Determining which genes are expressed and repressed in embryonic stem cells and neural progenitor cells

Elise Vu, Ramya Raviram, Naoki Kubo, Bing Ren

Department of Cellular and Molecular Medicine, University of California, San Diego

Introduction

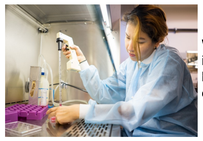
Embryonic stem cells (ESCs) are pluripotent. This essentially means that they have the ability to divide an unlimited amount of times to create more stem cells, or differentiate to produce more specialized cells like neural progenitor cells (NPCs). Unlike stem cells, the neural progenitor cells are only capable of dividing a limited number of times. However, these progenitor cells can differentiate into an array of neuronal as well as glial (surrounds neurons) cell types. In the process of neurogenesis, neurons are created in the brain. This process is extremely important when an embryo is growing. Neurogenesis also continues in some areas of the brain after being born and throughout an individual's life. The fully developed brain contains many specialized sections of function and neurons that have distinct associations as well as structure. For instance, the cerebellum, which controls voluntary movements, contains 101 billion neurons. During the growth of embryos, a variety of neurons located in the brain originate from regulated neurogenesis. Also, their neural stem cells differentiate during the process. Furthermore, in each cell, genes are either expressed (turned on) or repressed (turned off). Gene regulation is the process in which genes are turned on and off. When genes are repressed, they don't give the instructions for producing proteins. During growth, genes are expressed and repressed in different patterns to distinguish the appearance as well as function of brain cells from spinal cord cells.

Methods

In this lab, I used some experimental techniques to conduct my research, including cell culture and RNA-seq. I used cell culture to examine how mouse ESCs differentiate into mouse NPCs under a microscope. I also worked with RNA-seq to determine which genes are expressed and repressed and how this affects their functions in the cells.

Cell Culture

First of all, cell culture is a technique in which cells are grown under controlled circumstances *in vitro*, which means that the cells are living outside their preferred environment. Cells can be preserved under controlled circumstances after they've been removed from living tissue and placed in a culture dish. The conditions in which each cell type need to thrive are different. The conditions cells usually need to survive include a suitable dish with a medium that provides the necessary nutrients, factors that'll allow them to develop, gases, and regulation of the environment.



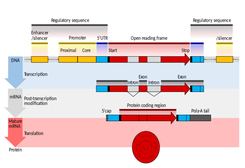
When performing cell culture, it's important to work in the hood, so that you don't contaminate the cells.

The cell culture plates are placed inside an incubator.

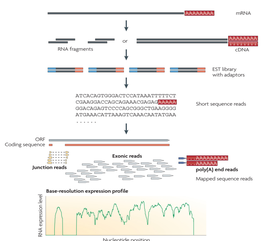
RNA-seq

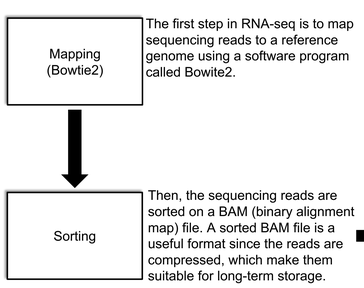
RNA-seq is used to show the presence and quantity of RNA in a biological sample. Its main use is finding gene expression changes between distinct cell populations. However, its original objective was to find which genomic loci are turned on in a cell during a certain timeframe over the whole expression range. During RNA-sequencing, genes are transcribed and spliced to create mRNA (messenger RNA) transcripts. The mRNA is isolated from the cell, broken up, and then copied into ds-cDNA. The ds-cDNA stands for double-stranded DNA, which is synthesized from a single-stranded RNA. After that, the ds-cDNA is sequenced. These sequences have the ability to be mapped back to a reference genome sequence. This determines which genome areas are undergoing the process of transcription. This information can help an individual locate where expressed genes are, their expression levels, and other splice elements.

Gene structure (eukaryotic protein-coding gene)



For the protein coding region (red), regulatory sequence manages the time and location expression occurs at. Promoter and enhancer regions (yellow) control the transcription of the gene into a pre-mRNA. The pre-mRNA is altered to get rid of introns (light grey) and add a 5' cap and poly-A tail (dark grey). The mRNA 5' and 3' untranslated regions (blue) maintain translation into the final protein product.

