

Comparative Epigenetic Analyses of Acute and Chronic Leukemia

Yan Zhang, Dianjing Guo

School of Life Sciences, The Chinese University of Hong Kong, Hong Kong, China
Email: djguo@cuhk.edu.hk

Received April 2015

Abstract

Comparative analysis of epigenetic alterations between acute vs. chronic leukemia, with an emphasis on histone modifications, was conducted. We focus on the promoter regions of the whole genomes as well as oncogenes. Our results revealed that obvious differential histone modifications pattern exists between the two subtypes. H3K27ac has a high tag density in the promoter region in both Dnd41 cell lines and K562 cell lines. H3K27ac and H3K4me1 have high correlation between the two cell lines of oncogenes. Similar results were also achieved in the promoter region of high expression genes in the Jurkat and K562 cell lines based on RNA-seq data. This suggests that H2K27ac and H3K4me1 are active regulators in leukemia cell lines.

Keywords

Leukemia, Epigenetics, Histone Modification, Promoter

1. Introduction

Epigenetics has been defined as the molecular factors and processes around DNA that are mitotically stable and regulate genome activity independent of DNA sequence [1]. Epigenetic research examines biochemical modifications of the genome and aims to determine the resulting consequences on gene function. The two most commonly studied epigenetic mechanisms are DNA methylation and histone modifications [2]. DNA methylation is the most widely studied epigenetic abnormality in tumor genesis. It refers to methylation of cytosine at CpG dinucleotides, and CpG dinucleotides are not randomly distributed throughout the human genome [3]. Loss of DNA methylation at CpG dinucleotides was the first epigenetic abnormality to be identified in cancer cells [4]. For instance, hypomethylation leading to activation of genes that are important in cancer includes promoter CpG demethylation in the overexpression of cyclin D2 and maspin in gastric carcinoma, MN/CA9 overexpression in human renal-cell carcinoma, S100A4 metastasis-associated gene in colon cancer and human papillomavirus 16 (HPV16) expression in cervical cancer [4].

Histone tails post-translational modifications are mediated by numerous enzymes, including acetylation, methylation, ubiquitination, sumoylation, biotinylation and crotonylation of lysines (K), and phosphorylation of serines (S) and threonines (T) [5]. The pattern of histone modifications, so called “histone code”, determines the configuration of chromatin, and adjusts the accessibility to effector proteins [6]. For example, tri-methylation of

lysine 4 on histone H3 (H3K4me3) is associated with active transcription, whilst H3K27me3 and H3K9me3 are the two chief repressive marks [7]. There is a large number of enzymes involved in the addition or removal of the covalent modifications, including histone acetyltransferases (HATs), deacetylases (HDACs), methyltransferases (HMTs), and demethylases (HDMs) [7] [8].

Concerning histone modifications, global loss of both acetylation of lysine 16 (H4K16ac) and of trimethylation of lysine 20 of histone H4 (H4K20me3) are commonly found in cancer cells, usually in association with DNA hypomethylation at repetitive DNA sequences [9]. In addition, altered methylation patterns of H3K9 and H3K27 have also been found in cancer cells [10]. Those modifications, which result in inappropriate gene silencing, are accomplished by histone-modifying enzymes [6]. Based on recent study, some researchers found a combinatory effect of different epigenetic modifications on gene expression in breast cancer subtypes, indicating that alterations of epi-modifications in the breast cancer subtypes not only affect genes involved in common cancer biology, but also regulate genes participating in subtype-specific functions [11].

Leukemia is a group of cancers that usually begins in the bone marrow and results in high numbers of abnormal white blood cells. Clinically and pathologically, leukemia is subdivided into a variety of large groups. The first division is between its acute and chronic forms: Acute leukemia (Acute lymphoblastic leukemia (ALL), Acute myelogenous leukemia (AML)) and Chronic leukemia (Chronic lymphocytic leukemia (CLL), Chronic myelogenous leukemia (CML)). Comparative analysis of epigenetic alterations between the ALL and CLL subtypes of leukemia, particularly in histone modifications and DNA methylation, may provide clue to reveal widespread distinct patterns of epigenetic alterations in leukemia and thereby gain important insight into tumorigenesis.

