

Research Article

The Contribution of Next Generation Sequencing Technologies to Epigenome Research of Stem Cell and Tumorigenesis

Jun Li¹, Tingting Wang², Xiuge Zhang³ and Xu Yang*

Department of Human Medical Research, Research & Cooperation Division, BGI-Shenzhen, Shenzhen 518083, China

Abstract

Epigenome contains another layer of genetic information, not as stable as genome. Dynamic epigenome can serve as an interface to explain the role of environmental factors. Stem cell and tumorigenesis are reported to be closely associated with epigenome modifications. Next generation sequencing (NGS) technologies have directly leaded to the recent advances in epigenome research of stem cell and cancer. DNA methylation and histone modification are two major epigenetic modifications. Four NGS-based approaches have been developed to identify these two epigenetic modifications, including whole genome bisulfite sequencing (WGBS), methylated DNA Immunoprecipitation Sequencing (MeDIP-Seq), reduced representation bisulfite sequencing (RRBS) and chromatin immunoprecipitation sequencing (ChIP-Seq). This paper reviews the recent advances of WGBS, MeDIP-Seq and RRBS for DNA methylation and ChIP-Seq for histone modification in the field of stem cell. The potential contribution of epigenetic modifications to tumorigenesis is also described. At present, the epigenome research still faces the defects of current sampling strategy and unknown network regulation pattern. In future, worldwide collaboration and latest sequencing technologies application are expected to solve these problem and offer new insight into epigenome research.

Keywords: Next generation sequencing; Epigenome; DNA methylation; Histone modification; Stem cell; Tumorigenesis.

Introduction

Genome sequencing has great positive effect on human disease research since its emergence. It enables researchers to explore and understand the mechanism of disease development on nucleic acid level. The effect of genome sequencing on human disease research has been obviously demonstrated by several international collaborative projects [1,4]. Human genome project, began in 1990 and completely accomplished in 2003, had constructed the first map of human genome, widely used as the reference sequence of subsequent human genome researches [1]. International Hapmap project, officially started in 2002 and initially published in 2005, firstly described the haplotype map of the human genome, revealing the common patterns of human genetic variants. The single nucleotide polymorphism (SNP) information in the Hapmap project is fundamental to explore common genetic variants affecting human health and disease [2,3]. Furthermore, the 1000 genome project launched in 2008 is expected to find more genetic variant information with larger samples and resources and build the most comprehensive catalogue of human genetic variations [4]. The project is designed to sequence 2,500 genomes of individuals from 27 populations and obtain comprehensive genetic variants contributing to the genetic diversity in human population, such as structural variants (SV) and copy number variants (CNV). The pilot study of the project was finished in 2010 and revealed unprecedented number and type of genetic variants [4]. The achievements of these large projects have switched on the "big science" mode of human disease research by collaboration of worldwide scientists. They are regarded as the milestones, setting the clear goal and reference for the subsequent numerous human disease researches based on genome sequencing.

However, as more genome sequencing researches emerged, it was found that the genetic variants of genome level were not enough to fully demonstrate and understand human disease mechanisms. It was speculated that there was another layer of information besides genome sequence to determine the state of human health and disease, based on two reasons below. First, as a multicellular organism, human body can produce a variety of cells corresponding to distinct functions. Since all human cells share the DNA sequence, information other than DNA sequence may occur to control cell development for a particular type to function in different tissues [5]. Second, the expression of gene in DNA sequence is regulated by environmentally induced changes, such as nutrient, toxins, drugs, infection, behavior and stress [6,7]. Genome sequencing can merely clarify the life diversity among individuals, populations and ethnic groups by detected genetic variants. However, the results of genome level research cannot help to explain the regulation mechanism of external factors to make the diversity occur, especially those similar genomes with different phenotypes. For example, monozygotic twins were born to have identical genome sequence, but would have different diseases with their growing up. Thus, it is expected to reveal the mystery by the study of another layer.

This further layer of information for regulating the differential gene expression was early described as 'epigenetic control' by Nanney in 1958 [8]. Although there is a little debate on the precise description of epigenetics, the fundamental definition of epigenetics refers to the heritable changes in cell or tissue specific gene expression with no alteration in the DNA sequences [6]. The heritable changes, inherited from cell to cell and generation to generation, are mostly established during the process of cellular differentiation and are steadily maintained

*Corresponding author: Xu Yang Ph.D, BGI-Shenzhen, Shenzhen 518083, China, E-mail: yangxu@genomics.cn

Received November 01, 2011; Accepted November 28, 2011; Published November 30, 2011

Citation: Li J, Wang T, Zhang X, Yang X (2011) The Contribution of Next Generation Sequencing Technologies to Epigenome Research of Stem Cell and Tumorigenesis . Human Genet Embryol S2:001. doi:10.4172/2161-0436.S2-001

Copyright: © 2011 Li J, 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.

through multiple cycles of cell division [9]. These heritable regulation mechanisms mainly include DNA methylation, histone modifications, nucleosome positioning, chromatin remodeling, genomic imprinting and ncRNA regulation. Multilevel epigenetic mechanisms constitute the system of regulating gene expression in cells. By cell specific regulations, those mechanisms are crucial for cellular developments, such as embryogenesis, cell differentiation [10]. Thus, the aberration in the epigenetic regulation system is reported to be associated with a wide range of diseases [11,14].

Similar to genome, epigenome contains another layer of genetic information, representing the overall epigenetic state of a cell. But epigenome is not as stable as genome, varying with influence of internal and external factors. According to alterations, various epigenomes can originate from one genome. Since most human diseases are well recognized to be jointly affected by genetic and environmental factors, the epigenome can consequently serve as a vital bridge of gene-environment interactions. Epigenome has been proved to play an important role in the development and function of cells, especially early embryo development [15,17]. The understanding of epigenome is clearly beneficial to human disease research. The increased epigenome researches in recent one decade have laid the good basis for understanding (Figure 1). Here, we will review recent epigenomic research advances of human disease. The review focuses on the application of next generation sequencing (NGS) technologies to demonstrate the contribution of epigenome to stem cell and tumorigenesis.

The NGS epigenome and stem cell research

The pace of epigenome researches increase with the progress of new techniques. Before the advent of NGS in 2005, DNA methylation microarray is usually used for genome-wide epigenome research, such as ChIP-chip, a technique combining chromatin immunoprecipitation (ChIP) with microarray technology (chip). However, NGS technologies have open a new chapter of epigenomic research as a powerful research tool for scientists [18,19]. NGS platforms (Roche 454 GS FLX, Illumina GA and HiSeq and Life Technologies SOLiD) are able to massively sequence a large quantity of sequence reads in parallel. Due to the characteristics of high-throughput data output, NGS has significantly accelerated the speed of scientific discoveries in epigenome research (Figure 1). The ability of massively parallel sequencing also allows researchers to first gain the comprehensive mapping of epigenome in different states. Compared with the previous techniques, NGS genome-wide epigenome mapping can reach unprecedented resolution through high-throughput data output. And several effective approaches based on NGS technologies are well developed and widely used [20,22,29]. These innovative advantages have made the stem cell research in the field of epigenomic blossom, but the review cannot cover all. In this section, we mainly focus on the application of four NGS-based approaches, including WGBS, MeDIP-Seq, RRBS and ChIP-Seq, to two primary forms of epigenetic marks, DNA methylation and histone modification (Table 1).

