NSD2, PCa, and Metastasis

Terrence P McGarty
White Paper No 158
February, 2019

There is an ongoing study of the genomic elements in the development and progression of prostate cancer, PCa. This report examines some recent work on the impact of NSD2, a methyltransferase, that can result in metastatic growth This is a study of epigenetic factors in PCa..

Notice

This document represents the personal opinion of the author and is not meant to be in any way the offering of medical advice or otherwise. It represents solely an analysis by the author of certain data which is generally available. The author furthermore makes no representations that the data available in the referenced papers is free from error. The Author also does not represent in any manner or fashion that the documents and information contained herein can be used other than for expressing the opinions of the Author. Any use made and actions resulting directly or otherwise from any of the documents, information, analyses, or data or otherwise is the sole responsibility of the user and The Author expressly takes no liability for any direct or indirect losses, harm, damage or otherwise resulting from the use or reliance upon any of the Author's opinions as herein expressed. There is no representation by The Author, express or otherwise, that the materials contained herein are investment advice, business advice, legal advice, medical advice or in any way should be relied upon by anyone for any purpose. The Author does not provide any financial, investment, medical, legal or similar advice in this document or in its publications on any related Internet sites.

Furthermore, this document contains references to and quotes from papers and documents under the premise of Fair Use in order to present ideas and understandings in context. The Author has attempted to make any and all references to such material separate from those of the author per se and has referenced the source expressly in all cases. These documents are for the dissemination of ideas and have no commercial intent. The Author would appreciate any communications relating to these documents and these should be sent to:

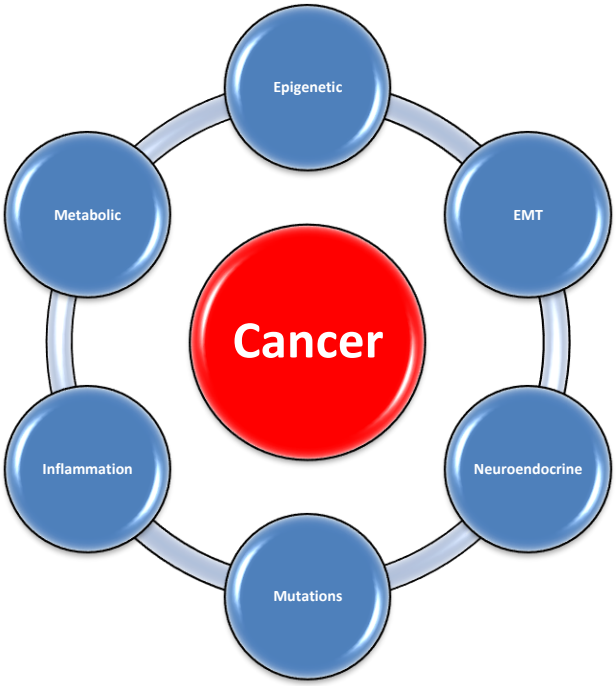
tmcgarty@telmarc.com.

Contents

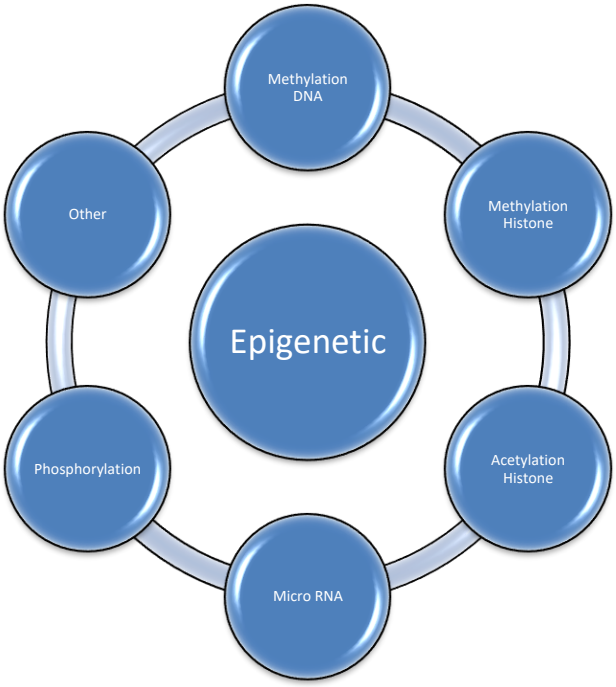
1 INTRODUCTION

Cancer is a complex systemic process whose evolution via metastatic progression is controlled by a plethora of steps, many of which may most likely be yet to be discovered. Each time a new target is identified, or proposed, there may be a rush to use this new simple paradigm as a single handed approach to mitigation. However, as we have noted again and again, cancer is a system problem, and in many ways it is the development of a "new organism" which is trying to take over its host, using the very cells that the host is dependent upon. As is said; nothing personal, it just wants to proliferate.

There are a multiplicity of actions resulting in cancer development and proliferation. We demonstrate a few graphically below, most of which we have examined in one way or another.



Let us consider the epigenetic element alone. It also has many subdivision as shown below.



In a recent paper by Aytes et al the authors have noted:

Deciphering cell-intrinsic mechanisms of metastasis progression in vivo is essential to identify novel therapeutic approaches. Here we elucidate cell-intrinsic drivers of metastatic prostate cancer progression through analyses of genetically engineered mouse models (GEMM) and correlative studies of human prostate cancer. Expression profiling of lineage marked cells from mouse primary tumors and metastases defines a signature of de novo metastatic progression. Cross-species master regulator analyses comparing this mouse signature with a comparable human signature identifies conserved drivers of metastatic progression with demonstrable clinical and functional relevance.

In particular, nuclear receptor binding SET Domain Protein 2 (NSD2)[1] *is robustly expressed in lethal prostate cancer in humans, while its silencing inhibits metastasis of mouse allografts in vivo. We propose that cross-species analysis can elucidate mechanisms of metastasis progression, thus providing potential additional therapeutic opportunities for treatment of lethal prostate cancer…*

In particular, we have demonstrated that the Nuclear receptor binding SET Domain Protein 2 (NSD2) is a robust marker of lethal metastatic prostate cancer and a key driver of prostate cancer metastasis, extending previous studies that have reported the relevance of NSD2 in prostate cancer.

[1] Also known as WHS; TRX5; KMT3F; KMT3G; MMSET; WHSC1; REIIBP.

NSD2 was discovered as the overexpressed product of the t(4;14)(p16.3;q32.3) translocation in multiple myeloma, and alternatively named Multiple Myeloma SET domain containing protein (MMSET), and was identified as a target gene on the 4p16 deletion for the Wolf- Hirschhorn Syndrome, and alternatively called Wolf-Hirschhorn Syndrome Candidate 1 (WHSC1. Previous studies have shown that genomic alterations occur in other cancer types in addition to multiple myeloma including pediatric leukemia and laryngeal tumors.

In prostate cancer, NSD2 has been shown to be up-regulated in advanced tumors coordinating with the activation of PI-3 kinase signaling, and to be a cofactor of androgen receptor. Notably, the role of NSD2 in cancer has been shown to be dependent on its activity as a histone methyltransferase for the histone H3 di-methyl K36 (H3K36me2). In the current study, we show that MCTP-39, a putative inhibitor of NSD2, inhibits prostate tumor growth in vivo.

However, several caveats preclude us from drawing the direct conclusion that MCTP-39 is acting to inhibit NSD2 activity in this context, including the potential activity of unknown metabolites and the potential lack of specificity of MCTP-39 given its relatively simple chemical structure.

Our intent herein is to explore the setting of this observation and explore it from an overall systems perspective.

Before proceeding it is worthwhile to lay forth the logic contained in this paper.

1. NSD2 is a gene product which is a methyltransferase acting on H3K36me2. This simply means that NSD2 is an enzyme which effects the methylation state of lysine on the 36 position on histone H3. Now it leaves open a large set of questions.

2. First, what activates NSD2, what is its pathway? One assumes that this is not a mutated NSD2 and thus it appears to be overly expressed. Is NSD2 unexpressed prior to this?

3. Second, as a methyltransferase, what is it targeting and why? There are very many H3 strings in chromosome. Which ones does it get attracted to and why. Is the process of specificity assisted by other factors such as specific histone code layout or underlying gene presence.

4. Third, after targeting the H3 tail element, what are the steps taken in the cell to lead to metastasis?

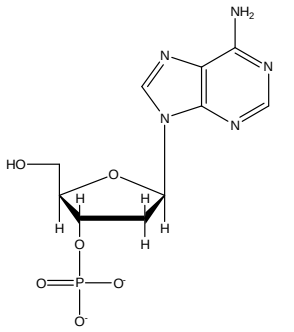
2 DNA, HISTONES AND EXPRESSION

The simplistic view of DNA is a long open and accessible strand of base pairs which can be accessed to produce mRNA via a transcription process. Unfortunately this is far from the fact. The DNA is warped around histones. These histones must open to allow the DNA to become accessible.

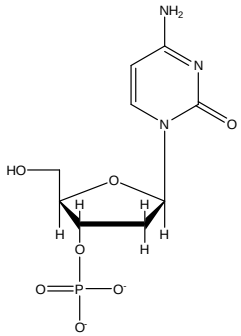
2.1 DNA

We start with the simple paradigm of DNA. This is a double stranded set of nucleotides held together via hydrogen bonds and separable when mRNA is being produced in a somewhat complex manner. The mRNA is then converted to proteins, most of the time and these proteins then do a great deal of the inner and extracellular work of communicating and activating cell processes. Simply stated, this should be an efficient system, however as we readily know the

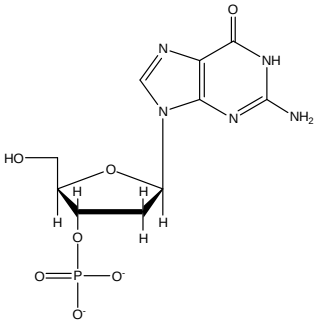
We begin with some simple facts about DNA and then we lead to the methylation of cytosine. But first, the basics of DNA. DNA is composed of just five basic elements; a ribose backbone with phosphates, and four different nucleotides (C, G, A and T). They align in a double stranded classic DNA pattern. The base pairs and their ribose/phosphate backbone parts are shown below.



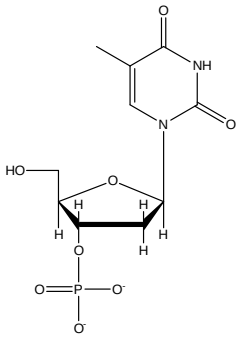
Adenine



Cytosine

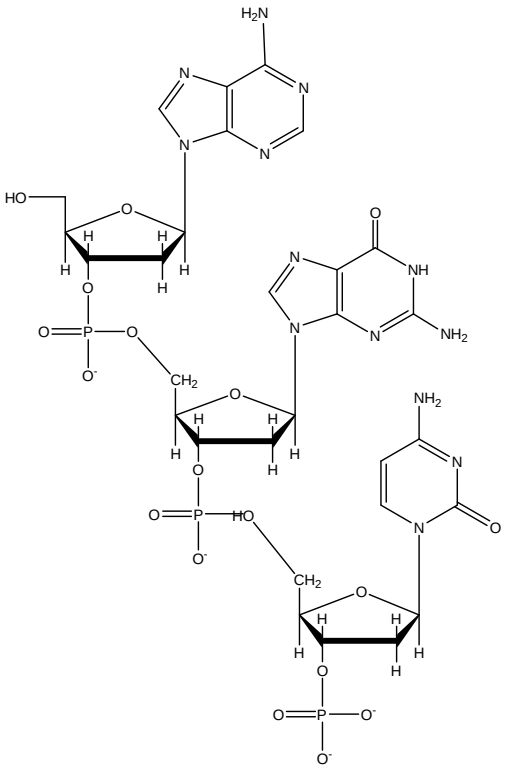


Guanine



Thymine

Now we connect these in the one side of the double helix as is shown below:



Then from here we can connect the A-T and G-C pairs which make up the DNA as we know it.

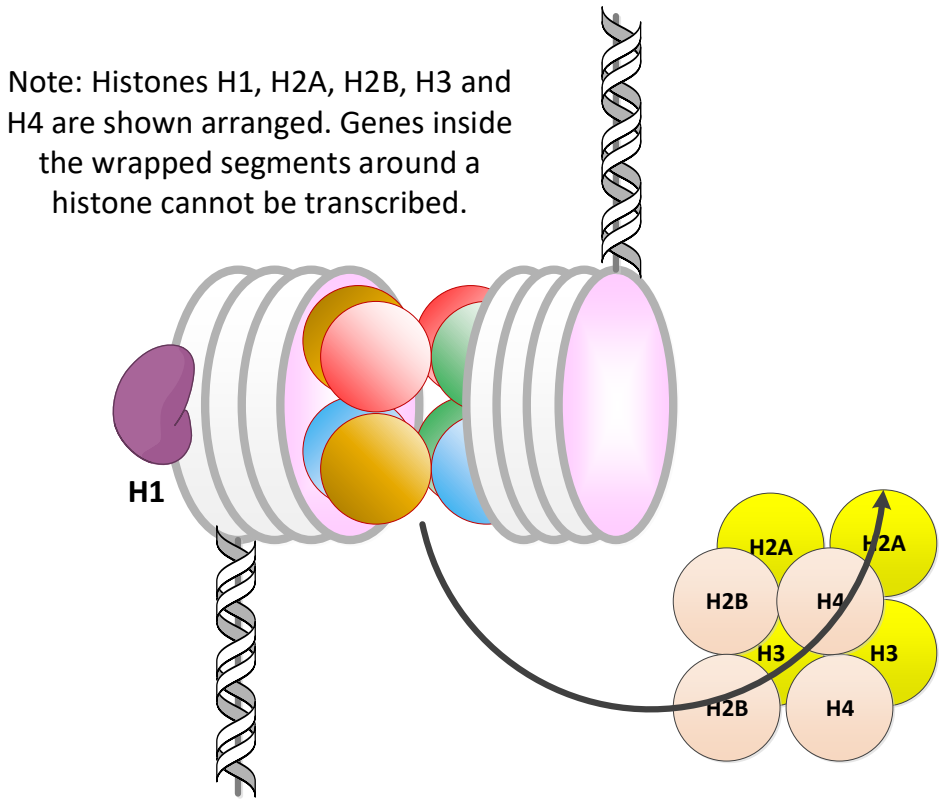
The key observation of Watson and Crick was the hydrogen bonding between base pairs. As Watson and Crick stated in 1953:

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fiber axis. They are joined together in pairs, a. single base from ***one chain being hydrogen-bonded to a single base from the other chain****, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds*

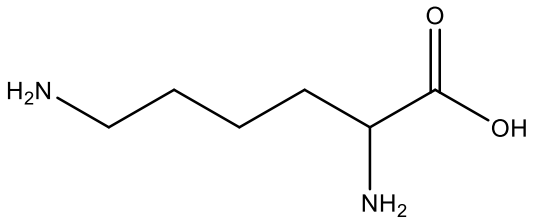
are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

2.2 HISTONES

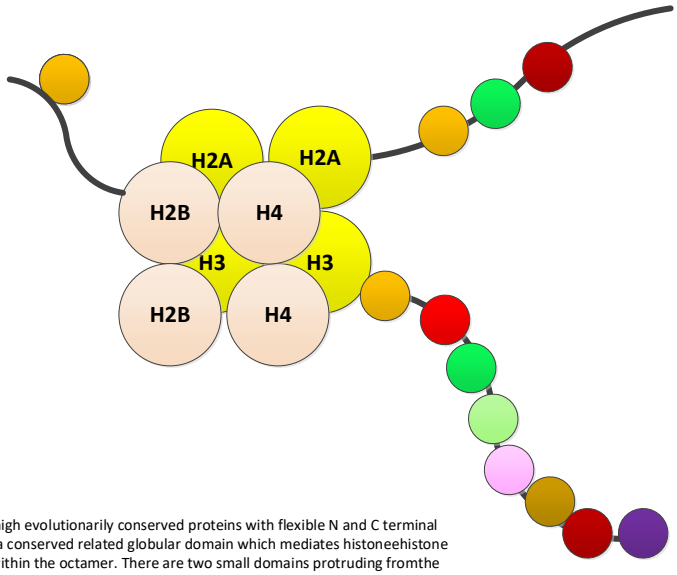
Cellular DNA is not stretched out but tightly wrapped around histones, a collection of eight protein "balls" which allow the DNA to be greatly compressed. In the graphic below we demonstrate this concept. The histones (H2A, H2B, H3, H4) are shown clustered with DNA wrapped around it. Specifically 147 base pairs are wrapped around. Between these histone clusters may be strands of non-clustered DNA. In a cluster the DNA cannot be accessed for reading and for the creation of mRNA. Thus histone sections can be viewed as silencing sections.



