Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature: _____ _____

Jacob Daniel Kagey **Date**

Role of chromatin in resilencing dynamics after drug-induced DNA demethylation

By

Jacob Daniel Kagey
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

_____

Paula M. Vertino, Ph.D.
Advisor

_____

Jeremy Boss, Ph.D.
Committee Member

_____

Tamara Caspary, Ph.D.
Committee Member

_____

William Kelly, Ph.D.
Committee Member

_____

Harold Saavedra, Ph.D.
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Role of chromatin in resilencing dynamics after drug-induced DNA demethylation

By

Jacob D. Kagey
B.A. Ohio Wesleyan University, 2004

Advisor: Paula M. Vertino, PhD

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in
Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

2010

Abstract
Role of chromatin in resilencing dynamics after drug-induced DNA demethylation
By Jacob D. Kagey

In cancer, large-scale epigenetic alterations are observed including the aberrant hypermethylation and silencing of a subset of CpG island containing tumor suppressor genes. At this juncture, the causes underlying the establishment and maintenance of this aberrant silencing remain incompletely understood. To study these processes, we have utilized *Target of Methylation-induced Silencing 1* (*TMS1*), a gene that is frequently methylated and silenced during carcinogenesis, as a model to investigate the relationship between the epigenetic regulation of *TMS1* and the DNA methylation status at the CpG island. We found that in normal cells the unmethylated CpG island is enriched for the active histone modifications histone H3 acetylation (H3ac), histone H3 lysine 4 di-methylation (H3K4me2), and histone H4 lysine 16 acetylation (H4K16ac) as well as positioned nucleosomes. Conversely, in cancer cells, the methylated CpG island associated with the repressive marks histone H3 lysine 9 di-methylation (H3K9me2) and histone H4 lysine 20 tri-methylation (H4K20me3). To investigate the role of DNA methylation in the establishment and maintenance of *TMS1* gene silencing we treated cells, methylated for *TMS1,* with the DNA methyl transferase inhibitor 5-aza-2'deoxycytidine (5-azaCdR). *TMS1* demethylation, induced by 5-azaCdR, was accompanied by gene re-expression, the loss of H3K9me2, and the re-acquisition of H3ac, H3K4me2, and RNA Pol II. Interestingly, H4K20me3 was unaffected by 5-azaCdR treatment. Furthermore, we kept these cells in culture following 5-azaCdR treatment and found that a sub-population of unmethylated *TMS1* alleles is stably maintained for over three months in the absence of drug. This stable, drug-induced state does not fully recapitulate the chromatin structure of the unmethylated *TMS1* locus, providing evidence for two independent repressive pathways negatively regulating *TMS1*. DNA methylation and H3K9me2 correlated with a lack of H3ac, H3K4me2, and Pol II at *TMS1*. The other repressive pathway was highlighted by the presence of H4K20me3 that prohibits H4K16ac at *TMS1*. To expand our understanding of the role that DNA methylation plays at *TMS1,* we next examined kinetics of DNA demethylation and remethylation following a transient treatment with 5-azaCdR on a genome-wide scale. We found CpGs can be categorized into three basic groups of demethylation/ remethylation kinetics. Those CpGs that are normally methylated in breast epithelial cells exhibited significant demethylation followed by a complete remethylation over a period of three months in culture. In contrast, those CpGs specifically methylated in cancer showed two distinct behaviors. One group demonstrated minimal demethylation followed by a partial but incomplete remethylation. This group was significantly enriched for genes occupied by SUZ12 in embryonic stem cells, suggesting these loci may be marked by histone H3 lysine 27 tri-methylation (H3K27me3) in normal tissues. The other group exhibited a substantial DNA demethylation followed by little to no *de novo* methylation. This latter class was not enriched for genes occupied by SUZ12 in embryonic stem cells. This finding suggests that the mechanisms underlying *de novo* methylation may differ between individual CpGs. Together, these data provide insight into how cancer specific methylation is established and may help to improve existing therapeutic strategies involving epigenetic-targeted agents.

Role of chromatin in resilencing dynamics after drug-induced DNA demethylation

By

Jacob D. Kagey
B.A. Ohio Wesleyan University, 2004

Advisor: Paula M. Vertino, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in
Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

2010

Acknowledgements

The work found in this thesis would not have been possible without the help, love, and support of numerous people.

I want to thank my parents, Richard and Betsy Kagey, for instilling in me a strong work ethic and, more importantly, a sense of humor. You have taught me that there are many situations in life when all you can do is laugh, which has been invaluable the past five years. You have always encouraged me to pursue a line of work that was both challenging and fulfilling. Without your unending love and support I would not be here today. And Mom, you will always be the original Dr. Kagey.

I have had many teachers who have nurtured my love of science and inspired me along the way. I am indebted to Mr. Kevin Froats, my fifth-grade science teacher; Dr. Wayne Longenecker, my AP Biology teacher; and Dr. Gary Hendrickson, my AP Calculus and Algebra II teacher. Their passion for their subject area, combined with their encouragement of my math and science education, helped me develop a strong foundation in and love of the natural sciences.

During my time at Ohio Wesleyan University, my love of science was transformed into a love of genetics and bench research. I want to thank Dr. Jed Burtt, in particular, for instilling in me a love of research through his sheer passion and love of Ornithology. He is a wonderful research mentor, fellow Detroit Tigers fan, and a good friend. I also thank Jed for being the first person to ever tear apart my writing. I thank Dr. Danielle Hamill for being a wonderful mentor, a voice of reason during trying times, and a great advisor. Danielle pushed me in directions that were not necessarily where I wanted to go, but ultimately were good for my career.

I would like to acknowledge Dr. Mark Nelson at the University of Arizona for providing me with my first experience in cancer research. When I arrived in Tucson, I had never even touched a pipette.

I am grateful to Lynn Togut and Carl McInroy for serving as my surrogate parents when I first moved to Atlanta. You were the only people I knew when I first moved south and you have always made me feel welcome and loved throughout the years.

The members of the Vertino lab—both past and present—have all been like a family to me for the past five and a half years. In particular, I would to thank Doris Powell for always being a good friend and great resource for everything from trouble-shooting a sequencing gel to finding a good dentist. I thank Dr. Priya Kapoor-Vazirani for always making me laugh and for being a great collaborator within in the lab. I would like to thank Pritty Patel for being a good friend and lab-mate—getting over all of the hurdles along the way has been much easier with your help. I thank Dr. Mike McCabe, not only for being an unending source of tools, help on home improvement projects, and advice, but also for also serving a great role model and mentor. Your drive and understanding of biology have been inspiring.

I would like to thank the members of my committee: Dr. Jerry Boss, Dr. Tamara Caspary, Dr. Bill Kelly, and Dr. Harold Saavedra. You all have been wonderful mentors and I have greatly enjoyed the discussions we have had throughout the years. I cannot thank you enough for your scientific input and support of my development as a scientist and as a person. Harold, I would like to thank you also for your mutual devotion to the Atlanta Hawks.

I am grateful for the comedic talent of Bill Simmons, Bill Maher, and J.E. Skeets and Tas Melas. Their continuous production of podcasts always provided me with a source of laughter. Together, you made the ridiculous number of plasmid mini-preps almost bearable—almost.

I am greatly indebted a number of people who have nurtured my development as a teacher during my time in graduate school. In particular I would like to thank Dr. Pat Marsteller, whose pedagogical guidance has been invaluable. I would also like to thank Dr. David Lynn, Jordan Rose, Dr. Adam Marcus, Dr. Lynn Zimmerman, Malaika Jordan, and Dr. Karen Schmeigal, who have provided me with the opportunity to develop as a teacher and also have provided wonderful feedback and discussion.

To my advisor, Dr. Paula Vertino, I want to thank you so much for taking me into your lab and being a wonderful mentor. You have always provided support for all of my teaching endeavors while, ensuring I never lost sight of the research at hand. I have received a wonderful education in your lab. You have been an unwavering supporter of mine and I cannot thank you enough for all of our scientific and life discussions (as well as help with my writing). You have been a great teacher and friend.

Lastly I would like to thank my wife, Emily. I could write another 200 page dissertation explaining what your love and support has meant to me over the past nine-plus years. In particular, thank you for making me smile and laugh. No matter how dire things seemed in lab, you have always been able to give me some perspective and make me laugh. This dissertation is full of moments when you helped pick me up off the floor and put me back on my feet. Many thanks and much love.

Contents

List of Figures and Tables

Figure 3: Model of how chromatin structure may protect the unmethylated *TMS1* locus from position effect variegation.
Figure 4: *TMS1* is regulated by two independent epigenetic repressive mechanisms.
Figure 5: CpGs fall into three distinct groups of remethylation kinetics following the treatment and removal of 5-azaCdR.

Chapter 1: Introduction

Introduction

Cancer genetics has been the focus of study for over a century. Advancements in cancer genetics have paralleled the advancements of molecular genetics, with our understanding of the two fields co-evolving. The work of Theodore Boveri in 1902 first suggested that tumor cells exhibit an unusual cytology, and might carry an irreparable and heritable defect (1). This idea was furthered by Peyton Rous who, in 1911, conducted a series of seminal experiments that identified a filtered substrate (the Rous Sarcoma Virus) capable of inducing sarcoma in healthy chickens. This was the first demonstration that normal healthy cells have the ability to transform into cancer cells (2). These concepts were developed prior to the identification of the molecular vehicle responsible for the transmission of heritable material from generation to generation, and further advancement was delayed until this discovery. In 1944 Avery, *et. al.* (3) published a seminal study that defined DNA as the genetic material. These experiments demonstrated DNA as the essential carrier of bacterial transformation. A decade later, work from James Watson, Francis Crick, Rosalind Franklin, and R.G. Gosling provided insight into the structure of DNA as well as the foundation of how that structure related to heritability (4, 5). Combining the early discoveries of Rous and Boveri with the basic understanding that DNA as the carrier of genetic material created the foundation necessary for major advancements within the field of cancer genetics. In the 1970s two classes of cancer related genes were identified: tumor suppressor genes (TSGs) and oncogenes (6, 7). Alterations leading to the inactivation of TSGs or the hyperactivity of oncogenes were demonstrated to be the heritable force driving carcinogenesis through the disruption of normal cellular processes these genes normally regulate. Initially, it was thought that

these alterations were achieved solely through mutations in the DNA sequence. In 1989, Greger, *et. al.* first demonstrated that TSGs can also be inactivated epigenetically in cancer through promoter hypermethylation while maintaining their wild type DNA sequence, and that this event was phenotypically identical to an inactivating mutation (8). This finding opened the door into a new field of cancer genetics, cancer epigenetics. The focus of this thesis is to further the understanding of the mechanisms surrounding the establishment and maintenance of epigenetic silencing of TSGs during carcinogenesis and the molecular alterations achieved through the use of epigenetic therapies.

Cancer

Cancer is a disease characterized by the uncontrolled proliferation of a group of cells distinguished by a series of acquired phenotypes including self-sufficiency in growth signals, insensitivity to anti-growth signals, the ability to avoid apoptosis, tissue invasion and metastasis, sustained angiogenesis, and limitless replicative potential (9). Cells acquire these alterations over time and the accumulation of these insults can result in transformation. Furthermore, these phenotypes can be acquired in different orders. Cancers are defined through their tissue of origin and exhibit tissue-specific characteristics in addition to the aforementioned general acquired phenotypes.

From a genetic point of view there are two basic underlying mechanisms in which cancer is established, familial and sporadic. Familial cancer refers to patients who have a germline mutation correlating with an increased risk of cancer development. The identification and understanding of TSG function was discovered through studies of a familial case of retinoblastoma. This ultimately led to the characterization of the well-

known TSG, *Retinoblastoma* (*Rb*) (10). Patients born with an inherited mutation in a TSG leads to an increased likelihood of developing cancer, which is inherited in a dominant fashion. Some examples of well-studied familial mutations that lead to an increased pre-disposition to cancer include *Rb, APC,* and *BRCA1.* Sporadic cancer arises wholly from the somatic acquisition of mutations and epigenetic insults. Sporadic cancers comprise ~90% of all cancers, and their prevalence is influenced by exposure to carcinogens (*i.e.* smoking, UV irradiation).

The prevalence of different cancer types is dependent on a number of factors including age, race, sex, ethnicity, geographical location, and diet. In females, breast cancer is the most prominent cancer diagnosed and is the 4th leading cause of cancer death. This highlights both the overall impact of breast cancer on the population, but also demonstrates recent advances in both the treatment and screening of breast cancer. In 2005 approximately 1.4 million Americans were diagnosed with cancer, including 211,240 cases of breast cancer. Additionally, 40,410 deaths were associated with breast cancer in the same year (11).

Cancer Genetics

At the molecular level, cancer is a genetic disease derived from the combination of increased oncogene activity and the loss of tumor suppressor gene activity. Proteins encoded from tumor suppressor genes or proto-oncogenes are often key regulatory factors governing normal cellular processes (9). In general, oncogenes function to promote cell growth and division, or confer resistance to apoptosis. Often oncogenes are analogously taught as the gas pedal of an automobile. Mutations in proto-oncogenes act

in a dominant fashion, in that only one copy of an oncogene is required to facilitate the phenotype. There are many different mechanisms leading to gain of function alterations in oncogenes including mutations, insertions, translocations, and duplications. The discovery of oncogenes can be traced back to the work of Peyton Rous. Through a series of experiments, Rous demonstrated that a filtered tumor substrate (cell free) was capable of inducing sarcoma in a healthy bird (2). Subsequent studies of the Rous Sarcoma Virus (RSV) demonstrated that the DNA responsible for the cellular transformation was not unique to the viral genome, but was related to a gene within avian genome and in fact, the proto-oncogene *src* (7). This finding allowed for the understanding of proto-oncogenes, which normally function in the regulation of cellular growth and development. When proto-oncogenes are altered they are capable of contributing to cellular transformation in a dominant fashion. Several other examples well studied of proto-oncogenes include *myc* and *ras*.

Tumor suppressor genes function as regulators of cell growth and division, DNA repair, apoptosis, etc. Analogously, TSGs are taught as the brakes of the car. Both copies of these genes (both sets of brakes) must be inactivated in order to contribute to cellular transformation. Early studies of TSG genes focused on cancer syndromes such as retinoblastoma in children, who acquired the disease through germline mutations or sporadically. This concept of TSGs was formalized through the work of Alfred Knudson, who found that patients with a family history of retinoblastoma typically acquired the disease earlier life and as a multi-focal disease. This was in contrast with patients who developed the disease sporadically tended to be older and develop a single focal disease (6). These findings lead Knudson to the two-hit hypothesis, which postulates that both

copies of TSGs must be inactivated for disease progression (**Figure 1**). Patients born with one mutant copy of *Rb* have an increased likelihood of developing cancer, since one copy of *Rb* is inactivated in all cells. From these works it was determined that both copies of the *Rb* gene must be inactivated for the promotion of cancer. If one copy remains functional it will still be able to function as a negative regulator (10). Furthermore, cancer cells could be suppressed when a wild type copy of the gene was re-introduced (12). Since the discovery of *Rb* many other TSGs have been identified, including *p53*, *APC, VHL* and, *CDH1* (13).

The inactivation of tumor suppressor genes can occur through a variety of mechanisms, including inherited or acquired inactivating mutations, deletions, or translocations (**Figure 2**). Although there are numerous mechanisms capable of inactivating TSG function, the manner in which each copy is inactivated is not imperative. Point mutations in the DNA sequence that disrupt normal gene function have often been found to inactivate tumor suppressor genes and contribute to cancer, as have deletion and insertion mutations. More recently it has been discovered that alterations in the epigenetic regulation of TSGs in cancer also have the potential to inactivate a tumor suppressor gene through transcriptional silencing (13). In fact a study of breast cancer patients with a mutation in the gene *BRCA1* were demonstrated to be phenotypically identical to patients who had *BRCA1* silenced epigenetically (14).

Epigenetics

Epigenetics is the study of heritable changes in gene regulation that are independent of alterations in DNA sequence. The term epigenetics is derived from the

Greek prefix ‘epi’ meaning over, giving the literally translation of epigenetics to mean over or above the genetics. Multi-cellular organisms arise from a single cell and during development exhibit a need for differential gene expression as all somatic cells house an entire copy of the genome (15). Although there are a number of “house keeping” genes expressed ubiquitously, many genes exhibit a need for differential regulation dependent on cell type, developmental stage, or environmental stress (16). This need for differential gene regulation is where epigenetic gene regulation comes into play. In 1940, C.H. Waddington first proposed the notion of epigenetic regulation during development through the observation of wing development in *Drosophila melanogaster*. Waddington used the term ‘epigenesis’ to describe the utilization of different genes in the development of different tissues in the adult organism (17). The general function of epigenetic regulation is the same across species, however variations exist in the molecular mechanisms by which this regulation is accomplished. In this introduction mammalian, in particular human, epigenetic regulation will be the focus.

