Understanding Cancer Through the Lens of Epigenetic Inheritance, Allele-Specific Gene Expression, and High-Throughput Technology

Maxwell P. Lee*

High Dimension Data Analysis Group, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, United States

Epigenetic information is characterized by its stable transmission during mitotic cell divisions and plasticity during development and differentiation. This duality is in contrast to genetic information, which is stable and identical in all cells in an organism with exception of immunoglobulin gene rearrangements in lymphocytes and somatic mutations in cancer cells. Allele-specific analysis of gene expression and epigenetic modifications provides a unique approach to studying epigenetic regulation in normal and cancer cells. Extension of Knudson's two-hits theory to include epigenetic alteration as a means to inactivate tumor suppressor genes provides better understanding of how genetic mutations and epigenetic alterations jointly contribute to cancer development. High-throughput technology has greatly accelerated cancer discovery. Large initiatives such as TCGA have shown that epigenetic components are frequent targets of mutations in cancer and these discoveries provide new insights into understanding cancer etiology and generate new opportunities for cancer therapeutics.

Keywords: epigenetics, cancer, allele, inheritance, therapy

INTRODUCTION

Epigenetics, first coined by Conrad Waddington in 1940s, was a conceptual model that describes the development process of forming a multicellular organism from a fertilized zygote (1). The concept had its root in the earlier studies in embryology and developmental biology. This epigenetic concept provided mechanisms that can bring about cellular changes in development and physiology but not involving changes of genetic materials. Although Mendel's work on genetic inheritance was well-recognized but the exact biochemical nature of genetic material was not known until a decade later when the double helix model of DNA was proposed in 1953 (2, 3). Following the discovery of the double helix structure of DNA, there was an explosion of studies to understand how DNA sequences were replicated and used as templates to synthesize mRNAs, and how mRNA sequences were translated to produce proteins, resulting in different cellular phenotypes and ultimately organism phenotypes (4). This was culminated as the central dogma of molecular biology in 1958 (5). A major focus of the biological research since that time was to elucidate the molecular mechanisms that underlie the differential gene expression programing in cellular differentiation in development, physiological response in daily activities, and pathological changes in diseases. The details emerging from these studies led to a general understanding of association among DNA methylation, gene expression, and physiological changes at the levels of

organisms and cells. The contemporary definition of epigenetics proposed by Holliday stated that epigenetics is the study of gene expression changes during cellular differentiation and mitotic inheritance of cellular gene expression pattern, which doesn't involve changes in DNA sequence (6). The expanded view of epigenetics includes many phenomena that can't be explained by Mendelian inheritance. Some prominent examples are X-chromosome inactivation and genomic imprinting in mammal and position effect variegation in *Drosophila* (7–11). Indeed, it was the study of these non-mendelian phenomena that largely initiated the identification and characterization of the biochemical components of epigenetic machinery. The current view of epigenetic system consists of DNA methylation, histone acetylation and methylation and other posttranslational modifications, chromatin remodeling complexes, and non-coding RNAs (12–15). Together, these epigenetic components control gene expression and form the basis of epigenetic memory that can be transmitted through mitotic cell division without DNA sequence changes.

There are numerous excellent reviews on epigenetics and cancer epigenetics. A few are cited here (12–14, 16, 17). In this short review, I will focus on a few selected topics that capture some aspects of epigenetics and epigenetic regulation in cancer from the perspective of epigenetic stability vs. plasticity and from the perspective of the allele-specific gene expression.

EPIGENETIC INHERITANCE AND PLASTICITY

Epigenetic information is characterized by its stable transmission during mitotic cell divisions and plasticity during development and differentiation. This duality differs from genetic information, which remains the same in every cell in an organism with the exception of a few cases such as immunoglobulin gene rearrangements in lymphocytes and somatic mutations in cancer. This duality is depicted in **Figure 1**. The two-states model provides a useful conceptual framework to think about epigenetic stability vs. plasticity.

Two classical examples of epigenetic phenomena are mammalian X chromosome inactivation and genomic imprinting (18, 19). Both are characterized by establishing active and inactive chromatin states in the two chromosomes in early embryogenesis, which are maintained during the life time of an organism. X chromosome inactivation involves gene expression silencing in one of the two X chromosomes, which ensures similar level of gene expression in both female and male cells. Which of the two X chromosomes to be inactivated is chosen randomly in the early embryogenesis. However, in the case of genomic imprinting, the inactivation occurs in specific genomic loci and the choice of which chromosome to be silenced is determined by the parental origin. Hundreds of imprinted genes have been identified. Some show gene silencing in all expressed tissues and during the entire life of the organism while others may display genomic imprinting only in selected tissues and affected by developmental stages and environmental exposure.
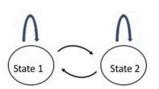

FIGURE 1 | Epigenetic stability vs. plasticity. The two states may represent any two conceptual epigenetic states such as an active chromatin vs. an inactive chromatin state or normal cell vs. cancer cell. The arc above the state represents maintenance of the state through events such as mitotic division whereas the arrows between the two states represent interconversion between the two states such as changes in chromatin structure during cellular differentiation, physiological response, or disease.

