Bioinformatics and stem cell research- A mini review

Author : Shiv Sanjeevi Sripathi

Categories : Softwares, Tools

Date : December 29, 2018

Stem cells are cells that can be differentiated into other types and are thus pluripotent with the ability to become cells of all lineages. Cells found in the blastocyst of embryos are called as embryonic stem cells or ESCs [1, 2] that are considered “gold” standard of pluripotency [3]. There are also adult stem cells found in several tissues for the purpose of repair such as mesenchymal stem cells that have been differentiated into various other tissues [4]. These pluripotent stem cells hold promise to aid in studying embryo development, differentiation of cells and regenerative medicine that aims at “personalized” medicine [3]. Another class of stem cell known as induced pluripotent stem cells (iPSCs) were created by expressing key transcription factors in adult cells [5]. The field of regenerative medicine has seen plenty of research articles on stem cells. While some stem cell types have been differentiated into specific cell types to cure several diseases such as neurodegenerative disorders, cancer, diabetes, heart disease etc. [2], iPSCs have been differentiated into retinal cells, endothelial cells, and neurons [6].

Where does bioinformatics fit in?

Bioinformatics is a merger of hardware, mathematics, networking, and databases to develop tools that can be used by a person interested in life sciences to process and analyze data [7]. Bioinformatic tools can potentially help in identifying its possible function, for example, KEGG can identify pathways, orthologs, and functions of sequences submitted [8]. The use of bioinformatics in stem cell biology initially revolved around self-renewal dynamics of adult stem cells [9] that later saw the application of molecular biology data along with the use of genome sequencing. With molecular profiling of single cells and systems biology that aid in modeling stem cell patterns, the field of bioinformatics can play a key role in stem cell biology [10].

A few tools:

Let’s discuss a few examples to gain a better understanding. The transcriptome of pluripotent stem cells has been first studied using DNA microarrays with classification algorithms that aid in distinguishing

among differentiated, multipotent and pluripotent stem cells [11]. In the case of larger datasets, the classification of pluripotent stem cells can be facilitated by the use of machine learning. One of such tools is an algorithm *PluriTest* that uses measurements of DNA microarrays to analyze pluripotent cells using bioinformatics models [12]. PluriNetWork can uncover mechanisms and molecules involved in pluripotency of stem cells using a combination of links to literature, gene ontology and automated analysis [2]. Mechanisms in stem cells such as regulation associated at a post-transcriptional level have been studied using next-generation sequencing techniques [3]. For example, the involvement of ZFP217, a zinc finger protein associated with chromatin in the regulation of pluripotency in human embryonic stem cells are shown with a MeRIP-Seq method [13].

Taken together, these and many other genome-wide molecular profiling studies have collectively contributed to our understanding of the multilayered regulation of pluripotency, and have further served as models to understand the regulation of cell type identity for other, less-investigated lineage [10]. A common curated system used a combination of social networking software as well as Wiki to combine research data, key genes and protein circuits to be used with ease and analysis with Cytoscape software [14]. Such a network is a common system composed of literature and details of transcription factors and signals that is tailor-made for a particular requirement [2].

The field of *epigenetics* makes an entry to analyze differences between ESCs and iPSCs as well as to study patterns seen with iPSCs such as their bias towards lineages of a donor [3, 10]. For instance, a study published in 2011 used a support vector machine learning algorithm based on methylation data of ESCs and iPSCs [15] that could identify the ESCs with precision but iPSCs at 61% sensitivity [3]. Regions of differential methylation were analyzed using ‘comprehensive high-throughput arrays for relative methylation’ (CHARM) to uncover promoters of factors for distinct lineages [16, 17].

Another application of bioinformatics in stem cell biology is to assess the differentiation ability of a stem cell using a “*scorecard*” approach. Bock et al, 2011 developed a deviation scorecard with methylation patterns and gene expression of human ESCs as they hypothesized that any deviation here could prevent differentiation to particular lineages. Differences in iPSC lines in comparison to ESCs were tabulated [15]. Several genes were listed as markers of germ layers, that when expressed at early stages indicate the differentiation potential, for example, hypermethylation of GRM (glutamate receptor) in motor neurons [3, 10].

An algorithm *TeratoScore* uses gene expression of teratomas to evaluate the differentiation ability of human pluripotent stem cells as they can differentiate into all three germ layers. The origin of a tumor, either pluripotent or tissue-specific cells can be classified by the tool [18]. Another tool *CellNet* uses gene expression profiles to give a prediction of a specific cell type in the query along with transcription factors [19]. The efficiency of differentiation of pluripotent stem cells can be predicted using a platform called KeyGenes that uses RNA-Seq or microarray data of human fetal tissues [20].