In mammals the two most widely studied epigenetic factors are DNA methylation and the post-translational modifications to histone tails (18). It is important to note that other factors also contribute to epigenetic regulation including the positioning of nucleosomes and the binding of non-histone proteins. The concerted effect of these epigenetic factors determines the transcriptional capability of specific loci. Chromatin, the combination of DNA and histone proteins, is broadly classified into two types, heterochromatin and euchromatin. Regions of the genome unfavorable to gene transcription are known as heterochromatin. Heterochromatin was first defined in 1928 as regions of the chromosome that remained condensed and darkly stained throughout the

cell cycle. Initially, it was hypothesized that heterochromatin was a region of the chromosome that did not contain any genes and therefore remained constitutively condensed (19). While indeed, there are regions of constitutive heterochromatin encompassing repeat elements, telomeres, and centromeres there are also regions of cell-type specific heterochromatin known as facultative heterochromatin. These regions include genes silenced in a particular cell type (20). Euchromatin encompasses regions in the genome that are only easily visible during chromosomal condensation during mitosis and are open and permissive to gene expression (21). Below the most well studied epigenetic entities that together characterize regions of the genome as either permissive or repressive to transcription are discussed in greater detail.

DNA Methylation

DNA methylation is not unique to mammals; plants, bacteria, and other organisms exhibit both CpG methylation and non-CpG methylation. In mammals, DNA methylation occurs at the 5' carbon of cytosine bases within the context of the CpG di-nucleotide (**Figure 3**) (22). CpG methylation is an essential component of epigenetic regulation within humans. X-chromosome inactivation, imprinted loci, repeat elements, and cellular differentiation are all processes that depend on DNA methylation for proper regulation. Early in development (embryogenesis) the genome undergoes an epigenetic erasure, removing the majority of the DNA methylation present from the sperm and egg genome. As the embryo develops *de novo* methylation occurs, invoking differential methylation patterns in the developing tissues, as well as the ubiquitous methylation of repeat elements (16). Methylation also plays a major role in human dosage compensation.

Humans accomplish dosage compensation through X-chromosomal inactivation in females by utilizing the combination of DNA methylation, non-coding RNAs, and histone modifications. Overall, CpG methylation is associated with transcriptional repression and is essential for the maintenance of differential gene expression in humans and maintaining genomic integrity.

Within the human genome the frequency of CpG di-nucleotides is lower than expected (23). However, 70% of genes contain a cluster of CpGs, often at the 5' regulatory region of the gene, known as CpG islands (24). Since this is the same region many transcriptional proteins bind, the CpG island serves as an epigenetic regulator of gene expression. In general, most CpG islands are unmethylated in normal cells and permissive to transcription. However, the methylation of certain CpG islands in differentiated cells provides a mechanism for the heritable silencing of genes unused by that particular cell type (25).

In humans, DNA methylation is mediated through DNMTs (DNA Methyl Transferases). *De novo* methylation, the new methylation of previously unmodified residues, plays an important role during development through the establishment of tissue specific methylation patterns, and is carried out by the enzymes DNMT3a/b. Furthermore; *de novo* methylation is capable of promoting differential gene regulation through the methylation and subsequent gene silencing of certain CpG islands (25). The methylated CpG di-nucleotide is a palindrome and provides a mechanism for the heritability of DNA methylation patterns. Maintenance methylation refers to the process of recapitulating the DNA methylation patterns following replication; this process is facilitated by DNMT1 (25).

At this point it seems that mammalian DNA methylation is a permanent modification (26). It should be noted that several recent studies have found evidence for the existence of DNA demethylating activity within human cells, and active DNA demethylation has been demonstrated in other organisms such as plants. A number of these studies have demonstrated DNA de-methylation activity through the use of a glycosylase. Other DNA de-methylating studies implicated MBD2 as an active de-methylase (27). However, this idea remains controversial, as other groups have reported an inability to repeat these findings.

Post-translational Histone Modifications

The human genome is approximately two meters in length and must be condensed and functionally organized in every cell. Approximately 146-bp of DNA is wrapped around an octomer of histone proteins (two each of H2a, H2b, H3, and H4) creating the basic repeating unit of chromatin, the nucleosome (28, 29). The manner in which the DNA is packaged with histones plays a role in the determining transcriptional potential of that particular region of DNA. In 1964, a study by Allfrey *et. al.* reported that histones were both acetylated and methylated following transcription. Although, the specific residues acetylated and methylated were unknown, both of these modifications were initially correlated with gene expression (30). Since this initial report, subsequent studies have revealed a much greater detail concerning the post-translational modifications of histone proteins. The N-terminal tails of histone proteins are subjected to considerable post-translational modifications including acetylation, methylation, phosphorylation, and ubiquitination. Certain modifications are associated with active transcription

(euchromatin) such as histone H3, lysine 4 di-methylation or tri-methylation (H3K4me2/3), acetylation of histone H3 on lysines 9/14 (H3ac), and acetylation of histone H4 lysine 16 (H4K16ac). Conversely, histone H3, lysine 9 di-methylation or tri-methylation (H3K9me2/3), histone 3 lysine 27 di-methylation or tri-methylation (H3K27me2/3), and histone H4 lysine 20 tri-methylation (H4K20me3) associate with regions repressive to transcription (heterochromatin) (18). Combinations of histone tail modifications are proposed to work together to create a histone code, which dictates higher order chromatin structure and the potential for that gene to be expressed (**Figure 4**) (31).

Enzymes that have the ability to add or remove specific post-translational modifications dynamically regulate the histone code. These complexes are specific for both the type of modification and the residue on the histone tail they modify. Histone acetylation is regulated through histone acetyl transferases (HATs) and histone de-acetylases (HDACs). HATs attach an acetyl group to lysine residues on histone tails, some examples of well-characterized HATs include Gcn5, which is responsible for the acetylation of H3K9/14, and hMOF, which is responsible for the acetylation of H4K16. Conversely, the de-acetylation of histone tails is accomplished through HDAC complexes such as the Co-REST and NuRD complexes (32).

Another dynamic histone modification is lysine methylation. Histone methyl transferases (HMT) facilitate the addition of methyl groups to lysine residues on the histone tails through the use of a SET domain (32). Unlike acetylation, histone tails can be mono-methylated, di-methylated, or tri-methylated. Some examples of HMT include EzH2, which is responsible for the trimethylation of H3K27, Suv4-20, which is

responsible for the di-methylation and tri-methylation of H4K20, Set7, which is responsible for the mono-methylation of H3K4, and Suv3/9 which is responsible for the mono-methylation of H3K9 (32, 33). Recently, several known histone de-methylases have also been reported including the lysine specific de-methylase (LSD1) (34). There are numerous other histone modifying enzymes, these examples are to serve as an example of the extent and specificity of histone modifying enzymes.

Like DNA methylation, patterns of histone modifications are though to be inherited during mitosis such that the histone code is heritable and each cell is able to recapitulate the expression profile present in the parent cell (18). The mechanism of this heredity is still controversial. Some favor a semi-conservative replacement of histones, similar to DNA replication, while others support the notion that entire nucleosomes are segregated to each DNA molecule and the newly deposited histones ascertain their histone modifications from surrounding histones that remained (35). Recent work has uncovered that the histone code is more complex than initially hypothesized. While the above marks are generally associated with either active or repressed domains, other marks have a more intricate role. For example, histone H3 lysine 36 tri-methylation (H3K36me3) is a functionally repressive mark, typically found in the body of active genes. It is thought that the presence of H3K36me3 prevents aberrant transcripts from re-initiating (36). Furthermore, in embryonic stem cells bivalent chromatin domains, which display both H3K27me and H3K4me, are observed. These bivalent domains are thought to represent genes that are poised for either activation or repression once cells undergo cellular differentiation, thus creating a more dynamic epigenetic regulation (37).

Functionally, the role each individual histone modification remains incompletely understood. Again, function is dependent on the particular modification. For example, H3K9me has been shown to interact with DNMTs and recruit DNA methylation (38). Another example of specific histone modification function is H4K16ac, which has been demonstrated to relax the DNA and histone contacts, results in the opening of the chromatin (39). Although a few studies have identified the specific role of certain modifications in specific contexts, the overall understanding of the contribution of individual histone modifications to the overall epigenetic regulation of chromatin domains remains poorly understood.

Nucleosome Positioning

One aspect of epigenetic regulation that has undergone a great deal of study recently is nucleosome positioning. A nucleosome that is present at the same genomic location throughout a population of cells is known as positioned. Positioned nucleosomes exist throughout the genome, but are not everywhere. Recently, it was estimated that half of positioned nucleosomes in the yeast genome could be accounted for strictly by DNA sequence (40). The remainder of positioning is attributed to interactions between histone proteins and specific DNA sequences, the binding of non-histone proteins, and neighboring positioned nucleosomes (41-43). DNA methylation does not alter the ability of histone proteins to bind DNA *in vitro*, but can alter nucleosome positioning *in vivo* through indirect mechanisms (*i.e.* through either the promotion or inhibition of non-histone protein binding) (43, 44).

Functionally, the positioning of a nucleosome has been shown to either maintain or prohibit gene activity depending on the context (42, 45). For example the *HPRT* locus demonstrates differential patterns of nucleosome positioning in both the active and silent state. In the inactive state the *HPRT* promoter exhibits one pattern of nucleosome positioning, while in the active state demonstrates a different and unique pattern, and both patterns are integral to their respective regulation. It is thought that the differential positioning provides accessibility to different protein binding sites (42).

Another hallmark of nucleosome positioning is the creation of a nucleosome-free region, initially characterized in yeast, which is a ~400bp region overlapping the transcriptional start site of active genes. The presence of RNA polymerase II (Pol II) and other transcriptional initiating factors are bound in this region and their presence is inhibitory to nucleosome occupancy. This phenomenon was recently demonstrated in humans (46). Typically, the nucleosomes immediately flanking the nucleosome-free region are strongly positioned (40).

Overall, nucleosome positioning has a role in both repressive and active gene regulation. Whether positioned nucleosomes are the consequence of other epigenetic factors, such as DNA methylation or histone modifications, or whether positioning leads to other downstream epigenetic events remains unresolved and may be context dependent.

Chromatin Domains and Insulators

DNA methylation, histone tail modifications, and positioned nucleosomes work together to organize the genome into chromatin domains, which are regions of structured

chromatin that regulate genes in a similar manner. Maintaining the integrity of different, neighboring chromatin domains is vital for proper gene regulation. The regulatory region located between two independently regulated chromatin domains is known as an insulator. Insulators demonstrate two related yet distinct functions, barrier activity and enhancer blocking activity. Individual insulators have been described as having either one or both of these functions. Barrier activity refers to the ability of an insulator to prevent the spread of heterochromatin from one domain to another, also known as position effect variegation (PEV). Enhancer blocking activity refers to the ability of an insulator to blocking a distal enhancer from functioning on a gene. Proteins often bind to insulator sequences and facilitate the barrier or enhancer blocking activity (47). In humans, several insulator proteins such as CTCF (CCCTC binding Factor) and USF-1 (Upstream Stimulatory Factor 1) have been described. The loss of an insulator protein binding can lead to the loss of function of that particular insulator (48).

Insulators were initially described through studies surrounding the *gypsy* retrotransposon in *Drosophila melanogaster*. It was found that *gypsy* insertion was capable of repressing the expression of a certain genes when inserted between the gene and its proximal enhancer. Suppression studies discovered that this insulator function was mediated through a DNA binding protein, *suppressor of (Hairy-Wing) (su (HW))*(49). The normal cellular function of *su (HW)* is to serve as a mediator of insulator function throughout the *Drosophila* genome through its binding to the insulator sequence. Subsequent studies have discovered other *Drosophila* insulator protein binding proteins allowing for the cell specific organizations of insulators (50).

In vertebrates the first insulator described was a DNase I Hypersensitive site (HS) found upstream of the *β–globin* in chickens (HS4) (51). The maintenance of this insulator has been shown to be essential for maintaining independent gene expression patterns between *β –globin* and neighboring genes. As with the gypsy insulator sequence, insulator activity within HS4 at the *β –globin* is mediated through the binding of a zinc finder protein, CTCF (52). Protein analysis revealed that CTCF has eight zinc fingers which mediate its binding to DNA. The combinations of eight zinc fingers provide CTCF with a multitude of potential DNA binding configurations and potential binding sequences. CTCF has been shown to bind insulators in a methylation sensitive manner providing another level of insulator regulation (53). Subsequent studies identified another mammalian insulator protein binding at HS4, such as USF-1 (54). The loss of insulator binding to a barrier between chromatin domains can lead to the spreading of chromatin domains and disruption of epigenetic regulation (48). More recent investigations into chromatin architecture have revealed that insulators not only co-localize within the nucleus, but that they also physically interact within the context overall nuclear organization, creating inter-chromosomal domains that are regulated in a similar fashion (55).

Epigenetics in Cancer

During carcinogenesis global alterations in epigenetic regulation are observed. DNA methylation patterns are distorted in cancer (56). Gama-Sosa *et. al.* reported in 1983 that global levels of CpG methylation in cancer cells was significantly lower than normal samples (57). It was later shown that this global hypomethylation correlated with

genomic instability, further contributing to cancer progression (58). Concurrent with global hypomethylation it was discovered by Baylin, *et. al.* that the *Calcitonin* gene was hypermethylated during carcinogenesis (59). Since this initial discovery a number of other CpG island containing genes have been demonstrated to be hypermethylated and silenced in cancer. A subset of these genes aberrantly methylated includes TSGs, the first such gene reported was *Rb* (8). Genome-wide methylation studies have shown distinct subsets of CpG islands are prone to frequent hypermethylation during cancer progression (60). A fraction of the CpG islands targeted for this hypermethylation are TSGs, and their hypermethylation correlates with gene silencing (61). Some commonly studied CpG island containing TSGs include *BRCA2*, *VHL,* and *p16 (62).* The aberrant methylation of TSGs is phenotypically identical to inactivating mutations, thus providing further supporting the notion that the method of TSG inactivation is extraneous (14). Although the molecular mechanisms underlying the epigenetic silencing of specific TSGs remains unclear, the work of several groups has found that underlying DNA sequence signatures and the presence of H3K27me3 in normal cells increased the likelihood of cancer-specific hypermethylation (63-67). Overall, the combination of global hypomethylation and the aberrant hypermethylation of certain TSGs denote that DNA methylation changes contribute to the progression of cancer in two diverse manners.

Epigenetic alterations in cancer also involve dramatic changes in both global and local histone modifications. On the global scale, tumors exhibit a decrease in H3K9me2, H4K20me2/3, and H4K16ac. This global decrease of the repressive marks (H3K9me3 and H4K20me3), normally enriched in constitutive heterochromatin, can synergistically function with the global hypomethylation, leading to not only genomic instability, but

also aberrant transcription from repeat elements (24, 68). Another hallmark of the cancer histone code is the repressive mark H3K27me3. A few well-studied TSGs have demonstrated the acquisition of H3K27me3 along with DNA methylation in cancer such as *CDH1* (69). However, the major focus of study of H3K27me3 has been surrounding the targeting of CpG islands for aberrant methylation. Several groups have demonstrated that the occupancy of the Polycomb Repressor Complex 2 (PRC2), (which contains the H3K27 HMT, EZH2) in embryonic stem cells is an accurate predictor of islands that are targeted for methylation and silenced during carcinogenesis. One currently hypothesis is that members of the PRC2 complex (responsible for depositing the H3K27me3 mark) interact with DNMTs thus targeting aberrant methylation during carcinogenesis (63, 66).

At individual aberrantly silenced loci there is an overall shift from an active to a more repressive histone code. In general, H3ac and H3K4me2/3 demarcate the active genes in normal cells. During gene silencing in cancer, these marks are typically replaced by the repressive marks H3K9me2/3 and /or H3K27me3. The combinatorial changes in DNA methylation and histone modifications create a heritable and repressive domain capable of inhibiting TSG activity (**Figure 5**). This gene silencing has also been shown to exhibit a loss of nucleosome positioning at certain loci such as *Mlh1,* and DNase I HSs formation, suggesting the alterations in DNA methylation and the histone code lead to changes in the overall chromatin packaging, creating a more heterochromatic domain (70, 71).

Although the end result of aberrant epigenetic silencing is understood at specific loci, the order of operations that lead to this silencing is still incompletely understood. There have been a few key studies providing evidence both that repressive histone

modifications such as H3K9me have the ability to interact with DNMTs and target *de novo* methylation (38, 72, 73). Conversely, CpG methylation has shown to recruit methyl-binding proteins, including HDACs, which in turn hypoacetylate associated nucleosomes (74). Methyl binding proteins can also recruit chromatin-remodeling complexes, including swi/snf, which results in the overall shift to a more repressive chromatin domain (75). Overall, these data point to a model in which there is substantial cross talk between DNA methylation and repressive histone modifications during epigenetic gene silencing in cancer.

Epigenetic therapies in Cancer

Epigenetic silencing of tumor suppressor genes provides a unique opportunity for therapeutic intervention. Unlike mutations or deletions, which physically disrupt the gene sequence, reactivation of epigenetically silenced tumor suppressor genes is possible because the wild-type DNA sequence remains. DNMT inhibitors, 5-aza-2'-deoxycytidine (5-azaCdR) and 5-azacytidine (5-aza) are clinically effective in their treatment certain cancer types and are currently the standard of care for myelodysplasitc syndrome (MDS). In addition, there are ongoing clinical trials utilizing these agents for the treatment of lukemias and solid tumors (76). Both 5-aza and 5-azaCdR interfere with the activity of DNMTs. This activity requires their incorporation into DNA as cytosine analogs during S phase. Once incorporated, they form a stable complex with DNMTs, trapping and targeting them for degradation. 5-aza is also capable of RNA incorporation; although to degree to which this influences clinical activity is unclear (77). Generally, 5-aza and 5-azaCdR are thought of as having similar effects on DNA methylation, although one study

did find differences in the re-expression profiles when treating the same cell line with the same concentrations of both drugs, however they did not investigate differences in DNA de-methylation (77). For the purpose of this introduction we will further discuss only 5-azaCdR.