Now each of these histone elements have tails of amino acids such a lysine, "K". Lysine is shown below:



Thus the histone tails can be seen as below where lysine and other amino acids are extended beyond the body of the histone.



Histones are high evolutionarily conserved proteins with flexible N and C terminal domains and a conserved related globular domain which mediates histoneehistone interactions within the octamer. There are two small domains protruding fromthe globular domain: an aminoterminal domain constituted by 20e35 residues rich in basic amino acids and a short protease accessible carboxyterminal domain [9e11]. Histone H2A is unique among the histones having an additional 37 amino acids carboxy-terminal domain that protrudes from the nucleosome

Now these tails may be affected by various chemical structure such as methyl like elements. When that happens the normal working may change. In addition these amino acids may already be methylated and this a protein may attack the methylation also changing the characteristics of the specific DNA within it.

3 METHYLATION

Methylation is one of the several epigenetic processes that can result in changes in gene expression. Methylation can suppress or activate expression via multiple paths. We briefly examine two of them; DNA and histone effects. It should be noted that methylation also can be accompanied by acetylation and other such factors each of which has their own effects.

3.1 DNA METHYLATION

DNA methylation is a process whereby the cytosine is changed by the insertion of a methyl group on the 5 carbon of the ring. It is a process which is epigenetic and can dramatically modify gene expression. In fact many of the methylation issue in humans are also common to plants, see the work by Zilberman. There has been a great deal of work demonstrating the impact of methylation on cancer progression; specifically the recent summary by Herman and Baylin, that of Palii and Robertson, that of Robertson and Wolffe, Strathdee and Brown, Calin and Croce, are all worth reviewing.

In this Chapter we examine methylation and its impact on several cancers. We will also examine briefly the causes of methylation as well as the therapeutics in use to modulate cancers that cause or persistence is supported by methylation related products, either directly or indirectly.

In the paper by Das and Singal, the authors define epigenetics in a quite clear manner:

Epigenetics can be described as a stable alteration in gene expression potential that takes place during development and cell proliferation, without any change in gene sequence.

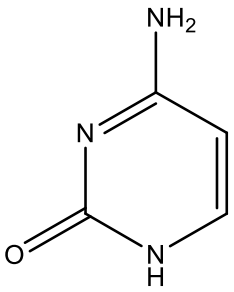
DNA methylation is one of the most commonly occurring epigenetic events taking place in the mammalian genome. This change, though heritable, is reversible, making it a therapeutic target.

Epigenetics has evolved as a rapidly developing area of research.

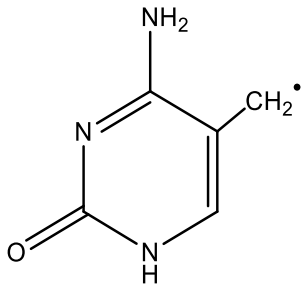
Recent studies have shown that epigenetics plays an important role in cancer biology, viral infections, activity of mobile elements, somatic gene therapy, cloning, transgenic technologies, genomic imprinting, developmental abnormalities, mental health, and X-inactivation

This is one of the clearest definitions of epigenetics and especially the linking of methylation to epigenetics. The classic Watson and Crick model, now some 60 years old, we had the paradigm of DNA, RNA and protein. It was the proteins which did the work. In the 1953 world the proteins stood one by one and the clarity of gene to protein was unquestioned. Yet as we have come to better understand the details, and the details always count, there are many interfering epigenetic factors that all too often get in the way. Methylation is but one of those factors.

Basic cytosine is shown below. It has two NH groups at opposite poles and single oxygen.



Now when the 5 carbon is replaced by a methyl group we obtain the form below. This is methylated cytosine.



Thus this small change in C, by adding the methyl group, can make for a dramatic difference in the expression of genes. For example a well-controlled gene for proliferation, such as PTEN, may have its control over-ridden by the methylation of Introns of CpG islands, namely collections of C, cytosine nucleotides, and G, guanine nucleotides. The introns may be down from the gene, they may even be on a promoter section. The impact could aberrant cell proliferation and growth.

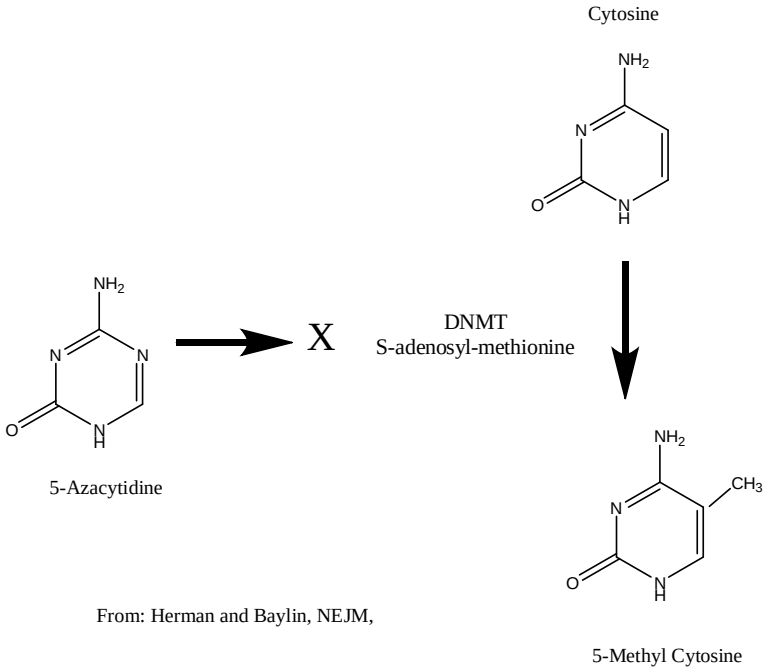
We examine the process; we then look at three types of cancers, a glandular, an epidermal, and a hematopoietic form and then examine some means used to control those cancers through the understanding or methylation and the control of it by therapeutics designed just for that purpose.

What is important about understanding methylation and especially all epigenetic changes is that it may perhaps be simpler to control them rather than a gene mutation. As Brower states:

The move from a purely genetic to an epigenetic model is crucial for prevention strategies. As numerous gene therapy trials have shown, it is very difficult to treat a genetic disease by re-activating the dormant, mutated gene or by replacing it with a non-mutated one. "Epigenetic changes are reversible, and therefore have an edge over genetics," says Mukesh Verma, an epigeneticist at the National Cancer Institute's division of cancer control and population sciences in Bethesda, Maryland. Furthermore, epigenetic changes in cancer occur before genetic mutations. "If you can prevent methylation of those tumour suppressor genes, you might have a valuable prevention strategy," says Baylin.

Thus if we see cancers when they are driven by methylation, then can we actually anticipate reversing the process by reversing the methylation changes. Thus with prostate cancer can we anticipate a preventative measure as one increasing certain methylation preventative therapeutics, can we do the same with say MDS, and can we attempt to do the same with say a melanoma. This is what we examine herein.

What is methylation? Simply, the attachment of a methyl group to the cytosine molecule creates a methylated C. This is not a complicated process but one which happens frequently and may have significant effects. Cytosine gets methylated and is converted to 5-methyl cytosine. This is accomplished by means of two enzymes as depicted below. This occurs when we have a C and G adjacent. It occurs to the C in that pair. We depict that transition below. Note also that by using 5-Azacytadine we can block that transition.



From: Herman and Baylin, NEJM,

Now there are the CpG islands. These are C, cytosine, and G, guanine, adjacent nucleotides which are connected via a phosphodiester bone between the two, and multiple collections of these paired nucleotides. The CpG island is then an area dense in these CG pairs connected by the phosphodiester bond, but the "island" may contain nucleotides other than the CG pairs, but generally are high in CG pair concentration, usually more than 50%.

One should note that the statistical probability of such large CG pairings would normally be quite low. One would anticipate equal probability for any nucleotide and any nucleotide pairing. Furthermore such a high concentration is statistically extremely rare but if often existentially quite common.

The CpG islands may be from 300 to over 3,000 base pairs in total length, and are frequently found in gene promoter regions. Thus when the CpG islands are methylated, namely the C is methylated, then the island gets silenced as does the corresponding gene. Namely methylation of CpG islands can result in gene silencing. This then becomes a critical issue if the gene is a

control gene such as PTEN, p53, or many of the critical pathway control genes. The CpG islands are also propagated to cell progeny during mitosis, thus a methylated island remains so in the cells progeny.

However understanding methylation of islands, and having a means to demethylate the islands may present a reasonable way to develop therapeutics for cancers resulting from methylated regions. We shall examine that shortly.

As Laird and Jaenisch state:

The normal pattern of 5-methylcytosine distribution DNA methylation in mammals is found as a covalent modification at the fifth carbon position of cytosine residues within CpG dinucleotides. Most of the CpG dinucleotides in the human genome are methylated.

However, 5-methylcytosine makes up less than 1% of all nucleotides, since CpG dinucleotides are under-represented about five-fold in the mammalian genome. The paucity of CpG dinucleotides in the mammalian genome is attributed to a higher mutation rate of methylated versus unmethylated cytosine residues.

CpG dinucleotides and 5-methylcytosine are unevenly distributed in the genome. Most of the genome is heavily methylated with a corresponding deficit in CpG dinucleotides. About 1 to 2% of the genome consists of islands of non-methylated DNA and these sequences show the expected frequency of CpG dinucleotides.

CpG islands are about 1 kb long and are not only CpG-rich, but generally G/C-rich as well and are found at the 5' end of genes. All known housekeeping genes and some tissue-specific genes have associated CpG islands.

We now want to discuss methylation and gene expression. Reference will be made to the work of Herman and Baylin, Jones and Takai, McCabe et al, Allis et al, and Issa and Kantarjian.

We begin with Herman and Baylin and their description of the diagram below:

In most of the mammalian genome, which is depicted here as exons 1, 2, and 3 of a sample gene (boxes 1, 2, and 3), introns of the gene (line between the exons), and regions outside the gene, the CpG dinucleotide has been depleted during evolution, as shown by the small number of such sites (circles).

Small regions of DNA, approximately 0.5 to 4.0 kb in size, harbor the expected number of CpG sites and are termed CpG islands. Most of these are associated with promoter regions of approximately half the genes in the genome (numerous circles surrounding and within exon 1 of the sample gene). In normal cells, most CpG sites outside of CpG islands are methylated (black circles), whereas most CpG-island sites in gene promoters are unmethylated (white circles).

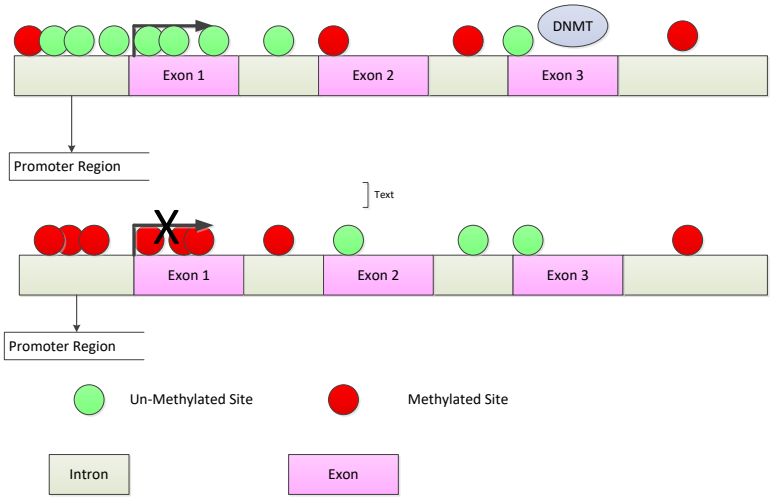
This methylated state in the bulk of the genome may help suppress unwanted transcription, whereas the unmethylated state of the CpG islands in gene promoters permits active gene

transcription (arrow in upper panel). In cancer cells, the DNA-methylation and chromatin patterns are shifted.

Many CpG sites in the bulk of the genome and in coding regions of genes, which should be methylated, become unmethylated, and a growing list of genes have been identified as having abnormal methylation of promoters containing CpG islands, with associated transcriptional silencing (red X at the transcription start site).

Although there are possible explanations and findings from ongoing investigations, it is not known why the DNA-methylating enzymes fail to methylate where they normally would and which of these enzymes are mediating the abnormal methylation of CpG islands in promoters.

We depict a modified version of their Figure below:



Thus methylation in this case blocks the expression of the targeted gene. Methylation may also progress to more dramatic changes. We discuss here the change of C to T, a serious change in a DNA base pair which can result in dramatic changes in gene expression.

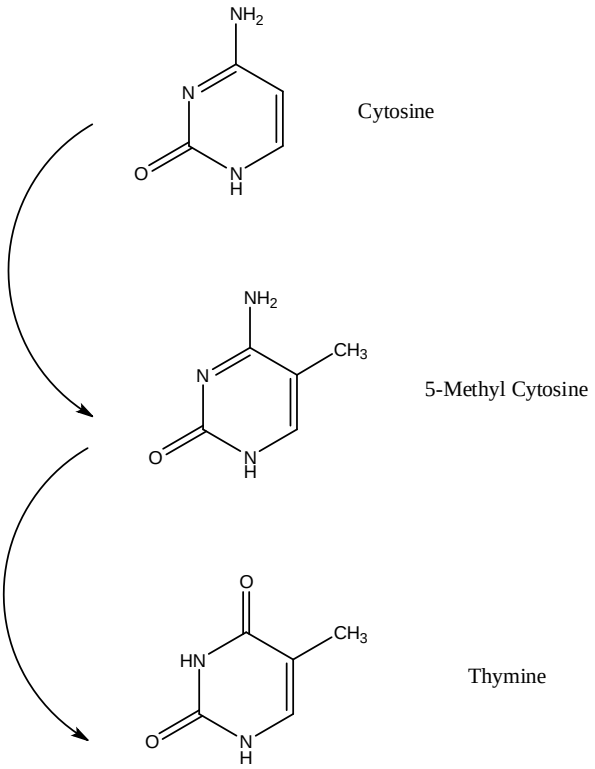
As Herman and Baylin state:

Although only four bases — adenine, guanine, cytosine, and thymine — spell out the primary sequence of DNA, there is a covalent modification of postreplicative DNA (i.e., DNA that has replicated itself in a dividing cell) that produces a "fifth base." Reactions using S -adenosyl-methionine as a methyl donor and catalyzed by enzymes called DNA methyltransferases (DNMTs) add a methyl group to the cytosine ring to form methyl cytosine.

In humans and other mammals, this modification is imposed only on cytosines that precede a guanosine in the DNA sequence (the CpG dinucleotide). The overall frequency of CpGs in the genome is substantially less than what would be mathematically predicted , probably because DNA methylation has progressively depleted the genome of CpG dinucleotides over the course of time.

The mechanism of the depletion is related to the propensity of methylated cytosine to deaminate, thereby forming thymidine. If this mutation is not repaired, a cytosine-to-thymidine change remains.

The depletion of CpG dinucleotides in the genome corresponds directly to sites of such nucleotide transitions, and this change is the most common type of genetic polymorphism (variation) in human populations.



