Multiple Mechanisms of DNA Methylation in Cancer Initiation and Development

Ziwei Wang
Jiangxi Medical College, Nanchang University, Nanchang, Jiangxi Province, 330031, China

* Corresponding author: angela@cas-harbour.org

Abstract. Epigenetic refers to DNA methylation, histone modification and histone variants, to change expression level of heritable molecular determinants without the changing DNA sequence. With the development of the sequencing and protein identification technology, a large number of studies from International Cancer Genome Consortium (ICGC) illustrates Epigenetic alterations play an important part in cancer initiation and metastasis. The proteomic and genomic studies find out the potential connection and mechanism between epigenetics and cancers. Nowadays, the epigenetic inhibitors are still the main therapies in clinical treatments. This article aims to find out the relationship between DNA methylation and cancer, and create a new version, where epigenetic alterations combines cancer cells to open a new gate of organ transplantation. In conclusion, using DNA methylation to transform the cancer cells into specific cells to build a new organ may solve the most common problem, cell pairings, in organ transplantation.

1. Introduction
Human cells are eukaryotic cells and contain very long DNA molecules packaged into a set of structure, chromosome, together with histone. The most important function of chromosome is to carry the entire genetic information and express the genome. Nucleosome is a basic unit of Chromosome in eukaryotes, consisting of a segment of DNA wound in sequence around the 8 histone proteins cores. During the past 17 years, sequencing of genome has been continued till 21st Dec 2017 and our genome sequencing result was released as GRCh38.p12. Each individual contains 21,000 protein coding genes, accounting for only 1%/ of the whole genome. Through studying neurosome coordinated replication, a previous study revealed that all the subunits of nucleosome can be modified to damage the organization and function of the chromatin[1].

Epigenetics, a series of heritable changes of gene expression, can influence gene transcription, genomic imprinting, genomic stability and X-chromosome inactivation[2]. There are 3 major layers of epigenetics: DNA methylation, histone modification, and histone variation with nucleosome positioning. Epigenetics alteration is dynamic and highly regulated by specific enzyme. All these modifications change the chromatin structure or interaction between the nucleosomes. However, these modifications are not 100% accurate. These genetic and epigenetic modifications alter the function of key genes which regulate fundamental process of cells, such as cell growth, survival and senescence. As a consequence, tumor cells are raised, uncontrolled and progressive.

This article focuses on the relationship between DNA methylation and cancer.
2. Methods to detect the epigenetics
With the development of next generation sequencing and tandem mass spectrometry, the genetic and proteomic information can be appreciated accurately. For DNA methylation sites only, because the majority of DNA methylation happens on CpG island, with the Bisulfite treatment, the methylated cytosine will still be cytosine while the unmethylated cytosine will become thymine. The combination of Bisulfite treatment and deep sequencing can detect the methylated rate of the CpG. The methylated rate of the CpG can indirectly indicates the methylation level. However, this method is limited in detecting DNA methylation.

Early in 1869, the DNA extraction was established by the chromatin immunoprecipitation (CHIP). In 2006, researchers established the next sequencing, which provides the transcriptome information. Through this process, the post-transcription modification and the mechanism of the non-coding RNA regulation can be known. Combination of CHIP and the deep sequencing tells the nucleosome positioning, transcription factor binding sites, the localization of histone and the DNA modification[3-7].

Tandem mass-spectrometry can not only be used to separate and identify the peptides and proteins but also to work out peptide sequences and quantify them with the help of labeling technology. By mapping the sample peptides to the transcriptome, whether the genome expression level is changed can be found out[3,8]. Through this method, the underlying mechanism of different modifications can be better understood. However, there are crosstalk between modifications at the same residues and different residues. Also the modifications are dynamic changes during development and/or differentiation. Therefore a high-resolution in vivo-genomic approach to detailing the dynamic changes events on any given nucleosome during gene expression was discovered.

3. DNA methylation related to the cancer
DNA methylation is catalyzed by a family of DNA methyltransferase (DNMT) which transfer a methyl group from S-adenyl methionine to a fifth carbon of a cytosine residues to form 5mC. And there are 4 types of DNA methylation and all of them can impact the genome stability, transposable elements and exposure. It can impact gene transcription in either positive or negative ways. A study done in 2015 focused on the DNA methylation statistics in mouse embryo stem cells. This study used Bisulfite treatment, with which the methylated cytosine would still be cytosine, thinking that the unmethylated cytosine would become thymine. Then researchers found that the majority of DNA methylation happened on GpG islands. And 78% of GpGs are methylated in mouse embryo stem cells[9]. Take all the methylations into account, around 70% happens within the mammalian promoters[10]. CpG islands methylation is strongly related to the transcription, replication and it is demethylated or hypermethylated during malignant tumor cells transformation.

