The Insight of Regulation Mechanism from DNA Methylation Based on Single Base Resolution Sequencing Data

College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, 150081

Abstract DNA methylation has an important role in regulation of transcript, cell development and differentiation, even disease. Dynamic of DNA methylation may be contribute to the multiple phenotypes. Many differential methylation regions were identified based on high-throughput methylation data between tissues, especially normal and tumors. However, DNA methylation patterns show high heterogeneity in the region from different tissues even with the same average methylation levels. Whether diverse methylation patterns reflect the complex and differential regulation mechanism of DNA methylation? Here, we reviewed the regulation mechanism and dynamics of DNA methylation and proposed the prospect of DNA methylation pattern polymorphism.

Keywords DNA methylation pattern; Dynamic; Polymorphism; High-throughput methylation data

1 Introduction
Gene expression is an important step in the life of the organism during converting genetic information into protein molecule showing active and function. In this process, the changes of gene expression may lead to disorder of protein function, cell differentiation, abnormal tissue formation even disease. DNA methylation is one of the epigenetic modification regulating gene expression and has important regulatory role in cell, tissue and the formation of a complex phenotype. Therefore, the regulatory mechanism of gene expression involved in DNA methylation and its role in the development of cancer and other diseases has been one of the important contents of epigenetic research.

DNA methylation occurs mainly at the cytosine of CpG dinucleotides fifth carbon atoms, p represents a linking deoxyxycytidine and deoxyguanosine nucleoside phosphate group, this is also known as 5- cytosine methyl - cytosine. DNA methylation in gene promoter as well as the close relationship between the transcriptional regulation of the promoter region is usually the main target for researchers (Ally et al., 2009). Abnormal methylation of the promoter region leading to silence is a common cause of epigenetic mechanisms even to cancer (Kikuchi et al., 2013). Thus, a large number of aberrant methylation of regions associated with cancers were identified in the past years according to the average level of CpG methylation sites, which are considered for the functional regions regulating gene expression (Noushmehr et al., 2010; Berman et al., 2012). However, the correlation between methylation levels in gene promoter region and gene is weak in genome-wide, only hypermethylation regions in cancer show the significant negative correlation weak (Aryee et al., 2013). Therefore, the average level of methylation of CpG sites in the region may only a small amount explanation of DNA methylation regulation mechanism about gene expression. Cytosine DNA methylation site region may contain different methylation pattern including consistent, bimodal or random methylation patterns that represent different epigenetic mechanisms eventually result in a different region shows the same average methylation levels. It has been shown, the same methylation level of regions showed a variety of DNA methylation patterns forming epipolymorphism both in normal or cancer tissue, (Landan et al., 2012). Highly heterogeneity of DNA methylation in cells may be the cause that lead to weak correlation between methylation and gene expression levels in the whole genome-wide. Distinguish complex patterns of DNA methylation in cells / tissues and quantitative assessment may contribute...
to better reflect the role of DNA methylation in gene expression, as well as a new understanding which is involved in cell differentiation and epigenetic mechanisms of tumorigenesis.

1.1 Regulational mechanism of DNA methylation
DNA methylation is participating in an important biochemical processes in normal development of higher organisms, the genome cytosine loci fifth carbon atom transfer catalyzed by a DNA methylation enzyme, adding a methyl group to form a 5-methyl cytosine (5mC). DNA methylation is removed during zygote formation, and then re-established through a process of continuous development of cell division. DNA methylation is heritable, and lead to genomic imprinting region, that is generally referred to as epigenetic regulation (Law and Jacobsen, 2010). So far, DNA methyltransferase enzymes include DNMT1, DNMT1b, DNMT1o, DNMT1p, DNMT2, DNMT3a, DNMT3b and transcript isoforms DNMT3L. There are two ways producing methylation including de novo methylation and methylation continued. DNMT3a and DNMT3b can able to make the double-stranded DNA unmethylated CpG methylation loci to a half and then the whole methylation which is an important de novo methylation transferase, involved in cell growth and differentiation and tumor-related gene methylation (Bedford and van Helden, 1987). As an important part of maintaining higher organisms normal development and cell differentiation, DNA methylation in cells require changing the gene expression pattern of cells to ensure programming and reprogramming performed. DNA methylation can cause chromatin structure, DNA conformation, DNA stability as well as DNA and protein interactions alter the way, to control gene expression. For example, CPS1 transcription start sites occur ultra methylation in HCC, the impact of the activity of the promoter region of the gene leads to the expression of silence CPS1 (Liu et al., 2011) in HCC. In addition, DNA methylation is also capable of inhibiting viral / mosaic expression of proto-oncogene or harmful elements in the genome, thus inhibiting the occurrence of cancer (Docherty et al., 2010). Because of the numerous regulatory processes, DNA methylation has been caused attention of epigenetic research.

