

Integrative Bioinformatics Approaches to Analyze Molecular Events in Pluripotency

Priyanka Narad* and Upadhyaya KC

Amity Institute of Biotechnology, Amity University, Noida, Uttar Pradesh, India

*Corresponding author: Priyanka Narad, Amity Institute of Biotechnology, Amity University, Noida, Uttar Pradesh-201301, India, Tel: 09999302904, E-mail: pnarad@amity.edu

Received date: July 21, 2014, Accepted date: August 23, 2014, Published date: August 30, 2014

Copyright: © 2014 Narad P, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Human embryonic stem cells (hESCs) have the capacity to proliferate almost indefinitely. Analysis of gene expression profiles of hESCs could offer an insight into the crucial genes concerned in maintaining pluripotency and the genes that may be involved in cell differentiation. Combining network and high throughput data enables to understand the role of epigenetic mechanisms, signalling pathways and transcription factors responsible in human pluripotency and screening of putative mechanistic relationships, validation of existing knowledge, yielding hypotheses, and suggestions for new experiments. Decoding the hub of factors and their associated interactions is an important beginning to understand the complexities associated with the hESCs and transfer the knowledge for creation of human induced pluripotent stem cells (hiPSCs). This review is aimed at increasing the elemental information of integrative bioinformatics approaches useful for studying self-renewal and differentiation processes using new holistic data on gene expression and epigenetic marks associated with cell pluripotency. Two basic approaches namely the representation as biological networks and semantic web technology have been described for the management of the ever increasing high throughput data. An integrated approach to disentangle the biological intricacies would be highly beneficial in the field of medical and health sciences.

Keywords: hESC; Pluripotency; Integrative Bioinformatics; Cytoscape

Background on Stem Cells

Stem cells are undifferentiated cell mass which have the capability of dividing and differentiating into diverse specialized cell types [1]. Embryonic stem cells (ESCs) are derived from the embryos generated following fertilization. The zygote divides into two and two into four and so on. With additional divisions, a multi-cellular mass of cells called a blastocyst is formed. The blastocyst is a hollow ball of cells with two layers, an outer trophoblast which in due course of time forms the placenta and inner cluster of cells called Inner Cell Mass (ICM). The fertilization can also be accomplished in vitro. It is feasible to pick up these embryonic stem cells with a pipette and shift them onto a petri dish for culturing. Embryonic stem cells have two basic properties: (i) Self renewal i.e. the ability to go through numerous cell divisions and (ii) Potency i.e. the potential of a stem cell to differentiate into various cell types [2,3]. The process of fertilization i.e. fusion of an egg and a sperm leads to a diploid totipotent cell. A totipotent cell is capable of turning into all cell types of an organism including the placental cells. Pluripotent cells can differentiate into all cell types of an organism except placental cells. Multipotent cells are capable of differentiating into a group of related cell types and unipotent cells can produce only one type of cells [3,4].

Over the last decade the pluripotent ESCs have generated a substantial research interest due to their potential therapeutic applications. Landmark advancement was made where it was demonstrated that under certain conditions normal somatic cells could be induced to become stem cells [5]. These specialized cells were termed as induced pluripotent stem cells (iPSCs). These cells meet the

basic characteristic features of the pluripotent cell, but it is still matter of conjecture whether iPSCs and embryonic stem cells are clinically similar. iPSCs were first reported in mouse in 2006 [5], and human iPSCs were initially reported in late 2007 [6]. iPSCs generated from the mouse cells have been shown to possess basic features of pluripotency like the ability to express stem cell markers, differentiating into various tissue types [7]. The same effects could also be accomplished using the human iPSCs and they are capable of generating all three germ layers. Research on stem cells is one of the most important areas of biology, but, as with numerous growing fields of scientific interest, research on stem cells raises scientific questions as fast as it would be generating new inventions. Two fundamental properties of stem cells relating to their long-term self-renewal i.e. why can ESCs divide for a year or more in the laboratory without differentiating, but most non-embryonic stem cells cannot do so under the same set of conditions; and what can be the factors that regulate stem cell proliferation and self-renewal [5]. In this review, we primarily summarize the insights gained by the application of Integrative Bioinformatics approaches like construction of biological networks and semantic web technology. In addition, we discuss the recent research undertaken to uncover determinants of ES cell identity encompassing genome wide transcription factor (TF) activity and chromatin modifications and DNA methylation.