Fig. 1. The methylation statistics in mouse embryonic stem cells. Through this figure, all C is accounting for 4.7% and the CWG is accounting for 1.2%. The majority of the DNA methylations occur for GpC islands[9].
The DNA hypermethylation of the promoters not only affects the gene expression but also affects the uncoding RNAs. Some of them play important roles in tumor cell proliferation and apoptosis control. Through those previous studies, the typical DNA methylation participate at CpG island can be summarized. The CpG island is part of the DNA sequence, larger than 200 bps, where the C+G content is greater than 50% and a ratio of observed to expected CpG is greater than 60%. Therefore, DNA methylation can be predicted through CpG density. CpG density refers to the number of the CpG sites within a window of 300bps upstream and downstream of the CpG of interest or the high risk tumor transferred gene sequences. If the CpG density is more than 30, the DNA sequence is hypomethylated. Contrastly, if the CpG density is less than 30, the DNA sequence is hypermethylated. Through this conclusion, the methylation level can be predicted.

There are three enzymes involved in the DNA methylation. Dnmt 1, de novo and Dnmt 3a/3b. DNMT 1 is a maintained methyl-transferase that recognise the DNA strands that one strand DNA is methylated while the other is not. This kind of strands is produced by DNA replication. Then the DNMT 1 will methylate the newly formed DNA strands[11]. Although DNMT 3a and DNMT3b are also involved in methylation hemi-methylation strand, their function is to establish DNA methylation during embrogenesis[12]. However, DMNT 1 and DMNT 3a/3b work complementary and simultaneously. MBD proteins, the reactors of methyl group, are also play an important role in DNA methylation. MBD1-4 and Mecp2 are reading methylcytosine for the recruitment of HDAC complexes to silent genes for repression. And the HDACs binds to active genes[13].
Table 1. The overview amond dietary influences of the methyl-CpG-Binding [14].

| Proteins | Transcripts | Protein-coding Isoforms (RefSeq) | Main binding | Main function | Main interaction |
|----------|-------------|---------------------------------|--------------|---------------|-----------------|
| MeCPa    | 21          | 2                               | Bind to a single mCpG | Neuro-development | HDAC            |
| MBD1     | 27          | 13                              | Methylated and unmethylated CpG | Cell division and differentiation | HDAC            |
| MBD2     | 5           | 2                               | Bind to highly methylated CpG | Stem cell pluripotency; gene expression repression | TSS, gene body NuRD complex |
| MBD3     | 13          | 2                               | Bind to 5-hydroxy-mCpG | Normal development and cell reprogramming | NuRD complex DNA damage repair, chromatin assemble |
| MBD4     | 10          | 5                               | Bind to 5mCpG/TpG Mismatch | Identification and excision of deaminated CpG | DNMT1 |
| MBD5     | 33          | 9                               | Bind to chromocenters and methylated pericentric heterochromatin | Development, neurogenesis, and neuronal gene regulation | Epigenetic machinery |
| MBD6     | 14          | 9                               | Bind to chromatin | Cell differentiation and proliferation | MBD2,MBD6 repress complex |

