Original Research Article **Bioinformatics**

International Journal of Pharma and Bio Sciences

ISSN 0975-6299

ANALYZING THE INTERVENING SEQUENCES OF PLURIPOTENCY ASSOCIATED GENES TO IDENTIFY CONSERVED SEQUENCE PATTERNS IN HUMAN USING INTEGRATIVE BIOINFORMATICS

PRIYANKA NARAD*[1], GULSHAN WADHWA[2], KAILASH C UPADHYAYA[3]

[1]Amity Institute of Biotechnology, Amity University Uttar Pradesh, Sector-125, NOIDA-201303, India.
[2] Ministry of Science and Technology, Department of Biotechnology, New Delhi, India.
[3]Amity Institute of Molecular Biology and Genomics, Amity University Uttar Pradesh, Sector-125, NOIDA-201303, India.

ABSTRACT

The identification of conserved regions in the set of genes responsible for maintenance of pluripotency in human is essential for identifying their mode of action. To understand the mechanism of transcriptional regulation in a complicated phenomenon like pluripotency, it is indispensable to identify and characterize the intervening region of the major genes responsible for maintenance of pluripotency. We identified 56 genes responsible for pluripotency using the available literature and searching databases. The region corresponding to 1000bp from the start site of the gene was retrieved for motif analysis. These set of sequences were submitted to the MEME suite for comprehensive analysis. We identified 5 conserved motifs which had occurrence in the intervening region of all the 56 sequences. This indicated that these motifs served an important role for all of these 56 genes. These motifs were compared to known transcription factor database such as HOCOMOCO, which gave us matches with TFs which have not been reported previously to be reported in early developmental stages in human. Expression analysis of these motifs across the developmental timeline indicate that they have a role in the early stage of development from the oocyte to the 6 celled stages and were down-regulated as soon as the cell enters the pluripotent state. As an inference, we propose that these TFs potentiate the genes responsible for pluripotency by maintaining their cell cycle and growth stages. Any malfunction in the expression profile of these TFs would probably block the genes to enter the pluripotent state.

KEYWORDS: hESCs, MEME, Transcription Factor, Pluripotency

PRIYANKA NARAD *
Amity Institute of Biotechnology, Amity University Uttar Pradesh, Sector-125, NOIDA-201303, India.

Received on: 02-11-2016
Revised and Accepted on : 17-01-2017
DOI: http://dx.doi.org/10.22376/ijpbs.2017.8.1.b582-587

INTRODUCTION

Stem cells are cells that can divide and differentiate into diverse specialized cell types. The properties of stem cell like self-renewal and potency make them the area of interest of many scientists. Given, their unique regenerative abilities, stem cells offer new potential for treating a variety of diseases. Because of these properties, Embryonic Stem Cells (ESCs) are an excellent system to study cellular differentiation in both normal and diseased states. Stem cell research is one of the most active areas in molecular biology and biomedicine. This is based on recent breakthroughs in generating 'induced pluripotent stem cells' (iPS cells) from somatic tissues. Such a ' reprogramming' of differentiated cells into 'pluripotent ' ones is possible by directly manipulating gene regulation, confronting the differentiated cell with artificial amounts of key transcription factors such as Oct4, Sox2 and Nanog. Scientists are just beginning to understand the signals inside and outside the cell that triggers the differentiation process[1, 2]. Numerous data has been generated on hESCs over the last few decades. Transcriptional program underlying human pluripotency is still ambiguous and opens an area for further research. The intervening region of the genes involved in the maintenance of pluripotency may contain several promoters that are usually located proximal to or overlapping the transcription initiation site. These sites also contain several sequence motifs with which transcription factors (TFs) interact in a sequence-specific manner. The combinations of the TF-binding motifs in promoters vary depending on the gene, so that an appropriate subset of genes can be expressed according to tissue types or developmental stages [3, 4]. Discovering sequence patterns in a large number of sequences still remains challenging. It is a difficult task to identify the motifs through experimental procedures. Hence, we employ an *insilico* approach to identify differentially expressed genes (DGEs) from microarray data pertaining to hESC. Next, we identify sequence patterns common to these genes by analysing the intervening region of them. Further we also performed gene ontology and comparison analysis on the motifs identified to find out their putative role in the maintenance of pluripotency in human. Finally, we study the developmental timeline from oocyte to the ESC state. This was crucial in the identification of the expression patterns for the known and identified motifs. The methodology and the results are discussed in the subsequent sections to highlight our findings.