2. Materials and Methods

2.1. Materials

Epigenetic and transcriptomic data sets of leukemia subtypes are from ENCODE (The Encyclopedia of DNA Elements), GEO (Gene Expression Omnibus) and EMBL-EBI (European Bioinformatics Institute) database:

We compare the alterations of five types of histone modifications (H3K4me1, H3K4me3, H3K27ac, H3K36me3 and H3K27me3) and DNA methylation. We use the embryonic stem (ES) cell line and normal blood cell line as control group.

2.2. Method

Reads generated from ChIP-Seq and RNA-seq were mapped to human reference genome (hg19) by Bowtie2 program [12]. To quantify RPKM expression (Reads Per Kilobase of transcript per Million reads mapped) at gene level, the DEGseq program is applied to assemble the reads with default parameters [13].

2.2.1. Calculation of Tag Destiny on the Genome, Promoter and Genes

For genome wide analysis, promoter regions up to 2 kb upstream of the transcription start sites (TSSs) were collected from Refseq [14]. Then we use “leukemia subtype” as key word to search the GeneCard database respectively. The analyses of oncogenes promoter can help us to find the real hallmark genes of leukemia subtypes.

ChIP-seq data can be used to locate promoter or CpG Island within few tens of base pairs. Tag densities at the special sites are a good indicator of epigenomic alterations, which makes it easier to quantify and compare epi-

Table 1. Description of leukemia subtypes cell lines.

Cell lines	Subtypes of leukemia	Description of cell lines
K562	CML	53-year-old female CML patient in blast crisis
Dnd41	ALL	13-year-old boy, with CDKN2B (p15INK4B) and CDKN2A (p16INK4A) deletions and TP53 (p53) mutation
Jurkat	ALL	14-year-old boy
HL-60	AML	36-year-old female

AML—Acute myelogenous leukemia, ALL—Acute lymphoblastic leukemia, CML—Chronic myelogenous leukemia.

genomic modifications at different DNA sites.

2.2.2. Analyses of Differential Histone Modification Sites

ChIPDiff program [15] was used for genome-wide quantitative comparisons of histone modification. R software package, including rtracklayer, GenomicRanges, Rsamtools and GenomicAlignments was used for statistical analysis.

3. Results

3.1. Differential Histone Modification between CLL and ALL at Chromosomal Level

Genome wide analysis suggested that there are obvious differential histone modifications in ALL and CLL cell line. The histone modification statistical analyses on the genome chromosome 1 were illustrated in **Figure 1**.

3.2. Comparison of Histone Modification in the Promoter Regions

As shown in **Figure 2**, Clear difference in histone modification was also observed in the promoter regions between ALL and CLL lines.

3.3. Density Analysis on the Promoter Regions of Leukemia Oncogenes

We collected 20 leukemia oncogenes and conducted the density analysis on their promoters. Similiarly, the histone modification patterns are quite different between the two cell lines (**Figure 3**).

3.4. Analyses of Gene Expression and Its Correlation with Histone Modifications

We compared the distribution of the histone modifications on 20 highly expressed gene in ALL and CLL. Of the 4 types of histone modifications, H3K27ac seems to be the major type (**Figure 4**).