Molecular Mechanism Underlying Pluripotency

Molecular framework of pluripotent state is a consequence of complex interplay of processes. The most significant of these are (i) signaling pathways, (ii) transcriptional regulation, (iii) epigenetic factors and (iv) miRNA, and their collective interaction help the cell to maintain the pluripotent state. A crosstalk between a series of signaling cascades and transcriptional regulation ensure the cell to

maintain pluripotency [8,9]. Human ESC requires TGF β /Actin and FGF signaling to maintain the capacity of self-renewal [10,11]. An important aspect of stem cell research is to gain an insight about the signaling pathways that are highly conserved for maintenance of pluripotency across species [12,13]. Signaling pathways important for self-renewal have been shown to be dependent on transcriptional activation, up and down regulation of many genes. Pluripotency is predominantly maintained by TGF β pathway through SMAD2/3, FGF2 signaling cascade that activates the MAPK, AKT and Wnt pathways acting through the β -catenin. Signaling through these pathways activate and express the three TFs: Oct4, Sox2, Nanog. These TFs in turn activate genes useful for maintenance of pluripotency and repress developmental genes. For example, TGF β directly targets and activates transcription of Nanog in hESC [14,15]. In 2006 Yamanaka and his colleagues [5] demonstrated the induction of a pluripotent state from embryonic or adult fibroblast cells of mouse by the introduction of four factors namely: Oct3/4, Sox2, c-Myc and Klf4 under ES cell culture conditions. They selected 24 candidate genes for factors that induce pluripotency in somatic cells based on their assay conditions. After examining the effect of withdrawal of individual factors from the cocktail of candidate genes, they finally concluded that these four TFs play important roles in generation of iPSCs. The reprogramming process could be identified as a stochastic process under which the epigenetic memory is erased and transcriptional circuitry is altered. Epigenetic mechanisms are fundamental to activate or repress gene during development. Activation marks such as Acetyl-H3-Lys-9, Acetyl H4 are important regulators of activation, whereas repressive marks such as Di-Trimethyl Histone are associated with repressed genes [16-18]. Methylation landscape of human early embryos indicates a selective transient global hypomethylation and hypermethylation of retrotransposable elements [19]. These results are in conformity with the earlier studies indicating that epigenetic memory must be erased from embryonic cells to achieve pluripotency providing possible explanation for global demethylation [20]. A large number of small molecules altering the chromatin structure, including demethylase and histone deacetylase (HDAC) inhibitors are required in the process of reprogramming [21]. These complex processes are known to involve multi-step gene activation and repression. Dynamic electronic representation and integration of data from various resources can help illuminate and analyze these molecular intricacies to an extent.

Complexity of Data and Resources

The advent of the post-genomic era has turned out to be a boon for the researchers all around the globe. The windfall is an augmented quantity of data for the analyses and applications in regenerative medicine. The wave of high-throughput technologies like RNA-seq to Next-gen sequencing in enabling data to be generated at an unprecedented rate in a non-curated state. The stem cell research has also emerged out of its infancy [22,23]. An array of molecular phenomena such as the chromatin modifications, DNA methylation patterns (epigenetics), role of non-coding RNAs are promising as an area of advanced biological research and a huge amount of data are being generated every day. This is in addition to the parallel work of directing the molecular mechanisms encompassing the transcription factors, signaling pathways and expression profiles. As of March, 2014, a text search of "Stem Cell Biology" at GQuery retrieve almost 33123 Pubmed results, 2225 nucleotide sequences, 605 Bioprojects, 10 genome sequencing projects, 212 genes, 2077 GEO profiles and 2434 GEO datasets only at NCBI which is considered as one of the major

bioinformatics resource. The goal has now shifted from an individual result to the capturing of the global picture of entire biological system using the approach of Integrative Bioinformatics [24].