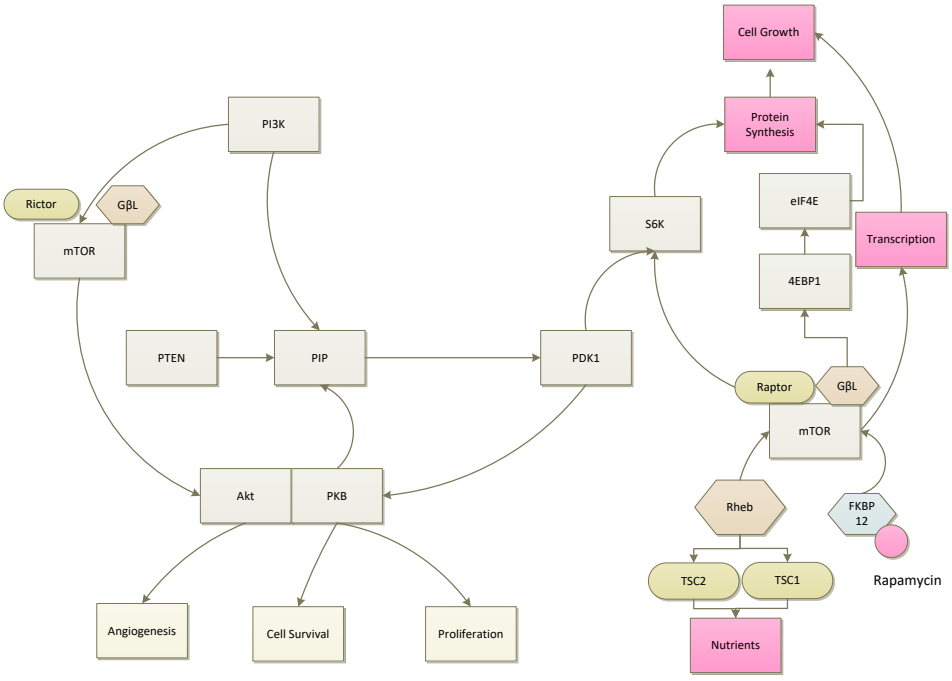
From Robertson (2001) we have some of the genes influenced by methylation or as he states:

CpG-island-associated genes involved in cell growth control or metastasis that can become hypermethylated and silenced in tumors.

We depict the Table below from Robertson on some of the genes impacted by this type of methylation. Most of these are significant regulatory genes.

Gene	*Function*
pRb	Regulator of G1/S phase transition
$p16^{INK4a}$	Cyclin-dependent kinase inhibitor
$p15^{INK4b}$	Cyclin-dependent kinase inhibitor
ARF	Regulator of p53 levels
hMLH1	DNA mismatch repair
APC	Binds β-catenin, Regulation of actin cytoskeleton?
VHL	Stimulates angiogenesis
BRCA1	DNA repair
LKB1	Serine/threonine protein kinase
E-cadherin	Cell - cell adhesion
ER	Transcriptional activation of estrogen-responsive genes
GSTPI	Protects DNA from oxygen radical damage
0^6-MGMT	Repair/removal of bulky adducts from guanine
TIMP3	Matrix metalloproteinase inhibitor
DAPK1	Kinase required for induction of apoptosis by y interferon
p73	Apoptosis structurally similar to p53

For example we show below some typical pathways and the above genes are seen targeted by methylation.



Methylation may then interfere with many of the genes in the above pathways.

The major question which is often asked is what causes methylation. In Allis et al on p 460 the authors discuss some of the putative cause of methylation and methylation related cancers. Although not confirmative it is consistent with clinical correlations as well.

As Issa and Kartarjian state:

Much remains to be learned about the causes of DNA methylation abnormalities in cancer; for the most part, methylation seems to be gene specific. In some cases, a rare methylation event appears in cancer because of selection , while in others methylation anomalies are downstream of an oncogenic event …

As McCabe et al state:

DNA methylation patterns in human cancer cells are considerably distorted. Typically, cancer cells exhibit hypomethylation of intergenic regions that normally comprise the majority of a cell's methyl-cytosine content . Consequently, transposable elements may become active and contribute to the genomic instability observed in cancer cells.

Simultaneously, cancer cells exhibit hypermethylation within the promoter regions of many CpG island-associated tumor suppressor genes, such as the retinoblastoma gene (RB1), glutatione S-transferase pi (GSTP1), and E-cadherin (CDH1). As a result, these regulatory genes are transcriptionally silenced resulting in a loss-of-function. Thus, through the effects of both hypo- and hyper-methylation, DNA methylation significantly affects the genomic landscape of cancer cells, potentially to an even greater extent than coding region mutations, which are relatively rare

McCabe et al continue:

Although the precise molecular mechanisms underlying the establishment of aberrant DNA hypermethylation remain elusive, recent studies have identified some contributing etiologic factors.

For example, chronic exposure of human bronchial epithelial cells to ***tobacco-derived carcinogens drives hypermethylation*** *of several tumor suppressor genes including CDH1 and RASSF2A.*

Stable knockdown of DNMT1 prior to carcinogen exposure prevented methylation of several of these genes indicating a necessary role for this enzyme in the molecular mechanism underlying hypermethylation.

The reactive oxygen species (ROS) associated with chronic inflammation is another source of DNA damage with the potential to affect DNA methylation as halogenated pyrimidines, one form of ROS-induced damage, mimic 5-methylcytosine and stimulate DNMT1-mediated CpG methylation in vitro and in vivo.

Indeed, study of the glutatione peroxidase 1 and 2 double knockout model of inflammatory bowel disease found that 60% of genes that are hypermethylated in colon cancers also exhibit aberrant methylation in the inflamed noncancerous precursor tissues. Although the mechanisms by which DNA damage mediates DNA methylation are not fully understood, O'Hagan and colleagues have examined the process with an engineered cell culture model in which a unique restriction site was incorporated into the CpG island of the E-cadherin promoter.

Thus the actual molecular mechanics leading to methylation are not fully understood but like most cancers inflammation appears to be a driving factor. What the cause of that inflammation may be is not yet clear.

As is stated in the paper by Miranda and Jones:

DNA methylation is a covalent modification in which the 5₀position of cytosine is methylated in a reaction catalyzed by DNA methyltransferases (DNMTs) with S-adenosyl-methionine as the methyl donor.

In mammals, this modification occurs at CpG dinucleotides and can be catalyzed by three different enzymes, DNMT1, DMNT3a, and DNMT3b.DNAmethylation plays a role in the long-term silencing of transcription and in heterochromatin formation.

As an epigenetic modification, DNA methylation permits these silenced states to be inherited throughout cellular divisions.

We continue with the discussion in Mirand and Jones as follows:

Silencing of genetic elements can be successfully initiated and retained by histone modifications and chromatin structure. However, these modifications are easily reversible making them make poor gatekeepers for long-term silencing. Therefore, mammalian cells must possess an additional mechanism for prolong silencing of these sequences. An important component of this process is DNA methylation. DNA methylation is a stable modification that is inherited throughout cellular divisions.

When found within promoters, DNA methylation prevents the reactivation of silent genes, even when the repressive histone marks are reversed. This allows the daughter cells to retain the same expression pattern as the precursor cells and is important for many cellular processes including the silencing of repetitive elements, X-inactivation, imprinting, and development.

We now present a key Figure from Miranda and Jones regarding the methylated reading of DNA. They state regarding the Figure below:

Chromatin structure of CpG islands and CpG poor regions in healthy cells and during cancer. In healthy cells, CpG islands are generally hypomethylated. This allows for an open chromatin structure. However, the CpG poor regions found in repetitive elements within the intergenic and intronic regions of the genome are methylated and thereby maintain a closed chromatin structure. In cancer and on the inactive X chromosome many CpG islands become methylated, forcing these regions into a closed chromatin structure.

When CpG islands located within promoters are methylated, the corresponding genes are persistently silenced. In contrast, the CpG poor regions become hypomethylated allowing for an open chromatin structure.

As Robertson states:

It is now clear that the genome contains information in two forms, genetic and epigenetic. The genetic information provides the blueprint for the manufacture of all the proteins necessary to create a living thing while the epigenetic information provides instructions on how, where, and when the genetic information should be used.

Ensuring that genes are turned on at the proper time is as important as ensuring that they are turned off when not needed.

The major form of epigenetic information in mammalian cells is DNA methylation, or the covalent addition of a methyl group to the 5-position of cytosine predominantly within the CpG dinucleotide. DNA methylation has profound effects on the mammalian genome.

Some of these effects include transcriptional repression, chromatin structure modulation, X chromosome inactivation, genomic imprinting, and the suppression of the detrimental effects of repetitive and parasitic DNA sequences on genome integrity.

Robertson then proceeds to detail the genes impacted by hypermethylation. We summarize them below:

Gene	*Function*
pRb	Regulator of G1/S phase transition
p16 INK4a	Cyclin-dependent kinase inhibitor
p15 INK4b	Cyclin-dependent kinase inhibitor
ARF	Regulator of p53 levels
hMLH1	DNA mismatch repair
APC	Binds b-catenin, Regulation of actin cyto-skeleton?
VHL	Stimulates angiogenesis
BRCA1	DNA repair
LKB1	Serine/threonine protein kinase
E-cadherin	Cell ± cell adhesion
ER	Transcriptional activation of estrogen-responsive genes
GSTP1	Protects DNA from oxygen radical damage
O6-MGMT	Repair/removal of bulky adducts from guanine
TIMP3	Matrix metallo proteinase inhibitor
DAPK1	Kinase required for induction of apoptosis by g interferon
p73	Apoptosis?, structurally similar to p53

Regarding PIN, the one which is most concern is the GSTP1 gene and its suppression allowing for DNA damage from inflammation and oxygenation damage.

In the context of cancer generation and progression, the epigenetic effect of hyper and hypo methylation is best described by Esteller:

The low level of DNA methylation in tumors as compared with the level of DNA methylation in their normal-tissue counterparts was one of the first epigenetic alterations to be found in human cancer.

The loss of methylation is mainly due to hypomethylation of repetitive DNA sequences and demethylation of coding regions and introns −regions of DNA that allow alternative versions of the messenger RNA (mRNA) that are transcribed from a gene. A recent large-scale study of DNA methylation with the use of genomic microarrays has detected extensive hypo-methylated genomic regions in gene-poor areas.

During the development of a neoplasm, the degree of hypomethylation of genomic DNA increases as the lesion progresses from a benign proliferation of cells to an invasive cancer.

Three mechanisms have been proposed to ex-plain the contribution of DNA hypomethylation to the development of a cancer cell:

(i) generation of chromosomal instability,

(ii) reactivation of transposable elements, and

(iii) loss of imprinting.

Under methylation of DNA can favor mitotic recombination, leading to deletions and translocations, and it can also promote chromosomal rearrangements. This mechanism was seen in experiments in which the depletion of DNA methylation by the disruption of DNMTs caused aneuploidy. Hypomethylation of DNA in malignant cells can reactivate intra-genomic endo-parasitic DNA.

3.1.1 Hypomethylation

As Laird and Jaenisch state:

Hypomethylation: Reduced levels of global DNA methylation have been reported for a variety of malignancies in the past decade. Gama Sosa and coworkers found that in a wide variety of tumors, hypomethylation not only correlated with transformation, but also with tumor progression . In their analysis, only 7% of 43 normal tissues had a 5-methylcytosine content below 0.8 mol%, whereas 10% of 21 benign tumors, 27% of 62 primary malignancies and 60% of 20 secondary malignancies had a 5-methylcytosine content below 0.8 mol%. On the other hand, Feinberg and coworkers did not find a further reduction in DNA methylation levels in the progression from benign to malignant colonic neoplasia, suggesting an early role for DNA hypomethylation in colorectal cancer

3.1.2 Hypermethylation

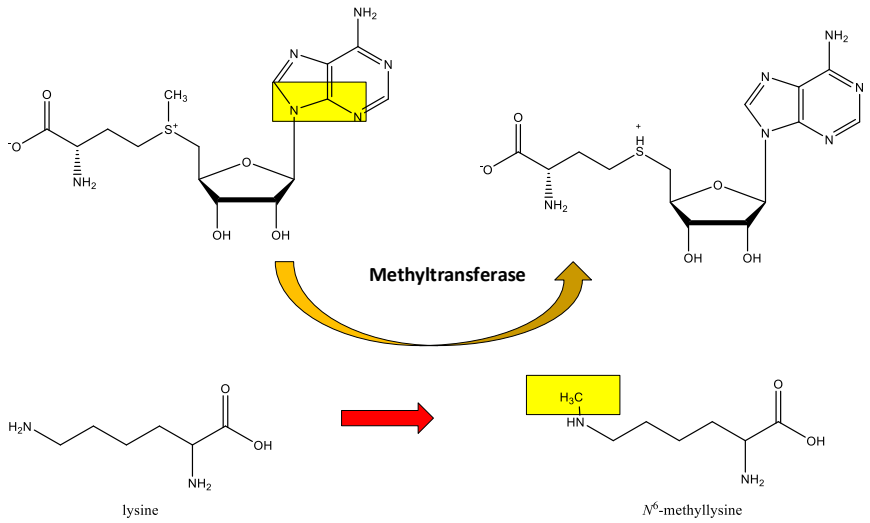
As again with Laird and Jaenisch we have:

Hypermethylation: There have also been many reports of regional increases in DNA methylation levels. Baylin and coworkers have found regional hotspots for hypermethylation on chromosomes 3p, 11p and 17p in a variety of human tumors. These include CpG island areas that are normally never methylated in vivo, but are found to be methylated in tumor tissues. This is reminiscent of the changes that occur at CpG islands at non-essential genes in tissue culture. Baylin's group has dissected the sequential order of hypermethylation events in an in vitro model for lung tumor progression. There is evidence for inactivation of tumor-suppressor gene function through hypermethylation of the Rb gene in sporadic retinoblastoma. Transient transfection experiments showed that specific hypermethylation in the promoter region of Rb could reduce expression to 8% of an unmethylated control. It is possible, therefore, that hypermethylation of tumor-suppressor genes leading to gene inactivation results in a selective growth advantage of the transformed cells.

3.2 HISTONE METHYLATION

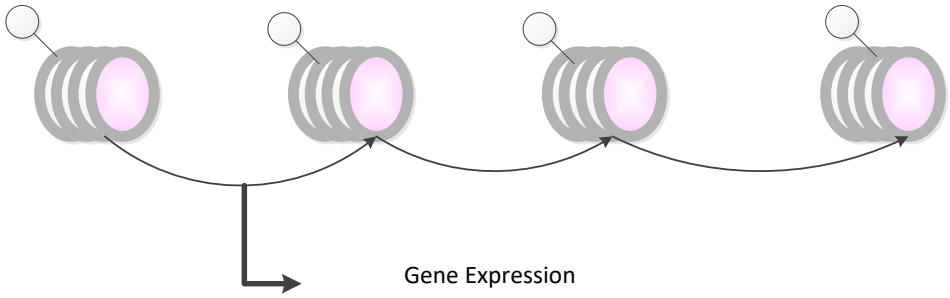
Histones can also be methylated. Here the effects can be materially more complex. The methylation can occur on the tails sticking out and each histone element has a tail and thus there interaction can be amongst the same histone tail or even more so between adjacent histone tails. This interaction results in the opening and closing of DNA, making for the expression or suppression of many genes.

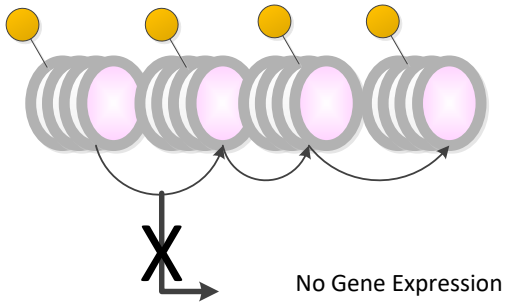
A typical methylation is enabled by an enzyme called a methyltransferase. We show this below.



Namely the lysine on a tail can be methylated on the distal end by transferring the methyl group using the enzyme methyltransferase. Lysine is not methylated and as such as acquired a substantially different set of attachment properties. A similar process can occur with acetylation. These processes are also reversible.