MATERIALS AND METHODS

IDENTIFICATION OF DGEs ASSOCIATED WITH HUMAN PLURIPOTENCY

The genes were identified from the gene expression data taken from the Gene Expression Omnibus[5] repository. The data include GSE21222[6] and GSE46872[7]. We used R 3.1.1 for windows for the identification of differentially expressed genes (DGEs). For the purpose of pre-processing we used RMA algorithm[8] and for identification of differentially expressed genes, we used limma package[9]. Both positive and negative regulators of pluripotency were taken into consideration. Once the differentially expressed genes were identified, we next sorted the functional association of these genes using 2 approaches namely literature curation and database search. During our literature review process, we searched for any information in the research/review papers pertaining to loss/gain of pluripotency associated with these genes. The second approach was using the database such as ENCODE[10], STRING[11] to establish the relatedness of these genes.

RETRIEVAL OF INTERVENING REGION OF DGEs

The gene sequence was retrieved from NCBI GENE database. Each gene was located in the genomic region and a 1000bp upstream the gene location were taken for all the 56 genes identified to be associated with pluripotency in human. The DGEs were subjected to further analysis and curation. This region contains the region toward the 5' end of the coding strand i.e. towards the start site, and intervening region between genes in some cases. A number of conserved sequences are present towards the upstream region of the gene. These conserved sequences contain specific elements that ensure the initial binding of transcription factors.

CONSERVED PATTERN ANALYSIS USING MEME SUITE

We submitted our set of 56 gene sequences as input to the MEME server[12]. We selected the normal mode for analysis. The site distribution was selected to identify the motifs which had at least one representation per sequence. The background model selected was 0 order model of sequences. We restricted the tool to identify the conserved pattern on the given strand only. The maximum width for the sequence was kept to be 20bp as most of transcription factor binding sites stretch to this length. From the MEME output, a query file containing the motif in MEME format was submitted to the MAST server[13]. This was repeated for each of the predicted motifs. For each sequence, the match scores were converted into various types of p-values and these are used to determine the overall match of the sequence to the group of motifs and the probable order and spacing of occurrences of the motifs in the sequence is displayed. The sequences were scanned against the eukaryotic promoter sequences database for human. Further, we ported all 5 motifs through the TOMTOM server[14]. Next we selected the target database to compare the motifs. The database category was selected as Human and database selected was HOCOMOCO Human (v10). The motifs were again submitted in MEME format to the GOMo[15] server. Human promoters were selected to analyse the submitted motifs to the server.

EXPRESSION ANALYSIS OF NOVEL TFs

The major processes related to pre-implantation in human development are still unclear. Human development program lasts for around ~6-8 days. During this time, the organization of chromatin and the pluripotent cells are established. Following the pattern of expression of the genes involved in pluripotency, we can find out the mechanism of activation/repression of pluripotency related genes. This could also provide an

insight into the epigenetic changes occurring through the developmental process. For this purpose, we selected the gene expression dataset GSE29397[16]. This sample dataset consists of expression profiles of developmental changes starting from as early as the oocyte stage to the ESC stage (24 samples). Gene expression profiling of human pre-implantation was followed with biological triplicates of oocyte to blastocyst stages, and compared with ESCs. We downloaded the raw dataset from GEO and R language software was used and data normalization was done using the RMA algorithm[17]. After the normalization, the data was transported to the excel file for establishing a line chart and graph.

RESULTS AND DISCUSSION

After pre-processing and normalizing the data, we identified the differentially expressed genes using limma. Next, the initial screening identified 56 genes/proteins which are associated with maintenance of human pluripotency [Table 1].

Table 1
This table highlights the genes identified that are associated with the maintenance of human pluripotency.