DNA methylation: DNA methylation is the most well studied epigenetic mechanism, referring to adding a methyl group at the carbon 5 position of cytosine through DNA methyltransferase (DNMT) enzymes to cytosine methylation in human genome. De novo methyl groups are catalysed by DNMT3A and DNMT3B enzymes to cytosine in newly synthesised DNA. Cytosine methylation, associated with gene silencing, is critical for hypermethylation in the promoter





 $\label{eq:table_transform} \begin{array}{l} \textbf{Table 1:} Four NGS-based methods to identify DNA methylation and histone modifications. \end{array}$

Name	Resolution	Theoretical coverage	Epigenetic mark	Applications
WGBS	Single nucleotide resolution	Whole genome	DNA methylation	Accurate measurement of genome-wide DNA methylation
MeDIP-Seq	High resolution	Highly methylated and high CpG density regions	DNA methylation	Rapid and large scale study of DNA methylation
RRBS	Single nucleotide resolution	CpG-rich regions	DNA methylation	Accurate DNA methylation study of targeted CpG-rich regions
ChIP-Seq	High resolution	Global binding sites of histones or transcriptional factors	Histone modification, transcrip- tional factors	Genome-wide study of histone modifications and transcriptional factors

The full names of NGS-based methods are represented as follows. Whole genome bisulfite sequencing (WGBS), Methylated DNA immunoprecipitation sequencing (MeDIP-Seq), Reduced representation bisulfite sequencing (RRBS) and chromatin immunoprecipitation sequencing (ChIP-Seq).

with CpG islands. The status of CpG sites in the genome is mostly methylated. But, CpG islands in the promoter regions in most human genes are not methylated [23]. DNA methylation is involved in a number of important processes such as maintaining genome stability, transcriptional silencing and genome imprinting. As a stable and heritable epigenetic mark, correct patterns of DNA methylation are crucial for normal development and lineage commitment [24,25]. Thus, the approaches based on NGS technologies to reveal the methylome are very crucial for human disease research. Three innovative NGS techniques are widely used in DNA methylation research, consisting of WGBS, MeDIP and RRBS.

• WGBS: Whole genome bisulfite sequencing (WGBS) is the gold standard method to detect and calculate DNA methylation level. NGS technologies enable WGBS to conduct DNA methylation study at single base resolution [26,28]. Treatment of DNA with sodium bisulfite will change unmethylated cytosine into thymine without alterations of methylated

cytosine. The ratio of change in the DNA sequences depends on the methylation status of individual cytosine. Thus, coupled with NGS technology, it allows for an unbiased genome-wide calculation of methylome status at single base resolution [26]. Due to WGBS's power to accurately measure DNA methylome, certain prevailing view has been challenged. For example, non-CG methylation refers to the pattern of methylation that does not occur in CpG dinucleotides, such as CT or CA. And DNA methylation was thought to primarily exist in CG islands. However, recent studies based on NGS technologies have revealed the disparate results [27,28]. As the first genome-wide map of methylated cytosines in a mammalian genome, Lister et al. [27] compared the human embryonic stem cells (hESCs) and fetal fibroblasts. The portion of non-CG methylation was much higher than expected through this study, for nearly one-quarter of all methylations identified in embryonic stem cells was found to be in a non-CG context. And non-CG methylations were enriched in gene bodies and depleted in protein binding sites and enhancers. Furthermore, non-CG methylation disappeared upon induced differentiation of the embryonic stem cells, and was restored in induced pluripotent stem cells. These interesting results strongly suggest that embryonic stem cells may rely on the high level of methylation in non-CG context for different regulatory patterns to affect gene regulation to maintain the pluripotency. It is also implied that there are alterations in epigenomic regulation mechanisms during the cell differentiation stages. As mentioned above, Laurent et al. [29] also reported the dynamic changes in the human methylome during differentiation by WGBS. Three cultured cell types were selected, including hESCs, a fibroblastic differentiated derivative of the hESCs and neonatal fibroblasts. And the mature peripheral blood mononuclear cells (monocytes) were set as a reference, for they were fully differentiated as an adult cell type. Developmental stage was reflected in both the level of global methylation and extent of non-CpG methylation. As representatives of progressive differentiation stages, hESCs have the highest level of methylation as a representative in the early stage of differentiation, while monocytes have the lowest level in the last stage, together with intermediate level of fibroblasts in the middle stage. Thus, epigenetic marks will dynamically regulate the development of various types of cells in different stages to function exactly.

In addition to hESCs, WGBS can also be used to study induced pluripotent stem cells (iPSCs). iPSCs are derived from somatic cells, epigenetically reprogrammed to lose tissue-specific features and gain pluripotency. Similar to hESCs, they can theoretically differentiate into any type of cells [30]. But the reprogramming mechanism of iPSCs is different from ESCs, so it is a hotspot to distinguish epigenome and genome betweem iPSCs and ESCs. Lister et al. [31] reported the first genome-wide DNA methylation profiles of iPSCs at single-base resolution. By comparison among the methylomes of human ES cells, somatic cells, and differentiated iPSCs and ES cells, the difference in DNA methylation status was found between iPSCs and ESC. Human iPSCs exhibited large aberrant epigenomics reprogramming, including somatic memory and aberrant reprogramming of DNA methylation. Moreover, it was revealed that errors in reprogramming CG methylation were transmitted at a high frequency by analyzing differentiation of iPSCs into trophoblast cells. The result proved that an iPSC reprogramming signature was maintained after differentiation. As an important regulatory mechanism in development, epigenetic reprogramming of DNA methylation occurs frequently during differentiation. The differentiation extent of iPSCs is intermediate between embryonic stem cells and somatic cells. It can be predicted that researches on epigenetic reprogramming will increasingly use WGBS to study iPSCs to reveal the accurate mechanisms.

WGBS can be engaged to study not only several types of stem cells mentioned above, but also adult somatic cells [28]. Wang et al. [32] studied the methylome of human peripheral blood mononuclear cells (PBMCs) by WGBS, and revealed the first Asian epigenome map of the same Asian individual whose genome was decoded in the YH project. Different from the result of Lister et al. [27] above, the portion of non-CG methylation in this study was minor, only <0.2% methylated non-CG sites. In addition, this study also revealed allele specific methylation between the two haploid methylomes, together with the previously generated whole genome sequencing data. From integrated results of different types of human cells in two methylome studies above, it could be clearly concluded that epigenomic statue is not stable to regulate the differentiation level in various types of cells. The conclusion has enlightened us to explore the contribution of non-CG methylation in maintaining and inducing cellular development, and implicated that non-CG methylation is not just existed in embryonic stem cells. With the characteristic of single base resolution, WGBS is expected to become a powerful tool in exploring the methylome differences of cells in various differentiated stages and tissue types.

MeDIP-Seq: Similar to WGBS, Methylated DNA Immunoprecipitation Sequencing (MeDIP-Seq) is a genomewide method to detect DNA methylation. However, different from sodium bisulfite treatment in WGBS, MeDIP-Seq is based on enrichment of methylated DNA sequence. The antibody especially recognizes genome-wide methylated cytosines, and the purified fraction of methylated DNA can be input to highthroughput DNA detection methods such as NGS [33]. Thus, this method is sensitive to the highly methylated and high CG density regions. Although lower resolution and less accuracy than WGBS, the characteristics of time saving and cost effective make it suitable for disease research in large sample size between cells and tissues. For example, the world largest ever epigenetics project, named as EpiTwin, was launched in 2010 by collaboration between Beijing Genomics Institute (BGI) and King's College London (TwinsUK). The EpiTwin project is to capture the subtle epigenetic differences between 5,000 twins throught MeDIP-Seq, and to explain why many identical twins don't develop the same diseases. Monozygotic twins are highly coincident in DNA sequence and consequently suitable to investigate the influence of epigenetic modifications on human diseases [34], such as autoimmune diseases [35,37].