The use of 5-azaCdR in the treatment of cancer has experienced two waves of clinical success. The first studies utilizing 5-azaCdR began in the 1960s, pre-dating any understanding of the molecular nature of its activity, when the compound was discovered to have anti-tumor effects, (78). Early clinical trials were conducted with high doses of 5-azaCdR and demonstrated some clinical success in the treatment of hematologic cancers (79). Although these early clinical trials demonstrated strong anti-tumor effects they were also accompanied by severe side effects including myelosupression, and the studies were abandoned. The pitfall of these studies was the correlation between dosage and anti-cancer effect (at higher dosage levels); patients were given the maximum tolerable dose leading to greater side effects. It was not until a study by Jones and Taylor, demonstrating that 5-azaCdR was more potent at inducing DNA de-methylating at lower doses that the use of 5-azaCdR in the clinic began a renaissance (80). At higher doses, 5-azaCdR inhibits DNA synthesis and can lead to cell cycle arrest, DNA damage, and, eventually cell death because its mechanism of action requires the incorporation into DNA (76). More recent clinical trials involving 5-azaCdR have utilized lower doses (with extended treatment cycles) and resulted in improved anti-tumor activity without toxic side effects. These clinical trials have found 5-azaCdR to improve response and survival of MDS patients when compared with traditional cytotoxic chemotherapy Thus, 5-azaCdR is now the standard of care for MDS. Currently, studies of low-dose treatment regimens with 5-

azaCdR are currently in clinical trials for leukemias and solid tumor types (76). Molecular analyses of bone marrow biopsies from patients treated with these agents have demonstrated that global and gene-specific DNA demethylation is achievable *in vivo* (81, 82). The degree of demethylation varies between patients. Whether this is an important indicator of clinical response remains controversial, and may ultimately depend upon the compartment being analyzed (*i.e*. repetitive element methylation versus methylated tumor suppressor genes) or the surrogate marker measured (DNA methylation versus gene expression) (83).

Treatment with 5-azaCdR has been demonstrated to lead to the demethylation and reactivation of epigenetically silenced TSGs (76). This demethylation is typically accompanied by a shift from a repressive to a more active histone profile at, including the loss of H3K9me2 and the re-acquisition of H3ac and H3K4me2 (70, 84, 85). However, studies by McGarvey *et. al.* have shown that this transition is not complete, as several marks were unaffected by treatment with 5-azaCdR (H3K27me3 and H3K9me3) suggesting their occupancy is not directly linked to the DNA methylation status (84). The principal hypothesis behind the efficacy of low doses of 5-azaCdR is that epigenetically silenced TSG are reactivated and can resume their function in monitoring key regulatory pathways resulting in these cells growth arrest or apoptosis. Studies in cell culture with 5-azaCdR have supported this hypothesis and shown that the reactivation of previously silenced pro-apoptotic genes re-establishes the ability of these cell to undergo programmed cell death (86). Clinically, this hypothesis has been more difficult to test since cancer cells that again become sensitive to apoptotic signals are unable to be assayed in the patient (76).

As discussed in the previous section (*Epigenetic alterations in cancer*) DNA methylation is not the only epigenetic factor to change during carcinogenesis, shifts in the post-translational histone modifications are also observed. One common theme among epigenetically silenced TSGs is the shift to a more repressive histone code, including hypoacetylation of histones concurrent with the addition of repressive marks (H3K9me2/3 and H3K27me3). Recently, HDAC inhibitors (HDACi) also have been the focus of developing novel epigenetic therapies. Treatment with HDACi leads to a global hyper-acetylation of histones, which results in the relaxation and opening of the chromatin (87). Treatment with HDACi also leads to the hyperacetylation of certain non-histone proteins; also substrates of HDACs, some of which are directly are involved with transcription (88). Cell culture studies found that the combinatorial treatment of cancer cells with 5-azaCdR and HDACi lead to the synergistic reactivation of silenced TSGs (89). Clinically, HDACi have demonstrated efficacy in cancer treatment both alone and in combination with 5-azaCdR (90-92).

TMS1 as a Model of Epigenetic Gene Silencing

To model the epigenetic silencing of tumor suppressor genes in cancer, a cell culture model was developed in which immortalized lung fibroblast cells and were stably transfected with a construct to over-expressing DNMT1. This over expression led to the aberrant hypermethylation at several specific CpG islands investigated, including *ESR1, Hic-1,* and, *CDH1* (93). Additionally, large scale CpG island methylation analysis was conducted via restriction landmark genomic scanning (RLGS) showed that not all CpG islands underwent hypermethylation in response to elevated levels of DNMT1, indicating

that there are differences in the inherent susceptibility of CpG islands to *de novo* methylation (94). Subsequent studies into this phenomenon found that both underlying DNA sequence signatures and histone modifications contribute to pre-determination of a CpG island's susceptibility to hypermethylation, thus creating islands that are either methylation prone or resistant (63, 94). One such CpG island that underwent hypermethylation in these cells was a novel gene named *Target of Methylation-mediated Silencing 1* (*TMS1*) (61, 93). Following its discovery, *TMS1* has been further pursued by our lab as a model of epigenetic gene silencing in cancer.

The *TMS1* gene resides on chromosome 16 and is composed of three exons (of which the second is alternatively spliced, creating two distinct mRNA isoforms. A ~1 kb CpG island overlaps the promoter region and transcription start site (**Figure 6**) (61). *TMS1* was methylated and transcriptionally silent in a number of breast cancer cells lines and in a subset of primary breast tumor samples. While not all samples cell lines exhibited a methylated *TMS1*, a direct correlation was observed between DNA methylation status and gene expression (**Figure 7**) (61). *TMS1* is also silenced in a subset of prostate, and lung cancers (61, 95, 96). Functional analyses of the TMS1 protein have revealed a role in both the apoptotic and inflammatory response pathways. *TMS1*'s role in inflammation has been shown to be through the activation of NFκB (97). In apoptosis *TMS1* was shown to play a role in the promotion of TNF-α mediated apoptosis mediated (95, 98). Recently, TMS1 has also been shown to have a role in detachment-induce apoptosis (anoikis) in normal breast epithelial cells, in part through the regulation of ERK phosphorylation. The role of TMS1 in anoikis was supported with the finding that transformed cells filling the breast epithelial lumen during ductal carcinoma *in situ*

(DCIS) lacked TMS1 expression, while those cells surrounding the lumen maintained expression. These findings support the notion that the epigenetic silencing of *TMS1* promotes resistance to anoikis thus contributing to the progression of DCIS (99). These initial investigations of the aberrant methylation of the *TMS1* CpG and subsequent alterations in phenotypes evidence that the epigenetic silencing of *TMS1* contributes to tumor progression.

Objectives

Although *TMS1* was found aberrantly methylated and silenced in a number of both breast cancer cell lines and primary tumor samples, the mechanisms responsible for the establishment and maintenance of this silencing remained unknown. To better understand the epigenetic mechanisms that keep *TMS1* unmethylated and expressing in normal cells and alterations that lead to silencing in cancer, preliminary studies were performed to define the basic chromatin landscape of the *TMS1* gene in methylated and unmethylated states. In normal breast epithelial cells, *TMS1* maintained an unmethylated CpG island flanked on both sides by regions of densely methylated DNA (100). A DNase I hypersensitivity assay found the unmethylated *TMS1* locus associated with 4 DNaseI hypersensitive sites (HSs). Two of these HSs (HS1 and HS3) mapped to the 5' and 3' boundaries of the CpG island (**Figure 8**) (100). HS2 mapped in the center of the island. It was discovered that GABα protein bound to HS2 in a methylation sensitive manner. GABα binding helps to maintain *TMS1* expression in unmethylated cells (101). HS4, the only HSs mapped outside of the CpG island, was found downstream of the island and currently its function is unknown. This study also found the unmethylated CpG island

was enriched for nucleosomes marked by H3ac (100). When the *TMS1* CpG island is aberrantly methylated, either due to the over expression of DNMT1 or during carcinogenesis, there is a dense and complete methylation across the island as well as the surrounding DNA. This methylation accompanied by with a loss of gene expression, three of the HSs, including all CpG island associated HSs, and enrichment of H3ac (100). These initial studies revealed that *TMS1* exhibited dramatic epigenetic differences dependent upon the DNA methylation status of the CpG island.

The above summarizes our knowledge of the epigenetic regulation of *TMS1* when the work of this thesis began. The goal of this thesis was to further the understanding of the epigenetic factors (histone modifications, nucleosome positioning) associating with the *TSM1* CpG island in both an unmethylated state in normal cells and in a methylated state in cancer (Chapter 2). Once the associative epigenetic factors surrounding the epigenetic regulation of *TMS1* were defined the next objective was to determine the function role individual factors plays in the regulation of *TMS1*. In particular, we further investigated the role of individual histone modifications, positioned nucleosomes, and DNA methylation during the epigenetic silencing of *TMS1* locus during carcinogenesis or in the maintenance of the unmethylated CpG island in normal cells. In the final objective, we expanded our findings at the *TMS1* locus to a genome-wide investigation of the role of DNA methylation in the establishment and maintenance of aberrantly methylated loci.

Figures

Figure 1.

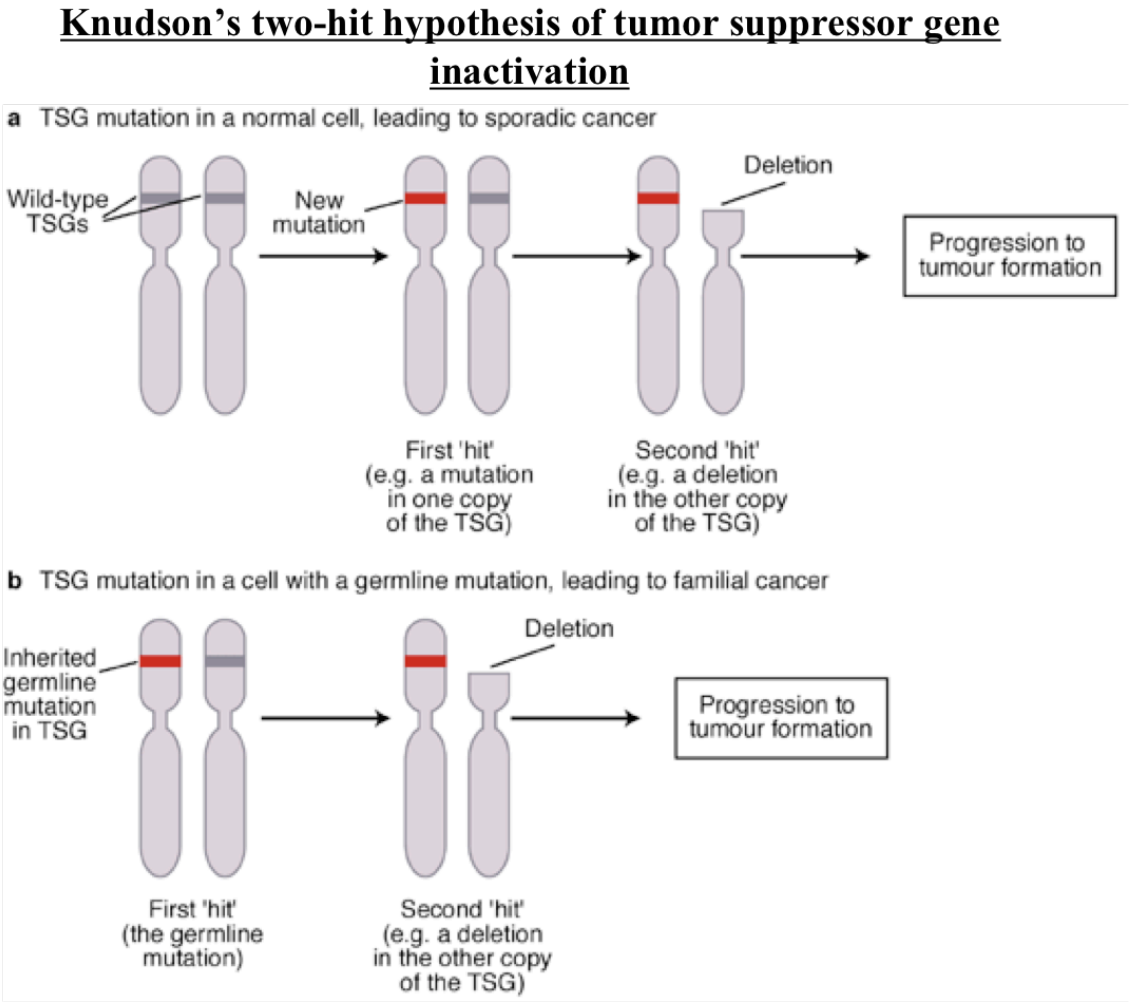


Figure 1. Knudson's two-hit hypothesis of tumor suppressor gene inactivation during carcinogenesis. A. Patients developing sporadic cancer must acquire mutations/deletions of both copies of a tumor suppressor gene to functionally contribute to carcinogenesis. **B.** Patients developing familial cancer need to only acquire a mutation/deletion in their one inherited wild type copy of the tumor suppressor gene. Figure adapted from http://journals.cambridge.org/fulltext_content/ERM/

Figure 2.

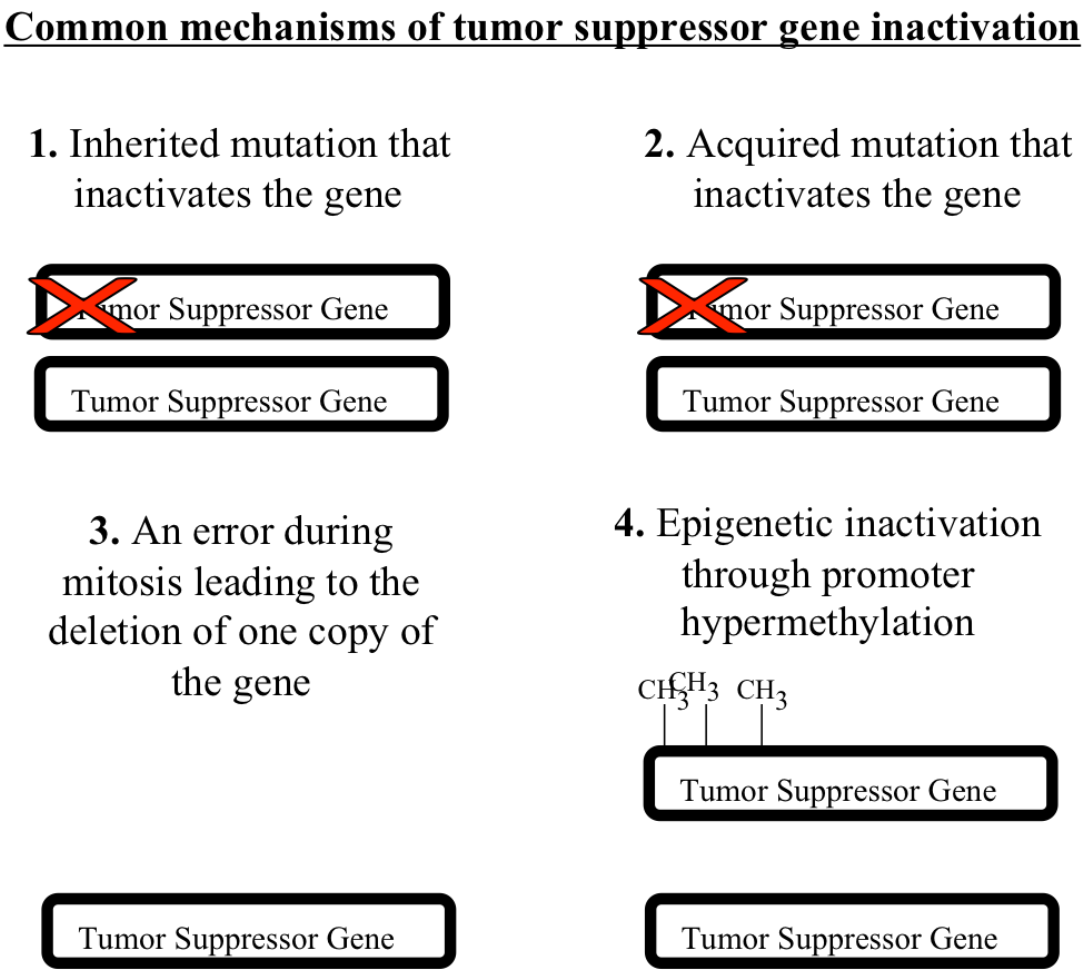


Figure 2. Mechanisms of tumor suppressor gene inactivation in cancer. The most commonly studied mechanisms of tumor suppressor gene inactivation are diagramed above. The particular mechanism of tumor suppressor gene of each allele is irrelevant.

Figure 3.