Either hypermethylation or abnormal hypomethylation are closely related to the development of cancer. The global hypomethylation in tumor tissue is one of the aberrant epigenetic changes in human cancer that was first discovered. Tumorigenesis in the development process, an abnormally low level of methylation of genomic DNA are increasing with cells from benign lesions to invasive cancer process (Fraga et al., 2004). Abnormal hypomethylation region mainly involved in cancer by three mechanisms: 1) the generation of chromosomal instability: DNA methylation leads chromosomal instability by promoting somatic deletion/ectopic recombination, or the promotion of chromosomal rearrangements e.g., by disrupting the DNA methyltransferase DNA methylation causes attenuation cause chromosomal instability (Karpf and Matsui, 2005). 2) Activation of transposable elements: DNA hypomethylation can activate aberrant retrotransposons, so that they are transcribed into the genome or ectopic expression of other regions around the affected gene. For example, DNA methylation unusually prompted low long interspersed nuclear element L1 insertion, resulting in abnormal CYMC proto-oncogene expression in colon and breast cancer (Rideout et al., 1990; Das and Singal, 2004). 3) Loss of imprinting: lower DNA methylation result in the loss of genomic imprinting. For example, IGF2 abnormality occurs in a variety of cancers hypomethylation result in loss of imprinting, but with increased cancer risk related to (Feinberg, 1999; Cui et al., 2003; Kaneda and Feinberg, 2005). In conclusion, abnormally low methylated regions is thought to induce cancer through activation of oncogenes or retrotransposable elements. On the other hand, hypermethylation in the promoter region of tumor suppressor gene in cancer is another important epigenetic changes, and has a specific cancer type (Esteller et al., 2001). Hypermethylation of promoter region leads to inactivation of tumor suppressor genes thus contributing to the incidence of cancer (Herman and Baylin, 2003), such as hypermethylation of BRCA1 promoter region of the gene silencing in primary cause of breast and ovarian cancer (Esteller et al., 2000). Ultrahigh methylated promoter regions can affect genes involved in the cell cycle, DNA repair, metabolism of carcinogens, interactions between cells, apoptosis and angiogenesis, these processes are closely related to the occurrence of cancer (Esteller, 2007).

Genome-wide methylation differences in the region, including the organization of methylation differences between regions (T-DMRs) (Rakyan et al., 2008),
methylation interregional differences in cancer and normal samples (C-DMRs) (Irizarry et al., 2009), the development process regional differences in methylation (D-DMRs) (Meissner et al., 2008), the reprogramming process differentially methylated region (R-DMRs) (Doi et al., 2009) as well as methylated regions within individual / inter differences (Intra / Inter- DMRs) (Bjornsson et al., 2008; Bock et al., 2008), were widely identified based on the statistical methods through comparing methylation levels between samples and be regarded as the genomic regions involved in transcriptional regulation has the potential to function. Currently, the algorithm on mining differentially methylated region has been matured, including t-test applying the comparison between two samples (Bibikova et al., 2006), a variety of the analysis of variance (Byun et al., 2009) and Kruskall-Wallis test (Eckhardt et al., 2006) and based on information entropy theory development the QDMR method (Zhang et al., 2011). QDMR method is not able to determine the sample and sample-specific restriction differentially methylated region.

Although DNA methylation by affecting the expression of proto-oncogene / tumor suppressor gene results in the occurrence of cancer is relatively common knowledge, but the relationship between aberrant methylation and gene expression in the area of cancer only in hypermethylation area shows a significant weak correlation, can not completely explain the regulation of DNA methylation.

1.2 Dynamic of DNA methylation

DNA methylation patterns in genome wide are basically a bimodal distribution, that a high level of most CpG methylation sites (> 85%) and most of the non-CpG island methylation (<10%). With a different DNA sequence, DNA methylation of genetic information from the global schema is not gametes. Almost all of the DNA methylation in the early embryonic period will be erased, and then re-establish a new bimodal pattern in the implantation period. This is similar to the recording operation is the key to clear the entire epigenetic markers, whose symbolic cleared in germline cells are totipotent in order to get to prepare. Bimodal methylation patterns, which once was established, would have been maintained at each cell division process, the implementation of a global regulatory mechanism of epigenetic inhibition.