GENE	GENE NAME	ENSEMBL ID
ZIC3	Zic Family Member 3	ENSG00000156925
ATBF1	Zinc Finger Homeobox	ENSG00000140836
NRG1	Pro-NRG1 ,neuregulin	ENSG00000157168
ACVR1	Activin A Receptor Type 1	ENSG00000115170
FAST	Troponin T3, Fast Skeletal Type	ENSG00000130595
LRP5	LDL Receptor Related Protein 5	ENSG00000162337
NEUROG1	Neurogenin	ENSG00000181965
POU5F1	POU Class 5 Homeobox	ENSG00000204531
TDGF1	Teratocarcinoma-Derived Growth Factor 1	ENSG00000241186
MYF5	Myogenic Factor 5	ENSG00000111049
SFRP5	Secreted Frizzled-Related Protein	ENSG00000120057
LRH1	Nuclear Receptor Subfamily 5 Group A Member 2	ENSG00000116833
GABPA	GA Binding Protein Transcription Factor Alpha Subunit	ENSG00000154727
HNF4A	Hepatocyte Nuclear Factor 4 Alpha	ENSG00000101076
NOGGIN	Noggin	ENSG00000183691
TGFB1	Transforming Growth Factor Beta	ENSG00000105329
PDGFRA	Platelet Derived Growth Factor Receptor Alpha	ENSG00000134853
PAX6	Paired Box 6	ENSG00000007372
GDF3	Growth Differentiation Factor 3	ENSG00000184344
SOX2	SRY-Box 2	ENSG00000181449
PRDM14	PR Domain 14	ENSG00000147596
HAND1	Heart And Neural Crest Derivatives Expressed 1	ENSG00000113196
SALL4	Spalt-Like Transcription Factor 4	ENSG00000101115
SOX7	SRY-Box 7	ENSG00000171056
CER1	Cerberus 1, DAN Family BMP Antagonist	ENSG00000147869
ZNF521	Zinc Finger Protein 521	ENSG00000198795
FRIZZLED	Secreted Frizzled-Related Protein	ENSG00000104332
GATA3	GATA Binding Protein	ENSG00000107485
FOXA2	Forkhead Box A2	ENSG00000125798
SKIL	SKI-Like Proto-Oncogene	ENSG00000136603
PTPN11	Protein Tyrosine Phosphatase, Non-Receptor Type 11	ENSG00000179295
LEFTY2	Left-Right Determination Factor 2	ENSG00000143768
REST	RE1 Silencing Transcription Factor	ENSG00000084093
SOX17	SRY-Box 17	ENSG00000164736
PDGF1	Platelet Derived Growth Factor Subunit A	ENSG00000197461
HIF1Alpha	Hypoxia Inducible Factor 1 Alpha Subunit	ENSG00000100644
SPHK1	Sphingosine Kinase	ENSG00000176170
WWP2	WW Domain Containing E3 Ubiquitin Protein Ligase	ENSG00000198373
XIST	X Inactive Specific Transcript (Non-Protein Coding)	ENSG00000229807
Follistatin	Follistatin	ENSG00000134363
c-myc	V-Myc Avian Myelocytomatosis Viral Oncogene Homolog	ENSG00000136997
ONECUT1	One Cut Homeobox	ENSG00000169856
STAT3	Signal Transducer And Activator Of Transcription	ENSG00000168610
ESX1L	ESX Homeobox	ENSG00000123576
S1P	Sphingosine-1-Phosphate Receptor	ENSG00000213694
PTEN	Phosphatase And Tensin Homolog	ENSG00000171862
KLF4	Kruppel-Like Factor 4 (Gut)	ENSG00000136826
SOST	Sclerostin	ENSG00000167941
S6K	Ribosomal Protein S6 Kinase B1	ENSG00000108443
PBX1	Pre-B-Cell Leukemia Homeobox 1	ENSG00000185630
L1TD1	LINE-1 Type Transposase Domain Containing 1	ENSG00000240563
MYST3	Lysine Acetyltransferase 6A	ENSG00000083168
EN1	Engrailed Homeobox 1	ENSG00000163064
HESX1	HESX Homeobox 1	ENSG00000163666
LHX1	LIM Homeobox 1	ENSG00000273706
FOXH1	Forkhead Box H1	ENSG00000160973

LEFTY1	Left-Right Determination Factor 1	ENSG00000243709
UTF1	Undifferentiated Embryonic Cell Transcription Factor 1	ENSG00000171794
ANOG	Nanog Homeobox	ENSG00000111704

The DGEs were subjected to further analysis and curation. To identify novel conserved elements, we used the upstream region of these sequences through the MEME suite[18]. We identified 5 sequence motifs of length ~20bp, which are present in the upstream region of all 56 genes [Fig.1].

