Cancer is the outcome of defective epigenetic copying of the pattern of selective gene activity in differentiated cells

Patrick A. Riley*

Totteridge Institute for Advanced Studies, The Grange, Grange Avenue, London N20 8AB, U.K. Email: rebc900@ucl.ac.uk

Abstract

The selective gene silencing which accounts for the specialization characteristic of differentiated cells is the result of reversible epigenetic modification of the genome. The process involved is heterochromatisation, a major component of which is brought about by DNA methylation with associated modifications of histones with consequent alterations in chromatin structure which control access to transcription sites. When cells divide the chromatin pattern is disturbed. To retain the differentiated state of the cell the chromatin pattern of the parent cell must be reproduced in the progeny. The mechanism involves the methylation of the replicated DNA strand to correspond with the methylation pattern of the parent strand. The complex machinery, which implicates many component elements, normally ensures that the pattern of gene silencing is accurately transmitted to the daughter cells. However, permanent derangement of the copying mechanism would result in inappropriate re-expression of previously silenced genes or failure of expression of previously active genes. It is the essence of the present proposal that the initiating event(s) of carcinogenesis are irreversible events, such as mutations, affecting the efficiency of the chromatin copying mechanism with resulting failure of fidelity of the vertical transmission of the gene activity pattern. Thus, in the affected clone, there will be a divergent range of anomalous gene expression with the emergence of sub-clones with abnormal structure and function, some of which will manifest properties characteristic of the malignant growth syndrome, such as invasiveness, metastasis and escape from proliferative control. This phase of carcinogenesis corresponds to what has been termed ‘progression’ and is characterised by genetic instability in the affected cell lineage.

Keywords: cancer, two-stage carcinogenesis, DNA methylation, epigenetic gene silencing, chromatin pattern, progression, initiation, malignant phenotype.

Introduction

The pathological diagnosis of malignancy rests on multiple cellular structural and functional abnormalities manifested as abnormal cytological features and deranged tissue architecture (1). The proposal advanced here is that these multiple pleiotropic abnormalities are due to derangement of the epigenetic process which maintains the normal pattern of gene expression when differentiated cells divide. The pattern-retaining mechanism involves a range of detailed processes, but a central feature is duplication of the pattern of DNA methylation. Methylated DNA is involved in maintaining the pattern of chromatin structure characteristic of differentiated cell types in which segments of the genome are unavailable for transcription (2). Because the stability of the organism is crucially dependent on the retention of the differentiated pattern, and therefore the accurate reproduction of the gene expression in daughter cells, it is probable that the duplication of the epigenetic pattern is normally subject to quality control involving the auditing of the process and elimination of defective division products.

The importance of mutations in carcinogenesis is universally accepted and it is widely recognised that...
cancer involves both genetic and epigenetic changes. Although the order in which these occur may vary, there is an advantage in proposing that mutational events precede epigenetic derangement as a way of explaining the apparent hypermutability of pre-cancerous cells (3, 4).

In general, the present proposal (5) is that carcinogenesis involves two stages: (a) Initiation, in which mutations, or other heritable derangements, affect the epigenetic mechanisms responsible for the vertical transmission of the pattern of gene activity characteristic of differentiated cells; and (b) progression, in which the failure of fidelity of the epigenetic mechanism results in genetic divergence of the affected clone. This process is schematically illustrated in Figure 1. Cells affected by agents that interfere with normal transmission of the epigenetically controlled pattern of gene expression give rise to a pool of pre-malignant cells. When these divide they produce sub-clones that exhibit divergent evolution with anomalous gene activities, some of which result in behavioural abnormalities characteristic of malignancy.

**DNA Demethylation**

In the early embryo there is a general demethylation of methylcytosine bases in the DNA which is partly through active demethylation by the action of DNA demethylases (6,7) and partly through DNA replication without methylation i.e. ‘passive’ demethylation (8,9). This seems to be a process that erases patterns of silenced genes which may be essential to obliterate the highly differentiated state of the sperm and egg. It is possible that similar processes are implicated in regaining ‘stem cell’ status in differentiated cells, as required in regenerating organisms.

**Generalised de novo Methylation**

In multicellular organisms a majority of DNA sequences are ‘permanently’ repressed. In mammals most of this silencing is brought about in the embryo
at the implantation stage by de novo DNA methylation by two key enzymes, DNA methyltransferase 3a and 3b (DNMT3a, DNMT3b) (10). This phase of DNA methylation is ubiquitous except for ‘protected’ potential transcriptional start sites (mainly CpG islands). The protection mechanism seems to involve binding to the potential transcription site of RNA polymerase II which recruits SET-domain proteins (HK methyltransferases) that methylate the lysine 4 of histone H3 (H3K4) on adjacent nucleosomes (11). The product (H3K4me) prevents binding by DNA methyltransferase 3L (DNMT3L) which is part of the de novo methylation complex (12) and thus excludes de novo methylation by DNMT3a and DNMT3b. This generalised and early DNA methylation appears to be essential for embryogenesis since DNMT3a and DNMT3b ‘knockouts’ in mice are lethal at the 8/9 days postcoitum stage (10).

