These cells expressed albumin and E-cadherin, stored glycogen and secreted albumin and urea. We observed, however, that these induced cells were not like primary hepatocytes. Analysis using our algorithm indicated that even though they had hepatocyte functions, the so-called induced hepatocyte-like cells were still classified as fibroblasts, because their fibroblast GRNs were not silenced. When we looked closely at their transcription factor signatures, we found high expression of genes characteristic of hindgut including prominent expression of CDX2—a master regulator of intestinal fate. This suggested that the putative induced hepatocytes were more like intestinal progenitor cells, or perhaps bi-potential endoderm progenitors, that had not been fully specified to the liver. Indeed, when we grew them under conditions that support formation of intestinal organoids,⁷ our converted fibroblasts made cystic organoid-like structures. We were indeed able to demonstrate their capacity to engraft into the colons of mice suffering from colitis. Thus, these cells were not truly induced hepatocytes, but were ‘semi-colon’ cells that can function like intestinal cells.

Stem cell biology is focusing on various routes for cellular engineering: somatic cell reprogramming, differentiation and conversion. We have now built an algorithm that has the capacity to provide a solid metric for comparing different cell types, and which promises to improve cellular engineering. The examples cited above indicate that the most efficient cell fate conversion is still somatic cell reprogramming, and that pluripotent cells are most faithfully converted to target tissues via *in vitro* differentiation. In contrast, the CellNet algorithm suggests that the major roadblock to successful conversion of one somatic cell to another is the challenge of silencing the initial GRN of the starting cell. With CellNet as a tool, we hope to refine future attempts at cellular

engineering to facilitate applications in research and clinical translation.

CONFLICT OF INTEREST

Dr Daley is a member of the Scientific Advisory Boards of the following companies, and receives consulting fees and/or holds equity in True North Therapeutics, Inc., Solasia, KK, Epizyme, Inc., Verastem, Inc., Advanced Cell Technologies, MPM Capital, LLP and Johnson & Johnson/Janssen R+D Division.

ACKNOWLEDGEMENTS

The symposium and publication of this supplement were sponsored by the Division of Hematology/Oncology at the Warren Alpert Medical School of Brown University and NIH Center of Biomedical Research Excellence (COBRE) for Stem Cells Biology at Rhode Island Hospital. The Daley Laboratory was supported by funds from the National Institutes of Health (P50HG005550 and R24DK092760).

REFERENCES

- 1 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; **126**: 663–676.
- 2 Hollenberg SM, Cheng PF, Weintraub H. Use of a conditional MyoD transcription factor in studies of MyoD trans-activation and muscle determination. *Proc Natl Acad Sci USA* 1993; **90**: 8028–8032.
- 3 Cahan P, Li H, Morris SA, Lummertz da Rocha E, Daley GQ, Collins JJ. CellNet: Network biology applied to stem cell engineering. *Cell* 2014; **158**: 903–915.
- 4 Bussmann LH, Schubert A, Vu Manh TP, De Andres L, Desbordes SC, Parra M *et al.* A robust and highly efficient immune cell reprogramming system. *Cell Stem Cell* 2009; **5**: 554–566.
- 5 Morris SA, Cahan P, Li H, Lu Y-F, Zhao AM, Sahalie J *et al.* Enhancing cellular engineering through network biology. *Cell* 2014; **158**: 889–902.
- 6 Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. *Nature* 2011; **475**: 390–393.
- 7 Sato T, Clevers H. Primary mouse small intestinal epithelial cell cultures. *Methods Mol Biol* 2013; **945**: 319–328.