Methylation sensitive restriction enzymatic analysis displayed that most of the non-methylated global demethylation genome before implantation embryo implantation period (Monk et al., 1987). Despite the global demethylation bed is considered to be re-established after the epigenome gamete formation plays an important role, but the exact function of the mechanism is still not been confirmed. After early embryonic DNA methylation clear, each individual in the implantation phase will establish a new model, which is mainly due to de novo methylation transferase DNMT3a, DNMT3b joint DNMT1 cause upregulation of global methylation (Okano et al., 1999). In this process, there is still a protection mechanism to avoid the presence of de novo methylation in the part of a special sequence --CpG island regions (Brandeis et al., 1994). Therefore, methylation patterns are bimodal methylation implantation period of dynamic change in response in somatic cell. Bimodal methylation patterns reflect different active mechanisms in the development process: 1) This part of the sequence undergo demethylation during gamete formation, namely re-methylated sequences (> 85%) to promote somatic cell survival; 2 ) remains not occur de novo methylation sequence (<10%) such as CpG island methylation occurs once may result in almost all regions of the genome of silence, which are keeping a relatively open structure (Cedar and Bergman, 2012). Recently, based on a single-base resolution of DNA methylation dynamics showed that DNA methylation and the dynamic configuration having a significant difference in mammalian brain tissue (Lister et al., 2013).

Studies have shown that DNA methylation was the critical step occurs during development, and has been maintained methylation patterns formed until needed further changes. Although based on the methylation level excavated a large number of tissue or cell lines between methylation regional differences, but there is no evidence on the organization/cells during the development of the dynamic process of the formation of these differentially methylated regions. On the other hand, compared to the epigenome of normal cells/tissues, the formation of cancer epigenome can be seen as micro-evolution of cells, in the process changing the methylation patterns follow ways that according to a kinds of regulatory rules methylation patterns change from one to another methylation pattern? Or a random mutation occurred independently
in parallel processes in multicellular many points? Also, or occurred during a single cell and then spread in a population of cells? These are still unresolved need to think about.

1.3 DNA methylation data of high throughput single base resolution

With the development of DNA methylation sequencing technology, DNA methylation data also undergoing change. In earlier studies, the main data is to obtain a small number of DNA methylation or restriction endonuclease gene marker genome scanning method by restriction does not apply to the genome-wide study (Shen and Maniatis, 1980; Kawai et al., 1993). DNA methylation co-immunoprecipitation (MeDIP) and its binding oligonucleotide array technology (MeDIP-chip) can be applied to genome-wide, to make up for lack of prior studies (Weber et al., 2005; Weber et al., 2007). Appearances of next-generation sequencing technology allow greatly improved throughput sequencing, and promote people to take advantage of this technology to develop new detection methods. Detection of DNA methylation is the standard technique bisulfite genomic sequencing (Bisulfite sequencing, BS-seq), bisulfite such de-methylated cytosines into uracil amino DNA did not occur, methylated cytosine remains unchanged, the desired fragment was amplified by PCR, the complete conversion of uracil thymine. Finally, the PCR products were sequenced and compared with sequences untreated, determines whether the CpG methylation sites. This method is a reliable and highly accurate method, it is possible on a single base accuracy by calculating the ratio of methylated CpG sites to quantify DNA methylation levels (Berman et al., 2012). BS-seq technology can detect most of the regions of the genome CpG sites or even the entire human genome, because of its advantages of single-base coverage accuracy by researchers to study DNA methylation patterns. BS-seq technology time-consuming and cost too much, at least 10 or more reads in order to obtain reliable data, requires a lot of cloning and sequencing plasmid extraction, the process is more complicated and expensive. BS-seq data are often covered with a lower number of depth and sample, and therefore need to set the threshold reaches its depth of coverage and depth of coverage precision balance. Reduce the proportion of bisulfite sequencing (Reduced representation bisulfite sequencing, RRBS) is an accurate, efficient, DNA methylation research methods economy is mainly for bisulfite sequencing and CpG islands in the promoter region, RRBS only can be achieved within the scope of its coverage precision of single-base resolution, and a high degree of reproducibility in the coverage area of multi-sample, suitable for a variety of the differences between analyzes (Meissner et al., 2008). RRBS technology because a sufficient sample size and depth of coverage, as methylation studies and cost-effective method adopted by researchers, has broad application prospects in the study of large-scale clinical samples.