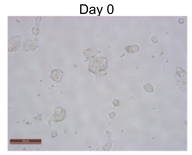
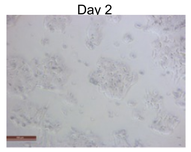
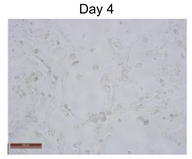
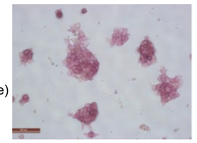
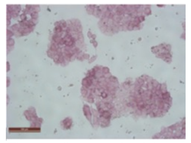
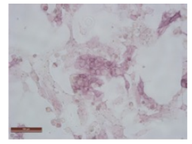




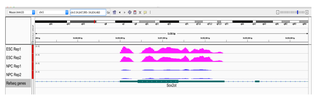
The first step in RNA-seq is to map sequencing reads to a reference genome using a software program called Bowite2.

Results

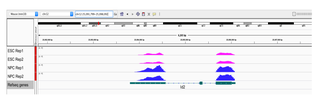
Cell culture data **Mouse Embryonic Stem Cells (mESCs) → Mouse Neural Progenitor Cells (mNPCs)**

	Day 0	Day 2	Day 4
Bright field (no staining)	 These ESCs form round, compact colonies. Lif (leukemia inhibitory factor) is added to the culture medium to maintain pluripotency.	 For the differentiation treatment, Lif is removed from culture medium and retinoic acid is added. Retinoic acid (a morphogen derived from vitamin A) allows the mESCs to transform into mNPCs.	 Unlike ESCs, these NPCs are branch-like and spread out.
AP staining (Alkaline Phosphatase)	 AP staining is used to differentially stain ESCs to distinguish them from NPCs. As you can see above, these cells are darkly stained, which means they're highly pluripotent.	 These cells are stained, but are lighter than the cells in the previous image. This means that the cells are somewhat pluripotent.	 The mouse embryonic stem cells have differentiated into mouse neural progenitor cells. Therefore, the cells are lightly stained and have low pluripotency.

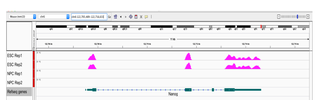
RNA-seq data **I used RNA-seq data from day 0 and day 4 of the culturing of ESCs and NPCs that were previously generated from the Ren lab. Then, I searched up key genes on IGV. After that, I identified which genes were on and off in ESCs and NPCs. I also looked at housekeeping genes.**



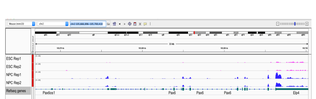
In the house mouse (*Mus musculus*) genome, the gene, Sox2, is highly expressed in embryonic stem cells (pink). However, this gene is lowly expressed in neural progenitor cells (blue). The product of the Sox2 gene is necessary for the regulation of stem cells in the central nervous system, and maintains gene expression in the stomach.



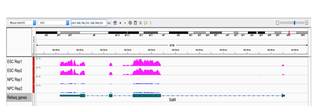
The Id2 gene is lowly expressed in ESCs, but highly expressed in NPCs. The protein encoded by this gene contributes to the negative regulation of cell differentiation.



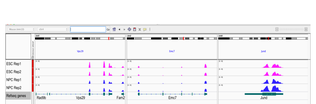
The Nanog gene is expressed in ESCs, but repressed in NPCs. The protein encoded by the Nanog gene is a DNA attaching homeobox transcription factor involved in the rapid reproduction of embryonic stem cells, renewal as well as pluripotency.



The Pax6 gene is lowly expressed in ESCs, but highly expressed in NPCs. This gene encodes a protein that has a homeobox, which functions as a transcription regulator. Not only that, the protein contributes to the development of neural tissues (specifically the eye).

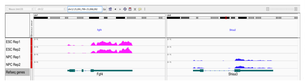


The Sall4 gene is highly expressed in ESCs, but lowly expressed in NPCs. The functions of this gene have been involved in a variety of embryonic development processes such as brain, heart as well as limb development. Not only that, this gene is a key pluripotency factor that's necessary for the regulation of stem cells.



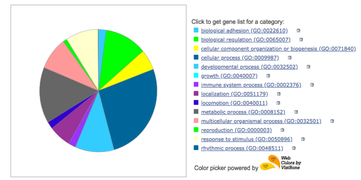
The genes, Vps29, Emc7, and Jund, are equally expressed in ESCs and NPCs. These genes are called housekeeping genes. Since housekeeping genes are involved in the regulation of cellular functions, they're likely to maintain constant expression levels in every cell and condition.

Here, I analyzed my RNA-seq data with DESeq2. This helped me determine which genes were changing the most (differential expression) among ESCs and NPCs.