Besides intensive research of DNA methylation, MeDIP-Seq can be applied for other fields, such as demethylation and 5-methylctosine (5mC). Demethylation is also very crucial for understanding the epigenetic mechanisms of human diseases. With both DNA methylation and demethylation, we could completely understand how these patterns of 5-methylcytosine are established and maintained. DNA demethylation is not as dynamic as methylation, as active DNA demethylation has been revealed to be merely observed during specific stages of development [38]. The existence of genomewide DNA demethylation has been reported in germ cells and early embryos [39]. Although the mechanisms of demethylation remain to be elucidated, few researchers have already begun to use MeDIP-Seq to study DNA demethylation. Chavez et al. [40] used MeDIP-Seq to analyze DNA methylation change during differentiation of hESCs to definitive endoderm. After analyzing the interplay between DNA methylation, histone modifications and transcription factor binding, demethylation was found to be mainly associated with regions of low CpG densities, in contrast to de novo methylation. Even so there are few reports of NGS applications on DNA demethylation research, its importance of DNA demethylation is expected to be gradually recognized as that of DNA methylation.

5-hydroxymethylcytosine (5hmC) is a lysine-modified base in various cell types in mammals at low level, generated by adding the hydroxymethyl group on the cytosine [41]. The formation of 5hmC is regulated by the enzyme reaction of of TET family [42,45]. Similar to the principle of 5mC antibody enrichment in DNA methylation study, MeDIP-Seq or other similar NGS-based techniques can also be applied to investigate the distribution and role of 5hmC in the genome by 5hmC-specific antibodies. As an important and novel mechanism of epigenetics, 5hmC was recently found in 2009 to be existed in embryonic stem cells, as well as human and mouse brains [42,45]. Pastor et al. [41] further used NGS-based approaches to present a genome-wide mapping of 5hmC in mouse embryonic stem cells (ESCs). It was found that 5hmC was strongly enriched in exons and near transcriptional start sites. The result suggested that 5hmC might regulate the transcription of ESCs, but its regulatory role is different from 5mC. Ficz et al. [46] used MeDIP-Seq to confirm the existence of 5hmC in mouse ESCs and its role during differentiation, and demonstrated the relationship of 5mC and 5hmC. 5hmC was found to be mainly associated with euchromatin, while 5mC was enriched at gene promoters and CpG islands. 5hmC could not occur alone, whereas it mostly depended on the existence of 5mC in the genome. It indicated that 5hmC contributed to enhance the transcription as the opposite role of methylation in inhibiting gene expression. During differentiation with decreased TET, the hydroxymethylation level at the ESC-specific gene promoters declined simultaneously with the enhanced methylation level and consequent gene silencing. However, the balance between 5mC and 5hmC was not simple, but different according to genomic regions. It was supposed by the research that the balance between pluripotency and differentiation was associated with the balance between 5mC and 5hmC. Researches have reported the distribution of 5hmC in many types of tissues, and its importance in the ESCs is being gradually recognized as mentioned above. However, researchers have just begun to be interested in this epigenetic mark of 5hmC, the limited information still remains to be investigated. We will know the biological roles of 5mC and 5hmC in ESCs and human diseases more clearly when more powerful methods have been developed to distinguish them discretely.

• **RRBS:** Reduced representation bisulfite sequencing (RRBS) is a fast and cost-effective method to provide qualified DNA methylation data, newly developed in recent years [47,49]. The first step is enzyme digestion by MspI, specifically cutting CCGG sites, and then is bisulfite treatment as the step in WGBS. Hence, RRBS can only cover CpG-rich regions such as promoter and other regulatory element, not genome-wide region as WGBS. It can still reach single base resolution as well as WGBS [48,50]. Thus, it is suitable to investigate the different methylated regions among samples for a broad scope of researches, such as medicine and biomarker [49,51].

As a recently developed NGS technique, few researches using RRBS have been published. Nevertheless, some researched have attempted to apply for biology and disease research [51,52]. For example, Wang et al. [51] applied RRBS to the human PBMC of the Asian individual from YH project, whose genome and epigenome has been systematically deciphered [28,32]. The result revealed that more than half of CpG islands and promoter regions were covered with a good coverage depth. Furthermore, the proportion of the CpG sites covered reached 80-90%, demonstrating good reproducibility of biological replicates [28]. Thus, it is a good choice for RRBS to focus on certain CpG-rich region of large samples to explore the DNA methylation differences. Besides, human disease can also be investigated by RRBS. Gertz et al. [52] used RRBS to study somatic DNA of six members in a threegeneration family. The result demonstrated the close relationship of genotype with DNA methylation. It was found that more than 92% of differential methylation between homologous chromosomes occurred on a particular haplotype, and 80% of DNA methylation differences could be explained by genotype. In addition, the study used transcriptional analysis to measure genes exhibiting genotypedependent DNA methylation, 22% of which had allele-specific gene expression differences. In general, this study highlighted the contribution of genotype to the pattern of DNA methylome. Along with the recognition of RRBS through increased publications, it will become a novel tool for DNA methylation research in many fields.

Histone modification: In addition to DNA methylation, histone modification is another type of epigenetic regulation mechanisms via chromatin change. DNA in the eukaryotic chromatin is wrapped around histone octamers, consisting of four highly conserved core histones, H2A, H2B, H3 and H4. Histones are subject to various posttranslational modifications, including but not limited to lysine, lysine and arginine methylation, serine and threonine phosphorylation, lysine acetylation, ubiquitination, sumoylation and ADP ribosylation. These modifications occur mainly within the histone amino-terminal tails [53]. The state of histone tails can contribute to alter the chromatin structure to determine the accessibility of the transcription machinery and other regulatory factors to DNA. Thus, histone modifications of the histone tails are important to regulate the level of chromatin condensation and gene expression [54]. Among various types of histone modifications, acetylation and methylation of specific lysine residues on N-terminal histone tails play a fundamental role in the formation of chromatin domains [53]. Acetylation is respectively established and removed by histone acetyltransferases and deacetylases. Likewise, methylation is regulated by histone methyltransferase and demethylase families. The contributing enzymes on methylation and acetylation specifically affect toward various histone proteins [55]. As the switch in on-off regulation of gene expression, lysine residues acetylation on histones is associated with gene activation, whereas methylation of lysine residues can result in either activation or silencing on gene expressions [56]. As an epigenetic mechanism, posttranslational modifications of histones are involved in the regulation of normal and disease-associated development. Due to technical restrictions, most of these posttranslational modifications of histones remain poorly understood. However, advances have been made obviously in recent years based on NGS application through ChIP-Seq approaches.