Now when we examine a histone complex it may appear as follows:





The top part shows no methylation and as such we have the DNA open between histone clusters. Now if we were to methylate them we may see the case shown below it, as above, where now we cannot insert a transcription factor and other necessary elements to effect transcription. Thus

As Helin and Dhanak note:

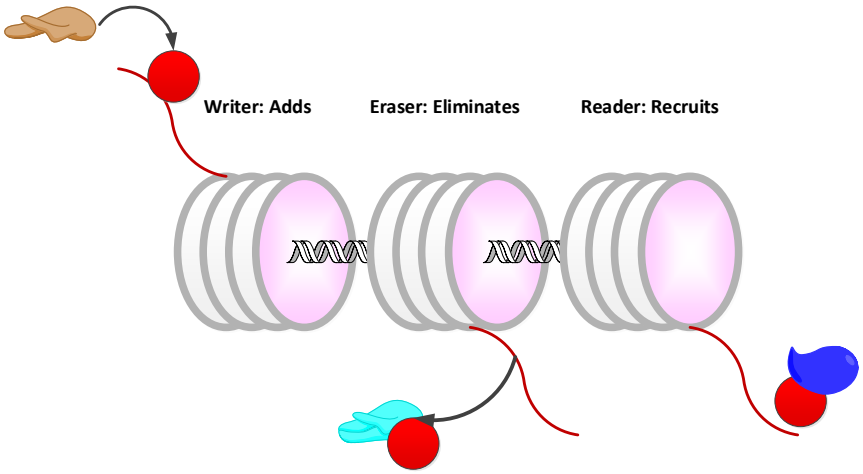
DNA is wrapped around histones (H2A, H2B, H3 and H4) to form nucleosomes. Nucleosomes are further compacted to form condensed chromatin. The compaction of DNA is in part regulated through post-translational modifications (PTMs) of the histone tails, which protrude from nucleosomes. Epigenetic regulators can in popular terms be divided into erasers, writers or readers of PTMs.

The erasers, such as histone deacetylases and histone demethylases, remove the PTMs and prepare the histones for other modifications.

The writers comprise enzymes such as histone acetylases, kinases, DNA and histone methyltransferases and ubiquitin ligases. The writers catalyse the PTMs on the DNA or the proteins, and may impose epigenetic heritability such as DNA methylation through copying and maintaining the modification.

Other modifications, such as histone acetylation, respond rapidly to environmental stimuli and are therefore more dynamic.

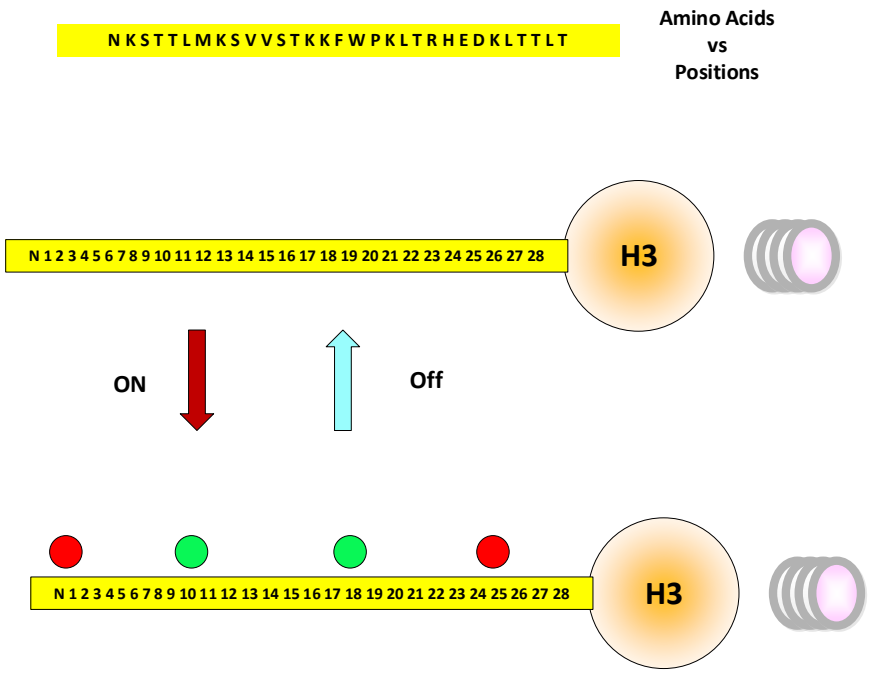
Readers of the post-translational modification include proteins with specific domains, such as bromo-, chromo-, tudor-, MBT-, PWWP-, WD40- and PHD-domains, which bind to the specific modification. The readers, which are often found in large protein complexes, interpret the modification and impose changes in chromatin structure.



The above depicts graphically these three processes. The work by Gut and Verdin further discuss this process as a temporal one, inherent in ageing of the genetic environment.

3.3 THE HISTONE CODE

The Histone Code was described by Strahl and Allis in 2000 and it can be simply explained as follows.



In the above we have a tail and tail locations and respective amino acids for each location. Now on the top there are no methylations or acetylations. We have then done so on the one below. We can assert that in the top condition we have the base state and then the one below some active state. Thus we go from off to on whatever that may mean. Thus as Strahl and Allis note in their presentation we have:

N	*1*	*2*	*3*	*...*	*27*	*28*	*Modification State*	*Associated Protein*	*Function*
	M						Methylated	SIRT	Silencing
		M					Methylated	SMC	Transcription
				M	M		Methylated	RCAF	Mitosis
	A						Acetylated	Bromodomain	Transcription
		A		P	M		Complex	TWIST	Silencing

Namely the histone code postulates what reaction will ensue when we have some form of epigenetic change on a specific tail of a specific histone and it indicates what protein is necessitated to effect this epigenetic change.

Now the histone code relates to the state of the tail as described by methylations or other related attachments and the resulting actions related thereto.

The above demonstrates the tail composed of a collection of amino acids and the extension of that from each of the histone elements. These tails allow for reactions which in turn result in changes of gene expression. As we shall see, the protein we are focusing on, NSD2, is a histone modifying protein and it targets a specific amino acid on the histone. In this case it targets H3K36me3. This nomenclature states:

1. Histone H3
2. K for lysine

3. Location 36 on the tail
4. methylated
5. tri methylated

Thus the notation can be specific as to the tuple:

{*histone:amino acid:location:modification:degree*} =- H3K36me3.

As Jenuwein and Allis had noted in 2000:

Chromatin, the physiological template of all eukaryotic genetic information, is subject to a diverse array of posttranslational modifications that largely impinge on histone amino termini, thereby regulating access to the underlying DNA. Distinct histone amino-terminal modifications can generate synergistic or antagonistic interaction affinities for chromatin-associated proteins, which in turn dictate dynamic transitions between transcriptionally active or transcriptionally silent chromatin states. The combinatorial nature of histone amino-terminal modifications thus reveals a "histone code" that considerably extends the information potential of the genetic code.

From Tollefsbol we have:

Equally important in the fine tuning control of chromatin organization is the interplay between the histone modifications, DNA methylation and ATP-dependent chromatin remodeling. The large number of histone modifications and the possible interplay between them led to the proposition of the so-called "histone code hypothesis" in which "multiple histone modifications, acting in a combinatorial or sequential fashion on one or multiple histone tails, specify unique downstream functions". This hypothesis led the scientific community to adopt some metaphors to describe it such that the code is written by some enzymes ("writers"), removed by others ("erasers"), and is readily recognized by proteins ("readers") recruited to modifications through the binding of specific domains.

Such a simplified version of the code is depicted below. Here we have depicted it differently. The top row is the histone element, the column the change which is made, and the cell entity is what happens when that change in epigenetic structure is made. This of course is a highly simplified result.

	H3K4	H3K9	H3K14	H3K27	H3K79	H4K20	H2BK5
Mono-meth	Active	Active		Active	Active	Active	Active
Di-meth		Repress		Repress	Active		
Tri-meth	Active	Repress		Repress	Active		Repress
Acetyl		Active	Active		Repress		

More complicated versions are available, In 2000 Strahl and Allis noted:

The `histone code' hypothesis. Histone modifications occur at selected residues and some of the patterns shown have been closely linked to a biological event (for example, acetylation and transcription). Emerging evidence suggests that distinct H3 and H4 tail modifications act sequentially or in combination to regulate unique biological outcomes. How this hierarchy of multiple modifications extends (depicted as `higher-order combinations') or how distinct combinatorial sets are established or maintained in localized regions of the chromatin fiber is not known. Relevant proteins or protein domains that are known to interact or associate with distinct modifications are indicated. The CENP-A tail domain might also be subjected to mitosis-related marks such as phosphorylation; the yellow bracket depicts a motif in which serines and threonines alternate with proline residues

From Tollefsbol (see Fig 4.1 Chapter 4) we have another slightly more complicated version:

Chromatin Modification	***Residues modified***	***Function regulated***
Acetylation	Lysine	Transcription, DNA repair, replication and condensation
Methylation (Lysine)	Lysine me1, me2, me3	Transcription, DNA repair
Methylation (Arginine)	Arginine-me1, Arginine-me2a Arginine-me2s	Transcription
Phosphorylation	Serine, Threonine, Tyrosine	Transcription, DNA repair and condensation
Ubiqutination	Lysine	Transcription, DNA repair
Sumoylation	Lysine	Transcription
ADP ribosylation	Glutamic	Transcription
Deimination	Arginine	Transcription
Proline isomerization	P-cis, P-trans	Transcription

In summary we can articulate this as follows:

1. A base state is present and in the base state the genes follow the base state expression.

2. A methyltransferase or equivalent is introduced. This means that it is activated by some means. We leave that to the side for the moment.

3. The methyltransferase targets a specific histone tail element. It then methylates that element.

4. The methylated tail then reconfigures the histone arrangement, opening or closing sections of DNA.

5. DNA expression is altered as a result of the change in the histone configurations. Proteins are produced which are then sent from the nucleus or kept there.

6. The new proteins commence the actions for which they function. Cells then proliferate, go through epithelial-mesenchymal transitions and the like.

Conceptually this is a simple process but in actuality there are a multiplicity of questions as to what and why.

4 METHYLATION AND CANCER

We now examine the impact of methylation on several cancers. We have selected three different types:

1. Glandular: This is prostate cancer. Many adenocarcinomas are typical of this type. Glands seem often to be the source of cancers and one could surmise it is because they are continually active cell sites with high mitotic activity.

2. Epidermal: We select melanoma as an example. This is especially interesting because it is a cancer which is often attributed to UV radiation, since the melanocytes are so close to the skin surface, a few dozen keratinocytes deep.

3. Hematopoietic: The majority of blood/bone generated cancers result from a variety of changes. CML is a classic model with a Philadelphia chromosome abnormality, a translocation. MDS, myelodysplastic syndrome, on the other hand, is a pre-cancerous state where hypermethylation is the driving factor. This is interesting in that unlike CML which has a clear genetic change, MDS has a clear hypermethylated state. It may result in a genetic change and thus AML but the progression may be mitigated by drugs which mitigate methylation.

We examine the literature on each as regards to methylation impact. As Helin and Dhanak have noted:

In the context of cancer, the discovery of genetic alterations in HMTs in several different tumour types has undoubtedly attracted much attention and provided additional support for the importance of epigenetic deregulation in a disease that is widely considered to be genetically driven. In some cases (such as the methyltransferase EZH2, discussed later), heterozygous point mutations in the catalytic SET domain lead to a gain of function of the wild-type enzyme, favouring trimethylation and the silencing of tumour suppressor genes and/or differentiation-specific genes.

*Similarly, in other cancers (**such as, increased expression of NSD2 in multiple myeloma**) chromosomal translocations result in increased expression of the methyltransferases, again leading to aberrant transcription and proliferation17. Conversely, lysine methylation induced by the HMT DOT1L results in sustained expression of several genes required for leukaemogenesis. Therefore, small molecule inhibitors, of for instance EZH2 or DOT1L, should be able to reduce or eliminate the site-specific lysine methylation introduced by the HMTs and reverse the oncogenic state*

4.1 PROSTATE

Prostate cancer is a complex malignancy of a glandular element. It may be indolent or highly aggressive, and at this time it is quite difficult to determine the difference based solely on pathology examination. One of the themes we shall see in methylation and cancers will be exogenous effects such as sunlight in melanoma, such as chemicals and radiation in MDS and

such as free radicals and infections in prostate cancer. These factors all seem to impact methylation.

In a recent (2013) paper by Vasiljevic et al they state:

Our data indicate CpG methylation of the first HSPB1 intron to be an important biomarker that identifies aggressive PCas otherwise regarded as low risk by current clinical criteria but that, biologically, require immediate active management.

This is a very powerful conclusion. It is a step to identifying indolent from aggressive. They continue:

Heat shock protein 27 (Hsp-27), encoded by the gene HSPB1 located on chromosome 7q11.23 has been shown in several independent studies to be a reliable biomarker of poor clinical outcome in human prostate cancer (PCa) as well as in human breast cancer, colorectal cancer and malignant melanoma.

Biologically, Hsp-27 is an anti-apoptotic protein that induces intracellular homeostasis and allows cellular repair and recovery after physical and chemical insults. Although Hsp-27 is constitutively expressed in most human cells, induced overexpression during carcinogenesis can lead to increased survival of the malignant cells.

Therefore, it is not surprising that studies link high expression of Hsp-27 to unfavorable prognosis in many cancer types. The prognostic potential has been confirmed in prostate cell lines14 as well as in prostate tissues where overexpression has been linked with hormone resistance and poor clinical outcome.

During early prostate carcinogenesis, expression of Hsp-27 protein becomes universally abrogated but may be re-expressed subsequently, in which case the malignancy develops an aggressive phenotype.

Although the specific factors controlling these changes are presently unknown, one plausible mechanism is DNA methylation (DNAme) of the HSPB1 gene. The majority of CpG dyads in the human genome are methylated with the exception of CG-rich regions called CpG islands.16 CpG islands mainly cover gene promoters and first exons and their hypermethylation is associated with repressed transcription of many tumor-suppressor genes.

Therefore, we test the hypothesis that the DNAme status of HSPB1, particularly the HSPB1 promoter, exon and intron regions, is an important determinant of PCa behavior.

Thereafter, we assess any potential relationship between DNAme and Hsp-27 protein levels. Our objectives are also to investigate the diagnostic biomarker potential, by comparing the methylation status of BPH vs PCa, and the prognostic potential of DNAme, by analyzing the association between the methylation and PCa-specific death in the well-characterized Transatlantic Prostate Group (TAPG) cohort.

They conclude:

In conclusion, HSPB1 is essentially unmethylated in BPH but with increasing neoplastic changes through to PCa, the gene becomes increasingly methylated, proceeding from the promoter in a 3' direction. In PCas with low Gleason score, higher methylation within the HSPB1 gene independently identifies patients with poor clinical outcome and hence is an objective biomarker identifying the immediate need for active intervention in the clinical management of this cohort of patients.

This is a powerful observation and sets the path for improved prognostics on PCa.

In an older paper by O'Shaughnessey et al they state:

PIN and prostate cancer lesions share a number of somatic genome abnormalities, including loss of DNA sequences at 8p and increased GSTP1 CpG island DNA methylation, among others. Finally, transgenic mouse strains prone to developing prostate cancers typically develop PIN lesions in advance of the appearance of invasive cancer.

We have discussed elsewhere the HGPIN issue regarding PCa and the questions raised by the assumed linear progression from HGPIN top PCa, except in certain cases where we hypothesize the removal of stem calls upon biopsy.

4.2 MELANOMA

Melanoma is a solid tumors which has the tendency to metastasize very rapidly. Melanoma is fundamentally a malignancy of the melanocytes and the melanocytes are often changed into a malignant state due to their proximity to the skin surface, at the basal layer of the epidermis, and the impact of UV light on their progress. We have argued elsewhere that methylation of portions of the DNA due to such factors as backscatter radiation may be a significant factor as well. The skin being so thin absorbs the radiation more strongly that the viscera and thus is millions of times more sensitive.