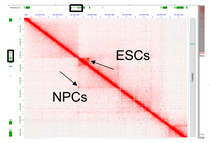


The Fgf4 gene is highly expressed in ESCs, but repressed in NPCs. Its log2FoldChange is -9.885997429. The protein encoded by the Fgf4 gene is part of the fibroblast growth factor family. Fgf4 family members are involved in an array of biological processes like embryonic development, cell growth, morphogenesis as well as tissue repair. On the other hand, the Shisa3 gene is repressed in ESCs, but highly expressed in NPCs. Its log2FoldChange is 7.186998005. The Shisa3 gene encodes a single-transmembrane protein. This protein regulates WNT and FGF signaling by obstructing the maturation as well as transport of their receptors to the surface.

Out of the 1,978 genes, 1,251 of them were up-regulated in NPCs and 727 of these genes were down-regulated in NPCs.



I used data from DESeq2 to formulate this pie chart on the Panther Classification System. Each section of the pie chart displays the biological processes that take place in certain genes. For instance, the Id2 gene performs a nitrogen compound metabolic process, which is represented by the black slice. This process consists of chemical reactions and pathways involving organic or inorganic compounds that contain nitrogen. Not only that, the Sox2 gene performs a single-multicellular organismal process (salmon-colored slice). This process occurs within a single, multicellular organism. Since the Shisa3 gene responds to a stimulus, it's represented by the cream yellow slice.



I used Hi-C data to identify promoter-enhancer interactions among ESCs and NPCs in the Sox2 gene. As you can see in this Hi-C map, enhancers are more active in ESCs than in NPCs. Also, the enhancers in NPCs are farther apart from the Sox2 gene than ESCs. This contributes to the up-regulation and down-regulation of genes in certain cells.

Conclusions

- Cell culture allows researchers to study the process of differentiation in embryonic stem cells to neural progenitor cells by growing these cells outside an organism.
- Also, researchers can use AP staining to determine levels of pluripotency in stem cells.
- RNA-seq has contributed immensely to the field of cellular and molecular medicine. It has changed my view as well as perspective of modern-day genetics. This experimental technique will continue to allow researchers to obtain new knowledge about genes, their expression levels, and the effects these expression levels have on the functions of our cells.
- In this lab, 10 genes (three of which were housekeeping genes) were compared among ESCs and NPCs to locate the differences in expression levels. Also, we were able to determine the functions of these genes and the biological processes that take place.
- Genes that contained negative log2FoldChange values were highly expressed in ESCs. This means that these genes are up-regulated. Meanwhile, the genes in NPCs were down-regulated, which means that the cell is decreasing the amount of a cellular element (for example, RNA or a protein) in response to an external stimulus.
- On the other hand, genes that contained positive log2FoldChange values were highly expressed in NPCs. In this case, genes in ESCs were down-regulated.
- However, some genes that contained log2FoldChange values close to zero were expressed in both types of cells.

Future Research

- In the future, I plan to manipulate these cells by deleting important genes or inserting genes into the DNA. Then, I can observe how this affects the process of differentiation through cell culture.
- Additionally, I'm determined to find out how altered genes (or mutations) contribute to certain hereditary diseases, such as Alzheimer's disease and cystic fibrosis.

References

1. Bartlett, Perry. *What is neurogenesis?*, Queensland Brain Institute, The University of Queensland, Australia, 10 July 2018.
2. Bryant, Steve. *Sox2 SRY (sex determining region Y)-box 2 [Mus musculus (house mouse)]*, National Center for Biotechnology Information, 24 July 2018. (used to find the functions of other genes as well).
3. Dundar, Friederike, et al. *Introduction to differential gene expression analysis using RNA-seq*, Applied Bioinformatics Core, Weill Cornell Medical College, Sept. 2015.
4. Eisenburg, Eli, and Erez Y. Levanon. *Human housekeeping genes, revisited*, Trends in Genetics, Cell Press, Oct. 2013.
5. Wikipedia contributors. *Cell culture*, Wikipedia, The Free Encyclopedia, 20 July 2018, en.wikipedia.org/wiki/Cell_culture.
6. Wikipedia contributors. *Gene structure*, Wikipedia, The Free Encyclopedia, 23 June 2018, en.wikipedia.org/wiki/Gene_structure.
7. Wikipedia contributors. *RNA-Seq*, Wikipedia, The Free Encyclopedia, 20 June 2018, en.wikipedia.org/wiki/RNA-Seq.

Acknowledgments

Sincerest thanks to Ramya Raviram and Naoki Kubo for welcoming and mentoring me in the lab (I felt like a real researcher!). Also, I would like to thank Bing Ren for giving me the opportunity to perform research in his lab. Lastly, I would like to thank Dr. Komives for organizing this enriching, but also fun program. This has been an unforgettable experience.