Studies of X chromosome inactivation and genomic imprinting played an instrumental role in establish DNA methylation and histone protein post-translational modifications and chromatin remodeling as the primary determinants of epigenetic state. There are many reasons why X chromosome and genomic imprinting are the excellent models to study epigenetics. The presence of a pair of active and inactive chromatin provides an ideal system to identify epigenetic marks that are specific for each epigenetic state but absent in the other epigenetic state. The DNA modification is relatively simple, involving methylation of the C5 in a cytosine (20). In mammalian genomes, the CpG dinucleotide occurs at much lower frequency than the other dinucleotides. This is because of the selective loss of CpG resulting from the conversion of 5-methylcytosine to thymine. However, there are genomic regions, where cluster of CpG dinucleotides are not methylated and consequently protected from the conversion, leading to the formation of CpG islands (CGIs) (21). About half of the mammalian genes contain CGIs, which are located near their transcription start sites.

Modifications of chromatin proteins are much more complex (12, 16). Both H3 and H4 histones undergo extensive post-translational modifications in their tails. These modifications include methylation, acetylation, phosphorylation, ubiquitination, etc. The combination of these modifications is referred to as the "histone code" (22), which carries the epigenetic information responsible for the maintenance of epigenetic state and dynamic change of epigenetic state.

From the perspective of epigenetic inheritance, DNA methylation state is maintained through DNA replication because semi-methylated DNA, the product of DNA replication, can be converted to fully methylated DNA by the action of DNMT1, which catalyze DNA methylation using semi-methylated DNA as substrate. DNA methyl transferase, DNMT3A, and DNMT3B, catalyze *de novo* methylation on DNA, thus providing a mechanism to acquire new DNA methylation marks to change chromatin state. However,

the effort of searching for enzymes that can catalyze DNA demethylation was unsuccessful until about 10 years ago. It led to the thinking in the past that perhaps DNA demethylation could be mediated only by passively losing half of the methylation during each cycle of DNA replication in the absence of DNMT1 activity. This was changed recently, when it was discovered that Ten-eleven translocation (TET) enzymes can catalyze demethylation of 5-methycytosine through sequential conversion of 5-methycytosine to 5-hydroxymethyl cytosine, to 5-formylcytosine, then to 5-carboxylcytosine, which can be converted to unmodified cytosine by terminal deoxynucleotidyl transferase (TDT) (23, 24).

Histone modifications are far more complicated than DNA methylation. But the general strategy is similar. There exist a pair of enzyme systems, histone post-translational modification "writers" and "erasers." For examples, histone acetyltransferase (HAT) serves as a writer whereas histone deacetylase (HDAC) serves as an eraser. Likewise, there are histone lysine methyltransferase (KMT) and histone lysine demethylase (KDM) to serve as writer and eraser, respectively. There are also protein arginine methyltransferases (PRMT), which act on arginine. Their opposing enzymes are peptidyl arginine deiminase (PADI). Each family also contains a large number of enzymes that can recognize specific substrate sequences. There is a third class of proteins called "readers" that can specifically bind to these post-translational modifications. For examples, bromodomain binds to acetylated lysine residue and chromodomain recognizes lysine methylation. Interestingly, histone acetyltransferase often contains bromodomain in addition to its activity to add acetylation to lysine. The multi-function structure of HAT enables it to catalyze acetylation in a processive manner to spread this post-translation mark (PTM) to the nearby nucleosome. This provides a potential mechanism for maintaining the PTM through mitotic division. Unlike DNA methylation, which produces semi-methylated DNA after DNA replication, the nucleosomes are randomly distributed into each of the two daughter cells after cell division. Half of the nucleosomes are derived from the parental cell and half are from newly deposited nucleosomes, which don't have PTMs. The ability of HAT to bind acetylated lysine and then catalyze addition of acetyl group to the nearby nucleosome allows the maintaining of this PTM through mitotic cell division.

The hallmark of epigenetics is the transmission of epigenetic marks through mitotic cell divisions. The duration of maintaining an epigenetic state varies. In the case of X chromosome inactivation and genomic imprinting, the active or inactive chromatin states are maintained throughout the lifetime. However, in most of cellular response to physiologic needs, the new epigenetic state is established and reversed back to normal state and the duration varies depending on particular physiology. DNA methylation mark is more stable while histone post-translational marks and other chromatin remodeling complexes display wide range of response time, serving different physiologic purposes.

ALLELE-SPECIFIC GENE EXPRESSION

X chromosome inactivation and genomic imprinting are characterized by the mono-allelic gene expression and epigenetic modifications. Mono-allelic gene expression also occurs in a number of other biological systems. In B lymphocytes, once an immunoglobulin gene rearrangement takes place on one chromosome, the rearrangement of the same gene from the other chromosome would be prevented. This phenomenum is termed as allelic exclusion, which ensures that an individual lymphocyte expresses a unique amino acid sequence of an immunoglobulin protein (25). Similar mechanism also operates in T lymphocytes for activating TCR genes (26). Another example of mono-allelic expression is the expression of human olfactory receptor genes. There are about 1,000 olfactory receptor genes, each of which is expressed from only one chromosome in a sensory neuron (27).