From Bennett we have:

A primary event in progression would be a cellular change that is clonally inherited, that contributes to the eventual malignancy, and that occurs independently rather than as a secondary result of some other oncogenic change.

*These events are either genetic (gene mutation, deletion, amplification or translocation), or epigenetic (a heritable change other than in the DNA sequence, **generally transcriptional modulation by DNA methylation** and/or by chromatin alterations such as histone modification). In clonal evolution of cancer, such a primary event would initiate a new, more progressed, clone with a growth advantage over its neighbors, or an alternative selective advantage such as migration…. The β-catenin pathway can be upregulated by several kinds of primary and secondary changes in melanoma. These include uncommon activating mutations of b-catenin (CTNNB1) itself, **methylation or mutation of APC**, overexpression of proto-oncoprotein SKI….*

In a recent paper by Mazar et al they report on melanoma as follows:

Here, we report that cell lines derived from malignant melanomas and melanoma patient samples have hypermethylated CpG islands in the 59-upstream regions of several miRNA coding genes, including that of miR-34b. We engineered two cell lines derived from metastatic melanoma to ectopically express miR-34b, and show that these cells exhibit reduced cell motility, decreased substrate attachment, and reduced invasion.

They continue:

The reduced expression of genes that are under the control of CpG island methylation is often reversed by treating the cells with the DNA methyl transferase inhibitor 5-Aza-29-deoxycytidine (5-Aza-dC). To assess the range and extent of miRNA expression under direct or indirect control of DNA methylation, we treated the melanoma cell line WM1552C (derived from a stage 3 malignant melanoma) with 5-Aza-dC and measured changes in miRNA gene expression using miRNA microarrays (see Methods). Several miRNAs, including miR-34b, -489, -375, - 132, -142-3p, -200a, -145, -452, -21, -34c, -496, -let7e, -654, and -519b, were found to be up-regulated

They conclude:

During melanoma formation, the initial genetic or epigenetic changes are thought to precede additional mutations and further epigenetic changes that affect the function of several signaling pathways. Aberrant DNA methylation patterns at the 59 noncoding region of the INK4a gene was discovered in melanoma , which is consistent with the involvement of epigenetic factors in melanoma development or progression.

Similarly, epigenetic silencing of PTEN expression occurs in certain malignant melanomas with no detectable mutation in the PTEN gene.

While the impact on melanoma development of epigenetic changes in several protein-coding genes is appreciated, there have been few reports of the impact of epigenetic regulation of noncoding RNAs, such as miRNAs.

The epigenetic modification of miR-34b may serve as a useful biomarker for early melanoma detection in humans, and therefore, one could propose to develop a novel sensitive miR- 34b epigenetic biomarker assay to screen skin biopsies in melanoma patients. Including a panel of non-coding RNA epigenetic markers in to widely used pathological and genetic markers will be advantageous for both patients and pathologists.

An investigation of miR-34b regulation and associated CpG island methylation in a large group of melanoma patient samples, in comparison with samples of matched normal tissues or melanocytic nevi, is both relevant and timely. Mir-34 group of miRNAs are known to be useful therapeutic target for various cancers...

The PTEN control of cell proliferation is well known. However here it is shown that methylation can suppress PTEN without a genetic modification. Methylation is thus a powerful tool that

surpasses genetic changes. Melanoma is an intriguing cancer because the effects of the environment are so well identified. Upon biopsy one can determine the extent of sun damage and ageing. Thus we can determine how much potential methylations effects are present as well.

4.3 MYELODYSPLASTIC SYNDROME

Myelodysplastic Syndrome is an uncommon hematological cancer mostly caused by excess exposure to radiation, chemicals such as benzene, and insecticides. The specific genetic causes are still a work in progress. However, there are certain therapeutics which address some of the pathway aberrancies which characterize the disease, specifically hypermethylation.

As Taferri and Vardiman state:

*According to the 2008 World Health Organization (WHO) classification system for hematologic cancers, the primary myelodysplastic syndromes are one of five major categories of myeloid neoplasms. The main feature of myeloid neoplasms is stem-cell–derived clonal myelopoiesis with altered proliferation and differentiation. The phenotypic diversity of these neoplasms has been ascribed to different patterns of dysregulated signal transduction caused by transforming mutations that affect the hematopoietic stem cell. There is increasing evidence that haploinsufficiency, **epigenetic changes**, and abnormalities in cytokines, the immune system, and bone marrow stroma all contribute to the development of the myelodysplastic syndromes.*

Thus MDS is both complex in presentation and complex in development. Melanoma and prostate cancer are more clearly characterized morphologically and generally in genetic development. The presentation may involve the white cells, red cells or platelets, or any combination thereof. It is often discovered as an incidental finding on a blood test with lowered amounts of one or several of the constituents. If it has progressed more it may also present in the bone biopsy with more than normal blasts, immature cells.

As DeVita et al state:

Myelodysplastic syndromes (MDSs) are a group of complex and heterogeneous clonal hematopoietic stem cell disorders whose defining characteristics are dysplasia of one or several hematopoietic cell lineages, hypercellular marrows, and blood cytopenias.

1 Although historically considered as a preleukemic state, most patients with MDS do not transform into an acute myeloid leukemia (AML), but will instead succumb to complications of persistent cytopenias. Indeed, the pathophysiology of MDS extends from immune-mediated mechanisms and excessive apoptosis resulting in marrow failure to arrest of maturation and proliferation resembling the mechanisms at play in AML.

2 The diverse pathophysiology of factors that contribute to the development of MDS is reflected in vast differences of patients' prognosis, which is increasingly recognized and reflected in the design of more elaborate systems of diagnosis, classification, and prognostication.

Let us begin with a simple set of statements regarding the micro RNA elements which are often seen at the heart of the disease. As Croce states:

Several of the miRNAs that have been described as suppressors have been found to be deleted or mutated in various human malignancies. For example, loss of miR-15a and miR-16-1 has also been observed in prostate cancer and multiple myeloma (TABLE 1). Members of the miR-29 family have been found to be deleted in a fraction of myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML) patients.

As Croce further states:

MicroRNAs as targets of epigenetic changes. The most studied epigenetic changes in cancer cells are the methylation of cytosines in the dinucleotide CpG in DNA62. Such 'methylable' sites, known as CpG islands, are preferentially located in the 5' region (which consists of the promoter, 5' uTR and exon 1) of many genes, are non-methylated in normal cells and are transcribed in the presence of the appropriate transcription factors. Methylation of the CpG islands of tumour suppressors results in their silencing and contributes to malignant transformation.

As mentioned above, the expression of miRNAs can be affected by genetic changes, such as deletion, gene amplification and mutation, and by transcription factors. In addition, the expression of miRNAs can be affected by epigenetic changes, such as methylation of the CpG islands of their promoters. Saito et al. reported that miR-127 is silenced by promoter methylation in bladder tumours and that its expression could be restored by using hypomethylating agents such as azacitidine.

This miRNA targets BCL6, an oncogene that is involved in the development of diffuse large b cell lymphoma. Therefore, the silencing of miR-127 may lead to the overexpression of bCL6. Other investigators have described additional miRNAs that are silenced by methylation in various cancers and that can be reactivated by hypomethylating agents.

As Das and Singal state:

Hypermethylation is associated with many leukemias and other hematologic diseases. *Many genes, such as the calcitonin gene, p15INK4B, p21Cip1/Waf1, the ER gene, SDC4, MDR, and so on, were seen to be hypermethylated in a variety of hematologic cancers.*

The calcitonin gene and p15 were hypermethylated in 65% of myelodysplastic syndromes*, and it was found that p15 methylation at diagnosis was associated with lower survival and transformation to acute myeloid leukemia.*

Also acquisition of p15 methylation at a later date signaled disease progression. These may suggest the role of p15 as a marker of leukemic transformation. Acute myeloid leukemia demonstrated frequent hypermethylation of ER, MYOD1, PITX2, GPR37, and SDC4

Thus MDS is closely related to methylation, and in effect is caused by methylation. In addition as we show below its management is also performed through an understanding of methylation and managing that process.

Understanding the impact of methylation in MDS recent efforts have led to certain therapeutics which have been of help.

As Issa and Kantarjian state:

Two nucleoside inhibitors of DNA methylation, azacitidine and decitabine, are now standard of care for the treatment of the myelodysplastic syndrome, a deadly form of leukemia. These old drugs, developed as cytotoxic agents and nearly abandoned decades ago were resurrected by the renewed interest in DNA methylation.

They have now provided proof of principle for epigenetic therapy, the final chapter in the long saga to provide legitimacy to the field of epigenetics in cancer. But challenges remain; we don't understand precisely how or why the drugs work or stop working after an initial response. Extending these promising findings to solid tumors faces substantial hurdles from drug uptake to clinical trial design.

We do not know yet how to select patients for this therapy and how to move it from life extension to cure. The epigenetic potential of DNA methylation inhibitors may be limited by other epigenetic mechanisms that are also worth exploring as therapeutic targets. But the idea of stably changing gene expression in vivo has transformative potential in cancer therapy and beyond.

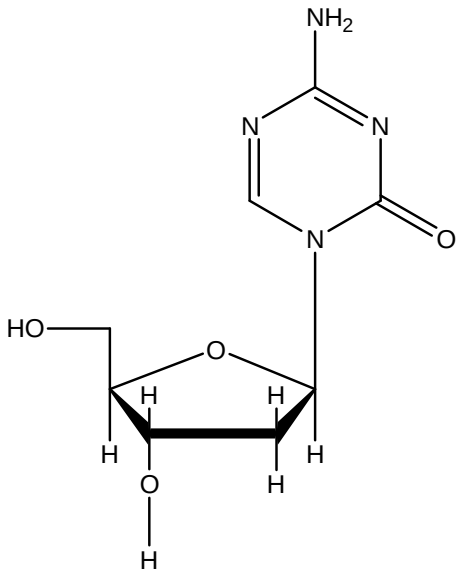
As Li has stated:

The strategies targeting DNA methylation. Epigenetic control of gene expression by DNA methylation has a great impact on cell proliferation and differentiation. Hypermethylation of promoter regions results in specific suppression of gene expression, including the expression of tumor suppressors, which could promote cancer development.

Conversely, demethylation of DNA may enhance cell apoptosis or reduce cell growth. This concept has been proven by a recently approved anticancer drug decitabine for the treatment of myelodysplastic syndrome. Decitabine (Dacogen; MGI Pharma) is a nucleoside analogue that inhibits DNA methylation.

It demethylates the p73 promoter and induces reexpression of p73, thus activating the caspase cascade and leading to leukemic myeloid cell death.26 DNA hypermethylation in tumor cells may be involved in resistance to interferon (INF)-induced apoptosis, and inhibition of DNA methylation may also enhance the therapeutic effect of INF. Treatment of cancer cells with specific DNA demethylating nucleoside analogue was shown to augment the effect of INF.

Now decitabine is shown below in detail. It is a cytosine derivative with several modifications. It functions in a manner similar to azacitidine. We have discussed that previously.



From Boumber et al we have the following regarding therapeutics for epigenetic drugs:

What Is Epigenetic Therapy? The understanding that epigenetic changes are prevalent in cancer and play a causative role in its biology has led to the development of new therapeutic approaches that target the epigenetic machinery. The first successful drugs developed as epigenetic agents were DNA methyltransferase inhibitors; these were followed by histone deacetylase inhibitors (HDIs).

Both classes of drugs aim at reversing gene silencing and demonstrate antitumor activity in vitro and in vivo. Several other classes of drugs have been developed that target various other components of the epigenetic machinery; one such class is the histone methyltransferases, with new drugs in this class currently in early preclinical development

The authors continue:

What Has Been Done? The inhibitors of DNA methylation used clinically are nucleoside analogues that get converted into deoxy-nucleotide-triphosphates (dNTPs) and become incorporated into DNA in place of cytosine during DNA replication. They trap all DNA methyltransferases and target them for degradation. At low doses these drugs do not inhibit proliferation; they reactivate gene expression and have shown clinical activity as anticancer agents. Azacitidine was the first hypomethylating agent approved by the FDA; its approval, in 2004, for the treatment of myelodysplastic disorders and leukemia, was followed by the approval, in 2006, of decitabine. Both drugs produce remissions or clinical improvements in more than 30% of patients treated. Features of responses have included the requirement for multiple cycles of therapy, slow response, and relatively few side effects. On the molecular level, demethylation, gene reactivation, and clonal elimination were observed in treated patients. The data in myelodysplastic syndrome (MDS) represent a proof-of-principle for epigenetic therapy for cancer, in particular in myeloid disorders.

From Boumber et al we have the following Table of many of the recent therapeutics:

Drug Class	Compound
DNMT Inhibitor	Azacitidine
	Decitabine
	S110
	CP-400
	Nanaomycin
HDAC Inhibitor	Vorinostat
	Romidepsin
	Panobinostat
	Valproic Acid
	Belinostat
HMT Inhibitor	Deazaneoplanocin
	Quinazoline
	Ellagic Acid
Histone demethylase inhibitor	Polyamine analogues
	Hydroxamate analogs
GAT inhibitor	Spermidinyl
	Hydrazinocurcumin
	Pyrazolone

As Stressman et al state:

Aberrant DNA methylation patterns play an important role in the pathogenesis of hematologic malignancies.

The DNA methyltransferase inhibitors azacytidine and decitabine have shown significant clinical benefits in the treatment of myelodysplastic syndrome (MDS), but their precise mode of action remains to be established. Both drugs have been shown the ability to deplete DNA methyltransferase enzymes and to induce DNA demethylation and epigenetic reprogramming in vitro. However, drug-induced methylation changes have remained poorly characterized in patients and therapy-related models.

We have now analyzed azacytidine-induced demethylation responses in myeloid leukemia cell lines. These cells showed remarkable differences in the drug-induced depletion of DNA methyltransferases that coincided with their demethylation responses. In agreement with these data, DNA methylation analysis of blood and bone marrow samples from MDS patients undergoing azacytidine therapy also revealed substantial differences in the epigenetic responses of individual patients.

Significant, transient demethylation could be observed in 3 of 6 patients and affected many hypermethylated loci in a complex pattern. Our results provide important proof-of-mechanism data for the demethylating activity of azacytidine in MDS patients and provide detailed insight into drug-induced demethylation responses.

The main problem with MDS is that there is not clear genetic pathway and causal relationship. As DeVita et al state:

No etiologic factor is identified in most patients with MDS. MDS is more frequent in men than women by a factor of 1.8. It has been associated with smoking and hair dyes, exposure to agricultural and industrial toxins, drugs (e.g., chloramphenicol), and occupational exposures to stone and cereal dusts. MDS has been associated with exposure to ionizing radiation (atomic bomb survivors in Japan, decontamination workers following the Chernobyl nuclear plant accident) and chronic exposure to low-dose radiation (radiopharmaceuticals). Some inherited hematologic disorders (Fanconi anemia, dy-skeratosis congenita, Shwachman-Diamond syndrome, Diamond- Blackfan syndrome) are also associated with a higher risk of MDS.

Thus there is no clear causal factor or factors recognized at this time.

In a recent paper by Suzuki et al the authors discuss some of the causes of methylation and in turn cancers. They state:

Evidence now suggests that epigenetic abnormalities, particularly altered DNA methylation, play a crucial role in the development and progression of human gastrointestinal malignancies. Two distinct DNA methylation abnormalities are observed together in cancer.

One is an overall genome-wide reduction in DNA methylation (global hypomethylation) and the other is regional hypermethylation within the CpG islands of specific gene promoters. Global hypomethylation is believed to induce proto-oncogene activation and chromosomal instability, whereas regional hypermethylation is strongly associated with transcriptional silencing of tumor suppressor genes.

To date, genes involved in regulation of the cell cycle, DNA repair, growth signaling, angiogenesis, and apoptosis, are all known to be inactivated by hypermethylation. Recently developed techniques for detecting changes in DNA methylation have dramatically enhanced our understanding of the patterns of methylation that occur as cancers progress. One of the key contributors to aberrant methylation is aging, but other patterns of methylation are cancer-specific and detected only in a subset of tumors exhibiting the CpG island methylator phenotype (CIMP).

