Functional Atlas of Mouse Embryonic Stem Cells Pluripotency and Early Differentiation

Huilei Xu[1,2], Ihor Lemischka[2], Avi Ma'ayan[1]

Short Abstract — **Mouse embryonic stem cells (mESCs) are derived from the inner-cell-mass of a blastocyst and can be maintained indefinitely in culture as well as differentiate into all adult cell types. This makes such cells an important model for regenerative medicine applications and a tool to study developmental disorders. While masses of genome-wide data profiling mESCs at different regulatory layers are rapidly accumulating, our understanding of the molecular circuitries governing cell fate of mESCs is still lacking. To this end, we constructed a database called ESCAPE (Embryonic Stem Cell Atlas from Public Evidence) by integrating data from many recent high-content mESC studies. We further developed a Boolean model from a subnetwork made of pluripotency and early differentiation regulators extracted from the database. The model was optimized to fit data from a Fluidigm single cell gene expression experiment and combinatorial RNAi perturbations followed by RT-PCR data. Collectively, the ESCAPE database and the Boolean model are capable of describing in details molecular regulatory mechanisms regulating self-renewal and early differentiation of mESCs.**

I. Background

Recent studies have shown that somatic cells can be "reprogrammed" into induced pluripotent stem (iPS) cells using simple combinations of transcription factors [1-2]. In order to harness the translational biomedical potential of embryonic stem (ES) and iPS cells we need to further characterize the molecular regulatory networks responsible for controlling pluripotency, self-renewal, iPS reprogramming, as well as commitment and differentiation into specific lineages. The regulation of mESCs self-renewal and pluripotency is mediated by a network of transcription factors auto-regulating their own expression as well as each other while suppressing the expression of differentiation genes. To further dissect and characterize such regulatory network, high throughput technologies such as cDNA microarrays, ChIP-seq, proteomics, and RNAi screens provide us with the ability to understand cell regulation at the system level with great molecular detail. However, integrating different datasets from multiple studies for extracting knowledge and form theories about coordinated regulatory mechanisms is still a challenge [3].

II. Results

To aid in addressing this challenge, we constructed an embryonic stem cell specific database called ESCAPE by collecting and integrating various data types including: 98,430 directed protein-DNA binding interactions extracted from ChIP-seq and ChIP-chip experiments; 125,531 indirect gene regulatory interactions from loss/gain-of-function studies followed by genome-wide mRNA expression profiling; 657 indirect protein-protein interactions from nine mass-spectrometry proteomics studies; and a list of 464 potential pluripotency genes from five RNA interference screens. Further, we constructed a Boolean model composed of 30 pluripotency and lineage-specific regulators extracted from the database and fitted the model to data from a Fluidigm single cell gene expression experiment. Finally, we validated the model by comparing in-silico gene knockdowns to combinatorial RNAi perturbations followed by RT-PCR.

III. Conclusion

The ESCAPE database is currently the most extensive resource consisting of multi-layered interactions in mESCs. It can be applied to generate systems-level views of gene regulatory networks and correlate genes across regulatory layers in an unbiased way. In addition, our Boolean model, including 30 interacting regulators, is significantly more comprehensive than previous models that attempted to simulate the dynamics of the gene regulatory network in mESCs. The experimental single cell gene expression data and combinatorial RNAi perturbations followed by RT-PCR data tightly anchors the model to experimental observations. Collectively, the different views of the database and the optimized dynamical model assisted us to elucidate the intricate relationships between an extended set of members of the pluripotency and early differentiation gene regulatory network in mESCs.

Acknowledgements: This work was funded by NIH grants 1P50GM071558-01A27398 and 1RC1GM091176-01.

[1]Department of Pharmacology and System Therapeutics, Systems Biology Center New York (SBCNY), New York, NY, 10029. E-mail: huilei.xu@mssm.edu, avi.maayan@mssm.edu

[2]Department of Gene and Cell Medicine, Mount Sinai School of Medicine, New York, NY, 10029. E-mail: ihor.lemischka@mssm.edu

References

[1] Takahashi, K. and S. Yamanaka, Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors Cell, 2006. 126(4): p. 663-676.

[2] Park, I.-H., et al., Reprogramming of human somatic cells to pluripotency with defined factors. Nature, 2008. 451(7175): p. 141-146.

[3] MacArthur, B.D., A. Ma'ayan, and I.R. Lemischka, Systems biology of stem cell fate and cellular reprogramming. Nat Rev Mol Cell Biol, 2009. 10(10): p. 672-681.