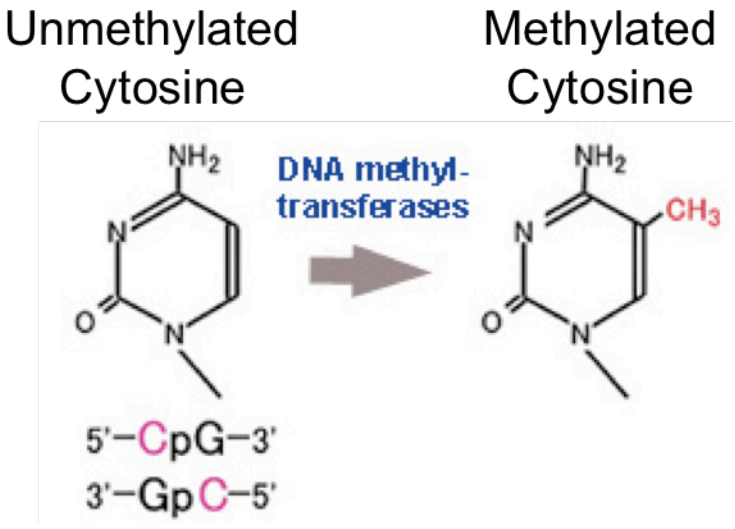


Figure 3. Mammalian DNA methylation occurs at the 5' carbon of cytosines found within the context of CpG. Figure adapted from Taylor 2006 (102).

Figure 4.

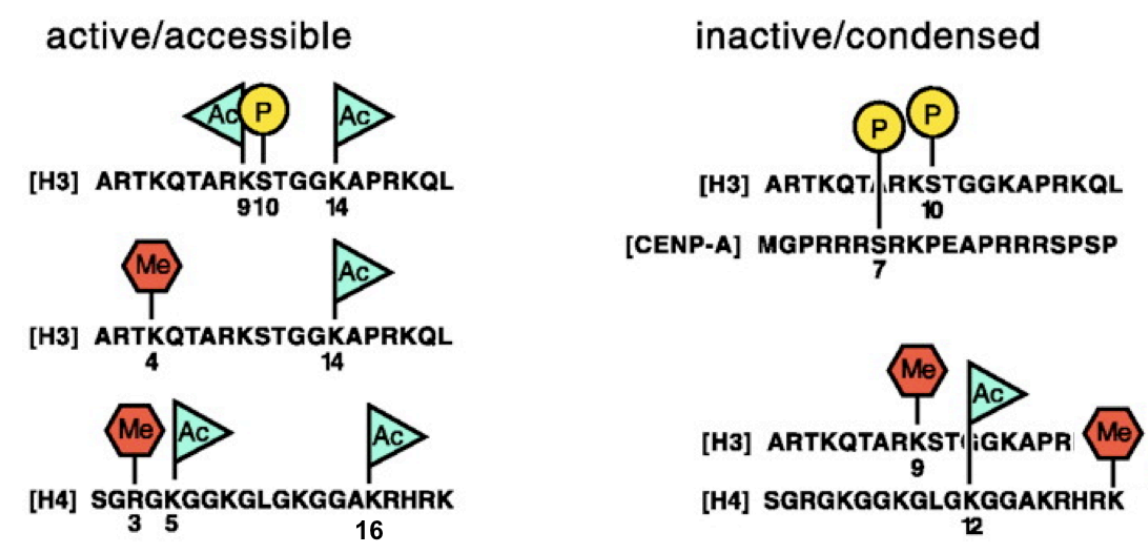


Figure 4. Subsets of histone modifications work in concert creating a histone code. Histone modifications help to create an epigenetic domain that is either permissive or inhibitory to gene transcription. Ac, acetylation; Me, methylation; P, phosphorylation. Figure adapted from Jenuwin *et. al.* (18).

Figure 5.

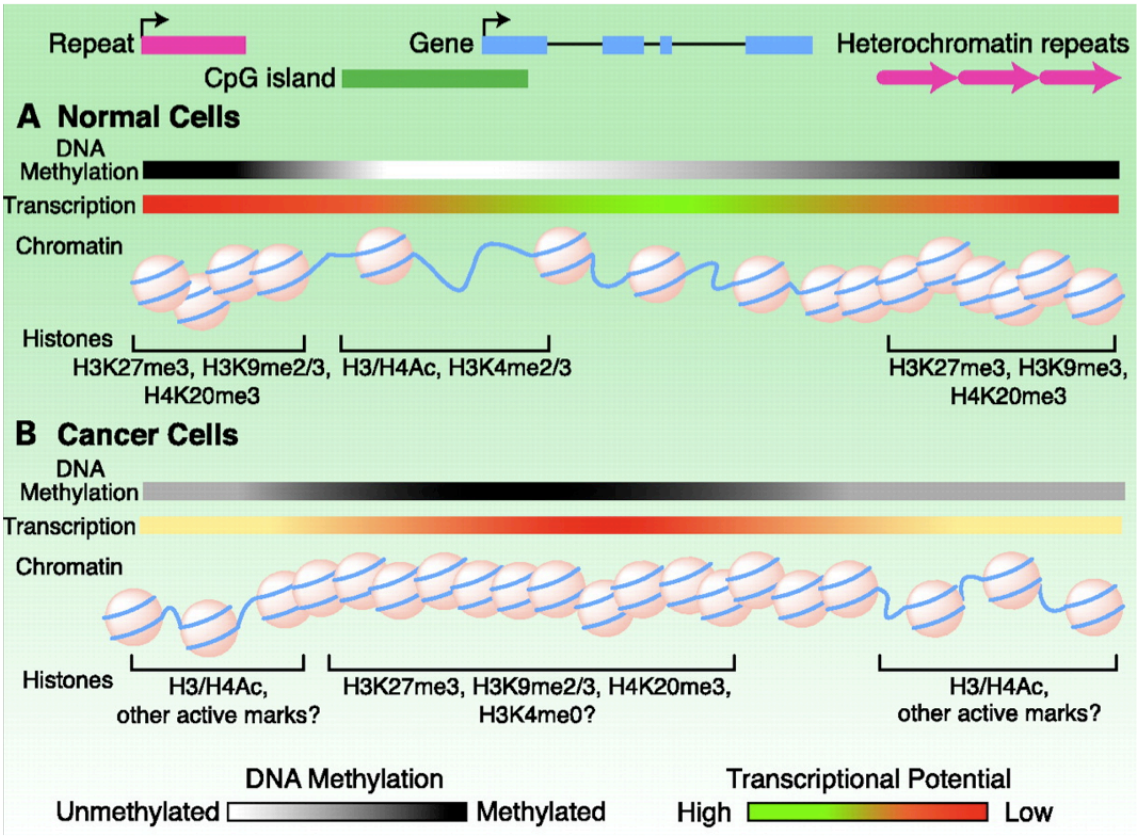


Figure 5. Alterations to DNA methylation and histone modifications result in the aberrant silencing of TSGs during carcinogenesis. A. Unmethylated CpG islands associated with active histone marks characterize normal cells. Also, heterochromatic regions and repeat elements are heavily methylated and marked by repressive histone modifications. **B.** In cancer a reversal of normal epigenetic regulation is observed. Subsets of TSG are hypermethylated and associate with repressive histone modifications, while repeat elements and heterochromatic regions are often hypomethylated leading to genomic instability. Figure adopted from McCabe *et. al.* (24).

Figure 6.

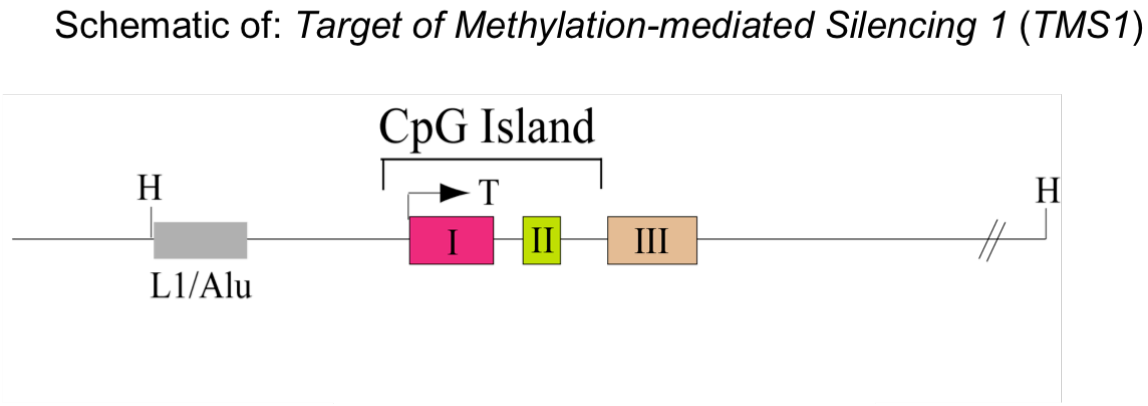


Figure 6. Schematic of the *TMS1* gene. *TMS1* is a CpG island containing TSG found on Chromosome 16. *TMS1* has 3 exons, of which exon 2 is alternatively spliced. The *TMS1* CpG island is approximately 1 kb in size and overlaps the transcription start site. Figure adapted from Kapoor-Vazirani *et. al.* (103).

Figure 7.

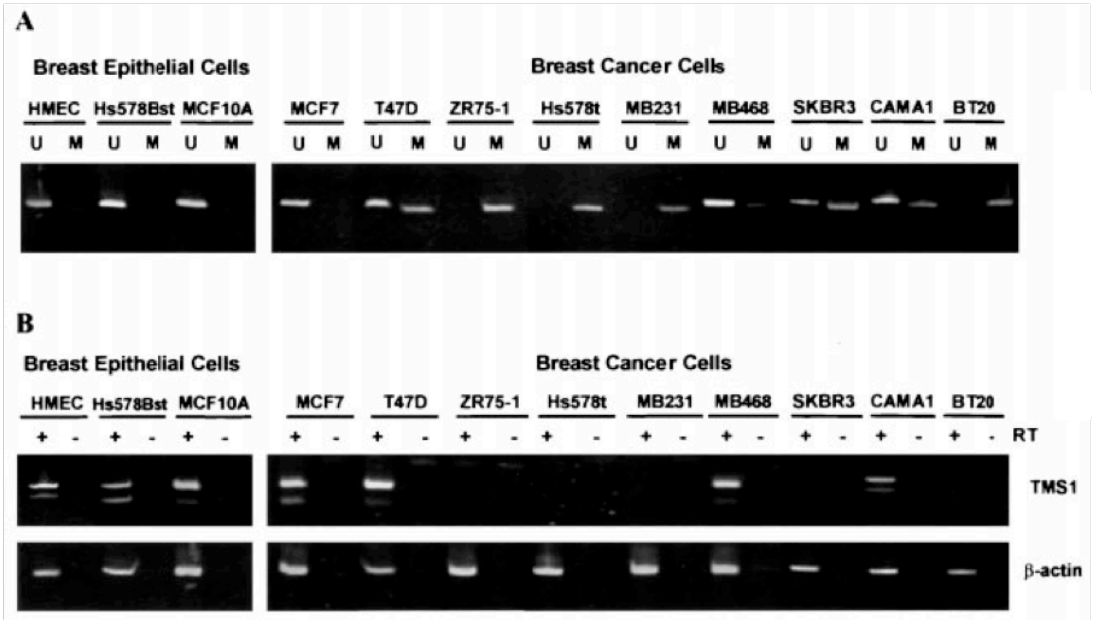


Figure 7. *TMS1* methylation status and gene expression demonstrate a direct correlation. A. Methylation Specific PCR methylation analysis of normal breast epithelial cells and breast cancer cells. U, unmethylated; M, methylated. **B.** *TMS1* expression was measured by RT-PCR. ß-actin expression was detected as a positive control. RT, reverse transcriptase. Figure from Conway *et. al.* (61).

Figure 8.

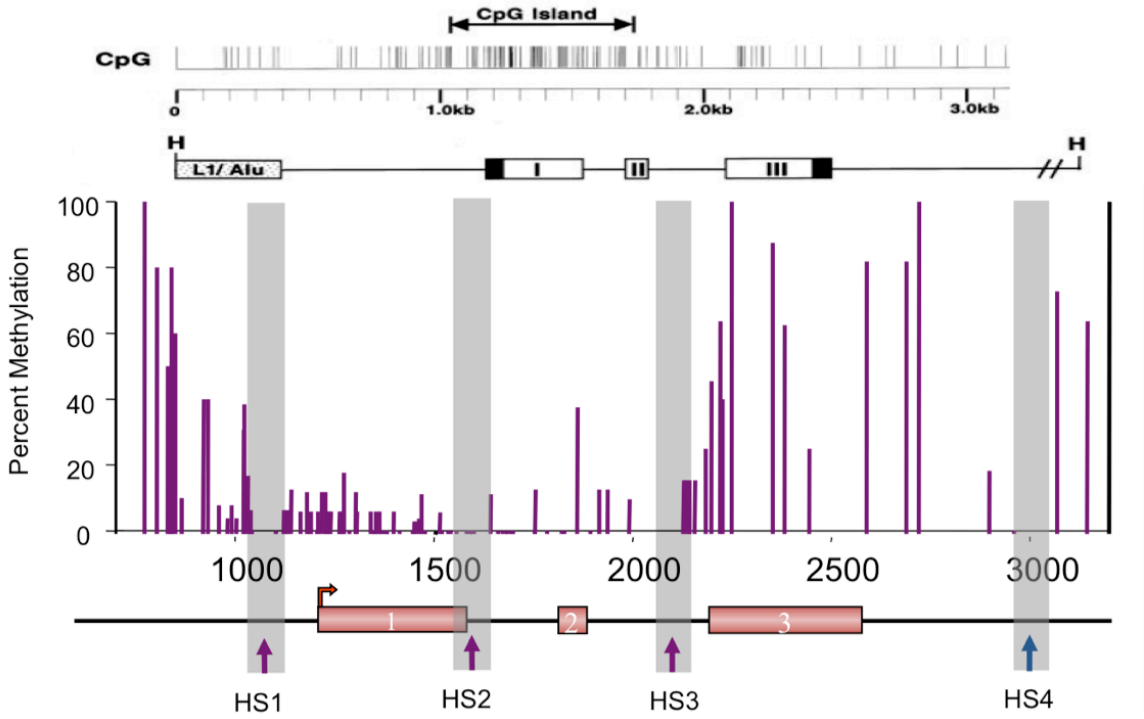


Figure 8. *TMS1* methylation boundaries correlate with DNase I hypersensitivity in normal cells. Percent *TMS1* methylation shown determined from bisulfite sequencing data. Four DNase I hypersensitive sites are represented by gray bars. HS, hypersensitive site. CpG density is found on the top of the diagram and gene schematic found on the bottom of the diagram. Figure adapted from data from Stimson *et. al.* (100).

Chapter 2: Role of hMOF-dependent H4 Lysine 16 acetylation in the maintenance of *TMS1/ASC* gene activity

Kagey, JD performed all nucleosome positioning experiments in MCF7 and MDA-MB231 cells, and in MCF7 cells with shRNA targeting *hMOF* and quantification of positioning. Wrote all sections of this manuscript pertaining to nucleosome positioning. Kapoor-Vazirani, P and Powell, DR performed all other experiments for this chapter. Kapoor-Vazirani, P wrote the remainder of the manuscript.

This work was been published in Cancer Research (2008; 68: (16).) and is included here as published.

Role of hMOF-dependent histone H4 lysine 16 acetylation in the maintenance of TMS1/ASC gene activity[1]

Priya Kapoor-Vazirani, Jacob D. Kagey, Doris R. Powell and Paula M. Vertino[2]

Department of Radiation Oncology and the Winship Cancer Institute,
Emory University, Atlanta, GA 30322

Running Title: Regulation of TMS1 by hMOF-dependent H4K16 Acetylation

Keywords: DNA methylation, gene regulation, histone modifications, chromatin, cancer

[1] This work was supported by an NIH Grant (2RO1-CA077337) to PMV, a CHIR Post-Doctoral Fellowship to PKV and a NSF PRISM Fellowship (#DGE0536941 and #DGE0231900) to JDK.

[2] To whom requests for reprints should be addressed, at the Winship Cancer Institute, Emory University 1365-C Clifton Rd NE, Atlanta, GA 30322. Phone: (404)778-3119; Fax: (404)778-5530; E-mail: pvertin@emory.edu

The focus of the work in this chapter was to identify differences in the epigenetic regulation of TMS1 the correlated with the methylation status of the CpG island. The findings from this chapter were used as a comparison for the studies, which altered various epigenetic factors, in subsequent chapters.

INTRODUCTION

Epigenetic mechanisms involve DNA methylation at cytosine residues and post-translational modifications of histone tails, both of which regulate gene transcription by altering chromatin structure and DNA-protein interactions (104, 105). In the human genome, most cytosines in CpG dinucleotides are methylated, except those in CpG islands, regions of the genome that contain a high density of CpG sites and encompass the promoter regions of more than half of the known genes (23). Methylation of promoter-associated CpG islands, whether developmentally programmed or occurring aberrantly during carcinogenesis, is associated with gene inactivation (106). Histone modifications can also act synergistically or antagonistically to define the transcription state of genes. Acetylation of histones H3 and H4 is associated with transcriptionally active promoters and an open chromatin configuration (107). Both dimethylation and trimethylation of histone H3 lysine 4 (H3K4me2, H3K4me3) have been linked to actively transcribing genes, although recent studies indicate that H3K4me2 marks CpG island-associated promoters regardless of the transcriptional status (108, 109). In contrast, methylation of histone H3 lysine 9 and 27 are associated with transcriptionally inactive promoters and condensed closed chromatin (110, 111). Interplay between the histone modifications and DNA methylation define the transcriptional potential of a particular chromatin domain.