The correlation of various types of epigenomics modifications and their roles in the regulation of gene expression has drawn lots of research attention recently. The spearsman correlation matrix of the histone modifications based on the 20 genes suggests that H3K4me1 and H3K27ac have a higher correlation between the two cell lines, and they are both active signals on the promoter area (**Figure 5**). For the other modifications, e.g. H3K27me3

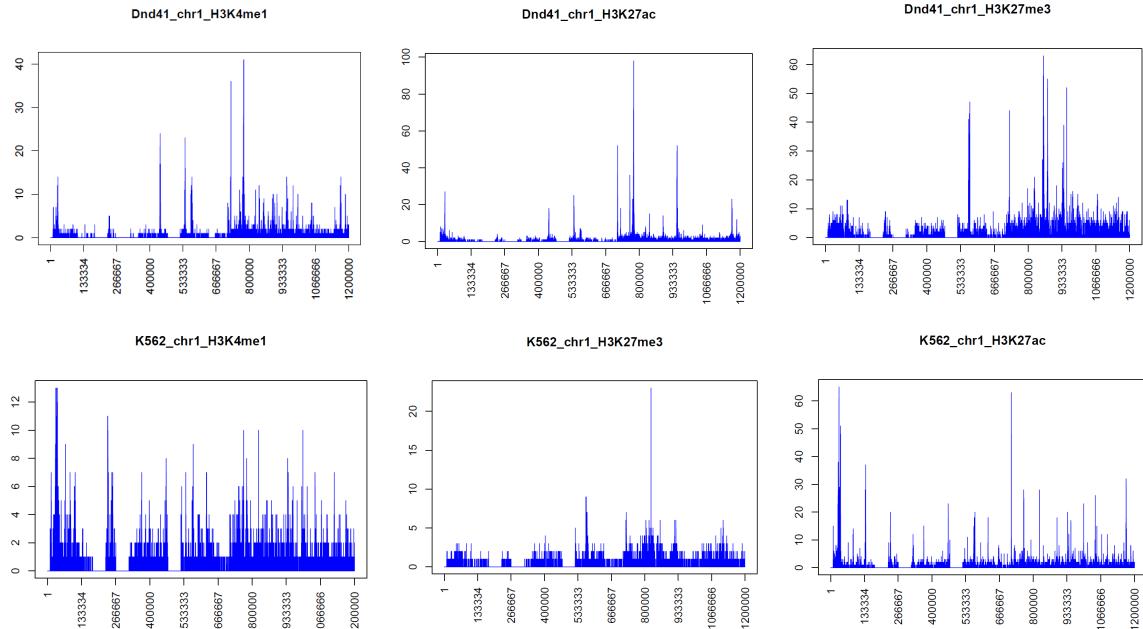


Figure 1. The histone modification patterns for H3K4me1, H3K27ac and H3K27me3 on chromosome 1 of ALL (upper panels) and CLL (lower panels).