Although the cause of altered patterns of DNA methylation in cancer remains unknown, it is believed that epidemiological factors, notably dietary folate intake, might strongly influence DNA methylation patterns.

Recent studies further suggest that polymorphisms of genes involved in folate metabolism are causally related to the development of cancer.

5 METASTASIS

The argument in the paper in question is that NSD2 drives metastasis. There are many ways to interpret this statement. One is a clinical presentation of the ultimate state of the proliferating cells. At the other extreme would be a detailed understanding of all the steps involved in metastasis and demonstrating that indeed each of these steps is present. The latter of course is both complex and at present incomplete.

As Mehlen and Puisieux have noted:

Metastasis occurs through a series of sequential steps in which tumour cells first migrate from the primary tumour, penetrate blood vessels and then colonize distant sites. It is a highly inefficient process. Indeed, very few of the tumour cells that gain access to the vasculature give rise to metastatic foci in a secondary organ. Recent data indicate that the mechanisms controlling metastasis can be regulated independently from primary tumour development. In vitro and in vivo, the metastatic potential of tumours is associated with an increased resistance to apoptosis.

Furthermore, the experimental modulation of apoptotic or anti-apoptotic factors influences metastatic efficiency. Anoikis and amorphosis are important barriers to metastasis. Anoikis is cell death induced by the disruption of cell attachment and cell–matrix interactions, whereas amorphosis is cell death stimulated by the loss of cytoskeletal architecture. Early survival of tumour cells after attachment to the secondary site and the development of micrometastases are crucial steps of the metastatic process.

Metastasis is the most common cause of cancer death. Most patients with metastatic disease respond transiently to conventional treatments. Further elucidation of the relationship between resistance to apoptosis of metastatic cancer cells and their chemoresistance should provide important clues to improve systemic therapies.

As Aytes et al note:

Metastasis is a complex process that culminates in the progressive accumulation of molecular alterations of cancer cells, which allow them to escape the confines of the tumor, survive during dissemination, and ultimately reside at distant sites, wherein requisite adaptive changes ensue in their new microenvironment. Therefore, it would be most informative to study the biological processes and molecular mechanisms underlying metastatic progression as occur in the context of the whole organism in vivo.

However, inherent challenges in accessing primary tumors and their metastases from cancer patients have made it difficult to study de novo metastasis formation. Moreover, most in vivo studies of metastasis have utilized transplantation models wherein cells or tumors are implanted

into host organisms, usually immune-deficient ones. While such investigations have advanced our understanding of metastasis mechanisms and have elucidated factors that promote organ tropism, they may not ideally model the cell-intrinsic mechanisms of de novo metastatic progression. Analyses of genetically engineered mouse models (GEMMs) can overcome these obstacles, since they enable access to tumors and their resultant metastases as they arise de novo during cancer progression in the whole organism

They continue:

NSD2 is a driver of metastatic prostate cancer progression. Among the candidate MRs (Master Regulator, a concept about the existence os a dominant gene product which controls the process), the highest level of MR activity as well as experimentally determined functional activity were observed for the histone methyltransferase, Nuclear receptor binding SET Domain protein 2 (NSD2). Notably, NSD2 is a putative cofactor of androgen receptor48 that has been previously implicated in advanced prostate cancer, and has been shown to collaborate with RAS signaling in other tumor contexts.

Therefore, we examined the expression of NSD2 at the mRNA and protein levels in nonmetastatic and metastatic contexts in both mouse and human prostate cancer. In the mouse prostate, we found that Nsd2 protein is expressed at low levels in non-metastatic tumors from the NP mice, while it is highly expressed in metastatic tumors from the NPK mice, as well as corresponding metastases from these mice.

Notably, Nsd2 was robustly expressed in nuclei of NPK tumors and lung metastases, coincident with high levels of Ki67, a marker of cell proliferation, strong expression of nuclear androgen receptor (AR), and robust expression of pan-cytokeratin (Pan-Ck). In human prostate cancer, we found that NSD2 expression is increased during cancer progression at both the mRNA and protein levels (Fig. 4b–e). In particular, expression of NSD2 mRNA levels were significantly higher in more advanced (Gleason ≥ 4 + 4; n = 104) versus earlier stage…

6 NSD2

NSD2 is also named MMSET and WHSC1. We will refer to it as NSD2 but the literature uses a multiplicity of names.

As NCBI notes[2]:

This gene encodes a protein that contains four domains present in other developmental proteins: a PWWP domain, an HMG box, a SET domain, and a PHD-type zinc finger. It is expressed ubiquitously in early development. Wolf-Hirschhorn syndrome (WHS) is a malformation syndrome associated with a hemizygous deletion of the distal short arm of chromosome 4. This gene maps to the 165 kb WHS critical region and has also been involved in the chromosomal translocation t(4;14)(p16.3;q32.3) in multiple myelomas. Alternative splicing of this gene results in multiple transcript variants encoding different isoforms. Some transcript variants are nonsense-mediated mRNA (NMD) decay candidates, hence not represented as reference sequences.

Now this may not be as clear as possible, It is better explained as Kuo et al have noted:

Histone lysine methylation signaling is a principal chromatin regulatory mechanism that influences fundamental nuclear processes. Lysine (K) residues can accept up to three methyl groups to form mono-, di-, and trimethylated derivatives (Kme1, Kme2, and Kme3, respectively). Methylated histone species are sensed and linked to downstream biological functions by methyl lysine-binding proteins, in a manner specified by the extent and sequence context of the methylation event.

The human genome encodes greater than fifty predicted protein lysine methyltransferases (PKMTs), of which several are deregulated in human disease. To understand how this diverse nuclear signaling network influences chromatin biology and disease, it is essential to elucidate the precise activity of individual PKMTs on their substrates. The vast majority of PKMT enzymes contain a conserved catalytic SET domain.

The methyltransferases are proteins which assist in transferring methyls to the tails of histones in this case. Now they continue:

NSD2 (also named MMSET and WHSC1) is a SET domain-containing PKMT implicated in diverse human diseases. For example, NSD2 haploinsufficiency is implicated in the developmental disorder Wolf Hirschhorn syndrome (WHS), which is characterized by growth and mental retardation, congenital heart defects, and antibody deficiencies, and NSD2-deficeint mice exhibit a spectrum of defects resembling WHS.

NSD2 is also implicated in the pathogenesis of the hematologic malignancy multiple myeloma (MM). MM is the second most common blood cancer, accounting for 1% of all cancers and 2%

[2] https://www.ncbi.nlm.nih.gov/gene/7468

of cancer deaths in the United States. 15%–20% of MM patients carry a translocation between chromosomes 4 and 14 [t(4;14)(p16.3;q32)], which places the transcription of two genes, NSD2 and FGFR3, under the control of strong IgH intronic Em enhancer and 30 enhancer, respectively, and leads to aberrant upregulation of these two genes. Notably, overexpression of NSD2, and not FGFR3, is thought to be important for t(4;14)-mediated myeloma pathogenesis.

It is interesting that there are many hematological disorders which are methylation related such as MDS as described before. However the previous cases are DNA methylation whereas these are histone related.

Thus, NSD2 plays an important role during mammalian development, and its overexpression is implicated in cancer. The physiologic catalytic activity of NSD2 is obscure, particularly because there are several conflicting reports in the literature.

Specifically, NSD2 and/or NSD2 isoforms have been proposed to generate numerous different histone marks, including trimethylation of H3 at lysines K4 (H3K4me3), K27 (H3K27me3), and K36 (H3K36me3), trimethylation of H4 at lysine K20 (H4K20me3), dimethylation of H4 at lysine 20 (H4K20me2), and dimethylation of H3 at lysine 36 (H3K36me2).

This NSD2 is one amongst many methyltransferases.

As Aytes et al noted in the work at question:

NSD2 has been reported to function as a histone methyltransferase that targets the histone H3 di-methyl mark on lysine 36 (H3K36me2). Accordingly, we found that silencing of NSD2 in either human or mouse cells resulted in a modest but reproducible reduction of the H3K36me2 mark, while not altering the mono-methyl marks on lysine 36 (H3K36me1) or other histone marks such as tri-methyl lysine 27 (H3K27me3) or lysine 9 (H3K9me3).

Furthermore, NSD2 silencing in either mouse NPK cells or human DU145 cells in vitro resulted in a 5–10 fold inhibition of colony formation ($p < 0.0001$, two-tailed Student's t-test), as well as significantly decreased invasion.

It was also noted in Nature by Pei et al that[3]:

p53-binding protein 1 (53BP1) is known to be an important mediator of the DNA damage response1, with dimethylation of histone H4 lysine 20 (H4K20me2) critical to the recruitment of 53BP1 to double-strand breaks (DSBs). However, it is not clear how 53BP1 is specifically targeted to the sites of DNA damage, as the overall level of H4K20me2 does not seem to increase following DNA damage.

It has been proposed that DNA breaks may cause exposure of methylated H4K20 previously buried within the chromosome; however, experimental evidence for such a model is lacking.

*Here we found that H4K20 methylation actually increases locally upon the induction of DSBs and that methylation of H4K20 at DSBs is mediated by the histone methyltransferase MMSET (**also known as NSD2 or WHSC1**) in mammals.*

Downregulation of MMSET significantly decreases H4K20 methylation at DSBs and the subsequent accumulation of 53BP1. Furthermore, we found that the recruitment of MMSET to DSBs requires the γH2AX–MDC1 pathway; specifically, the interaction between the MDC1 BRCT domain and phosphorylated Ser 102 of MMSET. Thus, we propose that a pathway involving γH2AX–MDC1–MMSET regulates the induction of H4K20 methylation on histones around DSBs, which, in turn, facilitates 53BP1 recruitment.

The above are one of several such methyltransferase functions. Now Lukas et al note as well:

Histone methylation has long been known to promote chromatin accumulation of the DDR-mediator 53BP1 through its Tudor domain61. However, both histone methylations associated with 53BP1 chromatin retention (H3K79me and H4K20me) have been considered as constitutive, which suggests that repositioning of nucleosomes after DNA damage exposes these otherwise hidden methylation marks.

Unexpectedly, H4K20 methylation was found to be increased at the DSB sites, through the methyltransferase MMSET (also known as WHSC1 and NSD2).

Double strand breaks, DSB, and often related to malignancies as well. They relate to BRCA and BRCA deficient cells can be subject to such breaks.

The ATM-mediated phosphorylation of MMSET is required for its binding to the phosphate-binding module of MDC1, suggesting that MDC1 is involved in yet another component of organizing chromatin flanking sites with DNA damage. How methylation cooperates with ubiquitylation in promoting BP1 chromatin retention remains puzzling; evidence suggests that the MMSET-mediated H4K20 methylation occurs independent of the RNF8–RNF168 pathway.

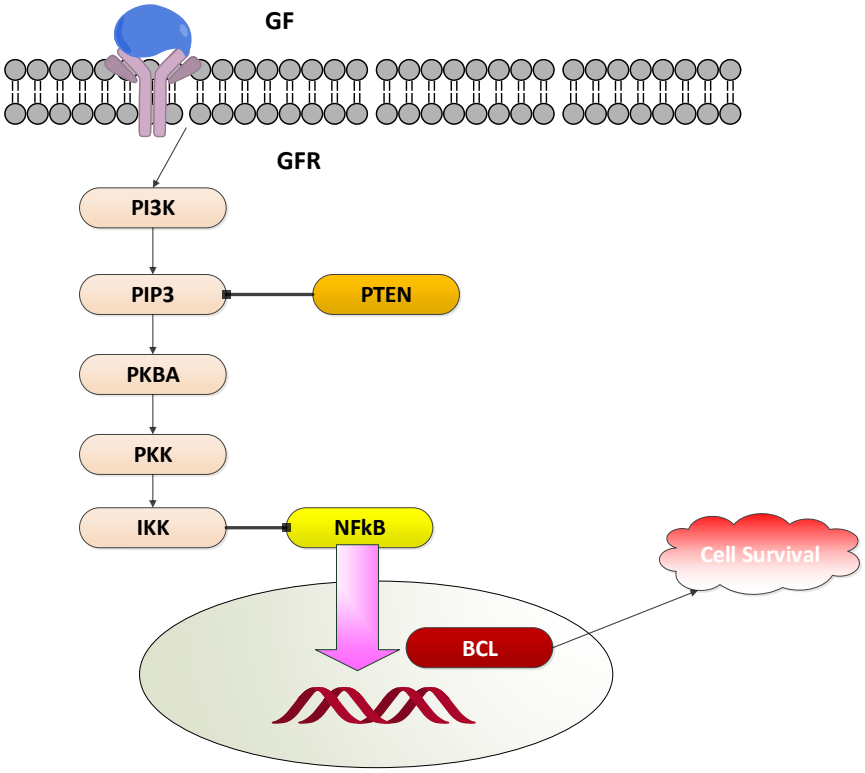
As Cucchiara et al note:

The nuclear receptor-binding SET domain-protein 2 (NSD2) is a histone methyltransferase that cooperates with the DBD of the AR. High levels of NSD2 are related to the expression of PSA (prostate specific antigen). A paper by Asangani et al. reported that high levels of NSD2 correlate with aggressive characteristics in PCa.

The mechanism of action is linked to the enhancer of zeste homolog 2 (EZH2), a component of Polycomb repressive complex 2 (PRC2). The enhancement of EZH2 leads to the transcriptional inhibition of miR-203, miR-31 and miR-26, which are repressors of NSD2. This complex mechanism facilitates an over expression of NSD2 with the generation of the active histone mark, H3K36me2.

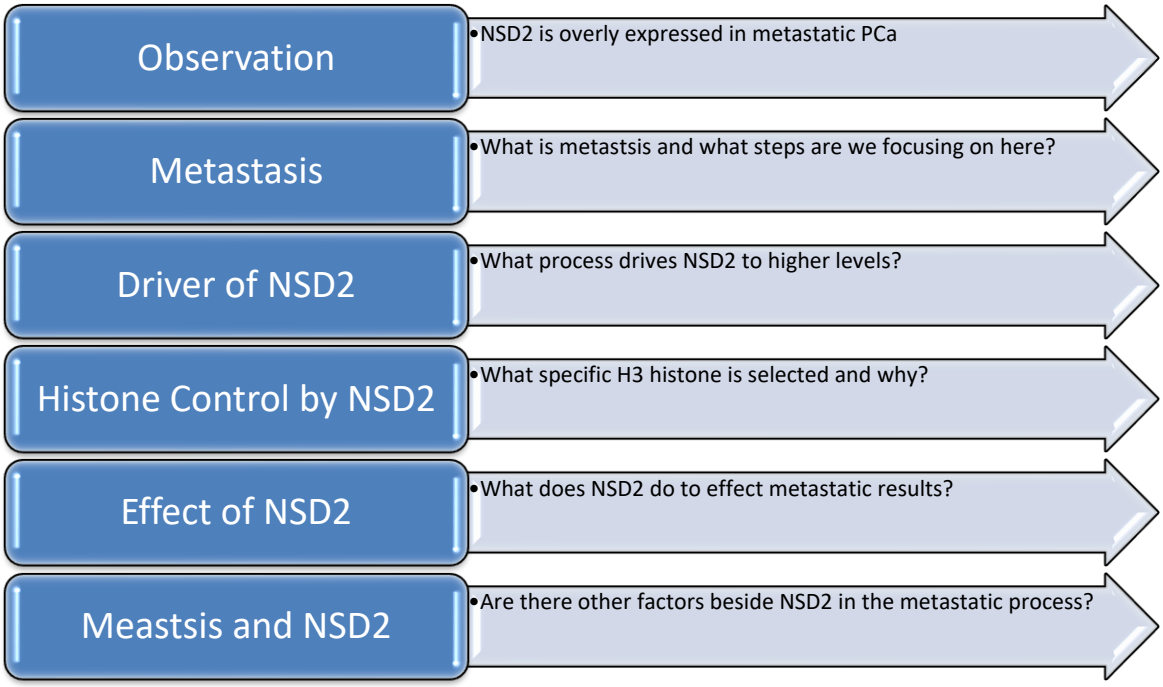
Moreover… NSD2 acts as a transcriptional coactivator of NF-kB for activation of target genes, such as IL-6, IL-8, VEGFA and survivin in CRPC cells.