The epigenetic landscape is drastically altered in human cancers. In cancer cells, global hypomethylation occurs at normally methylated CpG sites while hypermethylation occurs at select CpG islands (112). Aberrant CpG island methylation is accompanied by changes in the local histone modification patterns, including the hypoacetylation of histones H3 and H4, hypomethylation of histone H3 lysine 4 and hypermethylation of histone H3 lysine 9 and/or H3 lysine 27, resulting in gene silencing (112). Such epigenetic events contribute to carcinogenesis through the aberrant silencing of tumor suppressor genes. Recently, genome-wide studies showed that cancer cells undergo widespread alterations in histone modifications, including a global loss of acetylation at H4 lysine 16 (H4K16Ac) and trimethylation at H4 lysine 20 (H4K20me3) (68). Though the molecular alterations that occur in cancer cells are well studied, the precise order and mechanisms by which they precipitate gene silencing are still largely unknown.

Nucleosome positioning, or the preferential association of nucleosomes with specific genomic locations, is another feature with the potential to impact epigenetic regulation. Functionally, nucleosome positioning has been shown to maintain or prohibit gene activity depending on the context (42, 45). The translational positioning of nucleosomes has been attributed to interactions between histone proteins and specific DNA sequences, the binding of non-histone proteins, and neighboring positioned nucleosomes (41-43). DNA methylation does not alter the ability of histone proteins to bind DNA *in vitro*, but can alter nucleosome positioning *in vivo* through indirect mechanisms (43, 44). The relationship between histone modifications and positioning of nucleosomes is not well understood.

TMS1 (Target of Methylation-mediated Silencing), also known as ASC, is a proapoptotic signaling factor that is subjected to epigenetic silencing in human cancers (61, 95, 98, 113, 114). Originally identified in a screen for genes that were silenced in response to DNMT1 (DNA cytosine-5-methyltransferase-1) over-expression (61), subsequent studies have shown that TMS1 is silenced in conjunction with CpG island hypermethylation in a number of human tumors, including glioblastomas, breast, colorectal, and gastric cancers (96, 115-117). Loss of TMS1 function confers resistance to TNFa-induced apoptosis in breast cancer cells and its restoration suppresses cell growth (98). Treatment of cancer cells with the demethylating agent 5-aza-2'-deoxycytidine (5azadC) restores TMS1 expression, suggesting an important role for DNA methylation in TMS1 gene silencing (118).

The TMS1 CpG island is unmethylated in normal cells and breast cancer cells that express TMS1 (100, 118). In the active state, the TMS1 CpG island represents a distinct chromatin domain characterized by acetylated histones and three DNase I hypersensitive sites (HS) that span the CpG island. Two of these, HS1 and HS3, demarcate the 5' and 3' boundaries between the unmethylated CpG island DNA and densely methylated flanking DNA (100). Epigenetic silencing of TMS1 in breast cancer is accompanied by a localized remodeling of the CpG island-associated HS sites, hypoacetylation of histones, and hypermethylation of DNA (100). These data have led us to propose that the HS sites act in *cis*, perhaps by recruiting *trans*-acting factors, to block or actively oppose the spread of methylation into the CpG island.

To further understand the role of chromatin structure in the epigenetic silencing of tumor suppressor genes in cancer, we have used the TMS1 locus as a model to examine

the relationship between histone modifications, nucleosome positioning, and DNA methylation. We found that the TMS1 locus is characterized by distinct histone modification profiles in the active and inactive states. Interestingly, H4K16Ac and H4K20me3 exhibited unique localization across the active and inactive TMS1 locus, respectively. We further pursued the significance of H4K16Ac at TMS1 and determined that abrogation of H4K16 acetylation led to TMS1 silencing, which was accompanied by changes in the local nucleosome architecture. Our findings indicate that H4K16Ac plays a critical role in the maintenance of active gene transcription, and suggest that loss of H4K16Ac and transcriptional downregulation may be important steps in the epigenetic silencing of some tumor suppressor genes in cancer.

MATERIALS AND METHODS

Cell Culture-- IMR90 normal diploid fibroblasts and their SV40-transformed derivatives (IMR90/SV40) were obtained from the National Institute of Aging. The generation of IMR90/SV40 cells stably over-expressing DNMT1 (HMT.1E1) has been described (93). MCF7 and MDA-MB231 breast cancer cell lines and 293T embryonic kidney cells were obtained from the American Type Culture Collection and maintained in DMEM containing 2 mM glutamine and 10% FBS.

Chromatin Immunoprecipitation (ChIP)-- ChIP was carried out as described in the Acetyl-Histone H3 Immunoprecipitation Assay Kit (Cat. # 17-229) by Millipore. DNA recovered from ChIP was analyzed by real-time PCR. The reaction mixture (25 mL) contained 1 mL of the appropriately diluted DNA sample, 0.2 mM primers and 12.5 mL

of IQ SYBR Green Supermix (BioRad). The reaction was subjected to a hot start for 3 minutes at 95^{o}C and 50 cycles of 95^{o}C,10s; 55-65^{o}C, 30s; 72^{o}C, 30s. Melt curve analysis was done to verify a single product species. Starting quantities were determined relative to a common standard curve generated using MCF7 genomic DNA. Percent enrichment in each pulldown was calculated relative to input DNA. Primers pairs used for real-time analysis spanned the TMS1, CDH1 or ESR1 locus. Sequences are available upon request. Antibodies used were: rabbit IgG; (Santa Cruz; #SC-2027), histone H3 acetylated at lysine 9 and 14 (H3K9/14Ac; Millipore; #06-599), H4K16Ac (Abcam; #ab1762), H3K4me2 (Millipore; #07-030), H3K4me3 (Abcam; #ab8580), dimethylated histone H3 lysine 9 (H3K9me2; Abcam; #ab7312), trimethylated histone H3 lysine 9 (H3K9me3; Abcam; #ab8898), trimethylated histone H3 lysine 27 (H3K27me3; Millipore; #07-449), H4K20me3 (Abcam; #ab9053). Antibodies against hMOF and hMSL1 were gifts from Edwin Smith (Emory University).

Micrococcal Nuclease Digestion and Indirect End-labeling -- Isolation of intact nuclei and indirect end labeling was performed as described in (27) with minor modifications. Nuclei from 2 x 10^6 cells were digested with micrococcal nuclease (MNase) (10 – 200 u) at 25^{o}C for 10 min in RBS (10mM TrisHCl, 5mM $MgCl_2$, 0.5 M DTT, 0.3 mM sucrose, 0.4 mM PMSF). Reactions were stopped by incubation in 1% SDS, 20 mM EDTA, and 10 μg/ml RNaseA for 30 min at 37^{o}C, followed by incubation with 1 mg/ml Proteinase K at 50^{o}C overnight. DNA was recovered by phenol-chloroform extraction and ethanol precipitation. MNase-digested DNA (10 μg) was digested with 2 units each Hind III and Spe I at 37^{o}C overnight, separated on a 1.0% agarose gel and transferred to a nylon

membrane. Membranes were hybridized with a random-prime labeled Spe I-Xmn I probe anchored to the downstream Spe I site of TMS1. Approximately 100 pg Spe I-Hind III (2765 bp), Spe I- Bam HI (1001 bp) and Spe I-Eco RI (739 bp) fragments from the TMS1 locus were run on each gel as internal size markers. Membranes were washed to high stringency (2X SSC, 0.1% SDS at 65^{o}C), exposed to phosphor storage screens and analyzed by phosphoimage analysis using ImageQuant software (Molecular Dynamics).

RNAi Silencing-- For siRNA transfections, MCF7 cells were transfected with 100 nM siRNA (Dharmacon) against Lamin A/C, hMOF, or hMSL1 using Oligofectamine (Invitrogen). The sequences of hMOF and hMSL1 siRNAs are from Dou *et al.* (2005) and Smith *et al.* (2005), respectively. For lentiviral shRNA infections, 293T cells were transfected with 920 ng of pCMV-dR8.91 (viral packaging plasmid), 100 ng pMD2G-VSV-G (viral envelope plasmid) and 1 mg pLKO.1 or pLKO.1 containing hMOF shRNA (Open Biosystems; TRCN0000034875) using Lipofectamine (Invitrogen). Media containing lentiviral particles was collected at 40 and 64 hours post-transfection. Recipient cells were plated in 6 cm plates (1 x 10^6 cells) or 10 cm plates (2 x 10^6 cells) and infected with 0.1-0.5 mL supernatant in the presence of 8 mg/mL polybrene. Infected cells were placed under puromycin selection (0.5–1.0 mg/mL) and harvested at given time-points for subsequent analysis.

Western analysis-- Cells were lysed in RIPA buffer (50 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.5 mM EDTA, 1% NP-40, 0.5% sodium-deoxycholate, 0.1% SDS) containing protease inhibitors for 10 min on ice. Clarified lysates (100 mg) were electrophoresed on

a 12% SDS-PAGE gel and transferred to nitrocellulose membrane. The membrane was incubated with the appropriate 1^{o} antibody and HRP-conjugated 2^{o} antibody and subjected to chemiluminescence detection (Pierce). For detecting histones, cells were washed in PBS containing 5 mM sodium butyrate and resuspended in acid extraction lysis buffer (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 1.5 mM PMSF, 0.2 N HCl) for 30 minutes on ice. Lysates (100 mg) were electrophoresed on a 10% SDS-PAGE gel and blotted as described above. Antibodies used were: TMS1 (Protein Tech; #PTG10500-1-AP), GAPDH (Abcam; #ab8245), hMOF, hMSL1, H3K9/14Ac, H4K16Ac, estrogen receptor α (Santa Cruz; #SC-7207) and E-cadherin (BD Biosciences; #C20820).

Reverse Transcription-- Total RNA was isolated from MCF7 cells using the RNeasy Mini kit (Qiagen). RNA (6 mg) was pre-treated with DNase I and then reverse-transcribed using random hexamer primers and M-MLV reverse transcriptase. cDNA was amplified with primers against TMS1, hMOF or 18s rRNA using real-time PCR as described for the ChIP assays. Starting quantities were determined relative to a common standard curve generated using MCF7 cDNA. Primer sequences are available upon request.

Methylation-Specific PCR (MSP) and Bisulfite Sequencing-- MSP and bisulfite sequencing were performed as previously described (100). Primer sets used for MSP overlap a total of 6 CpG sites in the TMS1 CpG island and have been described (61).

Primers used for bisulfite sequencing correspond to primer set B in Stimson and Vertino (27).

RESULTS

The active TMS1 locus is marked by H3K4me while the inactive TMS1 locus is marked by H3K9me-- The TMS1 locus consists of 3 exons encompassing approximately 1.5 kb on chromosome 16 (Figure 1A) (61). A 1.0 kb CpG island spans the promoter region as well as exons I and II (Figure 1A). This region is unmethylated in cells expressing TMS1, such as IMR90 human diploid fibroblasts and MCF7 breast cancer cells, but is completely methylated in cells lacking TMS1 expression, such as DNMT1-overexpressing human fibroblasts (HMT.1E1) and MDA-MB231 breast cancer cells (Figure 1B) (61, 100, 118). To determine the relationship between histone modifications and DNA methylation at the TMS1 locus, we examined the distribution of histone H3 marks across a region covering approximately 1.5 kb upstream of the transcription start site to approximately 600 bp downstream of the termination sequence in MCF7 and MDA-MB231 cells by ChIP (Figure 1A, C). We found that the active TMS1 locus in MCF7 cells was enriched in nucleosomes marked by H3K9/14Ac, H3K4me2 and H3K4me3 but lacked H3K9me2 and H3K9me3 (Figure 1C). In contrast, the silent TMS1 locus in MDA-MB231 cells was hypoacetylated at H3K9/14Ac and lacked H3K4me2 and H3K4me3 but was enriched in H3K9me2 and H3K9me3 (Figure 1C). We also examined H3K27me3 at the TMS1 locus and found that there was little enrichment of this mark in either MCF7 or MDA-MB231 cells when compared to a negative control (IgG) and a positive control locus (eg. MYT1) (Figure 1C; data not shown) (119). An

inverse relationship between H3K4me and H3K9me was also observed in IMR90 and HMT.1E1 cells (data not shown). These results indicate that distinct histone modification profiles correlate with TMS1 gene activity. Specifically, methylation-associated silencing of TMS1 correlates with a shift from H3K4me to H3K9me.

Discreet Localization of H4K16Ac and H4K20me3 at the TMS1 locus-- We next examined the relationship between DNA methylation and histone H4 modifications. H4K16Ac and H4K20me3 are associated with active and constitutive heterochromatic regions, respectively (39, 120, 121). Recent studies have shown that cancer cells exhibit reduced levels of H4K16Ac and H4K20me3 overall relative to their normal counterparts (11). It was suggested that this global loss derived primarily from repetitive DNA sequences. However, the relationship between these marks and the epigenetic dysregulation of individual genes in cancer cells has not been addressed. To determine whether histone H4 modifications play a role in the epigenetic regulation of TMS1, we examined the profile of H4K16Ac and H4K20me3 in MCF7 and MDA-MB231 cells. We found that the TMS1 locus was enriched in H4K16Ac in MCF7 cells whereas MDA-MB231 cells were hypoacetylated at H4K16Ac (Figure 2A). Interestingly, the distribution of H4K16Ac in MCF7 cells showed discreet peaks of enrichment that corresponded almost precisely to the 5' and 3' boundaries of the unmethylated TMS1 CpG island. These H4K16Ac peaks were completely absent in MDA-MB231 cells (Figure 2A). Furthermore, while TMS1 lacked H4K20me3 in MCF7 cells, there was a prominent peak of H4K20me3 at the TMS1 promoter in MDA-MB231 cells (Figure 2A). To determine whether this pattern was selective for breast cancer cells, we also examined

the distribution of H4K16Ac and H4K20me3 in IMR90 normal diploid fibroblasts and HMT.1E1 cells (Figure 2B). Again, two prominent peaks of H4K16Ac flanked the TMS1 CpG island when it is unmethylated and active (IMR90), but were absent in when the locus is hypermethylated and silent (HMT.1E1). As well, a peak of H4K20me3 was observed in the TMS1 promoter region in HMT.1E1, but not in IMR90 cells. These data indicate that the epigenetic silencing of TMS1 is associated with the focal loss of H4K16Ac and the gain of H4K20me3.

Nucleosome positioning at the TMS1 gene -- We also examined nucleosome positioning within and around the TMS1 CpG island in MCF7 and MDA-MB231 cells. Isolated nuclei were digested with increasing amounts of MNase followed by digestion of the DNA with Hind III and Spe I to generate a 2.76 kb genomic fragment encompassing the TMS1 CpG island (Figure 2C). Indirect end-labeling analysis with a probe anchored at the 3' Spe I site allowed mapping of positioned nucleosomes. Nucleosomes were positioned at an approximate 200 bp intervals spanning the CpG island in MCF7 cells (Figure 2C). At similar levels of MNase digestion, MDA-MB231 cells exhibited more of a smear, indicating that nucleosomes are more randomly positioned throughout the CpG island in these cells (Figure 2C, Supplemental Figure 1). One exception to the positioning in MCF7 cells was the distance between MNase cut sites 3 and 4 that flank the transcription start site. This ~400 bp spacing may correspond to a nucleosome-free region at the transcription start site of TMS1 in actively transcribed cells. Within the regular array of positioned nucleosomes, differences existed in the degree of positioning. Preferential digestion by MNase was observed at cut sites 3, 5 and 7, suggesting that

nucleosomes associated with these sites are more strongly positioned (Figure 2C, Supplemental Figure 1). Sites 3 and 7 correspond to the boundaries between the unmethylated CpG island and methylated surrounding DNA as well as the observed peaks of H4K16Ac (100) (Figure 2A). The co-localization of H4K16Ac and strongly positioned nucleosomes suggests a possible link between the two.

Role of H4K16Ac at the TMS1 locus-- The finding that there are peaks of H4K16Ac and positioned nucleosomes flanking the CpG island when the TMS1 locus is unmethylated and expressed and that the locus lacks these features when TMS1 is methylated and silent raises the question of whether H4K16Ac plays a role in the regulation of TMS1. The majority of H4K16Ac in humans is mediated by hMOF, a member of the MYST family of histone acetyltransferases (122, 123). Downregulation of hMOF by RNAi has been shown to drastically decrease the global levels of H4K16Ac in human cells (122, 123). Western blot analysis showed no difference in the levels of hMOF between cells that express TMS1 (MCF7 and IMR90) and those that do not express TMS1 (MDA-MB231 and HMT.1E1), indicating that the lack of H4K16Ac at TMS1 in MDA-MB231 and HMT.1E1 cells is not due to differential expression of hMOF (data not shown).

To understand the role of H4K16Ac at the TMS1 locus, we knocked down hMOF in MCF7 cells with two independent siRNA molecules (Figure 3A). Consistent with previous studies (120), we found that downregulation of hMOF led to a global decrease in H4K16Ac, but had no affect on H3K9/14Ac levels (Supplemental Figure 2A). There was also a drastic reduction in the specific association of H4K16Ac, but not H3K9/14Ac, at the TMS1 locus in MCF7 cells downregulated for hMOF. There was a direct

correlation between the efficiency of hMOF knock-down by the two siRNAs and loss of H4K16Ac (Figure 3A and B). Knock-down of hMOF in MCF7 cells also led to a loss of nucleosome positioning across the CpG island (Figure 3C, Supplemental Figure 3). In contrast, there was no change in the levels or distribution of H3K4me2, H3K9me2, or H4K20me3 at the TMS1 locus upon hMOF downregulation (Supplemental Figure 4). Thus, hMOF downregulation results in a loss of H4K16Ac and impairs nucleosome positioning at the TMS1 locus, independent of changes in other histone modifications.