MBD2 was found in a large proportion of methylated regions and was associated with transcriptional silencing. A redistribution of MBD2 on methylated DNA occurred during oncogenic transformation, local DNA methylations transform frequently and independently[15,16]. Mutation in methyltransferase or reducers can not only be strongly related to developmental disorder, but also involved in epigenetic regulator mutation to cause cancer. For example, DNMT 3a has a chromatin-reader to modify the pump domain. This domain helps localize the enzyme to chromatin. And some cancer mutations affect this regulatory domain. The mutations like frameshift, non-sense, site mutation split and translocation can lead to AML, myeloproliferative disease(MD) and myelodysplastic syndrome (MDS). A study done in 2010 sequences the cancer genome. 25% of the patients with acute myeloid leukemia(AML) have DNMT 3A mutation. However, the underlying mechanism is still unclear. DNA methylation, a relatively stable chromatin modification, has a dynamic nature. By mapping the modification in pluripotent cells, differentiated cells has confirmed that an enzyme-activity within mammalian cells can alter this modification[10]. The ten-eleven translocation family of proteins is related to 5 methyl cytosine and its derivations. 5mC can be chemical modified at 2 sites in the amino group and the methyl group. The methyl group of 5mC can be modified by the additional hydroxy group catalyzed by TET enzymes to get 5-hydroxymethyl-cytosine(5hmC), which can also be chemical modified at 2 sites, the amino group and the hydroxymethyl group. The TET enzyme further oxidize 5hmC to form 5-formyl-cytosine(5fC) and then 5-carboxy cytosine(5aC). Eventually, the products of each pathways are recognized and cleaved off to replace with a naked cytosine mediated by TDG and/or SMUG1, both components of the base excision repair pathway. And these derivation are involved in transcriptional regulation. First, they are essential intermediates in the process of both active and passive DNA demethylation. Second, they can obstruct or enhance the MBD protein binding to chromatin. Third, 5hmC, a bivalent marker in promoters gene bodies or transcriptionally poised genes. So both TET family and 5hmC both have important role in both transcriptional activation and silencing. There are several examples that mutation in TET family are related to the cancer. The chromatin translocation, 4(10:11)(q22:Q23), can lead to the AML for inducing TET1 gene into the MLL gene. And the missence, nonsense, and frameshift of the TET2 are related to the AML, MPD and chromatin myelomonocytic leukemia(CMMCL). Moreover, patients caused by these mutation all have poor prognosis. TET2
mutations are mostly loss-of-function mutations, resulting in decreasing 5hmC levels and increase in 5mC levels in oncogenetic malignant cells to prevent them from apoptosis and cell cycle regulation[17]. Through this, the cell can be transformed to gain the immortal and enhanced self-renewal proteins.

![Fig. 4. Cancer mutation related to the epigenetic enzyme[3].](image)

In the past decades, DNA methylation is considered as a relatively stable chromatin modification. But in 2017, researchers used the high-resolution genome wide mapping of DNA methylation in pluripotent cells and differentiated cells. They found the extensive transcriptomic and epigenomic remodeling occur during Arabidopsis thaliana germination; and dynamic DNA methylation reconfiguration happened during seed development and germination. So the dynamic nature of DNA methylation can be confirmed, evidently signifying the existence of an enzymatic activity within mammalian cells that either erases or alters this chromatin modification. According to all the studies, the DNA methylation has a significant effect in tumor initiate and development. In other words, DNA methylation can indirectly contribute to the cells fate. So as a controller of the cell, the most of researches are inclined to tumor suppression or tumor cell apoptosis-inducing. However, DNA methylation can be controlled in lab to have specific alteration, which can help researchers to control the cell expression.

4. Discussion
In recent years, DNA methylation is used as a switch. However, based on clinical studies, the epigenetic alterations can only target a small group of genes due to the different epigenetics regulators in distinct tissues. In clinical therapy, a specific regulator can be targeted to enhance or inhibit one epigenetic change. Unfortunately, the treasure map of the mechanism involved in epigenetic regulators is not fully extended. How could these regulator recognize and target the specific tissues? Are these regulators the only way to cause the transformation or stop the transformation? Therefore, due to these unsolved questions, the epigenetic inhibitors or the enhancers can hardly be used in clinical cases at present days.

![Fig. 5. Epigenetic inhibitors in cancer treatment](image)
This figure introduces the development of epigenetic inhibitors and the current state in cancer therapies. First, the specific small-molecules are added in the cancer cell line in vitro. This could stop the proliferation and introduce the apoptosis. Second, this method is always coupled to the proteomic and genomic study to find out the potential mechanism. Third, the inhibitors are used in the vivo preclinical, the animal models to confirm whether they can bring the benefit as in vitro environment. Also this step can tell the pharmacokinetic properties and the side effects. Based on these, the inhibitors can be used in clinical studies. In recent years, DNMT and HDAC are extensively approved in clinical area[3].

What if DNA methylation is used in laboratory to alter the functions of cells? For years, researchers have been working out different methods to raise cells in laboratory environment. Many of those methods concentrated on the objective factors but seldom of them focused on the cells themselves. What if the researchers make use of DNA methylations to alter the metabolism of the cells in order to make them suitable for vivo environment? Researchers can easily raise enough amount of cells for study. Then the DNA can be altered back to the original status for the research. At the same time, this technique can be used in tissue reparation. Nowadays, clinical doctors use stem cells extract from fats and raise them in the lab. Once they reach the certain amount, those stem cells are injected back to the body and take the responsibility of tissue reparation. However, this method has an obviously malpractice: since the stem cells are extracted from the fat, they are not going to survive in other environment for a long time to finish the reparation. As a result, the treatment involves multiple stem cell injection. But if researchers can extract the cell from the certain tissue and alter its DNA and inject them back to the tissue for reparation, will it serve for a longer time to finish the reparation? Since every somatic cell contains the whole genome of human, the potential of these cells can be activated with some external stimuli. There is a possibility to raise the specific cells in lab with DNA alterations.