In addition, quantitative differences in the degree of gene expression between two alleles, marked with SNPs, are a widespread phenomenon, hereinafter referred to as allele-specific expression (ASE). We initially studied allele-specific gene expression using the Affymetrix SNP arrays and found extensive allelic variation in expression in the human genome (28). ASE differs from mono-allelic expression described in the previous sections. ASE showed differential gene expression between the two alleles in the range of 2–4 fold, which is in contrast to mono-allelic expression observed in imprinting and X chromosome inactivation. ASE is commonly affected by genetic polymorphisms near the gene and these polymorphisms play a regulatory role affecting gene expression (29, 30). This is particularly relevant since most of the GWAS identified SNPs are located in intragenic or intergenic regions. These SNPs impact phenotypes through gene expression regulation at the epigenetic level or post-transcriptional level.

UNDERSTANDING CANCER FROM THE PERSPECTIVE OF EPIGENETIC REGULATION AND ALLELE-SPECIFIC GENE EXPRESSION

It is well-established that cancer is caused by mutations that are acquired either from parents through germline inheritance or generated in somatic cells. Inactivating tumor suppressor genes and activating oncogenes both contribute to cancer development. In the case of inactivation of tumor suppressor genes, both alleles have to be inactivated in the cancer cell. This is best illustrated by Knudson's two-hits theory (31). The two-hits theory was postulated to explain why familial retinoblastoma develops earlier and bilateral while the sporadic retinoblastoma develops later and often unilateral. Based on the epidemiologic observation, Knudson hypothesized that retinoblastoma was caused by inactivation of both alleles of a tumor suppressor gene and in the case of familial syndrome one allele was inactivated in germline and the 2nd allele was inactivated in somatic tumor whereas in the case of sporadic cancer both alleles were inactivated in the somatic tumor. This paradigm can extend to include epigenetic alteration as a means to inactivate

tumor suppressor genes. Many tumor suppressor genes, such as BRCA1/2 and CDKN2A/B, are frequently silenced by DNA methylation and inactive chromatin marks (32–34).

Epigenetic alteration through germline inheritance can occur in familial cancer syndrome. Lynch syndrome is caused by mutations in the genes involved in DNA mismatch repair. Germline mutations in MLH1 and MSH2 causes the majority of Lynch syndrome. However, in several Lynch syndrome families, no mutations in mismatch repairs genes were found despite extensive effort of searching for causative mutations. Instead, heritable DNA methylation in the promoter regions of MLH1 or MSH2 was identified that silenced gene expression. Chan et al. analyzed a three-generation family using allele-specific methylation (ASM) and reported that methylation of MSH2 gene in the germ line cells correlated with the loss of the MSH2 protein in the colorectal adenocarcinomas (35). Besides silencing gene expression by DNA methylation, a somatic frameshift mutation was found in MSH2. The authors concluded that ASM in germline transmission was the first hit while the somatic mutation was the 2nd hit. In a separate study of a HNPCC family, the EPCAM gene had a germline deletion in the 3′ end and resulted transcription read-through into downstream MSH2 and an increase in DNA methylation in the promoter region of MSH2 (36). The deletion was co-segregated with ASM of the MSH2 promoter and the disease. Epimutation of the RB1 gene was recently found in a six generations retinoblastoma family (37). The germline methylation was inherited from the maternal chromosome. Interestingly, the detailed pedigree analysis also found a germline mutation that was transmitted through the paternal chromosome and showed incomplete penetrance. The authors concluded that both genetic mutation and epimutation contributed to the retinoblastoma in this family. A rare epimutation in the RB1 gene was also identified from another recent study (38). The authors showed that germline DNA methylation was associated with silencing of the RB1 gene expression.

Beckwith–Wiedemann syndrome (BWS), is another familial syndrome, which increases risk of developing multiple pediatric cancers. It has served as a model system for studying genomic imprinting and how abnormal genomic imprinting causes cancer. Multiple genetic and epigenetic mechanisms were identified that cause BWS, including mutation in CDKN1C (39), loss of imprinting in IGF2 (40), translocation involving KCNQ1 (41), and abnormal imprinting of a lincRNA, KCNQ1OT1 (42). All four genes are imprinted. IGF2 is normally expressed from the paternal chromosome but expressed from both chromosomes in tumors. KCNQ1OT1 is normally methylated on maternal chromosome but the methylation is frequently lost in BWS patient germline DNA. Allele-specific gene expression and allele-specific methylation analysis have played an instrumental role in elucidating various epigenetic mechanisms (43).

Two papers brought about wide appreciation of quantitative difference in gene expression between two alleles of APC in familial adenomatous polyposis (FAP) (44, 45). Yan et al. showed that 50% reduction in gene expression in APC was associated with predisposition to FAP. They studied six patients from two FAP families and didn't find any mutation in the APC gene. However, using ASE, they found 2-fold difference in gene expression between the two alleles in all 6 patients. Furthermore, tumors displayed loss of heterozygosity (LOH) and deleted specifically the high-expression allele.