We also determined the impact of hMOF knock-down on TMS1 expression. hMOF downregulation led to a concomitant decrease in TMS1 protein and mRNA levels (Figure 3A). Again, there was a direct correlation between the efficiency of hMOF knock-down by the two siRNAs and the magnitude of TMS1 repression (Figure 3A). In time course experiments, a good correlation was observed between the degree and timing of hMOF knock-down and TMS1 downregulation, supporting a direct relationship between the two (Figure 3D). TMS1 silencing followed hMOF downregulation, and was reversed when hMOF expression returned to baseline. Downregulation of hMOF with a lentivirus expressing a shRNA directed against hMOF also led to a decrease in TMS1 expression, in both MCF7 and IMR90 cells (Supplemental Figure 2B, C). As in MCF7 cells, the level of H4K16Ac across the TMS1 locus was drastically reduced in IMR90 cells upon hMOF downregulation (Supplemental Figure 2D). Together, these studies indicate that downregulation of hMOF and the associated loss of H4K16Ac and nucleosome positioning at the TMS1 locus are sufficient to precipitate silencing of TMS1. These studies provide functional relevance to the localization of H4K16Ac at the

active TMS1 locus and indicate that hMOF-mediated acetylation of H4K16 positively regulates TMS1 gene expression.

In humans, hMOF is a component of several histone modifying complexes including the hMSL, MLL1-WDR5 and hMSL1v complexes (120, 122). However, hMOF in the hMSL complex is thought to be responsible for the majority of H4K16Ac (122). Indeed, downregulation of hMSL1, another component of the MSL complex, has been shown to drastically reduce global H4K16Ac levels (122). To determine whether the MSL complex played a role at the TMS1 locus, we treated MCF7 cells with siRNA directed against hMSL1. Knock-down of hMSL1 led to a decrease in H4K16Ac levels, loss of the H4K16Ac peaks and suppressed TMS1 expression in a manner similar to that observed after downregulation of hMOF (Figure 4A, B). ChIP experiments showed that both hMSL1 and hMOF were enriched at the TMS1 locus in MCF7 cells, but not in MDA-MB231 cells (Figure 4C). Taken together, these data suggest that hMOF is acting as part of the MSL complex to mediate H4K16 acetylation at the TMS1 locus.

Effect of hMOF downregulation and loss of H4K16Ac on DNA methylation-- The peaks of H4K16Ac at the boundaries between the unmethylated CpG island and flanking methylated DNA in MCF7 cells, and the absence of these features in MDA-MB231 cells, where the CpG island is methylated suggests that the presence of this histone mark may play an important role in maintaining the integrity of the CpG island domain. To address this question, we determined the impact of hMOF downregulation on the methylation status of TMS1. MCF7 cells transfected with control or hMOF siRNAs were harvested five days post-transfection and the DNA subjected to methylation-specific PCR (MSP)

and bisulfite sequencing. Downregulation of hMOF had no affect on the methylation status of the CpG island within a 5 day time frame (Figure 5A). Sequencing of bisulfite-modified DNA showed no difference in the profile of DNA methylation across the TMS1 CpG island in MCF7 cells transfected with control or hMOF siRNA (Figure 5B). Downregulation of hMOF also had no impact on TMS1 CpG island methylation in IMR90 cells (data not shown). These results indicate that, at least within the short time frame analyzed, loss of H4K16Ac leads to TMS1 silencing without accompanying CpG island methylation.

To determine whether CpG island methylation would ensue over a longer time period after hMOF knock-down, we also used lentiviral shRNA infections to stably knock-down hMOF in MCF7 cells. However, we found that long-term downregulation of hMOF by this method was lethal. At day 6 post-infection, cells containing lentiviral hMOF shRNA were dead (data not shown). This observation supports previous findings indicating that MOF is required for cell viability (122-125).

H4K16Ac does not regulate all genes silenced by aberrant DNA methylation-- The above data indicate that loss of the H4K16Ac peaks at the TMS1 locus, through targeted disruption of hMOF, allows for silencing of TMS1. To determine whether this mechanism is operative at other genes that undergo epigenetic silencing in human cancers, we examined the distribution of H4K16Ac at two other genes, CDH1 and ESR1, which like TMS1, are subject to epigenetic silencing in human breast and other cancers (126, 127). Both CDH1 and ESR1 contain promoter-associated CpG islands and are unmethylated and expressed in MCF7 cells and are methylated and silent in MDA-

MB231 cells (Figure 6A) (126, 127). The CDH1 and ESR1 CpG islands were both enriched in H3K9/K14Ac in MCF7 but not MDA-MB231 cells, which is consistent with the tight association of H3Ac mark with actively transcribed genes (Figure 6B). H4K16Ac levels were very low across the CpG islands of both CDH1 and ESR1, as compared to that observed at TMS1 (compare Figure 6B to Figure 2A). More importantly, there was no difference in the levels of H4K16Ac between MCF7 and MDA-MB231 cells at either gene (Figure 6B). Consistent with a lack of association of H4K16Ac with these loci, downregulation of hMOF had no effect on the expression of either ESR1 or CDH1 (Figure 6C). These results indicate that hMOF and H4K16Ac may regulate a specific subset of genes.

DISCUSSION

Alterations in DNA methylation associated with the epigenetic silencing of tumor suppressor genes in human cancers have been well documented. However, the molecular mechanisms underlying this silencing process, including the role of histone modifications and chromatin structural features, remain poorly understood. Here, we have characterized the TMS1 gene, which is silenced in cancers by DNA hypermethylation. Consistent with other studies, we find that the active state is characterized by histones hypermethylated at H3K4 and hypomethylated at H3K9, whereas the inactive state exhibits histones hypomethylated at H3K4 and hypermethylated at H3K9 (128, 129). A recent genome-wide study comparing global histone modifications in normal tissues, cancer cell lines and primary tumors revealed that carcinogenesis is accompanied by a global loss of H4K16Ac and H4K20me3 (68). The study further suggested that this loss

was not associated with individual gene loci but rather to alterations at repetitive DNA sequences. Our findings indicate that gene-specific alterations in H4K16Ac and H4K20me3 also accompany carcinogenesis. The active TMS1 locus is characterized by two prominent H4K16Ac peaks and positioned nucleosomes that flank the boundaries of the unmethylated CpG island. Epigenetic silencing of TMS1 is characterized by a loss of these peaks, random positioning of nucleosomes and the appearance of a H4K20me3 peak at the transcription start site. hMOF-mediated acetylation of H4K16, and the hMSL complex, play an integral role in maintaining an open chromatin state at the TMS1 locus, as downregulation of H4K16Ac leads to loss of nucleosome positioning and decreased TMS1 expression.

H4K16Ac has been shown to localize to both heterochromatic and euchromatic regions and has been linked to gene activation (120, 122, 130). Here, we see that H4K16Ac localizes to peaks occurring precisely at the boundaries between the unmethylated CpG island and flanking methylated DNA in the active TMS1 gene. The peaks of H4K16Ac could promote gene expression in two ways. They may act to prevent or oppose the spread of DNA methylation into the regulatory regions contained within the CpG island. In *S. cerevisiae*, the silencing complex SIR2 binds along the telomeres and at the mating type locus to maintain these regions in a heterochromatic state (131, 132). Acetylation of H4K16 by Sas2p, the yeast ortholog of hMOF, at the boundary of these regions prevents the spreading of the SIR complex and thus heterochromatin, into nearby euchromatic regions. These observations indicate that SIR2 and H4K16Ac act in opposition to define the boundaries between active and inactive chromatin. Though we did not observe methylation of the CpG island upon H4K16Ac downregulation, it is

possible that complete deacetylation of H4K16 and/or additional factors is required to trigger CpG island methylation.

Alternatively, the H4K16Ac peaks at the active TMS1 locus may promote the binding of a factor that activates transcription or prevent the docking of a factor that represses TMS1 transcription. Deacetylation of H4K16 in this case would prevent the binding of the activator or promote the association of the repressor, resulting in TMS1 repression. Proteins involved in transcription have been shown to preferentially bind the acetylated or non-acetylated form of H4K16Ac. Bdf1, which associates with the TFIID transcription factor complex to promote transcription, binds H4K16 when it is acetylated (133). ISWI, an ATP-dependent remodeling complex that promotes transcription repression, specifically binds to the non-acetylated form of H4K16 (134).

Our studies indicate that unlike H3Ac and H3K4me, which are more widespread marks among actively transcribing genes, H4K16Ac has a more select set of target genes. The presence of H4K16Ac at these target genes may promote expression, and loss of H4K16Ac may be a pre-requisite for gene silencing. Indeed, the presence of SirT1, which antagonizes H4K16Ac and is the human homolog of yeast SIR2, at certain tumor suppressor genes is associated with gene suppression and ablation of SirT1 leads to a local increase in H4K16Ac levels and activation of these genes (135). While its presence at TMS1 was necessary to maintain TMS1 expression, H4K16Ac was not present at and did not regulate the CDH1 and ESR1 genes, both of which are silenced by aberrant CpG island methylation in cancers. These findings are in contrast to those of Pruitt *et al.* (46), who showed that inhibition of SirT1 restored CDH1 expression in cancer cells where the gene was methylated and silent, implying a role for H4K16Ac at the CDH1 locus. Genes

targeted by H4K16Ac likely include those that are involved in essential cellular functions, since knock-down of hMOF resulted in eventual cell death (124)(data not shown).

We find positioned nucleosomes throughout the CpG island in cells that are unmethylated and express TMS1. On the contrary, epigenetic silencing is associated with a loss of positioning. This is consistent with previous studies showing that CpG islands in the unmethylated and active state are associated with positioned nucleosomes whereas nucleosomes are randomly positioned across the silent locus (42, 45, 136). In the active TMS1 gene, positioning occurred at regular 200 bp intervals with one notable exception at the transcription start site, where there was an ~400 bp space suggesting the active TMS1 chromatin structure includes a nucleosome-free region at the transcription start site. Genome-wide analysis of nucleosome positioning in yeast has shown that active genes tend to have a nucleosome-free region overlapping the transcription start site that is flanked by two strongly positioned nucleosomes (40). Several mammalian genes have also been shown to exhibit a region thought to be devoid of nucleosomes coincident with transcription factor binding sites (137).

Interestingly, we further show that two strongly positioned nucleosomes flank the unmethylated CpG island and are coincident with the peaks of H4K16Ac (see Figure 2C), suggesting that the marking of these sentinel nucleosomes may play a role in maintaining the integrity of the CpG island domain and the positioning of the remaining nucleosomes in the CpG island. Indeed, knock-down of hMOF and the corresponding decrease in H4K16Ac led to a loss of positioning. Whether this is due to a direct effect of hMOF and associated factors on nucleosome positioning or is an indirect effect of the resulting

transcriptional downregulation is not clear. At this point, the factors that dictate preferential nucleosome placement are not well understood, but our data suggest that specific histone modifications, or the complexes that mark them, may play an important role.

H4K20me3 plays an important role in maintaining pericentric heterochromatin (121). Our data indicate that H4K20me3 localizes to a discreet peak upstream of the transcription start site of TMS1 in cases when the gene is methylated and silent, suggesting that H4K20me3 may also play a role in the repression of coding genes in euchromatic regions. Indeed, preliminary data indicates that downregulation of Suv4-20h2, the methyltransferase responsible for H4K20me3, stimulates TMS1 re-expression in cells where it is methylated and silenced, but only when DNA is first demethylated by 5azadC treatment (data not shown). The precise localization of H4K20me3 raises the question of how this mark is directed to the TMS1 locus. At pericentric heterochromatin, a model has been proposed wherein H4K20me3 is directed by pre-exiting H3K9me3 through the binding of HP1 and recruitment of the Suv4-20h enzymes (121). However, what little H3K9me3 was observed at the TMS1 locus was distributed throughout the promoter and coding regions whereas H4K20me3 was localized to a sharp peak near the transcription start (Figures 1B and 2A). The distinct distribution of H4K20me3 and H3K9me3 at the TMS1 locus suggests that additional mechanisms and/or factors may direct H4K20me3 to individual genes.

Recent studies have demonstrated that hMOF is downregulated in primary breast cancers and medullablastomas (125). In addition, loss of hMOF function in human cells leads to genomic instability, defects in cell cycle, chromosomal aberrations and impaired

DNA damage response (122, 123, 138). Our findings suggest that an additional consequence may be the epigenetic silencing of certain tumor suppressor genes. Although downregulation of hMOF-mediated H4K16Ac led to local changes in nucleosome positioning and silencing of TMS1 gene expression, it was not sufficient to drive the subsequent demethylation of H3K4 or the acquisition of more 'heterochromatic' features (H3K9me2, H4K20me3, DNA methylation) associated with the locus in the stably silent state observed in cancer cells. Thus, while H4K16 deacetylation and transcriptional down regulation may be requisite steps, there must be additional triggers necessary to achieve the more stable and heritable silencing associated with aberrant CpG island DNA methylation. Further studies to understand the precise mechanism by which H4K16 deacetylation mediates gene inactivation will allow development of potential therapeutic agents that prevent silencing of tumor suppressor genes inactivated by this mechanism in tumorigenesis.

Figures

Figure 1.

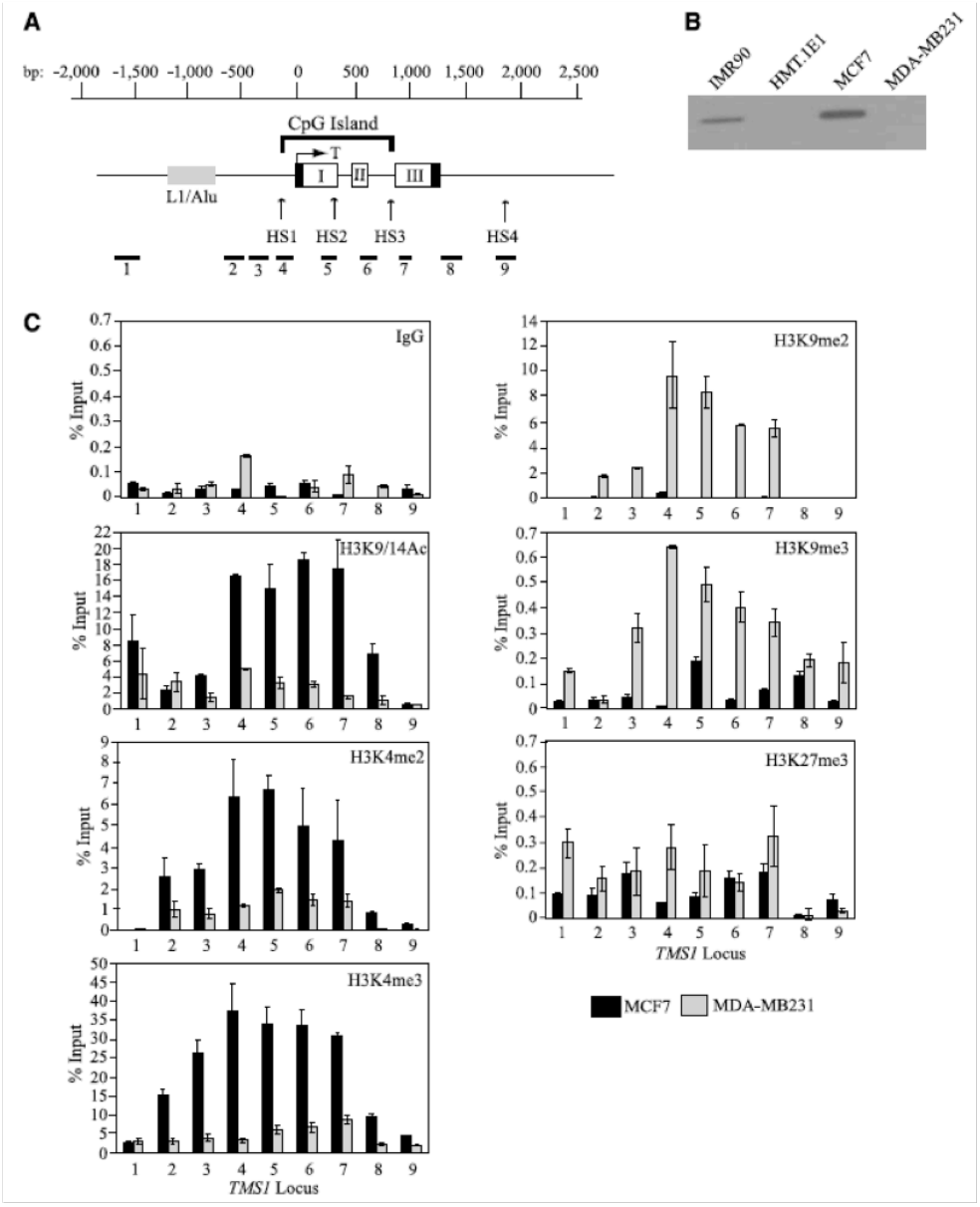


Figure 1. A) Schematic of the TMS1 genomic locus. The TMS1 gene consists of three exons (I, II and III). Non-coding regions are indicated by black boxes. The nucleotide positions are numbered with respect to the transcription start site (T) and are shown above the gene. The location of the CpG island is marked and spans from approximately -100 to +900 bp. The positions of the hypersensitive sites (HS1-HS4) and an upstream repeat element (L1/Alu) are shown. Primer sets used (1-9) for real-time PCR in ChIP assays are shown below the gene. **B)** Expression of TMS1. Protein lysates prepared from MCF7, MDA-MB231, HMT.1E1 and IMR90 cells were subject to western blot analysis with antibody against TMS1. **C)** Distribution of histone H3 modifications across the TMS1 locus. MCF7 and MDA-MB231 cell lines were subjected to ChIP with antibodies against the indicated histone modifications or a rabbit IgG (IgG) control. Immunoprecipitated DNA was amplified by real-time PCR with primer sets indicated in (A). Percent (%) input was determined as the amount of immunoprecipiated DNA relative to input DNA. Each ChIP was repeated at least three times, and although the immunoprecipitation efficiency varied between experiments, the profile of enrichment across the locus was consistent. Shown are the mean +/- standard deviation of triplicate determinations from a representative experiment.