As the stem cell research progresses, a number of relative repositories and research groups are also increasing at a high rate. We list out few of these which can be beneficial as an education resource. MEDLINE plus—Stem Cells/Stem Cell Transplantation is a database resource that contains information on latest news, health resources and clinical trials on stem cells. Stem DB is a comprehensive resource for the information relevant to the stem cell research. It provides the user to enter, update and retrieve information on stem cells. The current work and maintenance are funded by Euro SyStem. The Stem Cell Lineage Database (SCLD) provides user-editable lineage maps illustrating both endogenous development and the directed differentiation of human and mouse embryonic stem cells. These maps contain lineage relationships between individual cell types, gene expression profiles for cell type identification, and information on inductive stimuli that transition cells from one stage to another [25]. The Stem Cell Omics Repository (SCOR) was constructed as a centralized and query-able database for large-scale studies regarding human ES and iPS cells. SCOR is intended to be a resource where researchers can download data sets as well as query results for genes and proteins of interest [26].

Integrative Bioinformatics

Molecular biology is generating huge amounts of knowledge and high quality data after the successful completion of the Human Genome Project. A number of organisms have been sequenced and generated extensive knowledge on metabolic reconstructions and pathways, genome wide association and structural information, consensus patterns of regulatory regions; proteomics, transcriptomics, and metabolomics. Currently, there are numerous database systems and a wide range of tools available via internet which is directed at solving various biological tasks. The goal now is to list the individual parts of a system and assemble them to provide an omics view of the system. Integrative Bioinformatics can be considered as a new area of research using the tools of in silico science and electronic representations applied to life processes. Thus, the aim of integrative bioinformatics include integration of molecular biological, medical and other related data sets relevant for biological systems, design of metabolic, regulatory and expression networks and evaluation of original experimental data with in silico tools (Figure 1). The core approaches include semantic approaches and representation of the complex data in the form of a network [27,28].

Semantic Web Technology

A large number of data warehouses are available on the web-resources. There is a variety of information which can be collected from their diverse data sources. All the information is manually integrated to achieve the pre-defined goals. Sometimes, we even want to access original data and combine it with our own data generated. A solution to this is to design an integrated methodology or technology that can integrate all the data from various web-resources or databases into a well-defined link. A mere integration would not serve our purpose and a more advanced approach would be a standardization of the web of data. For instance, Pauling et al., [29] devised an approach for the reconstruction and comparison of transcriptional regulatory networks in prokaryote using semantic web technologies. Another example that can be mentioned here is that of Shah et al., [30] where

they have created Atlas – a data warehouse for locally storing information of biological sequences, molecular interaction, homology information, functional annotation of genes and biological ontologies. Excellent resource provided by Post et al., [31] has implemented the use of semantic web technology in biological data merged with the genomic data. Wasmuth et al., [32] were successful in providing a combinatory approach of the in vitro and in silico studies of key pathogenesis genes contained within the large surface coat gene super families from a broad array of eukaryotic pathogens.