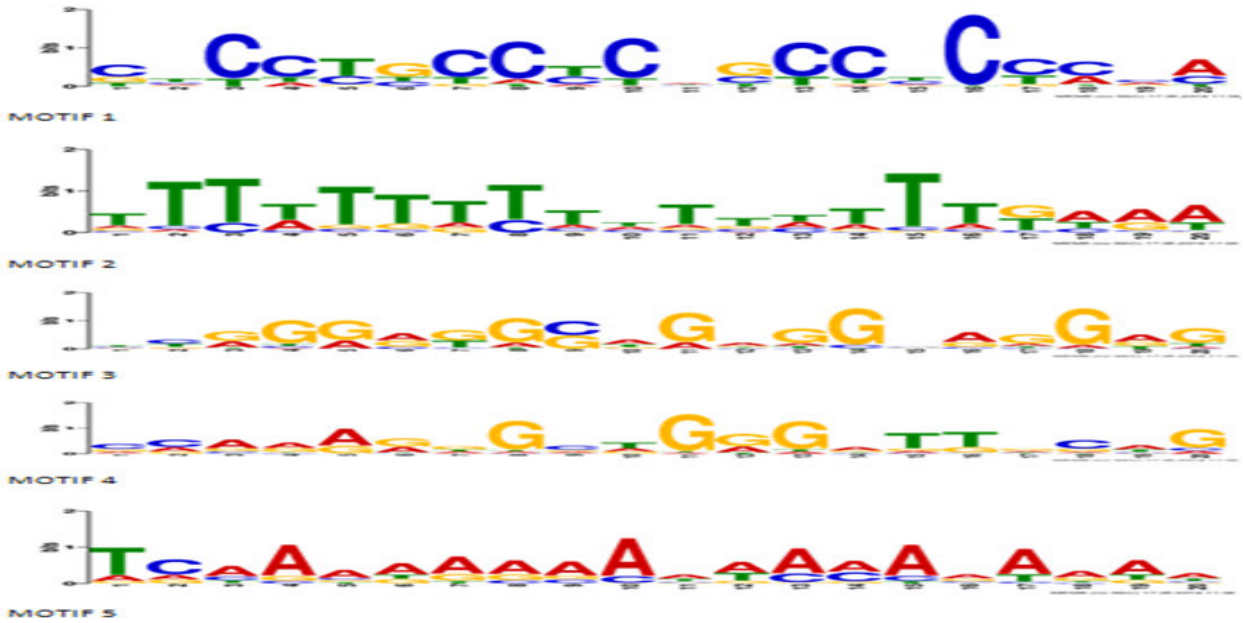

Figure 1
5 sequence motifs of length ~20bp, which are present in the upstream region of all 56 genes.

Next, we located the positions of these motifs using the MAST tool. We identified that all the motifs were located on the forward strand of the genes. To identify the functional annotation of these motifs we further used the TOMTOM module in the MEME suite. We selected a motif database for eukaryotic DNA called the Homo Sapiens Comprehensive Model Collection (HOCOMOCO)[19]. We got a substantial p-value for the identified motifs which were similar to known transcription factors [Table.2]. These transcription factors have not been associated previously with pluripotency.