• **ChIP-Seq:** Chromatin immunoprecipitation sequencing (ChIP-Seq) is a strong technique combining ChIP and NGS together for genome-wide DNA-protein interaction research. Capable of efficient genome-wide investigation on histone modification status, ChIP-Seq is suitable to identify histone

Citation: Li J, Wang T, Zhang X, Yang X (2011) The Contribution of Next Generation Sequencing Technologies to Epigenome Research of Stem Cell and Tumorigenesis . Human Genet Embryol S2:001. doi:10.4172/2161-0436.S2-001

modifications with specific DNA sequence [57]. In ChIP experiment, chromatin is first treated with sonication or MNase-digestion [58], and then enriched by specific antibody. After immunoprecipitation, NGS technologies can detect specific protein's binding sites. Compared with ChIP-chip, ChIP-Seq shows higher resolution and greater coverage, and can detect more peaks and narrower peaks with a better signal-to-noise ratio [57]. The high-resolution capability of identifying genome-wide histone modifications make it fit for human biology research [59]. For example, Terrenoire et al. [59] used ChIP-Seq to study histone modifications H3K9ac, H3K27ac and H3K4me3 in human metaphase epigenome. By comparison with histone modification levels across the interphase genome, H3K4me3 and H3K27ac were revealed to show a close correspondence. Oppositely, H3K27me3, a epigenetic mark associated with gene silencing, exhibited big differences. The study provided evidence for extensive epigenome remodeling at mitosis.

In the field of stem cell research, ChIP-Seq is also used due to its powerful ability of genome-wide histone modifications characterization. Larson et al. [60] used ChIP-Seq to study five histone modification marks (H3K4me2, H3K4me3, H3K27me3, H3K9me3, and H3K36me3) in mouse embryonic stem cells (ESCs). Coupled with a hidden Markov model (HMM), these marks were identified to be respectively existed in active, non-active and null domains. Each type of domains corresponded with distinct biological functions and chromatin structural changes during early cell differentiation. The study offered new insights into the role of epigenetics in long-range gene regulation. From the research examples above, we can conclude that ChIP-Seq is efficient and powerful to reveal the contribution of genome-wide histone modifications in epigenetic regulation mechanisms. More new insights are expected to be offered to make us understand the potential role of histone modifications in stem cell and human diseases deeper by ChIP-Seq.

Cancer epigenomics

It was the first time to be proved by scientists that epigenetic changes could be involved in both oncogenes and tumour suppressors in 1980s, which laid the cornerstone for our present acknowledgment of epigenetic markers as diagnostics and therapeutic biomarkers for cancer [61,62]. In 1983, Andrew Feinberg and Bert Vogelstein purified DNA from several human primary tumour tissues by methylationsensitive restriction enzymes and found lowered DNA methylation of specific genes in contrast to DNA from adjacent normal tissues. At that time, the predominant theory of tumorgenesis was the activation of oncogenes. However, Feinberg and Vogelstein's findings implied that DNA methylation alteration could lead to oncogene activation [61]. Later in the 1980s, tumour suppressor genes were widely recognized, which made it encouraging when relevant epigenetic changes were discovered in those tumour suppressor genes. For example, Greger et al. [62] demonstrated that an unmethylated CpG island at the 5' end of the retinoblastoma gene turned hypermethylated in tumour tissures from retinoblastoma patients, and they had the right to speculate that methylation could directly silence tumour suppressor genes . Later studies correlated the methylation of tumour suppressor genes to their actual silencing role in cancer, and proved that tumour suppressor genes could be reactivated by inhibiting DNA methylation [63].

Epigenetic modifications: DNA methylation, as the most wellstudied mechanisms in cancer epigenomics, is only one of many aspects of demonstrating the role of epigenetic alterations on tumorigenesis. Cancer epigenomcis involves the researches of all sorts of epigenetic alterations in cancer DNA sequence (Figure 2). Next, we will summarize the current advances of the hotspots of cancer epigenomic researches, DNA methylation, histone modification, chromatin remodeling to demonstrate the contribution of epigenomics to tumorigenesis.

DNA methylation: Human disease is closely associated with abnormality in DNA methylation pattern. DNA methylation will generally inhibit gene expression. For example, global hypomethylation in cancer genome usually results in genomic instability, and gene silencing of tumour suppressor genes is caused by hypermethylation in CpG islands of the promoter region [14]. The methylated promoter regions may directly prevent transcription factors, e.g. A P-2, c-Myc, E2F and NF-kB, from combining with promoters, leading to gene silence or low gene expression; at the same time, the methylated regulatory elements at the 5' end of the genes may specifically bind to the methyl CpG binding protein (MBP), indirectly inhibiting the forming of transcriptional complex; besides, DNA methylation can alter the conformation of chromatin to inactivate it. Whereas, non-methylation usually correlates with gene activation, and demethylation should be related to reactivation of silencing genes [64]. Thus, Aberrant DNA methylation regulations would lead to tumorgenesis. DNA methylation changes in cancer cells include the loss of methylation at normally methylated sequences (hypomethylation) and the gain of methylated sequences at sites usually unmethylated (hypermethylation) [65].

As two opposite forms of DNA methylation, hypermethylation and hypomethylation play distinct roles in tumorigenesis. Hypermethylation of the promoter CpG islands regions in the 5' end of cancer related genes in human tumour cell lines have been reported, such as tumour suppressor gene (p16) [66], metastasis suppressor gene (Nm23) [67], DNA repair gene (MLH1) [68], angiogenesis suppressor gene [69] and so on. Some genes are hypermethylated in many types of cancers, such as p16 [66]. However, other genes are associated with specific cancer. For example, GSTP1 has been reported to be hypermethylated only



Figure 2: Epigenetic regulation network in cancer. Cancer epigenomics involves the researches of DNA methylation, histone modification, chromatin remodeling which mutually affect each other in the process of tumorigenesis. The results can be gene silencing or reactivation, directly leading to altered gene expression and abnormal protein.

Page 6 of 10

in prostate cancer [70]. While hypomethylation has been reported in almost every human malignancy and prefers the repetitive sequences, transposable elements and proto-oncogenes in cancer, some studies indicate that hypomenthylation in cells can increase the expression of certain genes, such as RAS, c-myc and so on. The overall decrease in the level of 5 methyl cytosine can be worse if the tumour has become more malignant [71].

In recent studies, increasing evidences have pointed out the important role of DNA methylation in tumorigenesis. For example, Ummanni et al. [72] previously reported significant downregulation of ubiquitin carboxyl-terminal hydrolase 1 (UCHL1) in prostate cancer, but now showed that the underlying mechanism of UCHL1 downregulation in PCa was linked with the promoter hypermethylation. Furthermore, it was suggested that UCHL1 downregulation via promoter hypermethylation played an important role in various molecular aspects of PCa biology, such as morphological diversification and regulation of proliferation. Then, other experimental results demonstrated that methylation status of DNMT1 could influence the activities of several important tumor suppressor genes in cervical tumorigenesis and may have the potential to act as an effective target for treatment of cervical cancer [73]. Besides solid tumours, the same results can also be found in hematological malignancies. Deneberg et al. [74] observed a negative impact of DNA methylation on transcription in acute myeloid leukemia (AML). Genes targeted by Polycomb group (PcG) proteins and genes associated with bivalent histone marks in stem cells showed increased aberrant methylation in AML (p<0.0001). Furthermore, high methylation levels of PcG target genes were independently associated with better progression free (OR 0.47, p=0.01) and overall survival (OR 0.36, p=0.001). It is expected that methylation-related factors in tumorigenesis will still be the hotspot of cancer epigenome research.

• Histone modification: Histones are subject to posttranslational modifications by enzymes primarily on their N-terminal tails, but also in their globular domains. Such posttranslational modifications include methylation, citrullination, acetylation, phosphorylation, sumoylation, ubiquitination, and ADP-ribosylation. Here, we will mainly focus on relatively widespread methylation and acetylation.