An interesting question is whether DNA methylation causes inactive chromatin state or vice versa. An elegant study from Vogelstein's lab provided an important insight into answering this question (46). They generated double knockout of DNMT1 and DNMT3B, which eliminated most of DNA methylation in HCT116, a colon cancer cell line. The double knockout cells had slow growth rate in early passage cells but the late passage cells grow to comparable rate to the parental cells. This corresponded to a gradual increase in p16 methylation and consequently silencing of p16. The kinetics of chromatin marks changes was faster. Histone H4-acetylation increased as early as the passage 5 and H3K9-methylation appeared at the passage 22. But DNA methylation appeared at the passage 50. The p16 is heterozygous in HCT116, allowing tracking of both alleles for allele-specific analysis of gene expression, DNA methylation, and chromatin marks. The study revealed that only the wild type allele showed dynamic changes of epigenetic marks. These observations established that the order of epigenetic changes in this system was that it began with the gain of H4-acetylation and expression of p16, followed by H3K9 methylation and silencing of p16 expression and faster growth, and eventually cells fully re-gained DNA methylation and lose H4-acetylation mark and grew at comparable rate as the parental cells. The work also demonstrated the important role of silencing p16 in driving cellular proliferation.

ACCELERATING CANCER DISCOVERY WITH HIGH-THROUGHPUT TECHNOLOGY

The high-throughput analysis of gene mutations in human cancer was made possible after the human genome sequencing was completed in 2003. Some of the earliest studies that leveraged human genome sequence data to systematically identify mutated genes in human cancer were reported by the researchers from Johns Hopkins and Sanger Institute. These included the large scale analysis of coding sequences of human transcriptome in breast and colorectal cancer from Johns Hopkins in 2006 and 2007 (47, 48) and analysis of the coding exons of 518 protein kinase genes in multiple human cancers from Sanger Institute in 2007 (49). In 2005, NCI and NHGRI initiated The Cancer Genome Atlas (TCGA) initiative to comprehensively characterize genomic alterations in all major cancers. The pilot project was initially focused on glioblastoma multiforme (GBM) and ovary cancer, and it was extended to include more than 30 types of cancer in 2010. The first paper published from the TCGA initiative was the comprehensive analysis of mutation, DNA methylation, and gene expression of GBM in 2008 (50).

A very interesting study came from the comprehensive analysis of mutation in glioblastoma multiforme (GBM) by the Hopkins team in 2008, which led to the discovery of a recurrent mutation R132H in isocitrate dehydrogenase 1 (IDH1) and the mutation was shown to be associated with better survival (51).

The R132H mutation was always present as a heterozygous mutation, suggesting it functions as an oncogene. Detailed biochemical study showed that the IDH1 mutation generated 2-hydroxyglutarate (2HG) instead of alpha-ketoglutarate, which is the normal product of the wildtype enzyme (52). Excessive accumulation of 2HG contributed to the formation of gliomas, suggesting that 2HG acted as an onco-metabolite.

In 2010, TCGA team found that a subset of GBM has high CpG island methylation, which was termed as a glioma CpG island methylator phenotype (G-CIMP) (53). G-CIMP tumors were more prevalent among lower-grade gliomas and associated with IDH1 somatic mutations. The association between DNA methylation and IDH1 was not unique to GBM and it also occurred in acute myeloid leukemia (AML) (54). What was really intriguing was the finding that mutations in IDH1/2 and TET2 were mutually exclusive in AML. TET2 was known to catalyze the conversion of 5 methyl cytosine to 5 hydroxy methyl cytosine. This immediately suggested that IDH mutation inhibits TET2 activity. Indeed, expression of IDH1 mutant inhibited the production of 5 hydroxy methyl cytosine. The mechanism for the inhibition was because IDH mutants produced 2-hydroxyglutarate instead of 2-oxoglutarate. 2-oxoglutarate was co-factor for TET2 to catalyze hydroxy methylation whereas 2-oxoglutarate served as a competitive inhibitor to TET2. Therefore, either Tet2 mutation or IDH2 could cause accumulation of 5 methyl cytosine, generating the CpG island methylation phenotype. This explained why either IDH1/2 or TET2 mutation could block hematopoietic differentiation and cause proleukemogenic effect.

One of the emerging concepts from the high-throughput mutational analysis of human cancer genomes was the finding that chromatin components are the frequent targets of mutations in human cancer (55–57). Some examples are provided here. Recurrent mutations of the histone methyltransferase MLL2 were detected in 89% of follicular lymphoma (FL) and 32% of diffuse large B-cell lymphoma (DLBCL) (58). The histone H3K27 demethylase UTX was mutated in multiple human cancers (59). Mutations in EZH2, a histone H3K27 methyltransferase, was found in GCB subtype of DLBCL and follicular lymphoma (60). Mutations in DNA methyltransferase DNMT3A were identified in 25% of acute myeloid leukemia (AML) (61). Not only the histone modifiers were frequently mutated, but histone proteins were also the direct targets of mutations in human cancer. Both histone H3 variant H3.3 (H3F3A) and the histone H3.1 (HIST1H3B) were mutated in 30% of pediatric glioblastomas (62, 63). Interestingly, these mutations occurred at the specific sites, K27M and G34R/G34V, and were present in heterozygous. Detailed biochemical studies showed that H3K27M mutant acted in a dominant-negative manner to inhibit PRC2 activities and consequently reduced H3K27me3 level (64).