While DNA methylation plays a central role in stable long-term gene repression additional mechanisms mediate gene silencing (13) such as the polycomb complex binding (14) which also affects chromatin structure through tri-methylation of histone H3K27(15) although this mechanism is more readily reversible (13).

Repressor cued de novo DNA methylation
At about the time of gastrulation further de novo DNA methylation takes place affecting specific genes previously protected from inactivation. In this phase of development the DNA methylation is secondary to histone modification and appears to permit some flexibility in the permanency of gene silencing. The sequence of events, for which an example is the Oct-3/4 gene which is active in the early embryo (16), seems to affect genes that have a specific repressor bound to the promoter site (17). The repressor recruits a G9a-complex containing a histone de-acetylase (HDAC) and a H3K9 methyltransferase which tri-methylates H3K9 (18). The methylated H3K9 binds heteroprotein 1 (HP1) which results in the promoter site being included in heterochromatin and finally G9a recruits DNMT3a and DNMT3b bringing about de novo methylation and stable repression (19). The timing of DNA replication is influenced by the chromatin structure. Euchromatin-associated DNA (which includes all the housekeeping genes) is replicated early in S-phase, whereas heterochromatin-associated DNA is replicated late (20, 21). It is possible that late-replicating DNA is methylated either by DNMT1 if already methylated, or by recruitment of DNMT3a and DNMT3b where de novo methylation is involved.

It is evident that any interference with the operation of the gene silencing pattern will be detrimental to embryonic development and may be implicated in developmental and childhood malignancies, such as retinoblastoma, osteoblastoma, neuroblastoma, and nephroblastoma.

Chromatin structure and Maintenance of DNA methylation
In general DNA methylation operates by modification of chromatin structure by modulation of local structural features such as histone modification and nucleosome positioning (22,23). Methylated DNA is associated with heterochromatin, the structure of which limits (or excludes) accessibility to DNA of the transcription machinery.

Since the chromatin structure is disrupted during replication (24) the pattern needs to be re-established. The DNA methylation pattern is copied at the replication forks (25). Although some details remain obscure, the major process involved in maintaining genomic methylation involves DNA methyltransferase 1 (DNMT1) which is associated with DNA replication sites and restores bilateral methylation to hemimethylated DNA (26). In mammals, DNMT1 expression seems to be regulated through the retinoblastoma pathway involving retinoblastoma 1 (RB1) and retinoblastoma–binding protein 4 (RBBP4) (27). DNMT1 is recruited to replication sites by an interaction with proliferating cell nuclear antigen (PCNA) (28) and other chromatin-associated proteins, including ubiquitin-like plant homeodomain and RING finger domain 1 (UHRF1) which specifically binds to hemimethylated DNA (29), and lymphoid-specific helicase (LSH1) (30).

In addition there are processes that re-establish the associated histone modification pattern, although the details are not clear. It has been shown that unmethylated DNA is packaged in euchromatin nucleosomes with histones that are acetylated and include H3K4me. By contrast, methylated DNA is packaged with deacetylated histones with unmethylated H3K4 and H3K9me (31). Re-association with heterochromatin is determined by DNA methylation (32,33) which thus determines the stable long-term gene repression (silencing) observed in differentiated cell lineages.

Failure of maintenance of gene silencing
In the case of adult malignancies the present argument is based on the failure of replication of the epigenetic silencing pattern which permits the inappropriate re-expression of previously silenced genes with consequent tissue disruption. The failure
of fidelity of the vertical transmission of the gene silencing pattern results in increasing heterogeneity of gene expression enabling the ultimate emergence of cells expressing the phenotypic characteristics of the malignant syndrome, i.e. invasion, metastasis and escape from proliferative control. Deregulation of epigenetic mechanisms may result in both hyper and hypomethylation or may entail other processes that control the inheritance of chromatin patterning, but since the consequences of a failure to accurately replicate the epigenetic pattern are so devastating it would be anticipated that systems have evolved to check the accuracy of the pattern of somatic inheritance and to eliminate those cells in which a faulty pattern is detected. It is possible that this proof-reading process is embodied in the apoptosis mechanism. This notion is supported by evidence that conditional knockouts of DNMTs in differentiated somatic cells initiate apoptosis in a process dependent on p53 (34,35). Since the general argument is that cancer progression is brought about by deregulation of epigenetic mechanisms it is of note that that the histone variant macroH2A (36) fits the proposed model.

Failure of quality control mechanism
The normal regulation of epigenetic inheritance may be regarded as involving so-called ‘tumour suppressor’ genes. There are several genes that fall within this category. In particular, mutations of the p53 gene are commonly found in malignant tumours (37) and it is likely that p53 is implicated in the mechanism of stable clonal transmission. Another important gene in this category is the Rb gene (38). In the normal cell cycle, pRb becomes hyperphosphorylated during DNA synthesis and can act as a transcriptional regulator. It forms part of the methylation complex and is essential for normal development. Knockout mice die at 14 days of embryonic development (39).