Note the last statement. We have seen elsewhere that NSD2 drives RAS to prevent apoptosis and here we see it drives NF-kB to drive genes that do likewise. We briefly demonstrate that below:



7 OBSERVATIONS

The observations we will make refer to the questions in the figure below:



7.1 NSD2 EFFECTS

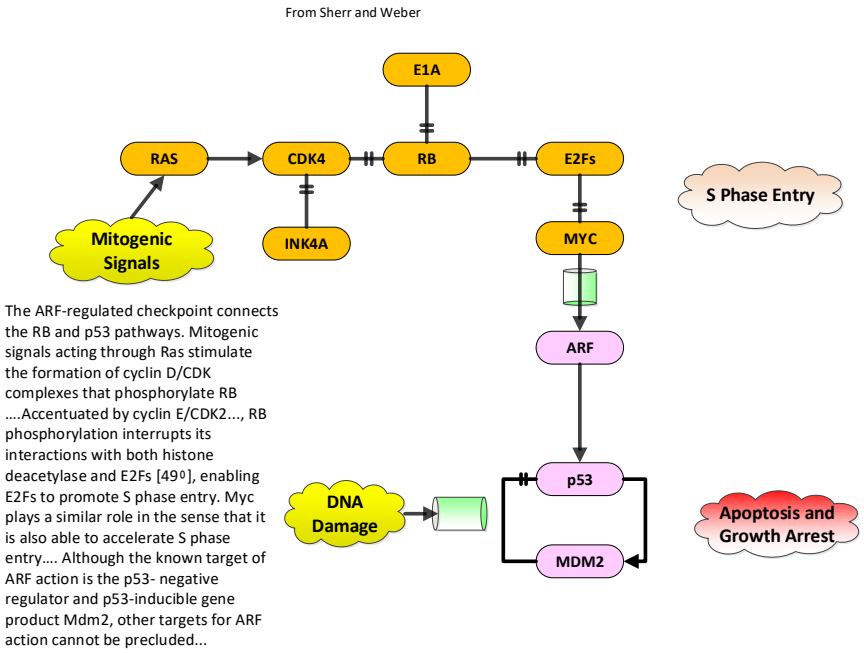
NSDs does affect other pathways. As Garcia-Carpizo et al note:

The NSD2 knock down interferes with the RAS-transcriptional program. We hypothesized that the dramatic changes in H3K36me2 content after NSD2 depletion would lead to important changes in gene expression. To test this, we assessed genome-wide gene expression by RNA-seq in H1299 cells infected with shNT, sh3 or sh5 and treated with doxycycline. Differentially upregulated and downregulated genes comparing sh3 or sh5 to shNT where selected at FDR < 0.05. Since NSD2 is involved in transcriptional activation, genes downregulated after NSD2 knock down are more likely to be NSD2 targets. Gene set Enrichment Analysis (GSEA)28 revealed that the RAS pathway was at the top of most significantly downregulated pathways.

NSD2 regulates the expression of genes marked with H3K36me2. Although previous reports described interactions between BRD4 and NSD2, there was a lack or enrichment of the cancer-acquired super-enhancer signature in genes changing expression after the NSD2 knock down, and genes contributing the most to the cancer-acquired super-enhancer signature in H1299 cells treated with JQ1 were not significantly changing after NSD2 knock down. These results suggest that in H1299, NSD2 is supporting the RAS-driven transcriptional pathway independently of

BRD4. To ascertain which genes are direct targets of NSD2, we conducted ChIP-seq of H3K36me2 in H1299 cells transduced with sh3 in the presence and absence of doxycycline.

Thus it appears that RAS pathways may be the likely target. We show a possible set of control elements below:

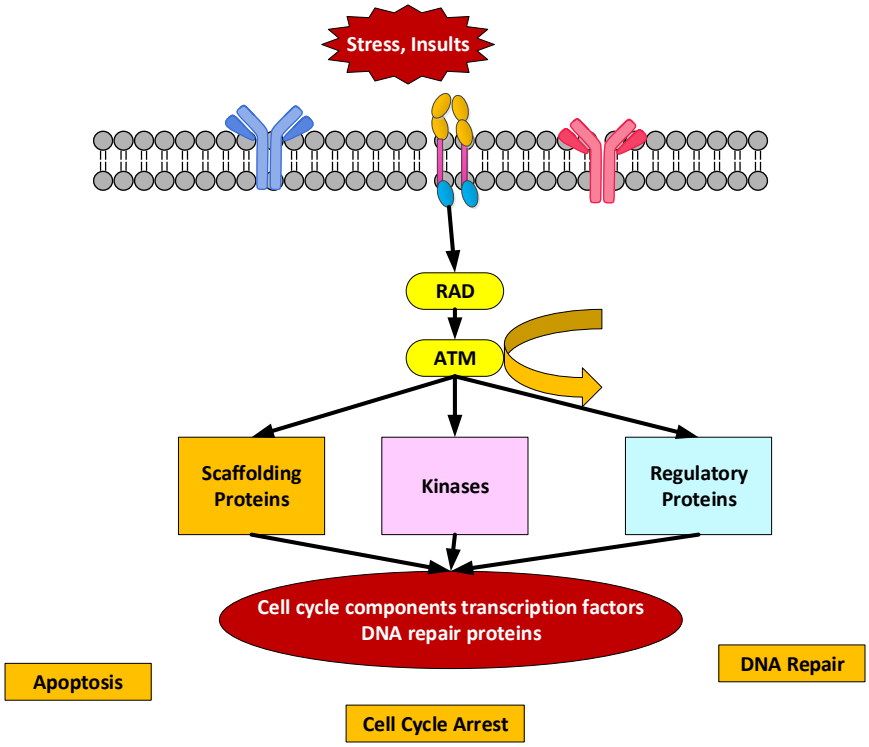


Thus NSD2 has some control on RAS and mitigates ultimately apoptosis.

7.2 DRIVERS OF NSD2

In the Appendix we see that BRCA1 and p53 are controllers of NSD2. That being the case we can see that either through some genetic defect or some gene alteration we have putative drivers. The detailed specifics do not appear to be in the current literature. Thus if they can be ascertained then it is possible to obtain better therapeutic targets.

As Abraham has noted regarding ATM in the graphic below:



Thus logically we can putatively state that stressors on the cell may activate the ATM pathway which in turn activates NSD2 and in turn stops apoptosis, cell arrest, and DNA repair. In a sense NSD2 is just an intermediary player in the overall process of metastasis.

7.3 HISTONE TARGETING OF NSD2

Histone targeting is the ability of a methyltransferase to target a specific histone. It does not appear that methyltransferase is a random targeting protein but that it appears to target specific gene sites. It is not clear how this is accomplished but one suspects the issues related to the histone code apply here.

7.4 METASTASIS EFFECTS

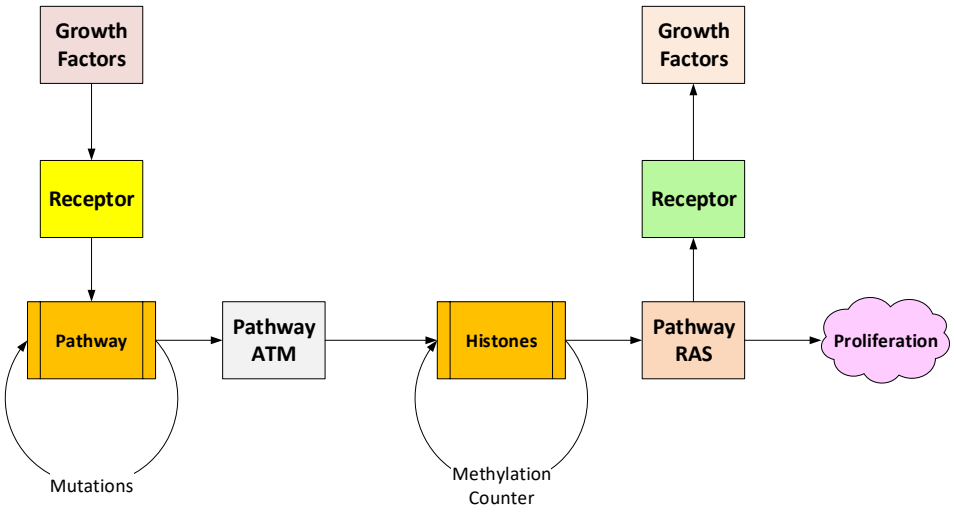
Now as we had discussed metastasis is allegedly facilitated by the actions of NSD2 and there is some indication as to how that may occur. However as histone targeting is in need of clarification so also is this issue. One suspects such factors as those related to growth factor and EMT characteristics apply[4],[5].

[4] https://www.researchgate.net/publication/330222973_EMT_and_Cancers

[5] https://www.researchgate.net/publication/329702571_Growth_Factors_Pathways_and_Cancers

7.5 SYSTEM DYNAMICS

It is possible to better understand the process by viewing it as a dynamic system, We may not have all of the elements but we do have a general understanding as to what needs to be included. A suggestion of a simple model is shown below:



Namely we have dynamic and stochastic parts denoted by pathway and histone changes and then their related effects on those elements. The control boxes then initiate proliferation and/or the release of many growth factors. It would be interested if this model can be fully developed. Use of such things as Petri Nets may be quite appropriate here.

The key observation is that there are well know separate parts but the complete collection of dynamic elements is still less than fully understood. The question would therefore be; is there a path to better therapeutics along this way?

7.6 PROTEIN DYNAMICS

The histone tails are segments of a protein. These tails get methylated and otherwise and thus the folding and attraction patterns change. It would be of interest t ascertain how these factors play in the opening and closing of the histone paths occur. There are various discussion in Allis et al regarding these issues but there does not appear to be an effective summary.

8 REFERENCES

1. Abraham, R., Cell cycle checkpoint signaling through the ATM and ATR kinases, Genes & Development 15:2177–2196 © 2001
2. Allis, C., et al, Epigenetics, 2nd Ed, CSH (Cold Spring Harbor) 2015
3. Armstrong, Epigenetics, Garland (New York) 2013
4. Asangani, I. A. et al. Characterization of the EZH2-MMSET histone methyltransferase regulatory axis in cancer. Mol. Cell 49, 80–93 (2013).
5. Aytes et al, NSD2 is a conserved driver of metastatic prostate cancer progression, Nature Communications, 2018, 9:5201
6. Bennett, D., How to make a melanoma: what do we know of the primary clonal events? Pigment Cell Melanoma Res. 21; 27–38, 2008
7. Bennett, R. L., Swaroop, A., Troche, C., & Licht, J. D., The role of nuclear receptor-binding SET domain family histone lysine methyltransferases in cancer. Cold Spring Harb. Perspect. Med. 7, a026708 (2017).
8. Berthelot and Terrat, Petri Nets Theory for the Correctness of Protocols, IEEE Trand Comm Dec 1982.
9. Calin, G., C. Croce, Chromosomal Rearrangements and Micro RNAs, Jrl Clin Inv 2007 pp 2059-2066
10. Cucchiara, et al, Epigenomic Regulation of Androgen Receptor Signaling: Potential Role in Prostate Cancer Therapy, Cancers 2017, 9, 9;
11. Das S., et al, MDA-9/syntenin: a positive gatekeeper of melanoma metastasis, Frontiers in Bioscience 17, 1-15, January 1, 2012.
12. DeVita, Hellman, and Rosenberg, Cancer: Principles & Practice of Oncology, 8e, Lippincott (New York) 2008.
13. DeVita, Hellman, and Rosenberg, Cancer: Principles & Practice of Oncology, 9e, Lippincott (New York) 2011.
14. DeVita, Hellman, and Rosenberg, Cancer: Principles & Practice of Oncology, 10th Ed, Lippincott (New York) 2019
15. Ellis et al, Epigenetics in cancer: Targeting chromatin modifications, Mol Cancer Ther 2009;8(6). June 2009
16. Ezponda, T. et al. The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. Oncogene 32, 2882–2890 (2013).
17. Garcia-Carpizo, V. et al. NSD2 contributes to oncogenic RAS-driven transcription in lung cancer cells through long-range epigenetic activation. Sci.Rep. 6, 32952 (2016).
18. Gut and Verdin, The nexus of chromatin regulation and intermediary metabolism, 24 October 2013 | Vol 502 | Nature | 489

19. Hassler and Egger, Epigenomics of cancer e emerging new concepts. Biochimie 94 (2012) 2219-2230

20. Heiner, Petri Nets for Systems and Synthetic Biology, M. Bernardo, P. Degano, and G. Zavattaro (Eds.): SFM 2008, LNCS 5016, pp. 215{264, 2008.

21. Helin and Dhanak, Chromatin proteins and modifications as drug targets, Nature | Vol 502 | 24 October 2013

22. Jenuwein and Allis, Translating the Histone Code, Science, Aug 10 2001, p 1074

23. Kuo et al, NSD2 Links Dimethylation of Histone H3 at Lysine 36 to Oncogenic Programming, Molecular Cell 44, 609–620, November 18, 2011

24. Laird, P., R. Jaenisch, DNA Methylation and Cancer, Human Molecular Genetics, Vol 3, 1487-95, 1999

25. Lukas et al, More than just a focus: The chromatin response to DNA damage and its role in genome integrity maintenance, Nature Cell Biology Volume 13 | Number 10 | October 2011

26. Marks et al, Cellular Signal Processing, 2nd Ed, Garland (New York) 2017

27. Maury and Hashizume, Epigenetic modification in chromatin machinery and its deregulation in pediatric brain tumors: Insight into epigenetic therapies, Epigenetics, 2017, Vol. 0, NO. 0, 1–17

28. Mehlen and Puisieux, Metastasis: a question of life or death, Nature Reviews, Vol 6 June 2006 p. 449

29. Mazar, J., et al, Epigenetic Regulation of MicroRNA Genes and the Role of miR-34b in Cell Invasion and Motility in Human Melanoma, PLOS One, September 2011, Volume 6, Issue 9

30. Miranda, T., P. Jones. DNA Methylation: The Nuts and Bolts of Repression, Cll Phys 2007 pp 384-390.

31. O'Shaughnessy, J, et al, Treatment and Prevention of Intraepithelial Neoplasia: An Important Target for Accelerated New Agent Development, Clinical Cancer Research, Vol. 8, 314–346, February 2002.

32. Pali, S., K. Robertson, Epigenetic Control of Tumor Suppression, Critical Reviews™ in Eukaryotic Gene Expression, 17(4):295–316 (2007).

33. Pei et al, MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites, 124 | Nature | VOL 470 | 3 February 2011

34. Peterson, Petri Nets, Computing Surveys, Vol. 9,No. 3,September 1977

35. Rando, Combinatorial complexity in chromatin structure and function: revisiting the histone code, Curr Opin Genet Dev. 2012 April ; 22(2): 148–155.

36. Robertson, K., A. Wolffe, DNA Methylation in Health and Disease, Nat Rev Gen 2000 pp 11-19

37. Robertson, K., DNA methylation, methyltransferases, and cancer, Oncogene (2001) 20, 3139-3155

38. Sankaran, S. M. & Gozani, O. Characterization of H3.3K36M as a tool to study H3K36 methylation in cancer cells. Epigenetics 12, 917–922 (2017).

39. Sankaran, S. M., Wilkinson, A. W., Elias, J. E. & Gozani, O. A PWWP domain of histone-lysine N-methyltransferase NSD2 binds to dimethylated Lys-36 of histone H3 and regulates NSD2 function at chromatin. J. Biol. Chem. 291, 8465–8474 (2016).