Histone acetylation is one of the most important modifications in cancer, which regulates the gene expression with reversibility. The histone acetyltransferases (HATs) acetylates conserved lysine amino acids on histone to improve the gene transcription (or the combination of transcriptional factors and regulatory elements). But, histone deacetylases (HDACs) removes acetyl groups from a ε -N-acetyl lysine amino acid on a histone to inhibit the gene transcription. As a major target for epigenetic therapy, HDACs are found overexpressed in different types of cancer. Actually, histone acetylation is essential to maintain the protein function and gene transcription. The imbalance of acetylation in cancer cells can change the structure of chromosomes and the level of gene expression, directly influencing the cell cycle, differentiation, apoptosis and tumorigenesis.

Recent advances in NGS enable genome-wide profiles of chromatin changes during tumorigenesis. Fraga et al. [75] have revealed a global loss of acetylated H4-lysine 16 (H4K16ac) and H4-lysine 20 trimethylation (H4K20me3) to lead to gene repression. Further, Wang et al. [76] used ChIP-seq method and found the fusion protein (AML1-ETO) generated by the t(8;21) translocation acetylated by the transcriptional coactivator p300 in leukemia cells isolated from t(8;21) AML patients, which followed by animal trails has indicates that lysine acetyltransferases represent a potential therapeutic target in AML. Lately, in order to investigate the epigenetic inactivation of the SFRP1 gene in Esophageal Squamous Cell Carcinoma (ESCC), Meng et al. [77] applied methylation-specific polymerase chain reaction (PCR), bisulfite sequencing, reverse-transcription (RT) PCR, immunohistochemistry and chromatin immunoprecipitation (ChIP) assay to detect SFRP1 promoter methylation, expression of the SFRP1 gene and histone modification in the SFRP1 promoter region. The SFRP1 promoter was found to be highly methylated in 95% (19/20) of the ESCC tissues and in nine ESCC cell lines. Furthermore, complete methylation of the SFRP1 gene promoter was correlated with its greatly reduced expression level.

In cancer cells, promoter CpG island hypermethylation is also associated with the combination of histone marks: deacetylation of histones H3 and H4, loss of histone H3 lysine K4 (H3K4) trimethylation, and gain of H3K9 methylation and H3K27 trimethylation [78,80]. H3K9 methylation and H3K27 trimethylation are also associated with aberrant gene silence in various types of cancer. By ChIP, Ballestar et al. [79] have found that the gene-specific profiles of Methyl-CpG binding proteins (MBDs) exist for hypermethylated promoters of breast cancer cells with a common pattern of histone modifications shared. It's interesting that Fujisawa et al. [81] found CpG sites in IL-13Rα2 promoter region were not methylated in all pancreatic cancer cell lines studied including IL-13Ra2-positive and IL-13Ra2-negative cell lines and normal cells. On the other hand, histones at IL-13R α 2 promoter region were highly acetylated in IL-13Ra2-positive but much less in receptor-negative pancreatic cancer cell lines. When cells were treated with HDAC inhibitors, not only histone acetylation but also IL-13Ra2 expression was dramatically enhanced in receptor-negative pancreatic cancer cells, which makes HDAC inhibitors new opportunity of target therapy.

In addition to methylation and acetylation, there are other kinds of modifications in histone, not so widely distributed as those mentioned above. However, all kinds of histone modifications are not separated but mutually linked in cancer cells. These histone modifications are integrated together to affect the histones of cancer cells. Consequently, the aberrant changes in the histone modifications will result in tumorigenesis.

Chromatin remodeling: Chromatin remodeling is the enzymedriven movement of nucleosomes, performed by chromatin remodeling complexes like SWI/SNF in human. Such can enable proteins such as transcription factors to bind to DNA wrapped around nucleosome cores. Genetic alterations of the genes involved in the chromatin remodeling process have been reported in many types of tumors recently [82,86]. For one study, the protein-coding exome has been sequenced in a series of primary clear cell renal carcinoma (ccRCC). Furthermore, it was reported that the SWI/SNF chromatin remodelling complex gene PBRM1 [4] was identified as a second major ccRCC cancer gene with truncating mutations in 41% (92/227) of cases. These data showed the marked contribution of aberrant chromatin biology [87]. For another study, the exomes of nine individuals with transitional cell carcinoma (TCC) have been sequenced. The study identified genetic aberrations of the chromatin remodeling genes (UTX, MLL-MLL3, CREBBP-EP300, NCOR1, ARID1A and CHD6) in 59% of our 97 subjects with TCC [82]. Dynamic chromatin remodeling is the base of

diverse biological processes, such as gene transcription, DNA replication and repair, chromosome separation and apoptosis. Together with these results, it is suggested that the aberrations of chromatin regulation might be a hallmark of cancer.

Aberrant chromatin remodeling may directly lead to the dysregulation of multiple downstream effector genes, consequently promoting the process of tumorigenesis [82]. For example, Nakazawa et al. [87] examined the histone H3 status in benign and malignant colorectal tumors by immunohistochemistry and western blotting, the results of which suggested that aberration of the global H3K9me2 level was an important epigenetic event in colorectal tumorigenesis and carcinogenesis involved with gene regulation in neoplastic cells through chromatin remodeling. Besides, different causes of chromatin remodeling may lead to different types of cancers. Much more researches should be carried on to determine the exact reasons and results.

Epigenetic marks as therapeutic targets: Epigenetic modifications are reversible, making them perfect therapeutic targets for cancer. Thus, cancer will be theoretically cured if the causal epigenetic aberrations are reversely corrected. According to this principle, many epigenetic drugs have been developed respectively corresponding to various epigenetic marks in recent decades. As hot epigenetic marks, DNA methylation and histone acetylation are extensively studied to successfully act as therapeutic targets.

First, the hypermethylation in CpG islands is commonly found in many types of tumours. DNA methylation inhibitor is the first one that is supposed to be available for cancer therapeutics. The remarkable discovery has been found that treatment with cytotoxic agents, 5-azacytidine (5-aza-CR) and 5-aza-2'-deoxycytidine (5-aza-CdR) would lead to the inhibition of DNA methylation that induces gene expression and causes differentiation in cultured cells [88]. 5-Aza-CR (azacitidine) and 5-aza-CdR (decitabine) have been approved by FDA for use in the treatment of myelodys-plastic syndromes, and promising results have also emerged from the treatment of hematological malignancies [89] or solid tumors [90]. There are some other possible DNA methylation inhibitors such as zebularine, which is orally administered and currently under investigation in many types of cancers. However, the demethylation drug have serious side effect of toxicity, which leaves a problem that seeks proper agents to act synergistically with the drugs. Luckily, clinical studies by Silverman et al. [91], Issa et al. [92] and other researchers generated a notable paradigm of oncology: therapeutic efficacy could be achieved at low drug doses. Such reduced doses were adopted in a large trial in patients with myelodisplastic syndrome (MDS) that would lead to leukaemia. It was revealed that the conversion time from MDS to frank leukaemia increased, as well as overall survival [93]. Now, two inhibitors-azacitidine (Vidaza; Celgene) and decitabine (Dacogen; Eisai)--have been approval by the FDA for MDS, and this improves the use of lowdose regimens not only for leukaemia, but also for solid tumours [94].