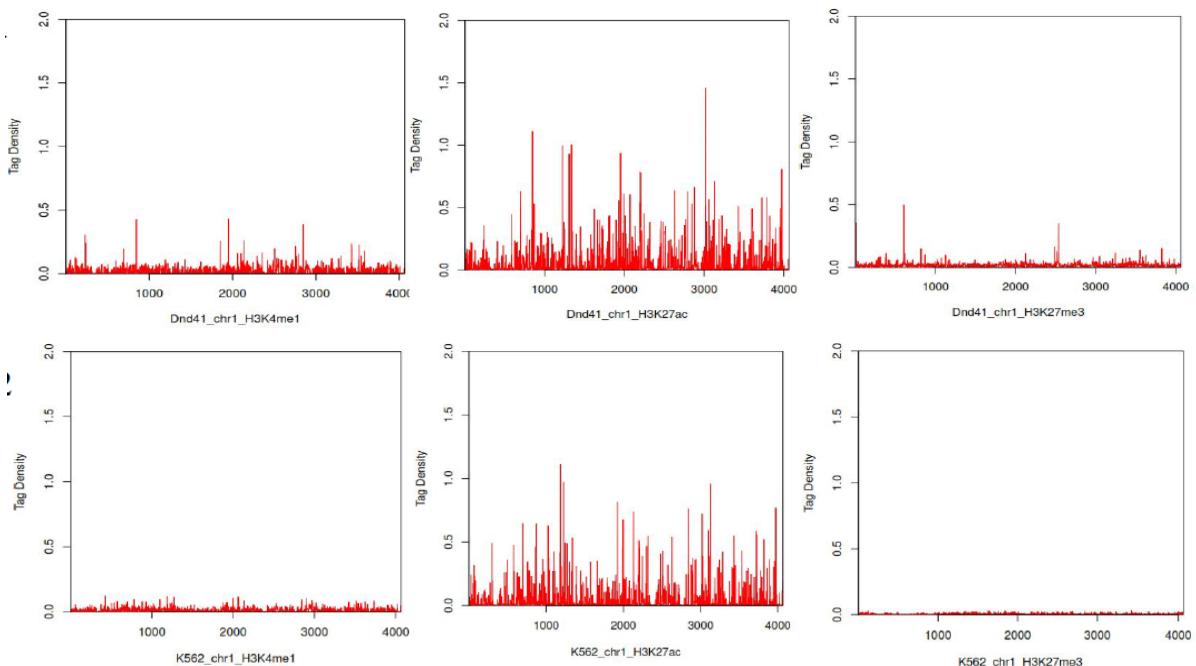


Figure 2. The histone modification pattern for chromosome 1 promoter region of ALL (upper panels) and CLL (lower panels).

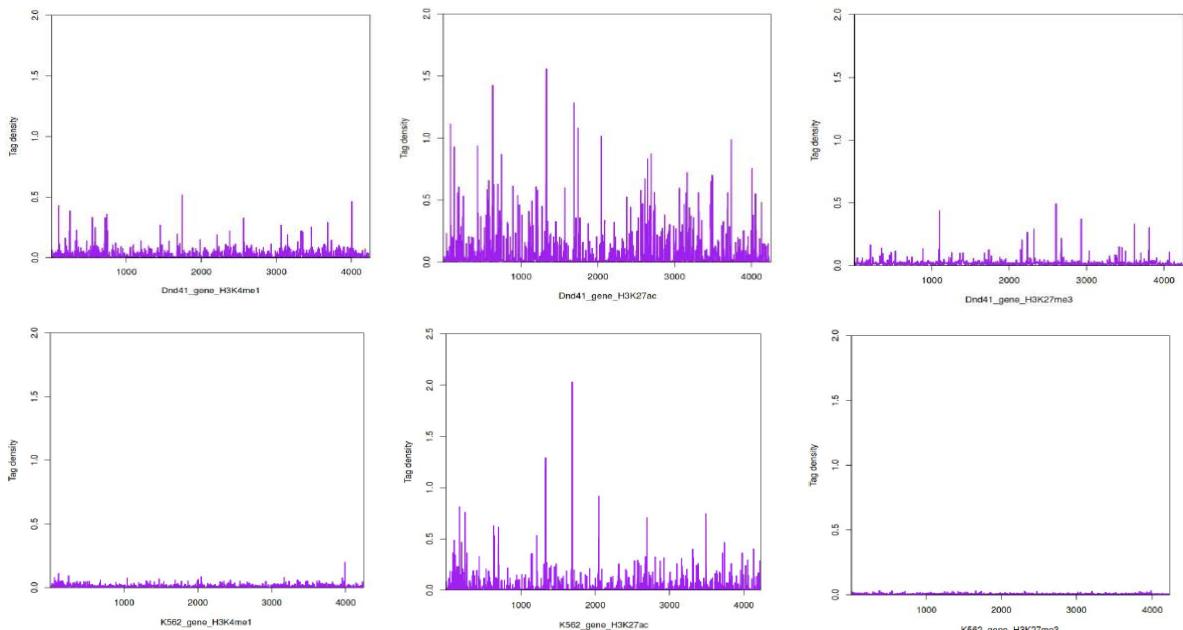


Figure 3. The H3K4me1, H3K27ac and H3K27me3 distribution on promoter of oncogenes in Dnd41 (upper panels) and K562 cell line (lower panels).

and H3K36me3 have the low tag density at the promoter region and low correlation between the two cell lines.