A data repository for stem cells called the *Cellfinder* looks at augmenting human embryonic stem cell registry (hESCreg) into a tool that facilitates the design of projects and analysis of the registry [21]. Additionally, a web-interface called *StemBase* contains SAGE (Serial Analysis of Gene Expression) data of mouse and human stem cells and allows for studying specific genes or markers [22].

Conclusion

This short review has highlighted a few of the tools that find use in stem cell research. The above-mentioned tools show that the field of bioinformatics holds much promise in analyzing stem cells using web interfaces and tools. With further inputs from the various “OMICS” that unravel the roles of molecules in a single cell, the use of bioinformatics can aid in analyzing fates of cells as well as potentially delve deeper into this exciting field of stem cells that are being pitched in as a panacea for several diseases that would help us realize an important goal of stem cell biology: a detailed glimpse into understanding the nuances of the cells vital for development and maintenance of life

References

1. Evans M. Discovering Pluripotency: 30 years of mouse embryonic stem cells. Nat Rev Mol Cell Biol. 2011;12(10):680- 6.
2. Babu PBR and Krishnamoorthy P. Applications of Bioinformatics Tools in Stem Cell Research: An Update. J Pharm Res. 2012;5(9),4863-6.
3. Nestor MW and Noggle SA. Standardization of human stem cell pluripotency using bioinformatics. Stem Cell Res Ther. 2013;4:37.
4. Pacini S. Deterministic and stochastic approaches in the clinical application of mesenchymal stromal cells (MSCs). Front Cell Dev Biol. 2014;12:50.
5. Takahashi K and Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126:663–6.
6. Bilic J and Izpisua Belmonte JC. Concise review: Induced pluripotent stem cells versus embryonic stem cells: close enough or yet too far apart? Stem Cells. 2012; 30:33–41.
7. Orozco A, Morera J, Jiménez S, Boza R. A review of Bioinformatics training applied to research in Molecular Medicine, Agriculture and Biodiversity in Costa Rica and Central America. *Brief Bioinform.* 2013;14(5):661–70.
8. Kyoto Encyclopedia of Genes and *Available on:* https://www.genome.jp/kegg/
9. Till JE, et al. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. Proc Natl Acad Sci U.S.A. 1964;51:29–36.
10. Bian Q and Cahan P. Computational Tools for Stem Cell Biology. Trends Biotechnol. 2016;34(12).
11. Müller FJ, Laurent LC, Kostka D, Ulitsky I, Williams R, Lu C, et al. Regulatory networks define phenotypic classes of human stem cell lines. Nature. 2008;455:401–5.
12. Müller F-J, Schuldt BM, Williams R, Mason D, Altun G, Papapetrou EP, et al. A bioinformatic assay for pluripotency in human cells. Nat Methods. 2011;8:315.
13. Aguilo F, et al. Coordination of m6A mRNA methylation and gene transcription by ZFP217 regulates pluripotency and reprog-ramming. Cell Stem Cell. 2015;17:689–704.
14. Narad P, Upadhyaya K and Som A. Reconstruction, visualization and explorative analysis of human pluripotency network. Network Biology. 2017;7:57-75.
15. Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, et al. Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell. 2011;144:439–52.
16. Kim K, et al. Epigenetic memory in induced pluripotent stem cells. Nature. 2010;467:285–90.
17. Kim K, et al. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. Nat. Biotechnol. 2011;29:1117–9.
18. Avior Y, et al. TeratoScore: assessing the differentiation potential of human pluripotent stem cells

by quantitative expres-sion analysis of teratomas. Stem Cell Rep. 2015;4:967–74.
19. Cahan P, et al. CellNet: network biology applied to stem cell engineering. Cell. 2014;158:903–15.
20. Roost MS, et al. KeyGenes, a tool to probe tissue differentiation using a human fetal transcriptional atlas. Stem Cell Rep. 2015;4:1112–24.
21. Borstlap J, Luong MX, Rooke HM, et al. International stem cell registries. In Vitro Cell Dev Biol – Animal. 2010;46(3-4):242-6.
22. Sandie R, Palidwor GA, Huska MR, et al. Recent developments in Stembase: a tool to study gene expression in human and murine stem cells. BMC Res Notes. 2009;2:39.

Sharing is caring. Spread the love!