Table 2
This table highlights the functional annotation of motifs identified matching them against known TFs

MOTIF	Known TF	Summary	TF binding sites in Gene Promoter
1	TBX15	This gene belongs to the T-box family of genes. This family encodes a phylogenetically conserved family of TFs that regulate a variety of developmental processes.	NRSFform1, NRSFform2, C/EBPalpha, IRF-7A, Spz1, AREB6, CHOP-10, Hlf, p53, Sox9
	TFDP1	This gene encodes a family of TFs that heterodimerize with E2F proteins to enhance their DNA-binding activity and promote transcription from E2F target genes. The encoded protein functions as part of this complex to control the transcriptional activity of numerous genes involved in cell cycle progression from G1 to S phase.	POU2F1a POU2F1 POU3F2 POU3F2(N-Oct-5a) POU3F2(N-Oct-5b)
2	CPEB1	This gene encodes a member of the cytoplasmic polyadenylation element binding protein family. These are involved in the regulation of mRNA translation, as well as processing of the 3' untranslated region, and play a role in cell proliferation and tumorigenesis.	HFH-1, Tal-1beta, E47, AREB6, MZF-1, Elk-1,ITF-2,LCR-F1.
	HOXC10	This gene belongs to the homeobox family of genes. These play an important role in morphogenesis in all multicellular organisms. The protein level is controlled during cell differentiation and proliferation, which may indicate this protein, has a role in origin activation.	E2F-3a, E2F-4, E2F-2, E2F-1, E2F, FOXO1a.
	ONECUT2	This gene encodes a member of the onecut family of transcription factors. The protein binds to specific DNA sequences and stimulates expression of target genes, including genes involved in melanocyte and hepatocyte differentiation.	GR-alpha, GR CUTL1, AREB6, MEF-2A, aMEF-2, FOXL1, Cdc5, Elk-1, GATA-1.
3	PURA	This gene product binds preferentially to the single strand of the purine-rich element termed PUR, which is present at origins of replication and in gene flanking regions in a variety of eukaryotes from yeasts through humans. Thus, it is implicated in the control of both DNA replication and transcription.	Max c-Myc Egr-1 Sp1 Max1 E2F-4 E2F-5 MZF-1 Oct-B1 E2F-3a
4	NR2F2	This gene encodes a member of the steroid thyroid hormone superfamily of nuclear receptors.	NF-kappaB1 MyoD RSRFC4
	PITX2	This protein plays a role in the terminal differentiation of somatotroph and lactotroph cell	AP-1 .ATF-2, c-Jun

		phenotypes, is involved in the development of the eye, tooth and abdominal organs.		
5	VDR	This gene encodes the nuclear hormone receptor for vitamin D3. Downstream targets of this nuclear hormone receptor are principally involved in mineral metabolism though the receptor regulates a variety of other metabolic pathways, such as those involved in the immune response and cancer.	PPAR-gamma2 PPAR-gamma1 PPAR-alpha	PPAR-

The identified TFs were analyzed through the oocyte, 2-celled, 4 celled, 6 celled, 8-10 celled, morula, blastocyst and embryonic stem cell state. We also took the expression values of known TFs like OCT4, NANOG, and SOX2 which have been known to play a role in the maintenance of pluripotency[20, 21]. Interestingly, we observed a drastic difference in the expression patterns of both the sets of TFs i.e. TFs known to be important for maintenance of pluripotency and motifs similar to TFs identified to be present in upstream regions of all the genes [Fig.2]. Most of the major TFs (OCT4, SOX2, NANOG, and KLF4) are showing very low expression at both the stages of development. As we progress further to the six-celled stage, we observed the first change in the expression pattern. Among the core pluripotency genes, NANOG, SOX2 and KLF4 are over-expressed at the six-celled stage. The expression of OCT4 is still under-expressed at the six-celled stage. At the eight-celled stage, we observe that OCT4 gets activated and now core circuitry is completely operational. The identified TFs through our experiment showed a high level of expression during the oocyte to the 4 celled stages, whereas at the onset of the pluripotency program their expression diminishes [Fig.2].