Tumour suppressor genes and regulation of DNMT1
One of the functions of p53 is the regulation of DNA methylation by binding to the specificity protein Sp1 to form a repressor complex (p53/Sp1) that interacts with the promoter region of DNMT1 and prevents transcription (40). Deficiency of p53 permits DNMT1 overexpression which may result in abnormal methylation patterns. pRb also has a regulatory function by binding to and inactivating the transcription factor E2F which is a positive regulator of the DNMT1 promoter (27). Hence, deficiency of pRb also contributes to DNMT1 overexpression. Overexpression of DNMT1 leads to anomalous DNA methylation including hypermethylation of tumour-suppressor gene promoters which is associated with cellular transformation. Silencing of genes necessary for normal function may result in expression of malignant characteristics, as demonstrated by the targeted methylation of the p16 suppressor gene in mice (41).

Viral Carcinogenesis
In addition to mutations, the introduction of viral genes can act by interfering with epigenetic mechanisms. Studies of the effect of carcinogenic viruses have shown that hepatitis B, hepatitis C, Kaposi’s sarcoma-associated virus, Epstein-Barr virus, and human papilloma virus regulate the expression of DNA methylating enzymes (42). In the case of HPV infection epigenetic alterations are induced by the E6 and E7 oncoproteins which influence the expression of DNMT1. The E6 and E7 oncoproteins of high risk HPV (HR-HPV) increase the expression and activity of DNMT1 (43); E6 by degrading p53 which binds to specificity protein 1 (Sp1) and chromatin-remodelling proteins to form a complex that binds to the promoter region of DNMT1 and prevents transcription of the methylase. Thus, by degradation of p53 there is overexpression of DNMT1 and hypermethylation of the promoter regions of important loci including those of tumour suppressor genes. Modulation of DNMT1 expression by E7 may occur by direct binding to DNMT1 which modifies the conformation of the enzyme and activates an interaction with DNA (44), or it may occur indirectly by binding to pRb leading to pRb degradation which releases the transcription factor E2F which regulates DNMT1 promoter activity resulting in overexpression of the methylase. The resulting derangements might be quite modest in their effects on clonal behaviour but could include activation of mechanisms leading to de novo methylation and/or active demethylation both of which could lead to progressive loss of epigenetic pattern.

Hypermutability
Among the major factors that determine the incidence of cancer are the size of the cell population at risk, the number of genes involved in generating the characteristics of malignant transformation, and the relevant mutation rate. Generally the recorded cancer incidence rates in humans accord with an approximately sixth power of age suggesting that six independent mutational events are necessary to bring about malignancy, i.e. six genes need to be affected. This poses a problem because estimates of the average mutation rate in humans is in the range $10^{-6}$.
per gene per year which underestimates by several orders of magnitude the probability of a normal cell undergoing transformation to a malignant cell, a difficulty that has been appreciated for a long time (45). A solution to the dilemma is to postulate that carcinogenesis involves hypermutability, and this has been suggested by a number of authors (3,4).

Two-stage Carcinogenesis
The basic idea is consistent with the accepted pathological division of carcinogenesis into two phases respectively known as initiation and progression (46), in which the initiating events occur against a background of normal mutation probability and the progression involves an enhanced mutation rate. This Two-Stage Model of carcinogenesis is eminently compatible with the present proposal that the second (progression) phase is due to epigenetic failure to retain the differentiated pattern of gene activity when a differentiated cell divides. In this scenario, initiation involves the mutation or interference with the action of genes instrumental in accurately duplicating the epigenetic gene activity patterns. Damage to or interference with the activity of these genes results in the failure of fidelity of vertical inheritance of gene activity patterns with the development of progressive clonal aberration. Calculations based on this model (47) are consistent with observed cancer incidence data.

General Predictions of the model
This theoretical scenario of the process of carcinogenesis makes certain predictions that are consistent with the observed facts including:

- All malignancies and pre-malignant tissue will manifest abnormalities of DNA methylation (49,50).
- Malignant and pre-malignant cells will exhibit chromatin abnormalities(51).
- Malignant and pre-malignant cells will exhibit divergent clonal evolution (52).
- Malignant and pre-malignant cells will exhibit uncoordinated and deranged metabolism with expression of genes associated with earlier developmental stages or of other tissue, manifested as slow growth and structural abnormalities (recognisable cytologically as bizarre cells).
- There will be no general immunological recognition of the anomalous proteins expressed by malignant and pre-malignant cells since they belong to the class of self-antigens.

The overall concept is consonant with long-held ideas regarding the nature of cancer including de-differentiation, cancer stem-cells and expression of embryological characteristics (53).

Acknowledgements
I thank my colleagues in the Quintox Group for useful discussions and encouragement. The work of the Totteridge Institute for Advanced Studies is supported by a generous grant from the Melanoma Research Society.

Abbreviations
DNMT DNA methyltransferase;
HDAC histone deacetylase;
HP1 heteroprotein 1;
HPV human papilloma virus;
LSH lymphoid-specific helicase;
PCNA proliferating cell nuclear antigen;
RBBP retinoblastoma-binding protein;
SP specificity protein;
UHRF ubiquitin-like and RING finger domain;

References