We also analyzed the RNA-seq data and calculated the RPKM value with DEGseq software package. Because the RNA-seq data of Dnd41 cell are not available, we used the other ALL cell lines Jurkat as the substitute. We first collect 20 high expression genes ($RPKM > 100$), which appear in Jurkat cell line and K562 cell line respectively. We then analyzed the distribution of histone modifications (Figure 6).

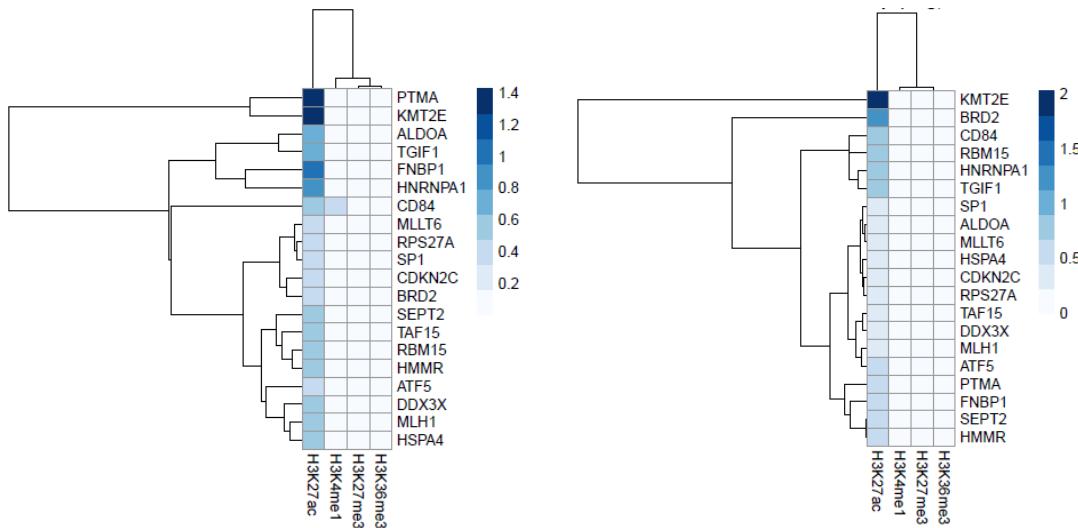


Figure 4. The heat map of 4 types of histone modifications in promoters of 20 highly expressed genes. Left: Dnd41 cell line; Right: K562 cell line.

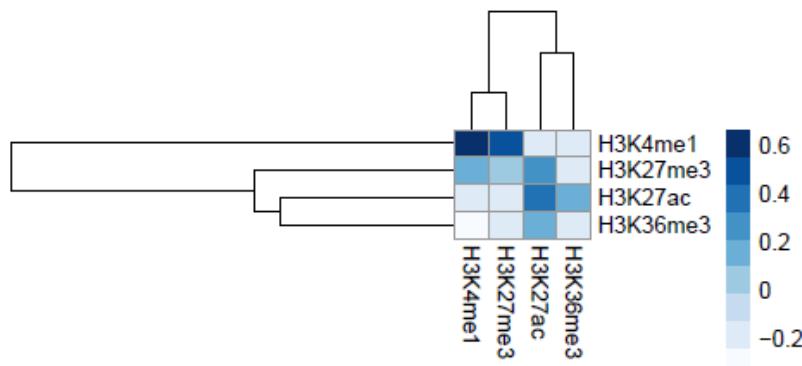


Figure 5. The spearman correlation matrix of histone modification in Dnd41 and K562 cell lines.