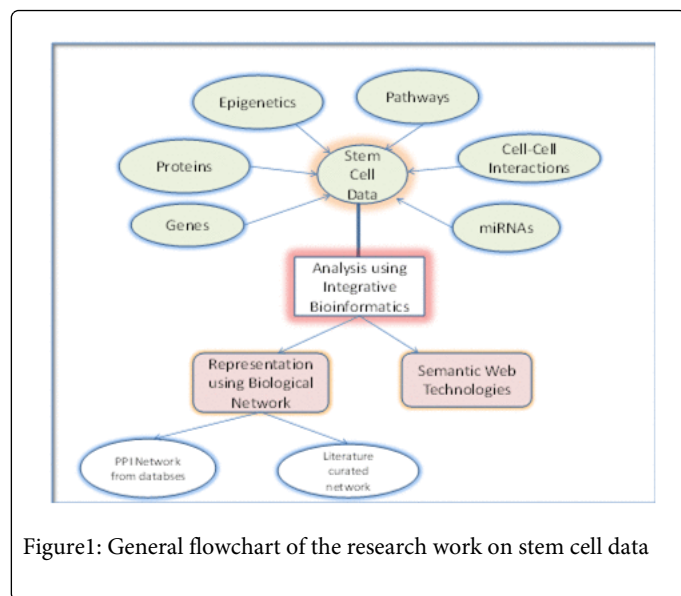


Figure1: General flowchart of the research work on stem cell data

Representation using Networks

The representation of molecular mechanisms underlying complex phenomenon where a large amount of cross-talk between the cellular components exist, has become an enormous task. Network representation aids in uncovering their intricacies and identifying functional modules. Any kind of network representation consists of nodes and edges; where nodes can be representative of molecular entities such as gene, gene products, miRNAs, metabolites and edges represent interactions that can be classified as up regulation, down regulation or neutral. Many aspects of a cellular function or an environment can be understood by the creation of biological representation as models. There are different types of states in hierarchy which can be depicted by network representations. The first type of representation is through metabolic reconstructions. They provide the stoichiometry values and enzyme-substrate reaction concentrations. The methods adopted for such a reconstruction analysis is the flux-balance analysis and ordinary differential equations. The other kind is the Virtual Cell Model. These types include the description of cellular events through the sequence information and linking it to reactions. The method for such type of representations is comparative genomics approach where genome scale integration of the data is done. Another type of representation is through the multi-cellular modeling approach. These include the cellular events such as cellular signaling and cell-to-cell adhesion. The methods adopted are of ODEs (Ordinary Differential Equations) and PDEs (Partially Differential Equations). Another type of a whole-organism representation is through the multi-organism models. These include the influences and dependencies between species such as host and pathogen. The methods adopted for the same are logical modeling

and Boolean networks [33,34]. Based on the fundamental concepts on biological representation, a large amount of work has been conducted to disentangle the pluripotent network. In particular Muller et al., [35] developed Plurinet which was an undirected network based on unsupervised clustering and classification approach. Another approach was proposed by Newman and Cooper [36], called AUTOSOME algorithm to generate a network consisting of 3421 genes. Computational approaches have been elemental to support the research for the determination of ES cell identity. A computational framework has been provided by Cirghu et al., [37] where they provide an approach for the systematic integration of the expression studies to identify the phenotype of any gene of interest. Xu et al., [38] provides an SVM classifier for predicting stemness of genes in mouse. Mikkelsen et al., [39] integrated the expression profiling with the ChIP-seq approach to dissect the complexities of achieving direct reprogramming. Another popular approach has been based on the application of Bayesian integration. Guan et al., [40] has been able to derive a functional network of laboratory mouse which can predict the functional linkages among the protein coding genes. In a nutshell, the advent of the next generation techniques like Chip-chip, chip-seq and RNA-seq clubbed with Integrative Bioinformatics approaches can help to explore the molecular events in pluripotency [41-43].

Tools for Visualization

A snap-shot of the state of a biological cell can be represented in the form of an electronic circuit. A network in the modern outlook is defines as a paradigm shift from independent modules to study of complex interactions. Open source software platforms have helped us design the virtual experiments and perform advanced scientific research. Lists of tools are available over the web browser including VisANT, Pathway Studio, Cell-Designer and Cytoscape to name a few. An electronic representation of biological networks can be achieved using these software platforms and they also provide plug-ins to perform analysis such as simulations. Cytoscape [44] is a well-known network visualization platform supporting a large set of features including standard and customizable network display information, ability to import/export to a large variety of interaction files, and network views which can be zoomed in. A large number of Plug-ins, more than 65 are available for different types of analysis such Gene Ontology and Gene expression data [45].