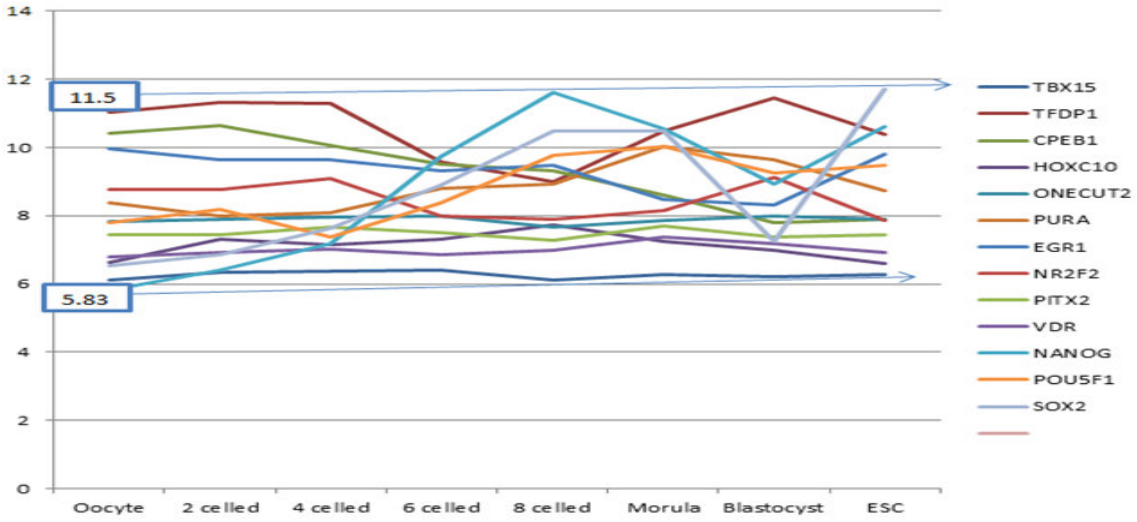

Figure 2
Expression profile across the developmental stages (Oocyte to the ESC stage) indicating a low limit of 5.83 and high limit of 11.5 expression intensity.

There are several aspects of the work which were intriguing to us. Firstly, the presence of the all 5 motifs across the identified differentially expressed genes. This indicated that these motifs served an important role for all of these 56 genes. Our first obvious inference was to assume that these motifs similar to TFs have a role in pluripotency maintenance in human. To prove our assumption we studied the expression pattern of all the TF genes out of the 56 genes identified. There were 32 TF genes present in the dataset. We took the expression pattern of these 32 TF genes associated with pluripotency and the identified novel 9 TFs through our work. Our assumption of these 9 TFs having similar expression pattern as the known 32 TF for pluripotency was proved to be incorrect. The expression pattern for pluripotency associated TFs showed that they were downregulated at the oocyte stage still the 8-celled stage and then was either upregulated/downregulated depending upon their role in maintenance of pluripotency. If the TF enhances pluripotency it was upregulated after the 8-celled stage up to the blastocyst, whereas, if the TF was shown to inhibit pluripotency they were downregulated till the blastocyst stage. So, we delved further into the facts, we observed that the novel motifs/TFs were mainly involved in the initial process of cell cycle development and growth[23, 24,25]. As an inference, we propose that these TFs potentiate the genes responsible for pluripotency by maintaining their cell cycle and growth stages. The work will be useful for the readers to understand the developmental process in more depth as it is important for the application of stem cells in diseases. Motifs are important elements which can depict conserved nature of the elements in the species.

CONCLUSION

The work was undertaken to identify basic sequence elements in the genes associated with human pluripotency. Ethical issues surrounding pluripotency has made it important for bio-informaticians to give cues to the experimentalist all over the world so that the lab work can be reduced. Towards this initiative we performed a very basic *insilico* molecular analysis. This hypothesis can be further subjected to experimental analysis. Further, we plan to analyze the region from the transcript start site of these genes and more number of genes which are experimentally verified to play a role in human pluripotency. This will benefit to identify more targets as well as TFs which may help promote pluripotency and be useful in regenerative medicine.

CONFLICT OF INTEREST

Conflict of interest declared none.