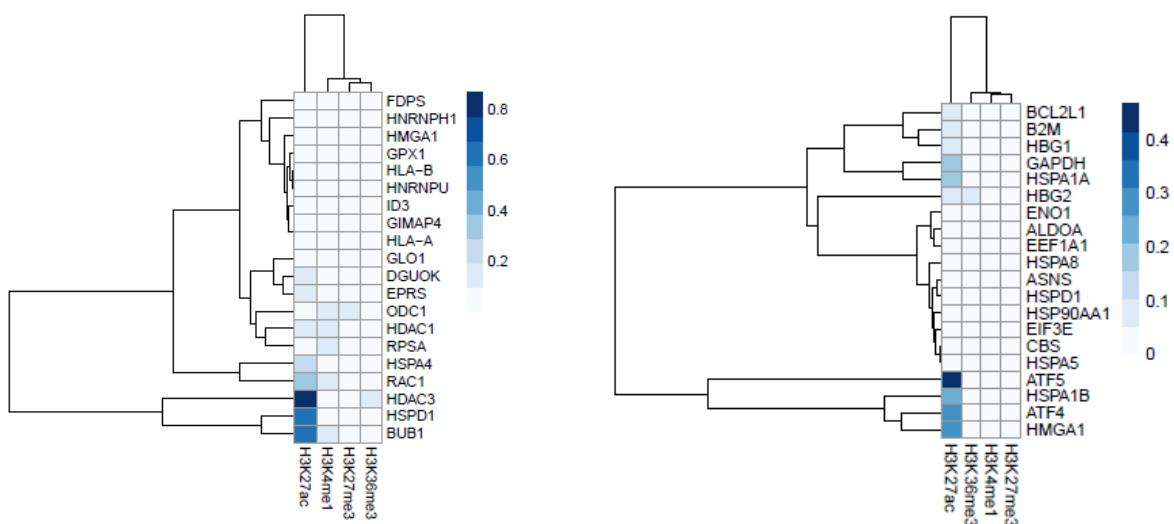


Figure 6. The heat map of 4 types of histone modifications in promoters of 20 highly expressed genes. Left: Jurkat cell line; Right: K562 cell line.

4. Conclusion

In this work, we found that H3K27ac has a high tag density in the promoter region in both Dnd41 cell lines and K562 cell lines. H3K27ac and H3K4me1 have high correlation between the two cell lines of oncogenes. Similar results were obtained in the promoter region of high expression genes in the Jurkat and K562 cell lines based on RNA-seq data. This suggests that H2K27ac and H3K4me1 are active regulator in leukemia cell lines. To date, the most thoroughly investigated therapeutic area in terms of epigenetics is cancer. Aberrant patterns of histone modifications have also been shown to have a role in oncogenesis [16]. Given the circumstances, comparative epigenetic analyses could provide a new approach to understand the mechanism of cancer and develop potential drug targets for anticancer drug discovery.