Developing a Gene Regulatory Model in Cytoscape

Expression analysis methods identify the active biological processes by Gene Ontology studies from the differentially expressed genes. Integration is beneficial since the fusion of two types of analysis is more likely to be authentic than a single source. Many sources exist to study the gene expression profile like Gene Expression Omnibus (GEO) and Array Express. They both are public repositories. Cytoscape is freely distributed under the open source GNU level General Public License. Networks are wonderfully built using Cytoscape and the same can be easily analyzed using its varied plug-ins. The protocol can be modular in nature and can be varied according to the user. The first step includes the fetching of the network data. This can be done using a GML file created by the user, ePath, BioNetBuilder etc. A plug-in called Agilent Search is also useful for getting the input values. Another kind is the building up of the network data based on literature curation. Under this category, the nodes and edges are generated from the experimental evidences in the published literature and then can be visualized in an electronic

representation using Cytoscape. The second step would be the refinement of the network data. Using the different layout options in Cytoscape, the most appropriate layout can be selected and the network can be refined. The third step would be the annotation of the network using the node, edge and network attribute browser. The nodes can be annotated using the identifiers from the major bioinformatics resources like NCBI, Unigene, ENSEMBL, iHOP, Uniprot, etc. The edges may represent the interactions between the nodes which in case of GRN may be activation, repression or interaction. The Gene Expression data matrix can be uploaded and subjected to the network feature analysis. A number of plug-ins are available to perform different types of analysis such as Network Analyzer, Vista Clara, etc. The last step is to identify Gene Ontology process and BinGO is an excellent plug-in for adding the GO processes to the network. When a complete GRN is constructed, significant biological inferences can be made regarding the biological problem designed.

Conclusions

Complexity of the stem cells and their reversible nature makes it difficult to project them with a unidirectional approach. Therefore, it is imperative for representation describing various molecular processes involved in stem cell niche in the form of an integrated network. The discovery of reprogramming of somatic cells was an enormous leap in the field of stem cell research for the biology of regenerative medicine. A latest research indicates the development of light sensitive retina using human stem cell. A number of neurodegenerative disorders find a ray of hope using the concept of iPS cells. Disentangling the complexity underlying the generation of iPS cells and the processes involved in the same would be a tremendous benefit for the biomedical sciences.