EPIGENETIC DYNAMICS IN CANCER TREATMENT

DNA methyltransferase inhibitors, 5-azacytidine (Vidaza) and 5-aza-2′-deoxycytidine (decitabine), are FDA approved drugs for myelodysplastic syndrome (MDS) and AML patients. Treatment with 5-azacitidine and decitabine increased overall survival in MDS patients than conventional care in phase III clinical trials (65, 66). The response rates are between 30 and 60%. The response in myeloid malignancies are better than lymphoid leukemia or solid tumors (67). This might be related to the observation that myeloid leukemia has a relatively low mutational burden but has mutations in the genes involved in controlling DNA methylation, such as TET2 or DNMT3A (68). Many studies were conducted to understand what the clinical factors and molecular alterations are associated with treatment response, only TET2 mutation was found to be weakly associated with clinical response to therapy (69, 70). DNA methyltransferase inhibitors have bi-modal activities. At low dose, they cause hypomethylation whereas at high dose, they are cytotoxic. Following treatment, there was global decrease in DNA methylation and hypomethylation was associated with better response (71). Hypomethylation of specific tumor suppressor genes such as CDKN2B was also observed, which was associated with reactivation of protein expression to a normal level (72). This is consistent with the mechanism of drug action.

Vorinostat (SAHA), belinostat (PXD101), and romidepsin are FDA approved histone deacetylase (HDAC) inhibitors for cutaneous T cell lymphoma (CTCL) patients (73). The response rates are between 30 and 40% of patients with CTCL (74–76). Panobinostat in combination with the proteasome inhibitor bortezomib is FDA approved for the treatment of drug-resistant multiple myeloma (77). However, the success of these drugs is limited to cutaneous T cell lymphoma and multiple myeloma, and they are not effective for solid tumors. Similar to DNA methyltransferase inhibitors, HDAC inhibitors also show bi-modal activities. Their efficacy is dose-dependent, and the drugs are cytotoxic at high dose (78). The drug targets are more complex since there are eleven HDACs and also many non-histone targets. The targets could be nuclear or cytoplasmic. The complexity makes it hard to predict what factors could determine how well patients respond to treatment.

Besides targeting DNA methyltransferases and histone deacetylase, recently identified mutations in histone modifiers and chromatin remodeling proteins offer new opportunities for targeted therapy (55, 79). These include development of JQ1 and I-BET that bind to acetyl lysine recognition motifs of bromodomain and extra-terminal (BET) of BRD4 (80, 81), which is involved in DNA translocation in several cancers and activation of MYC oncogene; development of an inhibitor of H3K79 *N*-methyltransferase (DOT1L), which is involved in leukemogenesis in mixed lineage leukemia (MLL) (82); development of a small molecule GSK2879552 that inhibits lysine demethylase 1 (LSD1) (83).

A major concern in cancer therapy, either chemotherapy or targeted therapy, is the development of resistance to cancer drugs. Many mechanisms contribute to drug resistance, including drug efflux and mutations in the targeted genes or related pathways. However, recent studies suggested that epigenetic alterations could provide another mechanism to acquire drug resistance, especially for slowly acquired resistance (84). The drug resistance involved activation of IGF1 signaling

pathway and chromatin alteration mediated by the histone demethylase KDM5A in a small population of drug-tolerant cells. Treatment with IGF1 receptor inhibitors or HDAC inhibitors can eliminate the drug-tolerant cells. The combination of chemo with HDAC inhibitors provide a potential new strategy to prevent development of drug resistance.

In conclusion, the studies of epigenetics and allele-specific gene expression and application of high-throughput technology provide powerful approaches to enhancing our understanding of cancer etiology and progression and also provide new opportunity for cancer therapeutics. There are some limitations when we try to understand cancer through the lens of epigenetic inheritance, allele-specific gene expression, and high-throughput technology. Germline epimutations are very rare events and some of which may be caused by yet unknown genetic variants. Although allele-specific gene expression can provide a unique perspective on the role of genetic variants on gene expression

regulation, we are often more interested in the combined gene expression contributed from both alleles and how the gene expression is associated with other biological phenomena. High-throughput technology is powerful for the discovery phase of the research. However, new findings should be rigorously validated by additional experiments.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