References

- [1] Skinner, M.K., Manikkam, M. and Guerrero-Bosagna, C. (2010) Epigenetic Transgenerational Actions of Environmental Factors in Disease Etiology. *Trends in Endocrinology & Metabolism*, **21**, 214-222. <http://dx.doi.org/10.1016/j.tem.2009.12.007>
- [2] Doherty, R., Farrelly, C.O. and Meade, K.G. (2014) Comparative Epigenetics: Relevance to the Regulation of Production and Health Traits in Cattle. *Animal Genetics*, **45**, 3-14. <http://dx.doi.org/10.1111/age.12140>
- [3] Bhat, A.A., Wani, H.A., Beigh, M.A., Bhat, S.A., Jeelani, S., Massood, A., et al. (2013) Epigenetic Promoter Methylation of hmlh1 Gene in Human Gut Malignancies: A Comparative Study. *Journal of Investigational Biochemistry*, **2**, 101-108. <http://dx.doi.org/10.5455/jib.20130409124009>
- [4] Feinberg, A.P. and Tycko, B. (2004) The History of Cancer Epigenetics. *Nature Reviews Cancer*, **4**, 143-153. <http://dx.doi.org/10.1038/nrc1279>
- [5] Deakin, J.E., Domaschenz, R., Lim, P.S., Ezaz, T. and Rao, S. (2014) Comparative Epigenomics: An Emerging Field with Breakthrough Potential to Understand Evolution of Epigenetic Regulation. *AIMS Genetics*, **1**, 34-54. <http://dx.doi.org/10.3934/genet.2014.1.34>
Wit, E. and McClure, J. (2004) Statistics for Microarrays: Design, Analysis, and Inference. 5th Edition, John Wiley & Sons Ltd., Chichester. <http://dx.doi.org/10.1002/0470011084>
- [6] Henrique, R., Luis, A. and Jerónimo, C. (2012) The Epigenetics of Renal Cell Tumors: From Biology to Biomarkers. *Frontiers in Genetics*, **3**, 94. <http://dx.doi.org/10.3389/fgene.2012.00094>
Giambastiani, B.M.S. (2007) Evoluzione Idrologica ed Idrogeologica Della Pineta di san Vitale (Ravenna). Ph.D. Thesis, Bologna University, Bologna.
- [7] Kouzarides, T. (2007) Chromatin Modifications and Their Function. *Cell*, **128**, 693-705. <http://dx.doi.org/10.1016/j.cell.2007.02.005>
Honeycutt, L. (1998) Communication and Design Course. <http://dcr.rpi.edu/commdesign/class1.html>
- [8] Shi, Y. (2007) Histone Lysine Demethylases: Emerging Roles in Development, Physiology and Disease. *Nature Reviews Genetics*, **8**, 829-833. <http://dx.doi.org/10.1038/nrg2218>
- [9] Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., et al. (2005) Loss of Acetylation at Lys16 and Trimethylation at Lys20 of Histone H4 Is a Common Hallmark of Human Cancer. *Nature Genetics*, **37**, 391-400. <http://dx.doi.org/10.1038/ng1531>
- [10] Nguyen, C.T., Weisenberger, D.J., Velicescu, M., Gonzales, F.A., Lin, J.C., Liang, G., et al. (2002) Histone H3-Lysine 9 Methylation Is Associated with Aberrant Gene Silencing in Cancer Cells and Is Rapidly Reversed by 5-Aza-2'-deoxycytidine. *Cancer Research*, **62**, 6456-6461.
- [11] Li, Y., Li, S., Chen, J., Shao, T., Jiang, C., Wang, Y., et al. (2014) Comparative Epigenetic Analyses Reveal Distinct Patterns of Oncogenic Pathways Activation in Breast Cancer Subtypes. *Human Molecular Genetics*, **23**, 5378-5393. <http://dx.doi.org/10.1093/hmg/ddu256>
- [12] Langmead, B. and Salzberg, S.L. (2012) Fast Gapped-Read Alignment with Bowtie 2. *Nature Methods*, **9**, 357-359. <http://dx.doi.org/10.1038/nmeth.1923>
- [13] Wang, L., Feng, Z., Wang, X., Wang, X. and Zhang, X. (2010) DEGseq: An R Package for Identifying Differentially Expressed Genes from RNA-seq Data. *Bioinformatics*, **26**, 136-138. <http://dx.doi.org/10.1093/bioinformatics/btp612>
- [14] Pruitt, K.D., Tatusova, T. and Maglott, D.R. (2007) NCBI Reference Sequences (RefSeq): A Curated Non-Redundant Sequence Database of Genomes, Transcripts and Proteins. *Nucleic Acids Research*, **35**, D61-D65. <http://dx.doi.org/10.1093/nar/gkl842>
- [15] Xu, H., Wei, C.-L., Lin, F. and Sung, W.-K. (2008) An HMM Approach to Genome-Wide Identification of Differential

- Histone Modification Sites from ChIP-seq Data. *Bioinformatics*, **24**, 2344-2349.
<http://dx.doi.org/10.1093/bioinformatics/btn402>
- [16] DeWoskin, V.A. and Million, R.P. (2013) The Epigenetics Pipeline. *Nature Reviews Drug Discovery*, **12**, 661-662.
<http://dx.doi.org/10.1038/nrd4091>