Second, reversing histone acetylation patterns back to normal through treatment with HDAC inhibitors have been proved to have antitumorigenic effects, including growth arrest, apoptosis and the induction of differentiation [95]. The antiproliferative effects of HDAC inhibitors are mediated by their ability to reactivate silenced tumor suppressor genes [96]. Suberoylanilide hydroxamic acid (SAHA), as an HDAC inhibitor, has been approved for clinical use as treatment of T cell cutaneous lymphoma and has gained the approval of FDA as vorinostat (Zolinza; Merck) [97]. Besides, romidepsin (Istodax; Celgene)

with the same remarkable efficacy in cutaneous T cell lymphoma has also been approved by FDA [98]. Although they are well tolerated with little toxicity, HDAC inhibitors as drugs have some side effects, including constitutional and gastrointestinal toxicity, cardiac trouble, myelosuppresion and others. However, the molecular mechanisms for drug response in these patients have not been determined yet. Several other HDAC inhibitors such as depsipeptide and phenylbutyrate are also under clinical trials [99].

Challenge and future of epigenome research

Major challenges: Benefit from the advent of NGS technologies, epigenome research has rapidly expanded in recent years. As described above, advances have been achieved in recent years. However, there are still two major challenges in epigenome research, respectively referring to sampling and integrated analysis of various epigenetic modifications [10]. Next, the review will discuss the two aspects in detail.

Epigenome research is expected to interpret the effect of epigenetic modifications caused by environmental factors. Thus, most epigenetic modifications are somatic and tissue or stage specific. Due to the dynamics of epigenetics, sampling is the first and critical step of epigenome research. To a large extent, mistakes in sample tissue selection will lead to the aborted and incorrect conclusion. For epigenome research of human disease, cancer is studied more intensively than other human diseases. That is attributed to the easier accessibility of cancer tissues after biopsy or surgery. However, as the obvious characteristic of cancer, tissue heterogeneity is still a problem in sampling for epigenome research. Many complex diseases, such as hypertension, don't exhibit tissue-specific pathogenesis. DNA samples from any tissues do not show significant difference. Thus, based on our current unclear understanding of pathogenesis, it is difficult to conduct epigenome research very well. Second, since the epigenome research of human disease is in the early stage, the study model is still robust and the exact sample size is also unknown. Third, due to tissue specificity, many types of tissues need to be collected to demonstrate the complete picture of epigenome. In general, the challenge of sampling arises from specific tissue selection, exact sample size and multiple tissue collection.

There are various types of epigenetic modifications, not limited to those described above in this review. First, it is necessary to explore every type of epigenetic modifications in the human genome. It is possible that most of them still remain to be found in future. Second, even if all epigenetic modifications have been revealed until now, there is still a long way for researchers to move. That is due to the network pattern of epigenetic regulations. Individual epigenetic modification does not work separately, but mutually to regulate gene expression of the whole genome. It is a large-scale project to clearly understand the subtle system of integrated regulations by epigenetic modifications.

Future direction: A decade ago, the human genome project (HGP) has been accomplished by collaborations of worldwide scientists. The constructed human genome map is a milestone for genome research in the history, providing a strong foundation for the following countless sequencing researches. Similarly, human epigenome map is essential to be constructed to promote the field of epigenome research. This large-scale scientific project can only be achieved by the way of HGP. Worldwide scientists must join in a global organization for collaborations to achieve this significant goal. Fortunately, many consortiums have been founded in recent years (Table 2). The human epigenome map is expected to be constructed in the near future.

Table 2: Worldwide epigenome consortiums

Page 8 of 10

Name	Website	Launched time	Main goals	
CIHR	http://www.cihr-irsc.gc.ca/	2000	 Transform health research in Canada by applying next-generation sequencing to more research on targeted priority and under-developed areas such as population health and health services research Transform research results into policies, practices, procedures, products and services 	
ENCODE	http://www.genome.gov/10005107	2003	 Funded by the National Human Genome Research Institute (NHGRI) Build a comprehensive parts list of functional elements in the human genome, including elements that act at the protein and RNA levels, and regulatory elements that control cells and circumstances in which a gene is active. 	
NOE	www.epigenome-noe.net/	2004	 Provide clear and visible benefits for the entire epigenetic community and as such supports both members and non-members with conferences, work-shops, training visits and shared resources. 	
EPITRON	www.epitron.eu	2005	Define the epigenetic treatment and identify novel drugs of cancer	
HEROIC	http://www.heroic-ip.eu/	2006	 Apply high-throughput sequencing technology to do a genome-wide epigenetic research, making a wider contribution to understand the primary genetic code of chromatin 	
BLUEPRINT	http://www.blueprint-epigenome. eu/	2007	 Develop novel technology for high-throughput epigenome mapping Provide reference epigenomes and epigenetic drug target 	
CANCERDIP	www.cancerdip.eu	2007	 Understand the mechanisms of DNA methylation deregulation in human cancer Identify epigenetic markers for clinical application 	
AEPIA	www.epialliance.org.au	2008	 Facilitate communication between Australasian research groups interested in epigenetic- based questions as well as clinicians, students and members of the public interested in epigenetics 	
Roadmap	www.roadmapepigenomics.org	2008	 Utilize next-generation sequencing technologies to map DNA methylation, histone modifications, chromatin accessibility and small RNA transcripts in stem cells and primary ex vivo tissues Produce a public resource of human epigenomic data to catalyze basic biology and disease-oriented research. 	
ICGC	www.icgc.org	2010	 Obtain a comprehensive description of genomic, transcriptomic and epigenomic changes in 50 different tumor types and/or subtypes which are of clinical and societal importance across the globe. 	

The consortium names are shortly listed in the table. The full names are represented as follows. Canadian Institutes of Health Research (CIHR), Encyclopedia of DNA Elements (ENCODE), Epigenome Network of Excellence (NOE), Epigenetic Treatment of Neoplastic Disease (EPITRON), High-throughput Epigenetic Regulatory Organisation in Chromatin (HEROIC), Medip Cancer Consortium(CANCERDIP), Australian Epigenetic Alliance (AEPIA), NIH Roadmap Epigenomics Mapping Consortium(Roadmap), International Cancer Genome Consortium (ICGC)

However, both genome and epigenome are desired to explain the mechanisms of complex life activities from the view of DNA level. Although the recent achievements can illustrate many phenomenons that were inexplainable in the past, more unsolved problems still remain to be explored. According to the central dogma, life is a systematic network with multidimensional activities. The activities on DNA level would interact with those in RNA and protein level. Thus, the researches on DNA level are obviously not enough. With various types of NGS technologies, it is possible to apply NGS in DNA, RNA and protein levels. The information in these levels is expected to be explored by NGS and integrated by bioinformatics to together reveal more discoveries in biology and human disease.

The rapid progress of sequencing technology has also contributed to the development of epigenome research. Third generation sequencing (TGS) technologies are expected to be commercial in recent several years. Compared to NGS, TGS exhibits many technical breakthroughs, such as small amount of samples, faster speed, less time, single cell sequencing and so on. These characteristics make TGS feasible to reveal unknown epigenetic mechanisms and speed up the epigenome research. The ability of single cell sequencing can largely solve the obtacle of tissue specificity in epigenome research. Combined with large-scale collaborations and latest sequencing technology, it is believed that epigenome research will contribute to explain one aspect of the complexity of nature and improve human health.