40. Strahl and Allis, The language of covalent histone modifcations, Nature |Vol 403 | 6 January 2000

41. Strathdee, G., R., Brown, Aberrant DNA Methylation in Cancer; Potential Clinical Interventions, Exp Rev Mol Med, 2002, pp 1-17

42. Tefferi, A., J. Vardiman, Myelodysplastic Syndromes, N Engl J Med 2009; V 361: pp 1872-85

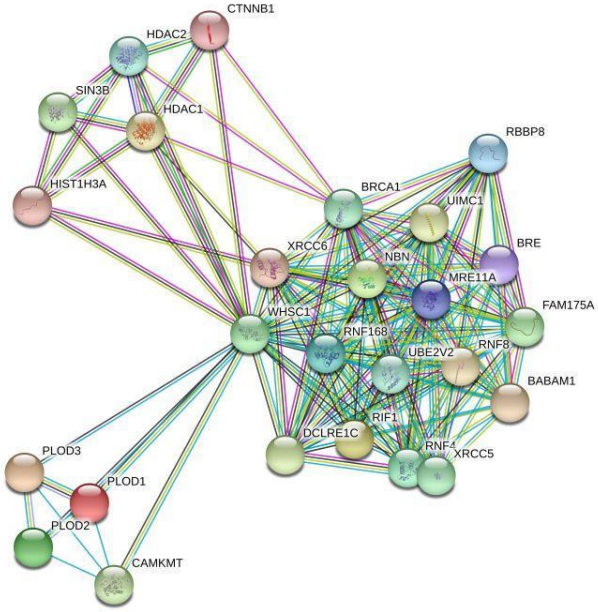
43. Tollefsbol, Epigenetics in Human Disease, Academic (New York) 2012.

44. Vasiljevic, N , et al, Association between DNA methylation of HSPB1 and death in low Gleason score prostate cancer, Prostate Cancer and Prostatic Disease (2013) 16, 35–40

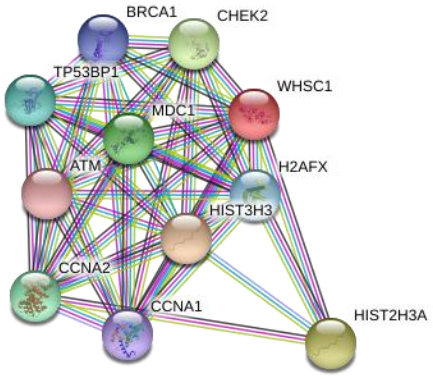
45. Zilberman, D., The Evolving Functions of DNA Methylation, Curr Opin in Plt Bio, 2008, pp 554-559.

9 PATHWAYS

We have putatively the following pathways where WHSC1 is the same as NSD2.[6]:



Also we have the simpler one below[7]:



Note in the above we have BRCA1 and p53 controlling WHSC1 or NSD2 expression.

ATM	The protein encoded by this gene belongs to the PI3/PI4-kinase family. This protein is an important cell cycle checkpoint kinase that phosphorylates; thus, it functions as a regulator of a

[6] http://version10.5.string-db.org/cgi/network.pl?taskId=Go5zVLjTF3Jw

[7] http://version10.5.string-db.org/cgi/network.pl?taskId=IQGIjnhA71nN

	wide variety of downstream proteins, including tumor suppressor proteins p53 and BRCA1, checkpoint kinase CHK2, checkpoint proteins RAD17 and RAD9, and DNA repair protein NBS1. This protein and the closely related kinase ATR are thought to be master controllers of cell cycle checkpoint signaling pathways that are required for cell response to DNA damage and for genome stability. Mutations in this gene are associated with ataxia telangiectasia, an autosomal recessive disorder.
PLOD1	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1; Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens. These hydroxylysines serve as sites of attachment for carbohydrate units and are essential for the stability of the intermolecular collagen cross-links (727 aa)
PLOD3	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3; Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens. These hydroxylysines serve as sites of attachment for carbohydrate units and are essential for the stability of the intermolecular collagen cross-links (738 aa)
RIF1	RAP1 interacting factor homolog (yeast); Required for checkpoint mediated arrest of cell cycle progression in response to DNA damage during S-phase (the intra-S- phase checkpoint). This checkpoint requires activation of at least 2 parallel pathways by the ATM kinase- one involves the MRN (MRE11A-RAD50-NBS1) complex, while the second requires CHEK2. RIF1 seems to act independently of both these pathways. Seems to play no role in either the G1/S or G2/M DNA damage checkpoints (2472 aa)
NBN	Nibrin; Component of the MRE11-RAD50-NBN (MRN complex) which plays a critical role in the cellular response to DNA damage and the maintenance of chromosome integrity. The complex is involved in double-strand break (DSB) repair, DNA recombination, maintenance of telomere integrity, cell cycle checkpoint control and meiosis. The complex possesses single-strand endonuclease activity and double-strand-specific 3'-5' exonuclease activity, which are provided by MRE11A. RAD50 may be required to bind DNA ends and hold them in close proximity. NBN modulate the DNA damage signal sensing by recru [...] (754 aa)
PLOD2	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2; Forms hydroxylysine residues in -Xaa-Lys-Gly- sequences in collagens. These hydroxylysines serve as sites of attachment for carbohydrate units and are essential for the stability of the intermolecular collagen cross-links (758 aa)
RNF4	Ring finger protein 4; E3 ubiquitin-protein ligase which binds polysumoylated chains covalently attached to proteins and mediates 'Lys-6'-, 'Lys-11'-, 'Lys-48'- and 'Lys-63'-linked polyubiquitination of those substrates and their subsequent targeting to the proteasome for degradation. Regulates the degradation of several proteins including PML and the transcriptional activator PEA3. Involved in chromosome alignment and spindle assembly, it regulates the kinetochore CENPH-CENPI-CENPK complex by targeting polysumoylated CENPI to proteasomal degradation. Regulates the cellular responses t [...] (190 aa)
RNF168	Ring finger protein 168, E3 ubiquitin protein ligase; E3 ubiquitin-protein ligase required for accumulation of repair proteins to sites of DNA damage. Acts with UBE2N/UBC13 to amplify the RNF8-dependent histone ubiquitination. Recruited to sites of DNA damage at double-strand breaks (DSBs) by binding to ubiquitinated histone H2A and H2AX and amplifies the RNF8-dependent H2A ubiquitination, promoting the formation of 'Lys-63'- linked ubiquitin conjugates. This leads to concentrate ubiquitinated histones H2A and H2AX at DNA lesions to the threshold required for recruitment of TP53BP1 an [...] (571 aa)
RBBP8	Retinoblastoma binding protein 8; Endonuclease that cooperates with the MRE11-RAD50-NBN (MRN) complex in processing meiotic and mitotic double-strand breaks (DSBs) by ensuring both resection and intrachromosomal association of the broken ends. Functions downstream of the MRN complex and ATM, promotes ATR activation and its recruitment to DSBs in the S/G2 phase facilitating the generation of ssDNA. Component of the BRCA1-RBBP8 complex that regulates CHEK1 activation and controls cell cycle G2/M checkpoints on DNA damage. Promotes microhomology-mediated alternative end joining (A-NHEJ) d [...] (897 aa)
MRE11A	MRE11 meiotic recombination 11 homolog A (S. cerevisiae); Component of the MRN complex, which plays a central role in double-strand break (DSB) repair, DNA recombination, maintenance of telomere integrity and meiosis. The complex possesses single-strand endonuclease activity and double-strand- specific 3'-5' exonuclease activity, which are provided by MRE11A. RAD50 may be required to bind DNA ends and hold them in close proximity. This could facilitate searches for short or long regions of sequence homology in the recombining DNA templates, and may also stimulate the activity of DNA li [...] (708 aa)
BRE	Brain and reproductive organ-expressed (TNFRSF1A modulator); Component of the BRCA1-A complex, a complex that specifically recognizes 'Lys-63'-linked ubiquitinated histones H2A

	and H2AX at DNA lesions sites, leading to target the BRCA1-BARD1 heterodimer to sites of DNA damage at double-strand breaks (DSBs). The BRCA1-A complex also possesses deubiquitinase activity that specifically removes ’Lys-63’-linked ubiquitin on histones H2A and H2AX (PubMed-17525341, PubMed-19261746, PubMed-19261749, PubMed-19261748). In the BRCA1-A complex, it acts as an adapter that bridges the interaction be [...] (415 aa)
CTNNB1	Catenin (cadherin-associated protein), beta 1, 88kDa; Key downstream component of the canonical Wnt signaling pathway. In the absence of Wnt, forms a complex with AXIN1, AXIN2, APC, CSNK1A1 and GSK3B that promotes phosphorylation on N-terminal Ser and Thr residues and ubiquitination of CTNNB1 via BTRC and its subsequent degradation by the proteasome. In the presence of Wnt ligand, CTNNB1 is not ubiquitinated and accumulates in the nucleus, where it acts as a coactivator for transcription factors of the TCF/LEF family, leading to activate Wnt responsive genes. Involved in the regulation [...] (781 aa)
HIST1H3A	Histone cluster 1, H3a (136 aa)
XRCC6	X-ray repair complementing defective repair in Chinese hamster cells 6; Single-stranded DNA-dependent ATP-dependent helicase. Has a role in chromosome translocation. The DNA helicase II complex binds preferentially to fork-like ends of double-stranded DNA in a cell cycle-dependent manner. It works in the 3’-5’ direction. Binding to DNA may be mediated by XRCC6. Involved in DNA non-homologous end joining (NHEJ) required for double-strand break repair and V(D)J recombination. The XRCC5/6 dimer acts as regulatory subunit of the DNA-dependent protein kinase complex DNA-PK by increasing the [...] (609 aa)
BABAM1	BRISC and BRCA1 A complex member 1; Component of the BRCA1-A complex, a complex that specifically recognizes ’Lys-63’-linked ubiquitinated histones H2A and H2AX at DNA lesions sites, leading to target the BRCA1-BARD1 heterodimer to sites of DNA damage at double-strand breaks (DSBs). The BRCA1-A complex also possesses deubiquitinase activity that specifically removes ’Lys-63’-linked ubiquitin on histones H2A and H2AX. In the BRCA1-A complex, it is required for the complex integrity and its localization at DSBs. Component of the BRISC complex, a multiprotein complex that specifically cle [...] (329 aa)
RNF8	Ring finger protein 8, E3 ubiquitin protein ligase; E3 ubiquitin-protein ligase that plays a key role in DNA damage signaling via 2 distinct roles- by mediating the ’Lys-63’- linked ubiquitination of histones H2A and H2AX and promoting the recruitment of DNA repair proteins at double-strand breaks (DSBs) sites, and by catalyzing ’Lys-48’-linked ubiquitination to remove target proteins from DNA damage sites. Following DNA DSBs, it is recruited to the sites of damage by ATM-phosphorylated MDC1 and catalyzes the ’Lys-63’-linked ubiquitination of histones H2A and H2AX, thereby promoting th [...] (485 aa)
HDAC1	Histone deacetylase 1; Responsible for the deacetylation of lysine residues on the N-terminal part of the core histones (H2A, H2B, H3 and H4). Histone deacetylation gives a tag for epigenetic repression and plays an important role in transcriptional regulation, cell cycle progression and developmental events. Histone deacetylases act via the formation of large multiprotein complexes. Deacetylates SP proteins, SP1 and SP3, and regulates their function. Component of the BRG1-RB1-HDAC1 complex, which negatively regulates the CREST-mediated transcription in resting neurons. Upon calcium s [...] (482 aa)
UIMC1	Ubiquitin interaction motif containing 1; Ubiquitin-binding protein (PubMed-24627472). Specifically recognizes and binds ’Lys-63’-linked ubiquitin (PubMed-19328070, Ref.35). Plays a central role in the BRCA1-A complex by specifically binding ’Lys-63’-linked ubiquitinated histones H2A and H2AX at DNA lesions sites, leading to target the BRCA1-BARD1 heterodimer to sites of DNA damage at double-strand breaks (DSBs). The BRCA1-A complex also possesses deubiquitinase activity that specifically removes ’Lys-63’-linked ubiquitin on histones H2A and H2AX. Also weakly binds monoubiquitin but wi [...] (719 aa)
DCLRE1C	DNA cross-link repair 1C; Required for V(D)J recombination, the process by which exons encoding the antigen-binding domains of immunoglobulins and T-cell receptor proteins are assembled from individual V, (D), and J gene segments. V(D)J recombination is initiated by the lymphoid specific RAG endonuclease complex, which generates site specific DNA double strand breaks (DSBs). These DSBs present two types of DNA end structures- hairpin sealed coding ends and phosphorylated blunt signal ends. These ends are independently repaired by the non homologous end joining (NHEJ) pathway to form co [...] (692 aa)
CAMKMT	Calmodulin-lysine N-methyltransferase; Catalyzes the trimethylation of ’Lys-116’ in calmodulin (323 aa)

SIN3B	SIN3 transcription regulator homolog B (yeast); Acts as a transcriptional repressor. Interacts with MXI1 to repress MYC responsive genes and antagonize MYC oncogenic activities. Interacts with MAD-MAX heterodimers by binding to MAD. The heterodimer then represses transcription by tethering SIN3B to DNA. Also forms a complex with FOXK1 which represses transcription (By similarity) (1162 aa)
FAM175A	Family with sequence similarity 175, member A; Involved in DNA damage response and double-strand break (DSB) repair. Component of the BRCA1-A complex, acting as a central scaffold protein that assembles the various components of the complex and mediates the recruitment of BRCA1. The BRCA1-A complex specifically recognizes 'Lys-63'-linked ubiquitinated histones H2A and H2AX at DNA lesion sites, leading to target the BRCA1-BARD1 heterodimer to sites of DNA damage at DSBs. This complex also possesses deubiquitinase activity that specifically removes 'Lys-63'-linked ubiquitin on histones H [...] (409 aa)
WHSC1	Wolf-Hirschhorn syndrome candidate 1; Histone methyltransferase with histone H3 'Lys-27' (H3K27me) methyltransferase activity. Isoform RE-IIBP may act as a transcription regulator that binds DNA and suppresses IL5 transcription through HDAC recruitment (1365 aa)
XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining); Single-stranded DNA-dependent ATP-dependent helicase. Has a role in chromosome translocation. The DNA helicase II complex binds preferentially to fork-like ends of double-stranded DNA in a cell cycle-dependent manner. It works in the 3'-5' direction. Binding to DNA may be mediated by XRCC6. Involved in DNA non-homologous end joining (NHEJ) required for double-strand break repair and V(D)J recombination. The XRCC5/6 dimer acts as regulatory subunit of the DNA-dependent protein kinase [...] (732 aa)
BRCA1	Breast cancer 1, early onset; E3 ubiquitin-protein ligase that specifically mediates the formation of 'Lys-6'-linked polyubiquitin chains and plays a central role in DNA repair by facilitating cellular responses to DNA damage. It is unclear whether it also mediates the formation of other types of polyubiquitin chains. The E3 ubiquitin-protein ligase activity is required for its tumor suppressor function. The BRCA1-BARD1 heterodimer coordinates a diverse range of cellular pathways such as DNA damage repair, ubiquitination and transcriptional regulation to maintain genomic stability. Reg [...] (1884 aa)
UBE2V2	Ubiquitin-conjugating enzyme E2 variant 2; Has no ubiquitin ligase activity on its own. The UBE2V2/UBE2N heterodimer catalyzes the synthesis of non-canonical poly-ubiquitin chains that are linked through 'Lys-63'. This type of poly-ubiquitination does not lead to protein degradation by the proteasome. Mediates transcriptional activation of target genes. Plays a role in the control of progress through the cell cycle and differentiation. Plays a role in the error-free DNA repair pathway and contributes to the survival of cells after DNA damage (145 aa)
HDAC2	Histone deacetylase 2; Responsible for the deacetylation of lysine residues on the N-terminal part of the core histones (H2A, H2B, H3 and H4). Histone deacetylation gives a tag for epigenetic repression and plays an important role in transcriptional regulation, cell cycle progression and developmental events. Histone deacetylases act via the formation of large multiprotein complexes. Forms transcriptional repressor complexes by associating with MAD, SIN3, YY1 and N-COR. Interacts in the late S-phase of DNA-replication with DNMT1 in the other transcriptional repressor complex composed o [...] (488 aa)