Transcriptional regulatory networks of single cells during in vitro hepatic differentiation of human pluripotent stem cells

Rathi D. Thiagarajan, Thomas Touboul, Robert Morey
University of California San Diego
Department of Reproductive Medicine
La Jolla, CA 92037

rthiagarajan@ucsd.edu

Louise C. Laurent
University of California San Diego
Department of Reproductive Medicine
La Jolla, CA 92037

llaurent@ucsd.edu

ABSTRACT

Directed differentiation of pluripotent stem cells (PSCs) into specific cell types in vitro begins with a relatively homogenous population of undifferentiated cells. Upon addition of specific growth factors, PSCs are capable of differentiating towards specific cell types of interest. However, in vitro differentiation often produces heterogeneous cell populations and low yields of the desired cell type, despite the fact that all of the cells are ostensibly exposed to the same conditions. By studying the different trajectories of single cells during in vitro differentiation and discovering the transcriptional regulatory networks required to specify and maintain cellular fate, we aim to understand the factors involved in in vivo differentiation and to develop strategies to improve the efficiency of in vitro differentiation.

Single cell RNA sequencing (RNA-seq) provides unprecedented resolution of gene activity among individual cells where previously, biological samples could only be profiled in "bulk". Recently, we demonstrated application of this technology to study rare cells that may give rise to cancer (Ramsköld D. et al. Nature Biotech. 2012). Using an established directed differentiation protocol, we used RNA-seq to obtain single base resolution of the single cell transcriptome during hepatic differentiation of human ESCs. In this project, we randomly collected eight single cells from each of three critical phases of hepatic differentiation: undifferentiated hESCs, definitive endoderm (day 3), and hepatic progenitor cells (day 8). Using XSEDE resources at the San Diego Supercomputing Center, we were able to access tools to analyze these large datasets in order to process raw data, perform QC, and data analysis using tools in the R statistical program, Linux based algorithms. Principal component analysis (PCA) clustering results indicated increasing heterogeneity among the 8 cells from a given stage as differentiation progressed, such that the undifferentiated cells were the most homogeneous and hepatic-like cells were the most heterogeneous. Inspection of transcriptional regulatory networks using the visualization software, Cytoscape, of transcription factors (TFs) identified from an unsupervised clustering analysis of single cells and in-silico TF motif predictions of distal promoters using FIMO from the MEME suite tools suggested combinatorial use of key lineage-specific TFs. These initial results highlight that single gene cell-type markers alone are not sufficient to characterize differentiation, and that detailed molecular characterization can highlight critical differences that can help describe a variety of cellular trajectories from a relatively uniform source population. This study demonstrates the application of next generation sequencing technology at single cell resolution and supercomputing resources at SDSC to understand in vitro differentiation and identify key TF networks involved in cell fate decisions.