References

- 1. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921.
- 2. International HapMap Consortium (2003) The International HapMap Project. Nature 426: 789-796.
- International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437: 1299-1320.
- 4. The 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467: 1061-1073.
- Deal RB, Henikoff S (2010) Capturing the dynamic epigenome. Genome Biol 11: 218-225.
- Faulk C, Dolinoy DC (2011) Timing is everything: the when and how of environmentally induced changes in the epigenome of animals. Epigenetics 6: 791-797.
- Zhang TY, Meaney MJ (2010) Epigenetics and the environmental regulation of the genome and its function. Annu Rev Psychol 61:439-466.
- 8. Nanney DL (1958) Epigenetic control systems. PNAS 44: 712–717.
- 9. Lund AH, van Lohuizen M (2004) Epigenetics and cancer. Genes Dev 18: 2315-2335.
- Ku CS, Naidoo N, Wu M, Soong R (2011) Studying the epigenome using next generation sequencing. J Med Genet 48: 721-730.
- Portela A, Esteller M (2010) Epigenetic modifications and human disease. Nat Biotechnol 28: 1057-1068.

Citation: Li J, Wang T, Zhang X, Yang X (2011) The Contribution of Next Generation Sequencing Technologies to Epigenome Research of Stem Cell and Tumorigenesis . Human Genet Embryol S2:001. doi:10.4172/2161-0436.S2-001

Page 9 of 10

- Urdinguio RG, Sanchez-Mut JV, Esteller M (2009) Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. Lancet Neurol 8: 1056-1072.
- Ballestar E (2011) Epigenetic alterations in autoimmune rheumatic diseases. Nat Rev Rheumatol 7: 263-271.
- 14. Esteller M (2008) Epigenetics in cancer. N Engl J Med 358: 1148-1159.
- Hammoud S, Douglas TC (2011) The emerging role of the sperm epigenome and its potential role in development. Biennial Review of Infertility 358: 181-194.
- 16. Carrell DT, Hammoud SS (2010) The human sperm epigenome and its potential role in embryonic development. Mol Hum Reprod 6: 37-47.
- 17. Chan D, Jacquetta Trasler (2011) The Sperm Epigenome. Sperm Chromatin 1: 95-106.
- Mardis ER (2008) The impact of next-generation sequencing technology on genetics. Trends Genet 24: 133-141.
- Morozova O, Marra MA (2008) Applications of 'next generation sequencing technologies in functional genomics. Genomics 92: 255-256.
- Hirst M, Marra MA (2010) Next generation sequencing based approaches to epigenomics. Brief Funct Genomics 9: 455-465.
- 21. Satterlee JS, Schubeler D, Ng HH (2010) Tackling the epigenome: challenges and opportunities for collaboration. Nat Biotechnol 28: 1039-1044.
- Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, et al. (2010) Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell Stem Cell 6: 479-491.
- 23. Laird PW (2010) Principles and challenges of genome-wide DNA methylation analysis. Nat Rev Genet 11: 191-203.
- 24. Senner CE (2011) The role of DNA methylation in mammalian development. Reprod Biomed Online 22: 529-535.
- Liang P, Song F, Ghosh S, Morien E, Qin M, et al. (2011) Genome-wide survey reveals dynamic widespread tissue-specific changes. BMC Genomics 12: 231.
- 26. Li N, Ye M, Li Y, Yan Z, Butcher LM, et al. (2010) Whole genome DNA methylation analysis based on high throughput sequencing technology. Methods 52: 203-212.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, et al. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462: 315-322.
- 28. Li Y, Zhu J, Tian G, Li N, Li Q, et al. (2010)The DNA methylome of human peripheral blood mononuclear cells. PLoS Biol 8: e1000533.
- Laurent L, Wong E, Li G, Huynh T, Tsirigos A, et al. (2010) Dynamic changes in the human methylome during differentiation. Genome Res 20: 320–331.
- Kim K, Doi A, Wen B, Ng K, Zhao R, et al. (2010) Epigenetic memory in induced pluripotent stem cells. Nature 467: 2852-2890.
- Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, et al. (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471: 68-73.
- 32. Wang J, Wang W, Li R, Li Y, Tian G, et al. (2008) The diploid genome sequence of an Asian individual. Nature 456: 60-65.
- Bock C, Tomazou EM, Brinkman AB, Müller F, Simmer F, et al. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 28: 1106-1114.
- Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, et al. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 102: 10604-10609.
- Ballestar E (2010) Epigenetics lessons from twins: prospects for autoimmune disease. Clin Rev Allergy Immunol 39: 30-41.
- 36. Javierre BM, Fernandez AF, Richter J, Al-Shahrour F, Martin-Subero JI, et al. (2010) Changes in the pattern of DNA methylation associate with twin discordance in systemic lupus erythematosus. Genome Research 20: 170-179.
- Baranzini SE, Mudge J, van Velkinburgh JC, Khankhanian P, Khrebtukova I, et al. (2010) Genome, epigenome and RNA sequences of monozygotic twins discordant for multiple sclerosis. Nature 464: 1351-1356.

- Wu SC, Zhang Y (2010) Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol 11: 607-620.
- Chen ZX, Riggs AD (2011) DNA methylation and demethylation in mammals. J Biol Chem 286: 18347-18353.
- 40. Chavez L, Jozefczuk J, Grimm C, Dietrich J, Timmermann B, et al. (2010) Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. Genome Res 20: 1441-1450.
- Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, et al. (2011) Genomewide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473: 394-397.
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala HM, et al. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324: 930–935.
- Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, et al. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 466: 1129-1133.
- 44. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, et al. (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell 8: 200-213.
- Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324: 929-930.
- 46. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, et al. (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473: 398-402.
- Smith ZD, Gu H, Bock C, Gnirke A, Meissner A (2009) High-throughput bisulfite sequencing in mammalian genomes. Methods 48: 226-232.
- 48. Gu H, Bock C, Mikkelsen TS, Jäger N, Smith ZD, et al. (2010) Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. Nature Methods 7: 133-136.
- 49. Gu H, Smith ZD, Bock C, Boyle P, Gnirke A (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat Protoc 6: 468-481.
- Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, et al. (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. Nat Biotechnol 28: 1097-1105.
- 51. Wang L, Sun J, Wu H, Liu S, Wang J, et al. (2011) Systematic assessment of reduced representation bisulfite sequencing to human blood samples: A promising method for large-sample-scale epigenomic studies. J Biotechnol.
- 52. Gertz J, Varley KE, Reddy TE, Bowling KM, Pauli F, et al. (2011) Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. PLoS Genet 7: e1002228.
- Bártová E, Krejcí J, Harnicarová A, Galiová G, Kozubek S (2008) Histone modifications and nuclear architecture: a review. J Histochem Cytochem 56: 711-721.
- 54. Biancotto C, Frigè G, Minucci S (2010) Histone modification therapy of cancer. Adv Genet 70: 341-386.
- 55. Jenkins TG, Carrell DT (2011) The paternal epigenome and embryogenesis: poising mechanisms for development. Asian J Androl 13: 76-80.
- 56. Imai K, Ochiai K (2011) Role of histone modification on transcriptional regulation and HIV-1 gene expression: possible mechanisms of periodontal diseases in AIDS progression. J Oral Sci 53: 1-13.
- 57. Park PJ (2009) ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet 10: 669-680.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, et al. (2007) Highresolution profiling of histone methylations in the human genome. Cell 129: 823-837.
- Terrenoire E, McRonald F, Halsall JA, Page P, Illingworth RS, et al. (2010) Immunostaining of modified histones defines high-level features of the human metaphase epigenome. Genome Biol 11: R110.
- Larson JL, Yuan GC (2010) Epigenetic domains found in mouse embryonic stem cells via a hidden Markov model. BMC Bioinformatics 11: 557.