Figure 2.

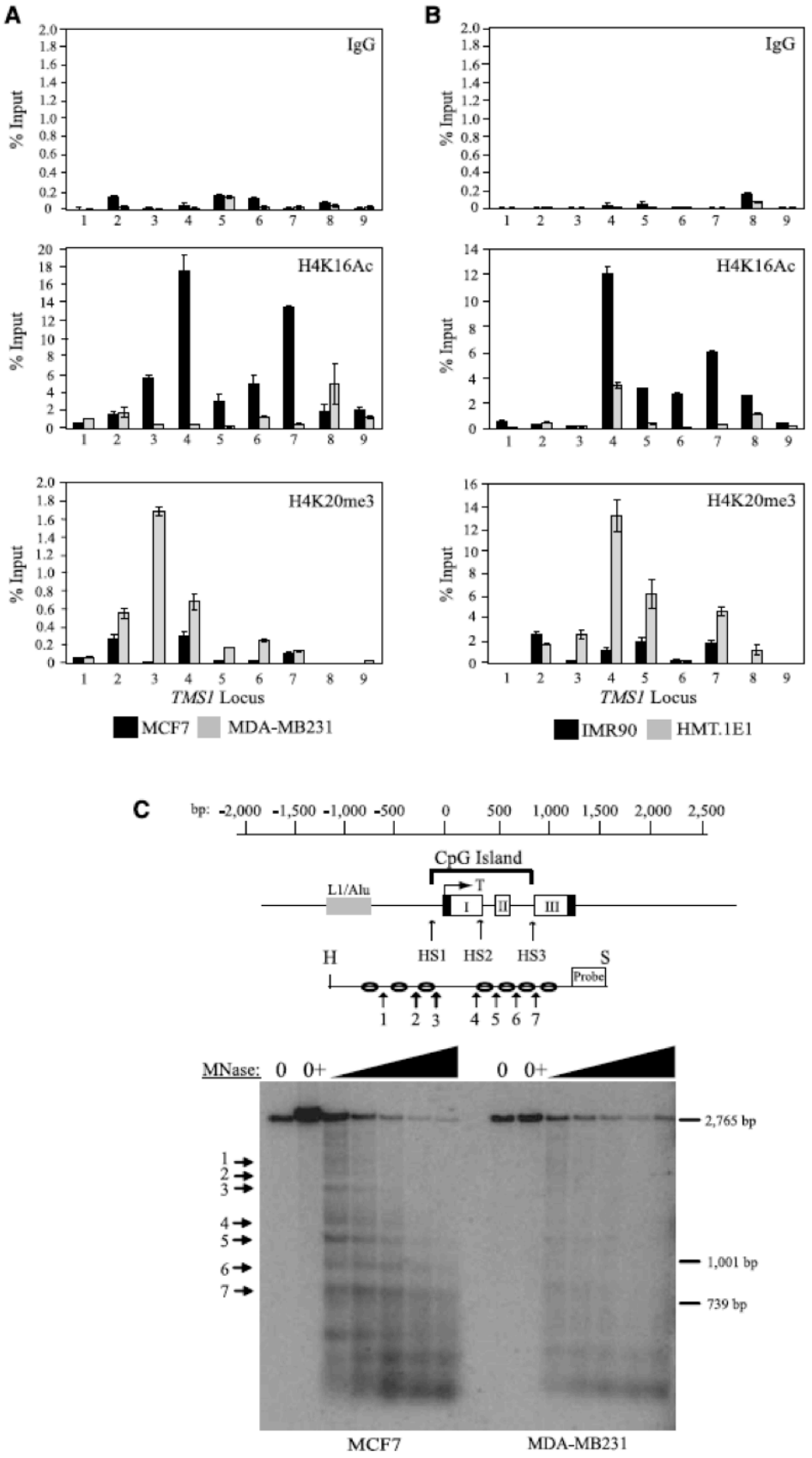

A
B
IgG
H4K16Ac
H4K20me3
% Input
TMS1 Locus
MCF7
MDA-MB231
IMR90
HMT.1E1
C
bp: -2,000 -1,500 -1,000 -500 0 500 1,000 1,500 2,000 2,500
CpG Island
L1/Alu
T
I
II
III
HS1
HS2
HS3
H
S
Probe
MNase: 0 0+ 0 0+
2,765 bp
1,001 bp
739 bp
MCF7
MDA-MB231

Figure 2. Localization of H4K16Ac and H4K20me3 across the TMS1 locus. **A)** MCF7 and MDA-MB231 breast cancer cells or **B)** IMR90 and HMT.1E1 cell lines were subjected to ChIP with antibodies against rabbit IgG (IgG) or the indicated histone modifications. Immunoprecipitated DNA was amplified by real-time PCR with primer sets indicated in Figure 1A. Percent (%) input was determined as the amount of immunoprecipiated DNA relative to input DNA. Each ChIP was repeated at least three times, and although the immunoprecipitation efficiency varied between experiments, the profile of enrichment across the locus was consistent. Shown are the mean +/- standard deviation of triplicate determinations from a representative experiment. **C)** Nucleosome positioning at the TMS1 locus. Nuclei from MCF7 (left) or MDA-MB231 (right) cells were incubated in MNase digestion buffer alone (0), digestion buffer plus $CaCl_2$ (0+), or digestion buffer plus $CaCl_2$ and 10-200U MNase. MNase-digested DNA (10 mg) was digested with Hind III (H) and Spe I (S), separated on a 1% agarose gel, and subject to Southern blot analysis using a probe anchored to the 3' Spe I site. Preferential MNase cut sites are depicted with arrows. Shown is the relative migration of a 2765bp, 1001bp and 739bp Spe I-anchored fragments from the TMS1 locus that were included as internal markers. A representative experiment is shown.

Figure 3.

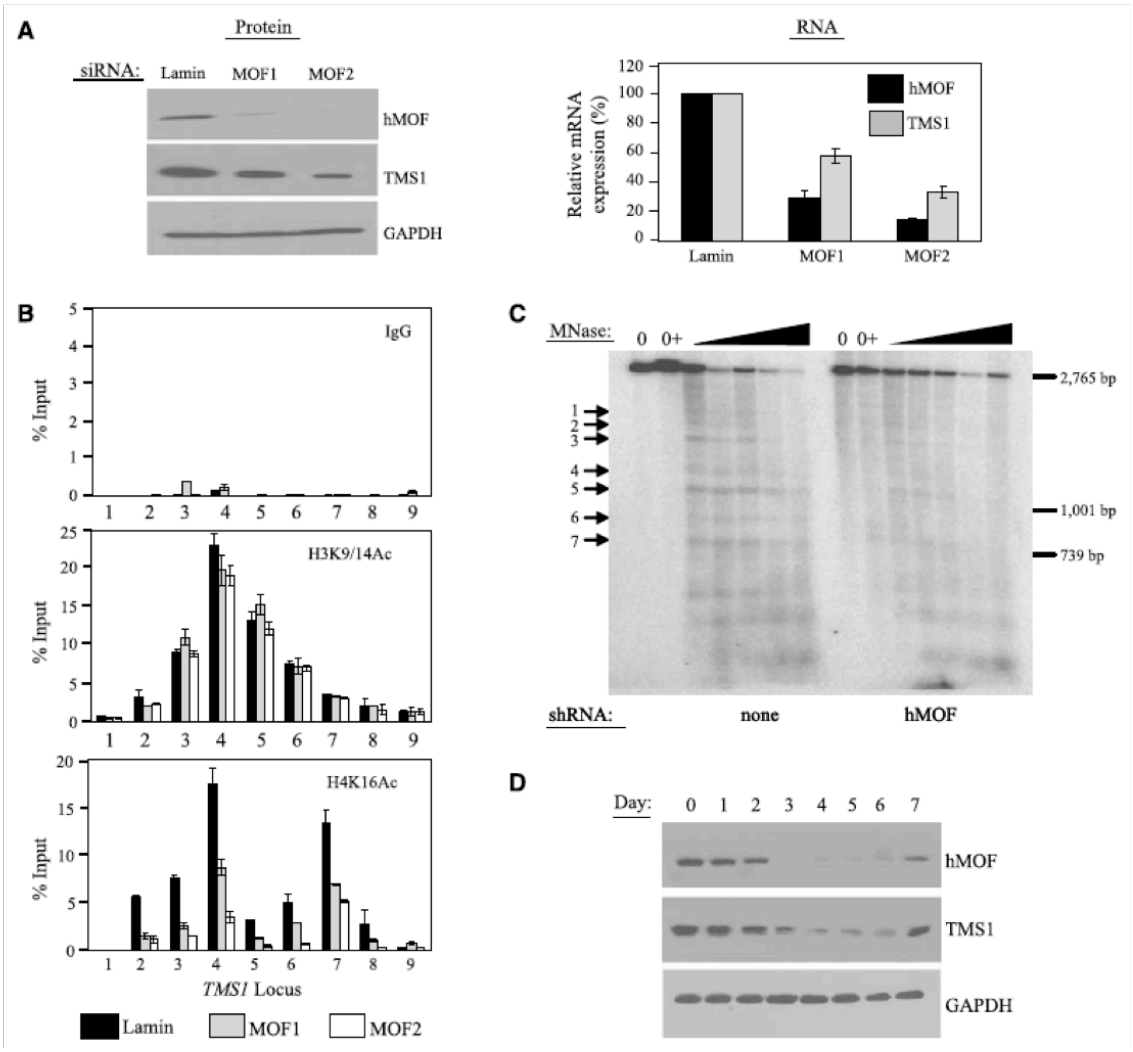

A
Protein
siRNA:
Lamin
MOF1
MOF2
hMOF
TMS1
GAPDH
RNA
Relative mRNA expression (%)
hMOF
TMS1
Lamin
MOF1
MOF2
B
IgG
H3K9/14Ac
H4K16Ac
% Input
TMS1 Locus
Lamin
MOF1
MOF2
C
MNase:
0 0+
0 0+
2,765 bp
1,001 bp
739 bp
shRNA:
none
hMOF
D
Day:
0 1 2 3 4 5 6 7
hMOF
TMS1
GAPDH

Figure 3. Effect of hMOF downregulation on the TMS1 locus. **A)** MCF7 cells were transfected with 100 nM of siRNA targeting hMOF (MOF1 or MOF2) or Lamin A/C (Lamin) (irrelevant control) and harvested four days post-transfection. Cells were analyzed for hMOF, TMS1 or GAPDH (loading control) protein expression by western blot analysis (left) or for hMOF, TMS1 or 18s (internal control) RNA expression by real time PCR (right). The levels of expression of hMOF and TMS1 mRNA are expressed relative to that obtained in cells treated with Lamin siRNA, after normalization to 18s. Shown is the mean +/- standard deviation of three independent experiments. Experiments were also done using scrambled non-targeting siRNA as a control and similar results were observed. **B)** MCF7 cells were transfected as in A) and ChIP was performed with antibodies against rabbit IgG (IgG), H3K9/14Ac or H4K16Ac. Immunoprecipitated DNA was amplified by real-time PCR with primer sets indicated in Figure 1A. Data represent the percent of input DNA recovered. Each ChIP experiment was repeated at least twice with reproducible results. Shown are the mean +/- standard deviation for triplicate determinations from a representative experiment. **C)** MCF7 cells infected with an empty pLKO.1 vector (none) or a pLKO.1 expressing hMOF shRNA (hMOF) were analyzed for nucleosome positioning exactly as described in Figure 2C. Preferential MNase cut sites (1-7) are indicated with arrows. Shown is the migration of a 2765bp, 1001bp and 739bp Spe I fragments from the TMS1 locus that were included as internal markers. A representative experiment is shown. **D)** MCF7 cells were transfected with hMOF siRNA and processed at 0 to 7 days post-transfection for western blot analysis using the indicated antibodies.

Figure 4.

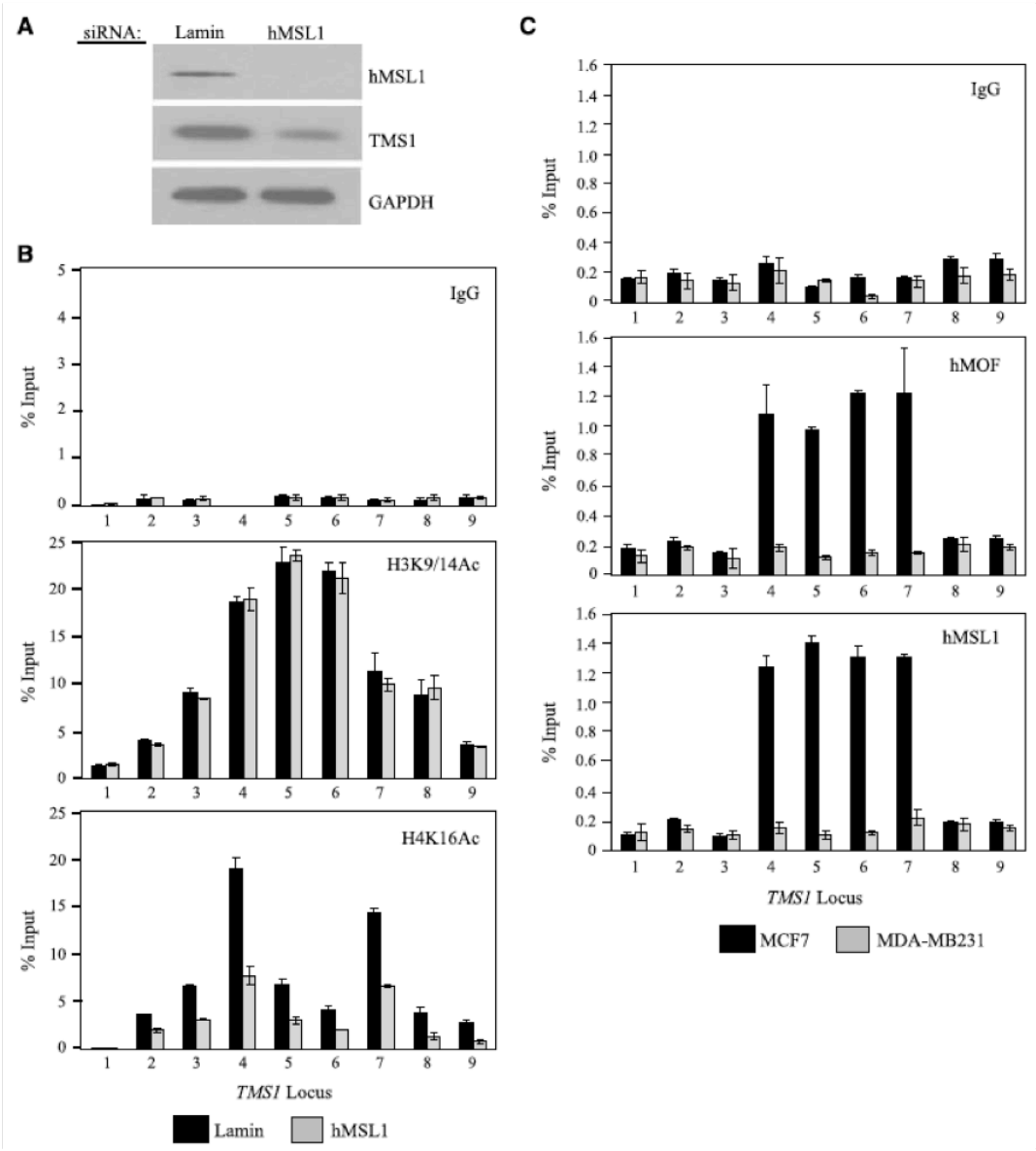

A
siRNA:
Lamin
hMSL1
hMSL1
TMS1
GAPDH
B
IgG
H3K9/14Ac
H4K16Ac
% Input
TMS1 Locus
Lamin
hMSL1
C
IgG
hMOF
hMSL1
% Input
TMS1 Locus
MCF7
MDA-MB231

Figure 4. Effect of hMSL1 silencing and localization of the MSL complex at the TMS1 locus. MCF7 cells transfected with 100 nM of Lamin A/C (Lamin) or hMSL1 siRNA were harvested four days post-transfection and analyzed for **(A)** hMSL1, TMS1 and GAPDH protein expression by western blot analysis and **(B)** chromatin immunoprecipitation with antibodies against rabbit IgG (IgG), H3K9/14Ac and H4K16Ac, exactly as described in Figure 3C. Similar results were obtained when a scrambled siRNA was used as a negative control. Shown are the mean +/- standard deviation of triplicate determinations from a representative experiment. **C)** Chromatin from MCF7 and MDA-MB231 cells were subjected to immunoprecipitation with antibodies against rabbit IgG (IgG), hMOF or hMSL1 as described in Figure 3C. Each ChIP experiment was repeated at least twice with reproducible results. Shown are the mean +/- standard deviation of triplicate determinations from a representative experiment.