REFERENCES

1. Waddington CH. The epigenotype. *Endeavour.* (1942) 1:18–20.
2. Watson JD, Crick FH. Genetical implications of the structure of deoxyribonucleic acid. *Nature.* (1953) 171:964–7. doi: 10.1038/171964b0
3. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature.* (1953) 171:737–8. doi: 10.1038/171737a0
4. Cobb M. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol.* (2017) 15:e2003243. doi: 10.1371/journal.pbio.2003243
5. Crick FH. On protein synthesis. *Symp Soc Exp Biol.* (1958) 12:138–63.
6. Holliday R. DNA methylation and epigenetic inheritance. *Philos Trans R Soc Lond B Biol Sci.* (1990) 326:329–38. doi: 10.1098/rstb.1990.0015
7. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature.* (1991) 351:153–5. doi: 10.1038/351153a0
8. DeChiara TM, Robertson EJ, Efstratiadis A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell.* (1991) 64:849–59. doi: 10.1016/0092-8674(91)90513-X
9. Ferguson-Smith AC, Cattanach BM, Barton SC, Beechey CV, Surani MA. Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature.* (1991) 351:667–70. doi: 10.1038/351667a0
10. Goldschmidt RB. Chromosomes and genes. *Cold Spring Harb Symp Quant Biol.* (1951) 16:1–11. doi: 10.1101/SQB.1951.016.01.003
11. Lyon MF. Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature.* (1961) 190:372–3. doi: 10.1038/190372a0
12. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. *Nat Rev Genet.* (2016) 17:487–500. doi: 10.1038/nrg.2016.59
13. Feinberg AP, Tycko B. The history of cancer epigenetics. *Nat Rev Cancer.* (2004) 4:143–53. doi: 10.1038/nrc1279
14. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet.* (2003) 33(Suppl. 1):245–54. doi: 10.1038/ng1089
15. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet.* (2002) 3:415–28. doi: 10.1038/nrg816
16. Berger SL. Histone modifications in transcriptional regulation. *Curr Opin Genet Dev.* (2002) 12:142–8. doi: 10.1016/S0959-437X(02)00279-4
17. Feinberg AP. The epigenetics of cancer etiology. *Semin Cancer Biol.* (2004) 14:427–32. doi: 10.1016/j.semcancer.2004.06.005
18. Barlow DP, Bartolomei MS. Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol.* (2014) 6:a018382. doi: 10.1101/cshperspect.a018382
19. Brockdorff N, Turner BM. Dosage compensation in mammals. *Cold Spring Harb Perspect Biol.* (2015) 7:a019406. doi: 10.1101/cshperspect.a019406
20. Bird AP. Functions for DNA methylation in vertebrates. *Cold Spring Harb Symp Quant Biol.* (1993) 58:281–5. doi: 10.1101/SQB.1993.058.01.033

21. Bird AP. CpG-rich islands and the function of DNA methylation. *Nature.* (1986) 321:209–13. doi: 10.1038/321209a0
22. Jenuwein T, Allis CD. Translating the histone code. *Science.* (2001) 293:1074–80. doi: 10.1126/science.1063127
23. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature.* (2010) 466:1129–33. doi: 10.1038/nature09303
24. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science.* (2011) 333:1300–3. doi: 10.1126/science.1210597
25. Storb U, Ritchie KA, O'Brien R, Arp B, Brinster R. Expression, allelic exclusion and somatic mutation of mouse immunoglobulin kappa genes. *Immunol Rev.* (1986) 89:85–102. doi: 10.1111/j.1600-065X.1986.tb01474.x
26. Khor B, Sleckman BP. Allelic exclusion at the TCRbeta locus. *Curr Opin Immunol.* (2002) 14:230–4. doi: 10.1016/S0952-7915(02)00326-6
27. Serizawa S, Miyamichi K, Sakano H. One neuron-one receptor rule in the mouse olfactory system. *Trends Genet.* (2004) 20:648–53. doi: 10.1016/j.tig.2004.09.006
28. Lo HS, Wang Z, Hu Y, Yang HH, Gere S, Buetow KH, et al. Allelic variation in gene expression is common in the human genome. *Genome Res.* (2003) 13:1855–62. doi: 10.1101/gr.1006603
29. Cheung VG, Conlin LK, Weber TM, Arcaro M, Jen KY, Morley M, et al. Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat Genet.* (2003) 33:422–5. doi: 10.1038/ng1094
30. Knight JC, Keating BJ, Kwiatkowski DP. Allele-specific repression of lymphotoxin-alpha by activated B cell factor-1. *Nat Genet.* (2004) 36:394–9. doi: 10.1038/ng1331
31. Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA.* (1971) 68:820–3. doi: 10.1073/pnas.68.4.820
32. Belinsky SA, Nikula KJ, Palmisano WA, Michels R, Saccomanno G, Gabrielson E, et al. Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proc Natl Acad Sci USA.* (1998) 95:11891–6. doi: 10.1073/pnas.95.20.11891
33. Dobrovic A, Simpfendorfer D. Methylation of the BRCA1 gene in sporadic breast cancer. *Cancer Res.* (1997) 57:3347–50.
34. Esteller M. Epigenetic lesions causing genetic lesions in human cancer: promoter hypermethylation of DNA repair genes. *Eur J Cancer.* (2000) 36:2294–300. doi: 10.1016/S0959-8049(00)00303-8
35. Chan TL, Yuen ST, Kong CK, Chan YW, Chan AS, Ng WF, et al. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nat Genet.* (2006) 38:1178–83. doi: 10.1038/ng1866
36. Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, et al. Heritable somatic methylation and inactivation of MSH2 in families with

Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat Genet.* (2009) 41:112–7. doi: 10.1038/ng.283

37. Quiñonez-Silva G, Dávalos-Salas M, Recillas-Targa F, Ostrosky-Wegman P, Aranda DA, Benítez-Bribiesca L. Monoallelic germline methylation and sequence variant in the promoter of the RB1 gene: a possible constitutive epimutation in hereditary retinoblastoma. *Clin Epigenet.* (2016) 8:1. doi: 10.1186/s13148-015-0167-0

38. Gelli E, Pinto AM, Somma S, Imperatore V, Cannone MG, Hadjistilianou T, et al. Evidence of predisposing epimutation in retinoblastoma. *Hum Mutat.* (2019) 40:201–6. doi: 10.1002/humu.23684

39. Hatada I, Ohashi H, Fukushima Y, Kaneko Y, Inoue M, Komoto Y, et al. An imprinted gene p57KIP2 is mutated in Beckwith-Wiedemann syndrome. *Nat Genet.* (1996) 14:171–3. doi: 10.1038/ng1096-171

40. Rainier S, Johnson LA, Dobry CJ, Ping AJ, Grundy PE, Feinberg AP. Relaxation of imprinted genes in human cancer. *Nature.* (1993) 362:747–9. doi: 10.1038/362747a0

41. Lee MP, Hu RJ, Johnson LA, Feinberg AP. Human KVLQT1 gene shows tissue-specific imprinting and encompasses Beckwith-Wiedemann syndrome chromosomal rearrangements. *Nat Genet.* (1997) 15:181–5. doi: 10.1038/ng0297-181

42. Lee MP, DeBaun MR, Mitsuya K, Galonek HL, Brandenburg S, Oshimura M, et al. Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc Natl Acad Sci USA.* (1999) 96:5203–8. doi: 10.1073/pnas.96.9.5203

43. Lee MP. Allele-specific gene expression and epigenetic modifications and their application to understanding inheritance and cancer. *Biochim Biophys Acta.* (2012) 1819:739–42. doi: 10.1016/j.bbagrm.2012.02.007

44. Yan H, Dobbie Z, Gruber SB, Markowitz S, Romans K, Giardiello FM, et al. Small changes in expression affect predisposition to tumorigenesis. *Nat Genet.* (2002) 30:25–6. doi: 10.1038/ng799

45. Yan H, Yuan W, Velculescu VE, Vogelstein B, Kinzler KW. Allelic variation in human gene expression. *Science.* (2002) 297:1143. doi: 10.1126/science.1072545

46. Bachman KE, Park BH, Rhee I, Rajagopalan H, Herman JG, Baylin SB, et al. Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. *Cancer Cell.* (2003) 3:89–95. doi: 10.1016/S1535-6108(02)00234-9

47. Sjöblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, et al. The consensus coding sequences of human breast and colorectal cancers. *Science.* (2006) 314:268–74. doi: 10.1126/science.1133427

48. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. *Science.* (2007) 318:1108–13. doi: 10.1126/science.1145720

49. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. *Nature.* (2007) 446:153–8. doi: 10.1038/nature05610

50. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature.* (2008) 455:1061–8. doi: 10.1038/nature07385

51. Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science.* (2008) 321:1807–12. doi: 10.1126/science.1164382

52. Dang L, White DW, Gross S, Bennett BD, Bittinger MA, Driggers EM, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature.* (2009) 462:739–44. doi: 10.1038/nature08617

53. Noushmehr H, Weisenberger DJ, Diefes K, Phillips HS, Pujara K, Berman BP, et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell.* (2010) 17:510–22. doi: 10.1016/j.ccr.2010.03.017

54. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell.* (2010) 18:553–67. doi: 10.1016/j.ccr.2010.11.015

55. Audia JE, Campbell RM. Histone modifications and cancer. *Cold Spring Harb Perspect Biol.* (2016) 8:a019521. doi: 10.1101/cshperspect.a019521

56. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. *Cell.* (2012) 150:12–27. doi: 10.1016/j.cell.2012.06.013

57. Garraway LA, Lander ES. Lessons from the cancer genome. *Cell.* (2013) 153:17–37. doi: 10.1016/j.cell.2013.03.002

58. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature.* (2011) 476:298–303. doi: 10.1038/nature10351

59. van Haaften G, Dalgliesh GL, Davies H, Chen L, Bignell G, Greenman C, et al. Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. *Nat Genet.* (2009) 41:521–3. doi: 10.1038/ng.349

60. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet.* (2010) 42:181–5. doi: 10.1038/ng.518

61. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med.* (2010) 363:2424–33. doi: 10.1056/NEJMoa1005143

62. Schwartzentruber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature.* (2012) 482:226–31. doi: 10.1038/nature10833

63. Wu G, Broniscer A, McEachron TA, Lu C, Paugh BS, Becksfort J, et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat Genet.* (2012) 44:251–3. doi: 10.1038/ng.1102

64. Lewis PW, Müller MM, Koletsky MS, Cordero F, Lin S, Banaszynski LA, et al. Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science.* (2013) 340:857–61. doi: 10.1126/science.1232245