 Feinberg, A. P., Vogelstein, B (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature 301: 89-92.

62. Greger V, Passarge E, Hopping W, Messmer E, Horsthemke B (1989) Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum Genet 83: 155-158.

- Laird PW, Jackson-Grusby L, Fazeli A, Dickinson SL, Jung WE, et al. (1995) Suppression of intestinal neoplasia by DNA hypomethylation. Cell 81: 197-205.
- Fuks F (2005) DNA methylation and histone modifications:teaming up to silence gene. Curt Opin Genet Dev 15: 490-495.
- 65. Plass C (2002) Cancer epigenomics. Hum Mol Genet 11: 2479-2488.
- Darbro BW, Lee KM, Nguyen NK, Domann FE, Klingelhutz AJ (2006) Methylation of the p16 (INK4a) promoter region intelomerase immortalized human keratinocytes co-cultured with feeder cells Oncogene 25: 742I-7433.
- 67. Hartsough MT, Clare SE, Mair M, Elkahloun AG, Sgroi D, et al. (2001) Elevation of breastcarcinoma Nm23-H1 metastasis suppressor gene expression and reduced motility by DNA methylation inhibition. Cancer Res 61: 2320-2327.
- Hitchins MP, Ward RL (2007) Erasure of MLH1 methylation in spermatozoa implications for epigenetic inheritance. Nat Genet 39: 1289.
- Margetts CD, Astuti D, Gentle DC, Cooper WN, Cascon A, et al. (2005) Epigenetic analysis of HIC1, CASP8, FLIP, TSP1, DCR1, DCR2, DR4, DR5, KvDMR1, H19 and preferential 11p15.5 maternal-allele loss in von Hippel-Lindau and sporadic phaeochromocytomas. Endocr Relat Cancer 12: 161-172.
- 70. Meiers I, Shanks JH, Bostwick DG (2007) Glutathione S-transferase pi (GSTP1) hypermethylation in prostate cancer: review 2007. Pathology 39: 299-304.
- Wilson AS, Power BE, Molloy PL (2007) DNA hypomethylation and human diseases. Biochim Biophys Acta 1775: 138-162.
- Ummanni R, Jost E, Braig M, Lohmann F, Mundt F, et al. (2011) Ubiquitin Carboxyl-Terminal Hydrolase 1 (UCHL1) is a Potential Tumour Suppressor in Prostate Cancer and is frequently Silenced by Promoter Methylation. Mol Cancer 10: 129.
- Zhang Y, Chen FQ, Sun YH, Zhou SY, Li TY, et al. (2011) Effects of DNMT1 silencing on malignant phenotype and methylated gene expression in cervical cancer cells. J Exp Clin Cancer Res 30: 98.
- 74. Deneberg S, Guardiola P, Lennartsson A, Qu Y, Gaidzik V, et al. (2011) Prognostic DNA methylation patterns in cytogenetically normal acute myeloid leukemia are predefined by stem cell chromatin marks. Blood 118: 5573-5582.
- 75. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, et al. (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet 37: 391–400.
- Wang L, Gural A, Sun XJ, Zhao X, Perna F, et al. (2011) The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. Science 333: 765-769.
- Meng Y, Wang QG, Wang JX, Zhu ST, Jiao Y, et al. (2011) Epigenetic Inactivation of the SFRP1 Gene in Esophageal Squamous Cell Carcinoma. Dig Dis Sci 56: 3195-3203.
- Fahrner JA, Eguchi S, Herman JG, Baylin SB (2002) Dependence of histone modifications and gene expression on DNA hypermethylation in cancer. Cancer Res 62: 7213-7218.
- Ballestar E, Paz MF, Valle L, Wei S, Fraga MF, et al. (2003) Methyl-CpG binding proteins identify novel sites of epigenetic inactivation in human cancer. EMBO J 22: 6335-6345.
- Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, et al. (2006) The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439: 871-874.
- Fujisawa T, Joshi BH, Puri RK (2011) Histone modification enhances the effectiveness of IL-13 receptor targeted immunotoxin in murine models of human pancreatic cancer. J Transl Med 8: 9:37.
- 82. Gui Y, Guo G, Huang Y, Hu X, Tang A, et al. (2011) Frequent mutations of

This article was originally published in a special issue, **Epigenetics**, **stem cells and tumorigenicity** handled by Editor(s). Dr. Yue Zhang, Harvard Medical School, USA; Yujing Li, Emory University School of Medicine, USA; Yanhong Ji, Xian Jiaotong University, China chromatin remodeling genes in transitional cell carcinoma of the bladder. Nat Genet 43: 875-878.

Page 10 of 10

- van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, et al. (2009) Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nat Genet 41: 521-523.
- Wiegand KC, Shah SP, Al-Agha OM, Zhao Y, Tse K, et al. (2010) ARID1A mutations in endometriosis-associated ovarian carcinomas. N Engl J Med 363: 1532-1543.
- Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, et al. (2010) Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature 463: 360-363.
- Varela I, Tarpey P, Raine K, Huang D, Ong CK, et al. (2011) Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature 469: 539-542.
- Nakazawa T, Kondo T, Ma D, Niu D, Mochizuki K, et al. (2011) Global histone modification of histone H3 in colorectal cancer and its precursor lesions. Hum Pathol.
- Constantinides PG, Jones PA, Gevers W (1977) Functional striated muscle cells from non-myoblast precursors following 5-azacytidine treatment. Nature 267: 364-366.
- Lubbert, M. (2000) DNA methylation inhibitors in the treatment of leukemias, myelodysplastic syndromes and hemoglobinopathies: clinical results and possible mechanisms of action. Curr Top Microbiol Immunol 249: 135-164.
- Momparler RL, Eliopoulos N, Ayoub J (2000) Evaluation of an inhibitor of DNA methylation, 5-aza-20 -deoxycytidine, for the treatment of lung cancer and the future role of gene therapy. Adv Exp Med Biol 465: 433-446.
- Silverman LR, Mufti GJ. (2005) Methylation inhibitor therapy in the treatment of myelodysplastic syndrome. Nat Clin Pract Oncol 2 Suppl 1: S12-23.
- 92. Issa JP, Kantarjian H (2005) Azacitidine. Nature Rev Drug Discov Suppl: S6-7.
- Fenaux P, Mufti GJ, Hellstrom-Lindberg E, Santini V, Finelli C, et al. (2009) Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. Lancet Oncol 10: 223–232.
- Baylin SB, Jones PA. (2011) A decade of exploring the cancer epigenome biological and translational implications. Nat Rev Cancer 11: 726-734.
- Sharma S, Kelly TK, Jones PA (2009) Epigenetics in cancer. Carcinogenesis 31: 27-36.
- Carew JS, Giles FJ, Nawrocki ST (2008) Histone deacetylase inhibitors: mechanisms of cell death and promise in combination cancer therapy. Cancer Lett 269: 7–17.
- Duvic M, Talpur R, Ni X, Zhang C, Hazarika P, et al. (2007) Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T -cell lymphoma (CTCL). Blood 109: 31–39.
- Olsen EA, Kim YH, Kuzel TM, Pacheco TR, Foss FM, et al. (2007) Phase Ilb multicenter trial of vorinostat in patients with persistent, progressive, or treatment refractory cutaneous T -cell lymphoma. J Clin Oncol 25: 3109–3115.
- Cortez CC, Jones PA (2008) Chromatin, cancer and drug therapies. Mut Res 647: 44-51.