5. Conclusion
With the next generation sequencing development and proteomic techniques, the connection between epigenetics alteration and cancer can be confirmed within details. However, the principle in oncogenic transformation is still unclear. Besides, the general statements of crosstalk between gene mutation and epigenetic pathways are still in discussion. However, apparently most of the studies are using DNA methylation to stop the cancer grow and development. In the second place, the main features of tumor cells involve immortal state, the high speed of fission, and pluripotent state that the ability to give rise to the majority of human cells. Based on this, DNA methylation can be used to rebuild the tumor cells, leading them to the differentiation into the target cells and revealing organ reconstruction. This can open a new gate of organ transplantation. The researchers can dig deeper in this field. As the day comes, multiple individual therapies in epigenetics and the organ reconstruction can evolve and surprise the public.

First and foremost, I would like to show my deepest gratitude to my teachers and professors in my university, who have provided me with valuable guidance in every stage of the writing of this thesis. Further, I would like to thank all my friends and roommates for their encouragement and support. Without all their enlightening instruction and impressive kindness, I could not have completed my thesis.

References
[1] Allis C. D., Jenuwein T., and Reinberg D. Epigenetics. Spring Harbor, NY: Cold Spring Harbor laboratory Press. (2007)
[2] Kanwal R., Gapta K., Gapta S. Cancer Epigenetics: An introduction. In: Verma M(eds). Cancer Epigenetics. Methods in Molecular Biology(Methods and protocols), Human Press. New York NY, 1238, 3-25(2015)
[3] Dawson M. A., Kouzarides T. Cancer Epigenetics: From Mechanism to Therapy. Elsevier Inc. Jun. 06, 2012, 150(1) 12-27(2012)
[4] Sega E., Windom J. From DNA Sequence to transcriptional behaviors: a quantitative approach. Nat, Rev, Genet. 10 443-56(2009)
[5] De wit E., D Laat W. A decade of 3C technologies: Insights into nuclear organization. Genes, Dec. 26, 2012, (26) 11-24 (2012)
[6] Farhach P. J. Insights from genomic profiling of transcription factors. Nat. Rev. Genet. 10, 605-16(2009)
[7] Caird P. W. Principles and challenges of genowide DNA methylation analysis. Nat, Rev, Genet. 11, 191-203(2010)
[8] Stunnenberg H.G, Vermenlen M. Towards cracking the epigenetic code using a combination of high-throughput epigenomics and quantatitve mass-spectrometry-based proteomics. Bioessays. 33, 547-51(2011)
[9] Li Z., Dai H., Martos S. N., Xu B., Gao Y., Li T., Zhu G., Schones D. E., Wang Z. Distinct roles of DNMT1-dependent and DNMT1-independent methylation patterns in the genome of mouse embryonic stem cells. Genome Biol. 2015 Jun 2,16,115(2015)
[10] Baylin S. B., Jones P. A. A decades of exploring the cancer epigenetics-biology and translational implications. Cancer. 11, 726-34(2011)
[11] Li E, Bestor T.H, Janisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 69,915-26.
[12] Okano M, Bell DW, Harbor RA, Li E. DNA methyltransferase Dnmt 3a and Dnmt 3b are essential fro de novo methylation and mammalin development. Cell. 99,247-57(1999)
[13] Perissi V1, Jepsen K, Glass CK, Rosenfeld MG. Deconstructing repression: evolving models of co-repressor action. Nat Rev Genet. 2010 Feb, 11(2):109-23(2010)
[14] Gigek CO, Chen ES, Lopes Filho GJ, Smith MA. The Methyl-CpG-Binding Domain(MBD) Protein Family: An overview and Dietary influences. Handbooks of Nutrition, Diet, and Epigenetics. Pp.1-15.
[15] Hainer SJ, McCannell KN, Yu J, Ee LS, Zhu LJ, Rando OJ, Fazzio TG. DNA methylation directs genomic localization of Mbd2 and Mbd3 in embryonic stem cells. Elife. 2016 Nov 16,5. pii: e21964(2016)
[16] Menafra R, Brinkman AB, Matarese F, Franci G, Bartels SJ, Nguyen L, Shimbo T, Wade PA, Hubner NC, Stunnenberg HG. Genome-wide binding of MBD2 reveals strong preference for highly methylated loci. PLoS One. 2014 Jun 13, 9(6):e99603(2014)
[17] Cimmino L, Abdel-Wakab O, Levine RL, Alfantis I. TET family proteins and their role in stem cell differentiation and transformation. Cell stem cell. Biol. 4,590-97.