65. Fenaux P, Mufti GJ, Hellstrom-Lindberg E, Santini V, Finelli C, Giagounidis A, et al. 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.* (2009) 10:223–32. doi: 10.1016/S1470-2045(09)70003-8

66. Lübbert M, Suciu S, Hagemeijer A, Rüter B, Platzbecker U, Giagounidis A, et al. Decitabine improves progression-free survival in older high-risk MDS patients with multiple autosomal monosomies: results of a subgroup analysis of the randomized phase III study 06011 of the EORTC Leukemia Cooperative Group and German MDS Study Group. *Ann Hematol.* (2016) 95:191–9. doi: 10.1007/s00277-015-2547-0

67. Issa JP, Kantarjian HM. Targeting DNA methylation. *Clin Cancer Res.* (2009) 15:3938–46. doi: 10.1158/1078-0432.CCR-08-2783

68. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med.* (2013) 368:2059–74. doi: 10.1056/NEJMoa1301689

69. Bejar R, Lord A, Stevenson K, Bar-Natan M, Pérez-Ladaga A, Zaneveld J, et al. TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients. *Blood.* (2014) 124:2705–12. doi: 10.1182/blood-2014-06-582809

70. Itzykson R, Kosmider O, Cluzeau T, Mansat-De Mas V, Dreyfus F, Beyne-Rauzy O, et al. Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. *Leukemia.* (2011) 25:1147–52. doi: 10.1038/leu.2011.71

71. Yang AS, Doshi KD, Choi SW, Mason JB, Mannari RK, Gharybian V, et al. DNA methylation changes after 5-aza-2'-deoxycytidine therapy in patients with leukemia. *Cancer Res.* (2006) 66:5495–503. doi: 10.1158/0008-5472.CAN-05-2385

72. Daskalakis M, Nguyen TT, Nguyen C, Guldberg P, Köhler G, Wijermans P, et al. Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2'-deoxycytidine (decitabine) treatment. *Blood.* (2002) 100:2957–64. doi: 10.1182/blood.V100.8.2957

73. Falkenberg KJ, Johnstone RW. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov.* (2014) 13:673–91. doi: 10.1038/nrd4360

74. Duvic M, Talpur R, Ni X, Zhang C, Hazarika P, Kelly C, et al. Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T-cell lymphoma (CTCL). *Blood.* (2007) 109:31–9. doi: 10.1182/blood-2006-06-025999

75. Piekarz RL, Frye R, Turner M, Wright JJ, Allen SL, Kirschbaum MH, et al. Phase II multi-institutional trial of the histone deacetylase inhibitor romidepsin as monotherapy for patients with cutaneous T-cell lymphoma. *J Clin Oncol.* (2009) 27:5410–7. doi: 10.1200/JCO.2008.21.6150

76. Whittaker SJ, Demierre MF, Kim EJ, Rook AH, Lerner A, Duvic M, et al. Final results from a multicenter, international, pivotal study of romidepsin in refractory cutaneous T-cell lymphoma. *J Clin Oncol.* (2010) 28:4485–91. doi: 10.1200/JCO.2010.28.9066

77. San-Miguel JF, Hungria VT, Yoon SS, Beksac M, Dimopoulos MA, Elghandour A, et al. Panobinostat plus bortezomib and dexamethasone versus placebo plus bortezomib and dexamethasone in patients with relapsed or relapsed and refractory multiple myeloma: a multicentre, randomised, double-blind phase 3 trial. *Lancet Oncol.* (2014) 15:1195–206. doi: 10.1016/S1470-2045(14)70440-1

78. Azad N, Zahnow CA, Rudin CM, Baylin SB. The future of epigenetic therapy in solid tumours–lessons from the past. *Nat Rev Clin Oncol.* (2013) 10:256–66. doi: 10.1038/nrclinonc.2013.42

79. Jones PA, Issa JP, Baylin S. Targeting the cancer epigenome for therapy. *Nat Rev Genet.* (2016) 17:630–41. doi: 10.1038/nrg.2016.93

80. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, et al. Selective inhibition of BET bromodomains. *Nature.* (2010) 468:1067–73. doi: 10.1038/nature09504

81. Nicodeme E, Jeffrey KL, Schaefer U, Beinke S, Dewell S, Chung CW, et al. Suppression of inflammation by a synthetic histone mimic. *Nature.* (2010) 468:1119–23. doi: 10.1038/nature09589

82. Daigle SR, Olhava EJ, Therkelsen CA, Majer CR, Sneeringer CJ, Song J, et al. Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell.* (2011) 20:53–65. doi: 10.1016/j.ccr.2011.06.009

83. Mohammad HP, Smitheman KN, Kamat CD, Soong D, Federowicz KE, Van Aller GS, et al. A DNA hypomethylation signature predicts antitumor activity of LSD1 inhibitors in SCLC. *Cancer Cell.* (2015) 28:57–69. doi: 10.1016/j.ccell.2015.06.002

84. Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell.* (2010) 141:69–80. doi: 10.1016/j.cell.2010.02.027