High-throughput DNA methylation data allows one to study tissue/cell line specific for methylated regions from single-base accuracy while being able to study methylated regions associated with cancer and cell differentiation during the development of DNA methylation variation patterns. Disease with high-throughput methylation database, e.g. NGSmethDB (Geisen et al., 2014) and Diseasemeth (Lv et al., 2012), can promote mining disease epigenetic markers. However, previous studies have generally through the region methylation sites proportions quantitative methylation level, and then between the average level of methylation of CpG sites within the region comparison sample DNA methylation level differences are ignored the complexity of the difference in the cell population of different cells in methylation patterns. Gilad et al., studies the polymorphism of methylation patterns in the diversity of tissue / cell lines in methylated regions consisting of four sites from RRBS data and found with the same level of methylation region shows a high degree of apparent polymorphism (Landan et al., 2012). The results suggest that there may be different mechanisms of epigenetic methylation region even though at the same level, and therefore only consider the level of methylation may only get limited results.

2 Discussion

DNA methylation has an important regulatory role in cell differentiation and tumorigenesis which has been known for a long time. The next generation of high-throughput single nucleotide resolution sequencing data allow the further research of regulatory mechanisms of DNA methylation on a larger scale and unprecedented resolution, to generate new knowledge about DNA methylation. Providing bisulfite
sequencing data based on a single base resolution in more tissues, the methylation status in single-base cytosine sites of tissues or species, allowing the methylation status of individual binding sites from the cell level and a quantitative model to re-assess the complexity of their models in the cell population, and further more accurately interpret DNA methylation in epigenetic regulation mechanisms. However, the current challenge is how the cellular level to complex patterns of DNA methylation quantitative estimates and research relationship between gene expression, and in the process of cell differentiation or dynamic changes of cancer, including: 1) the complexity of the impact site of cytosine methylation status in different cell within the same organization, the distance near the cytosine loci factors of DNA methylation patterns, how these factors into account complex model to quantify; 2) how to determine DNA methylation patterns consistent with the degree of complexity of the region as well as screening for cancer/tissue-specific DNA methylation patterns complex; 3) how to build a suitable test complex patterns of DNA methylation dynamics cells microevolution model, which is subject to the specific measure phenotypic selection constraints or cancer cell differentiation process.

Acknowledged
We thank the support of Students Research and Training Program of Heilongjiang Province (grant number 201410226023), the national Natural Science Foundation of China (grant number 61402139), the Health Department Project of Heilongjiang Provincial (grant number 2012-799).

Reference
Bedford M.T., and Van Helden P.D., 1987, Hypomethylation of DNA in pathological conditions of the human prostate, Cancer Res, 47; 5274-5276 http://dx.doi.org/10.1158/0008-5472.CAN-10-1304
Esteller M., 2007, Cancer epigenomics: DNA methylomes and histone-modification maps, Nat Rev Genet, 8; 286-298 http://dx.doi.org/10.1038/nrg2005


http://dx.doi.org/10.1158/0008-5472.CAN-03-4061


http://dx.doi.org/10.1093/nar/gkt1202


http://dx.doi.org/10.1056/NEJMra023075


http://dx.doi.org/10.1038/ng.298


http://dx.doi.org/10.1158/0008-5472.CAN-05-2959


http://dx.doi.org/10.1158/0008-5472.CAN-05-1961


http://dx.doi.org/10.1093/nar/21.24.5604


http://dx.doi.org/10.3389/fgene.2013.00271


http://dx.doi.org/10.1038/ng.2442


http://dx.doi.org/10.1038/nrg2719


http://dx.doi.org/10.1126/science.1237905


http://dx.doi.org/10.1016/j.ajpath.2010.03.023


http://dx.doi.org/10.1093/nar/gkr1169


http://dx.doi.org/10.1038/nature07102


http://dx.doi.org/10.1016/j.ccr.2010.03.017

Okano M., Bell D.W., Haber D.A., and Li E., 1999, DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, Cell, 99: 247-257.

http://dx.doi.org/10.1016/S0092-8674(99)00016-6


http://dx.doi.org/10.1101/gr.077479.108

Rideout W.M., 3rd, Coetzee G.A., Olumi A.F., and Jones P.A., 1990, Methylat...