References

1. Bambakidis NC, Butler J, Horn EM, Wang X, Preul MC, et al. (2008) Stem cell biology and its therapeutic applications in the setting of spinal cord injury. *Neurosurg Focus* 24: 20.
2. Puri MC, Nagy A (2012) Concise review: Embryonic stem cells versus induced pluripotent stem cells: the game is on. *Stem Cells* 30: 10-14.
3. Preynat-Seauve O, Krause KH (2011) Stem cell sources for regenerative medicine: the immunological point of view. *Semin Immunopathol* 33: 519-524.
4. Yu J, Thomson JA (2008) Pluripotent stem cell lines. *Genes Dev* 22: 1987-1997.
5. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126: 663-676.
6. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131: 861-872.
7. Reya T, Morrison SJ, Clarke MF, Weissman IL (2001) Stem cells, cancer, and cancer stem cells. *Nature* 414: 105-111.
8. Dalton S (2009) Exposing hidden dimensions of embryonic stem cell cycle control. *Stem Cell* 4: 9-10.
9. Wray J, Kalkan T, Smith AG (2010) The ground state of pluripotency. *Biochem Soc Trans* 38: 1027-1032.
10. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318: 1917-1920.
11. Vallier L, Alexander M, Pedersen RA (2005) Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* 118: 4495-4509.
12. Warmflash A, Arduini BL, Brivanlou AH (2012) The molecular circuitry underlying pluripotency in embryonic stem cells. *Wiley Interdiscip Rev Syst Biol Med* 4: 443-456.
13. Sato N, Sanjuan IM, Heke M, Uchida M, Naef F, et al. (2003) Molecular signature of human embryonic stem cells and its comparison with the mouse. *Dev Biol* 260: 404-413.
14. Hao J, Li TG, Qi X, Zhao DF, Zhao GQ (2006) WNT/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev Biol* 290: 81-91.
15. Xu RH, Sampsel-Barron TL, Gu F, Root S, Peck RM, et al. (2008) NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Stem Cell* 3: 196-206.
16. Minard ME, Jain AK, Barton MC (2009) Analysis of epigenetic alterations to chromatin during development. *Genesis* 47: 559-572.
17. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448: 553-560.
18. Zhao XD, Han X, Chew JL, Liu J, Chiu KP, et al. (2007) Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Stem Cell* 1: 286-298.
19. Guo H, Zhu P2, Yan L3, Li R3, Hu B4, et al. (2014) The DNA methylation landscape of human early embryos. *Nature* 511: 606-610.
20. Lee HJ, Hore TA2, Reik W3 (2014) Reprogramming the methylome: erasing memory and creating diversity. *Stem Cell* 14: 710-719.
21. Meshorer E, Misteli T (2006) Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* 7: 540-546.
22. Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, et al. (2008) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol* 26: 795-797.
23. Macarthur BD, Ma'ayan A, Lemischka IR (2009) Systems biology of stem cell fate and cellular reprogramming. *Nat Rev Mol Cell Biol* 10: 672-681.
24. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30: 207-210.
25. Hemphill EE, Dharia AP, Lee C, Jakuba CM, Gibson JD, et al. (2011) SCLD: a Stem Cell Lineage Database for the annotation of cell types and developmental lineages. *Nucleic Acids Res* 39: 525-533.
26. Ho Sui SJ, Begley K, Reilly D, Chapman B, McGovern R, et al. (2012) The Stem Cell Discovery Engine: an integrated repository and analysis system for cancer stem cell comparisons. *Nucleic Acids Res* 40: 984-991.
27. Niida A, Smith AD, Imoto S, Tsutsumi S, Aburatani H, et al. (2008) Integrative bioinformatics analysis of transcriptional regulatory programs in breast cancer cells. *BMC Bioinformatics* 9: 404.
28. Flower DR, Attwood TK (2004) Integrative bioinformatics for functional genome annotation: trawling for G protein-coupled receptors. *Semin Cell Dev Biol* 15: 693-701.
29. Pauling J, Röttger R, Tauch A, Azevedo V, Baumbach J (2012) CoryneRegNet 6.0--Updated database content, new analysis methods and novel features focusing on community demands. *Nucleic Acids Res* 40: 610-614.
30. Shah SP, Huang Y, Xu T, Yuen MM, Ling J, et al. (2005) Atlas - a data warehouse for integrative bioinformatics. *BMC Bioinformatics* 6: 34.
31. Post LJ, Roos M, Marshall MS, van Driel R, Breit TM (2007) A semantic web approach applied to integrative bioinformatics experimentation: a biological use case with genomics data. *Bioinformatics* 23: 3080-3087.
32. Wasmuth JD, Pszeny V, Haile S, Jansen EM, Gast AT, et al. (2012) Integrated bioinformatic and targeted deletion analyses of the SRS gene superfamily identify SRS29C as a negative regulator of Toxoplasma virulence. *MBio* 3.
33. Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, et al. (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437: 1173-1178.
34. Ideker T, Galitski T, Hood L (2001). A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet* 2; 343-372.

35. Müller FJ, Schuldt BM, Williams R, Mason D, Altun G, et al. (2011) A bioinformatic assay for pluripotency in human cells. *Nat Methods* 8: 315-317.
36. Newman AM, Cooper JB (2010) AutoSOME: a clustering method for identifying gene expression modules without prior knowledge of cluster number. *Bioinformatics* 11: 117.
37. Cinghu S, Yellaboina S, Freudenberg JM, Ghosh S, Zheng X, et al. (2014) Integrative framework for identification of key cell identity genes uncovers determinants of ES cell identity and homeostasis. *Proc Natl Acad Sci U S A* 111: 1581-1590.
38. Xu H, Lemischka IR, Maayan A (2010) SVM classifier to predict genes important for self-renewal and pluripotency of mouse embryonic stem cells. *Syst Biol* 4: 173.
39. Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, et al. (2008) Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454: 49-55.
40. Guan Y, Myers CL, Lu R, Lemischka IR, Bult CJ, et al. (2008) A genomewide functional network for the laboratory mouse. *PLoS Comput Biol* 4: e1000165.
41. Huang H, Hu ZZ, Arighi CN, Wu CH (2007) Integration of bioinformatics resources for functional analysis of gene expression and proteomic data. *Front Biosci* 12: 5071-5088.
42. Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44-57.
43. Huang da W, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1-13.
44. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498-2504.
45. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, et al. (2007) Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* 2: 2366-2382.