REFERENCES

1. Oeckinghaus, A., & Ghosh, S. (2009). The NF-κB family of transcription factors and its regulation. Cold Spring Harbor perspectives in biology, 1(4), a000034.
2. Fagnocchi L, Mazzoleni S, Zippo A, Integration of Signaling Pathways with the Epigenetic Machinery in the Maintenance of Stem Cells, *Stem Cells Int,*. 2016.
3. Mitchell, P. J., & Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science, 245(4916), 371-378.
4. Novina, C. D., & Roy, A. L. (1996). Core promoters and transcriptional control. Trends in Genetics, 12(9), 351-3552.
5. Edgar, R., Domrachev, M., & Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research, 30(1), 207-210.
6. theunissen TW, Powell BE, Wang H, Mitalipova M et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell 2014 Oct 2;15(4):471-87.
7. Hanna J, Cheng AW, Saha K, Kim J et al. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl AcadSci U S A 2010 May 18; 107(20):9222-7.
8. Wu, Z., & Irizarry, R. A. (2004). Preprocessing of oligonucleotide array data. Nature Biotechnology, 22(6), 656-658.
9. Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research, gkv007.
10. ENCODE Project Consortium. (2004). The ENCODE (ENCyclopedia of DNA elements) project. Science, 306(5696), 636-640.
11. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., ... & Kuhn, M. (2014). STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic acids research, gku1003.
12. Bailey, T. L., Williams, N., Misleh, C., & Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic acids research, 34(suppl 2), W369-W373.
13. Timothy L. Bailey and Michael Gribskov, "Combining evidence using p-values: application to sequence homology searches", Bioinformatics, 14(1):48-54, 1998.
14. Fabian A. Buske, Mikael Bodén, Denis C. Bauer and Timothy L. Bailey, "Assigning roles to DNA regulatory motifs using comparative genomics", *Bioinformatics*, 26(7), 860-866, 2010.
15. Shobhit Gupta, JA Stamatoyannopolous, Timothy Bailey and William Stafford Noble, "Quantifying similarity between motifs", *Genome Biology*, 8(2):R24, 2007.
16. Vassena, R., Boué, S., González-Roca, E., Aran, B., Auer, H., Veiga, A., & Belmonte, J. C. I. (2011). Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. Development, 138(17), 3699-3709.
17. Clough, E., & Barrett, T. (2016). The Gene Expression Omnibus Database.*Statistical Genomics: Methods and Protocols*, 93-110.
18. Timothy L. Bailey, Mikael Bodén, Fabian A. Buske, Martin Frith, Charles E. Grant, Luca Clementi, Jingyuan Ren, Wilfred W. Li, William S. Noble, "MEME SUITE: tools for motif discovery and searching", Nucleic Acids Research, 37:W202-W208, 2009.
19. Kulakovskiy, I. V., Medvedeva, Y. A., Schaefer, U., Kasianov, A. S., Vorontsov, I. E., Bajic, V. B., & Makeev, V. J. (2013). HOCOMOCO: a comprehensive collection of human transcription factor binding sites models.Nucleic acids research, 41(D1), D195-D202.
20. De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., and Daley, G. Q. (2015). Hallmarks of pluripotency. *Nature*, 525(7570), 469-478.
21. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., and Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, *122*(6), 947-956.
22. Clotman, F., Jacquemin, P., Plumb-Rudewiez, N., Pierreux, C. E., Van der Smissen, P., Dietz, H. C., & Lemaigre, F. P. (2005). Control of liver cell fate decision by a gradient of TGFβ signaling modulated by Onecut transcription factors. *Genes & development*, *19*(16), 1849-1854.
23. Attwooll, C., Oddi, S., Cartwright, P., Prosperini, E., Agger, K., Steensgaard, P., & Helin, K. (2005). A novel repressive E2F6 complex containing the polycomb group protein, EPC1 that interacts with EZH2 in a proliferation-specific manner. Journal of Biological Chemistry, 280(2), 1199-1208.
24. Chapman, D. L., Garvey, N., Hancock, S., Alexiou, M., Agulnik, S. I., Gibson-Brown, J. J., ... & Papaioannou, V. E. (1996). Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development.Developmental Dynamics, 206(4), 379-390.
25. Festuccia N, Dubois A, Vandormael-Pournin S, Tejeda EG, Mouren A, Bessonnard S, Mueller F, Proux C, Cohen-Tannoudji M, Navarro P. Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nat. Cell Biol*., 2016, 18(11):1139-48.