Figure 5.

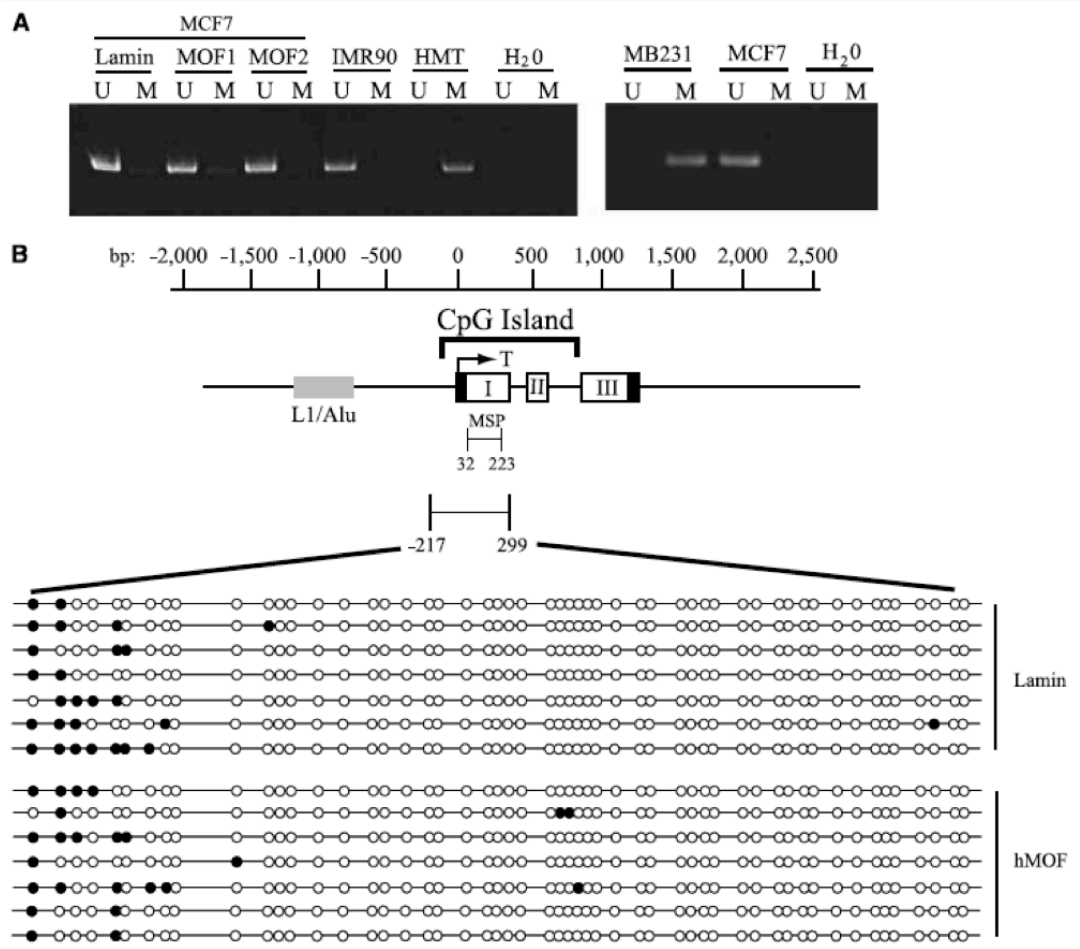


Figure 5. Effect of hMOF downregulation on TMS1 DNA methylation. **A)** MCF7 cells were transfected with siRNA (100 nM) against hMOF (MOF1 and MOF2) or an irrelevant control Lamin A/C (Lamin) and harvested 4 days post-transfection. Genomic DNA was isolated, modified by sodium bisulfite and amplified by methylation specific PCR (MSP) with primer sets specific for either methylated (M) or unmethylated (U) DNA. DNA from IMR90 and MCF7 cells served as a control for unmethylated DNA, while that from HMT.1E1 and MDA-MB231 cells served as a control for methylated DNA. The TMS1 region (32-223) amplified by MSP is shown in Part B (MSP). **B)** DNA from MCF7 cells transfected with Lamin A/C (Lamin) or MOF2 (hMOF) siRNA was modified with bisulfite and amplified with a primer set that spans 53 CpG sites in the TMS1 CpG island (100). Products were subcloned and sequenced. Each row indicates the sequence of an independent clone where methylated (black circles) and unmethylated (white circles) CpG sites are indicated.

Figure 6.

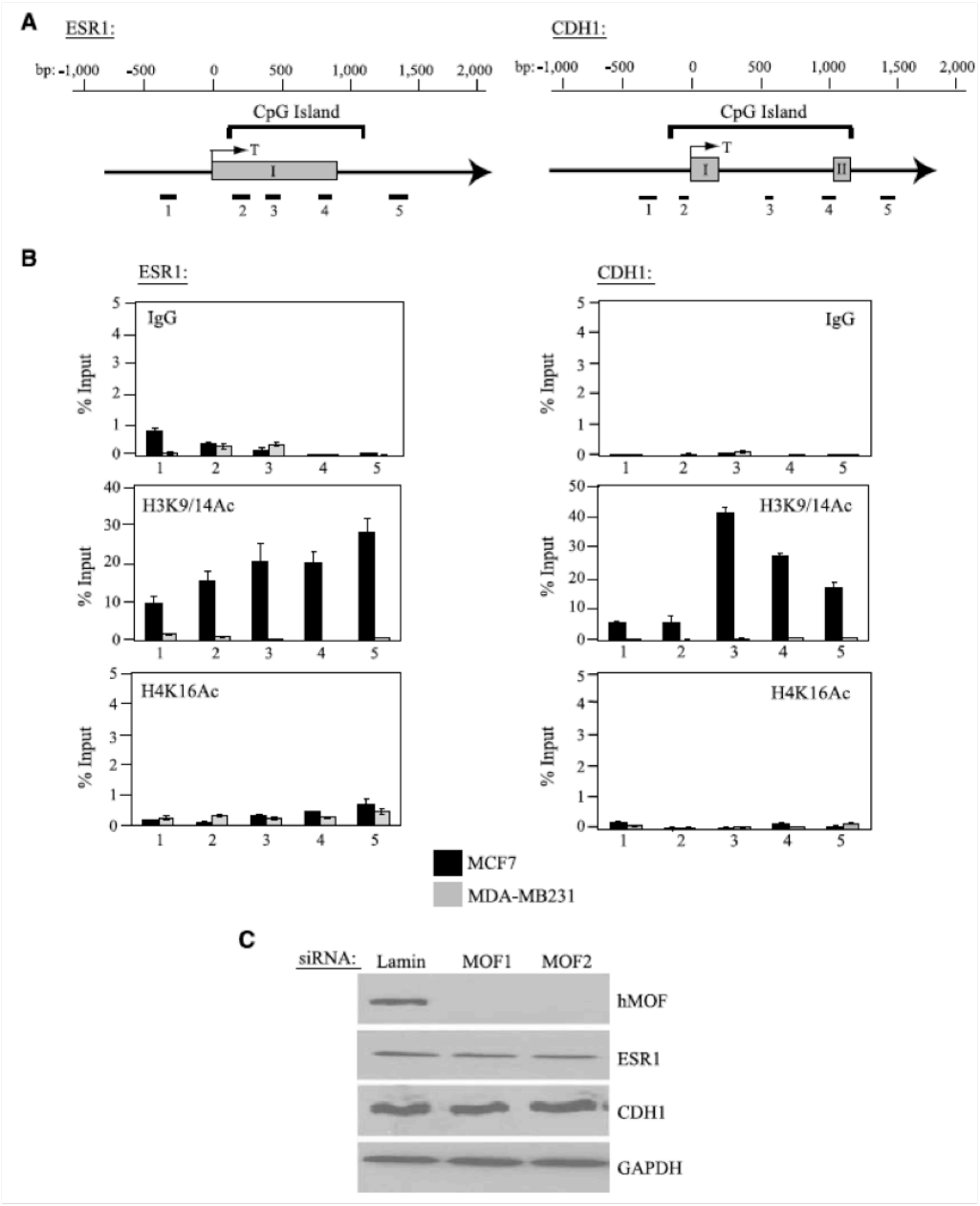
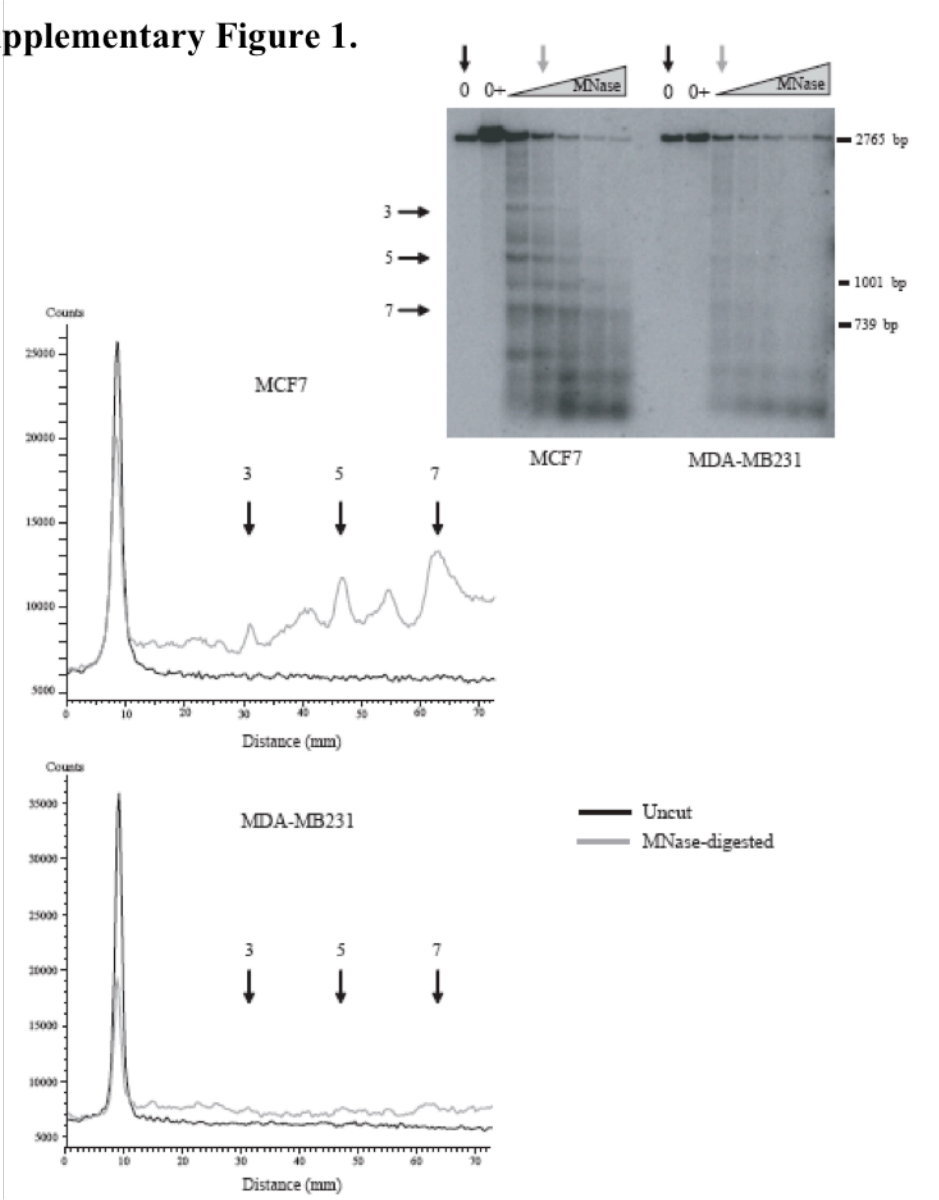
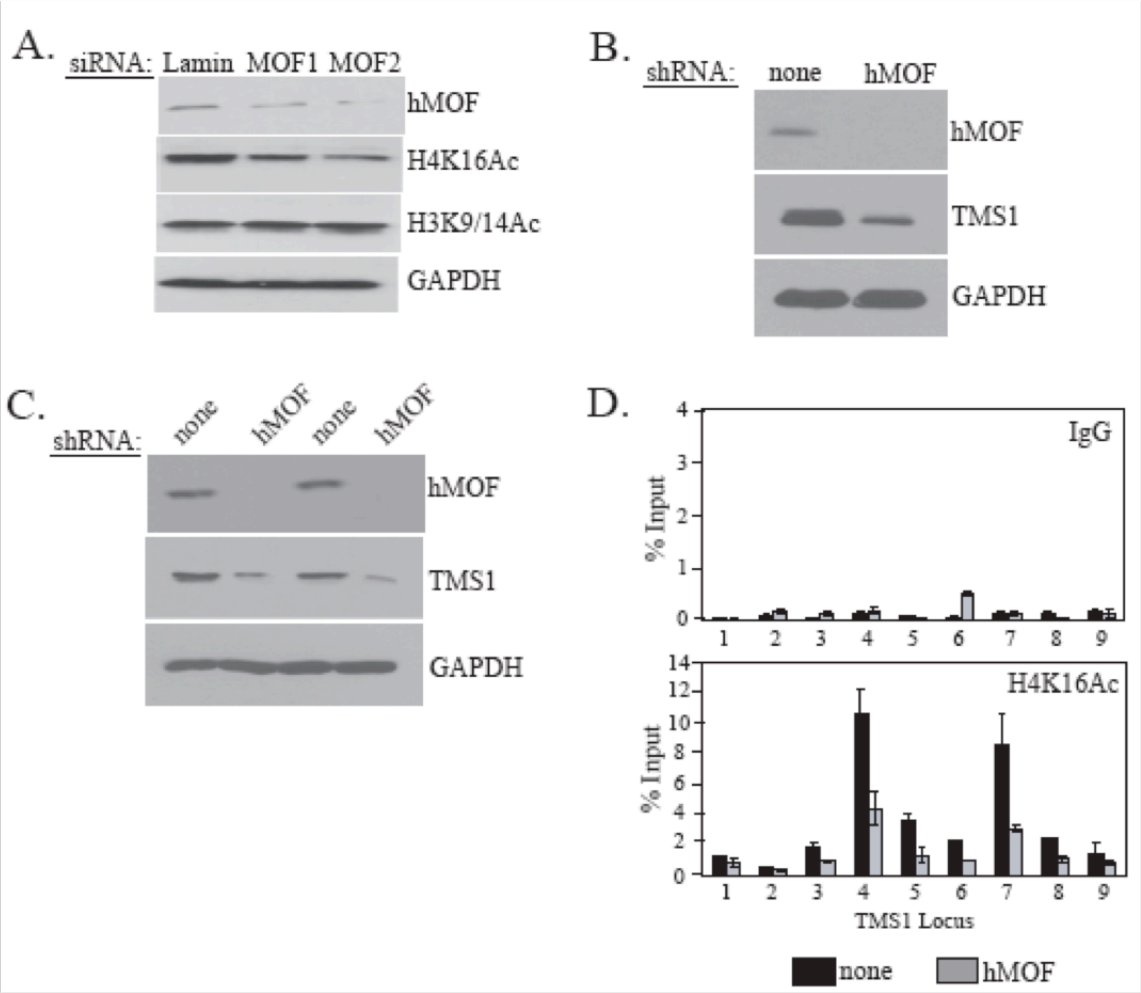

A
ESR1:
CDH1:
bp: -1,000 -500 0 500 1,000 1,500 2,000
CpG Island
T
I
II
1 2 3 4 5
B
ESR1:
CDH1:
IgG
H3K9/14Ac
H4K16Ac
% Input
MCF7
MDA-MB231
C
siRNA: Lamin MOF1 MOF2
hMOF
ESR1
CDH1
GAPDH

Figure 6. Role of H4K16Ac in the regulation of the ESR1 and CDH1 loci. **A)** Schematic of the region encompassing the CpG islands of the ESR1 and CDH1 gene. Nucleotide positions with respect to the transcription start site (T) and the CpG island are indicated above each gene diagram. Exons are shown in grey boxes. The regions amplified by primer sets (1-5) used in real-time PCR are shown below each gene. **B)** Localization of H3K9/14Ac and H4K16Ac at the ESR1 and CDH1 CpG island. ChIP analyses were performed for MCF7 or MDA-MB231 cells with the indicated antibodies or a negative control (IgG), followed by real-time PCR of regions depicted in A). Each experiment was conducted at least three times and although the immunoprecipitation efficiency varied between experiments, the profile of enrichment across each locus was consistent. Shown are the mean +/- standard deviation of triplicate determinations from a representative experiment. **C)** MCF7 cells were transfected with siRNA (100 nM) against Lamin A/C (Lamin) (irrelevant control) or hMOF (MOF1 and MOF2) and the expression of ESR1 and CDH1 protein was determined by western blotting. The same blot was also exposed to anti-hMOF and anti-GAPDH antibodies.



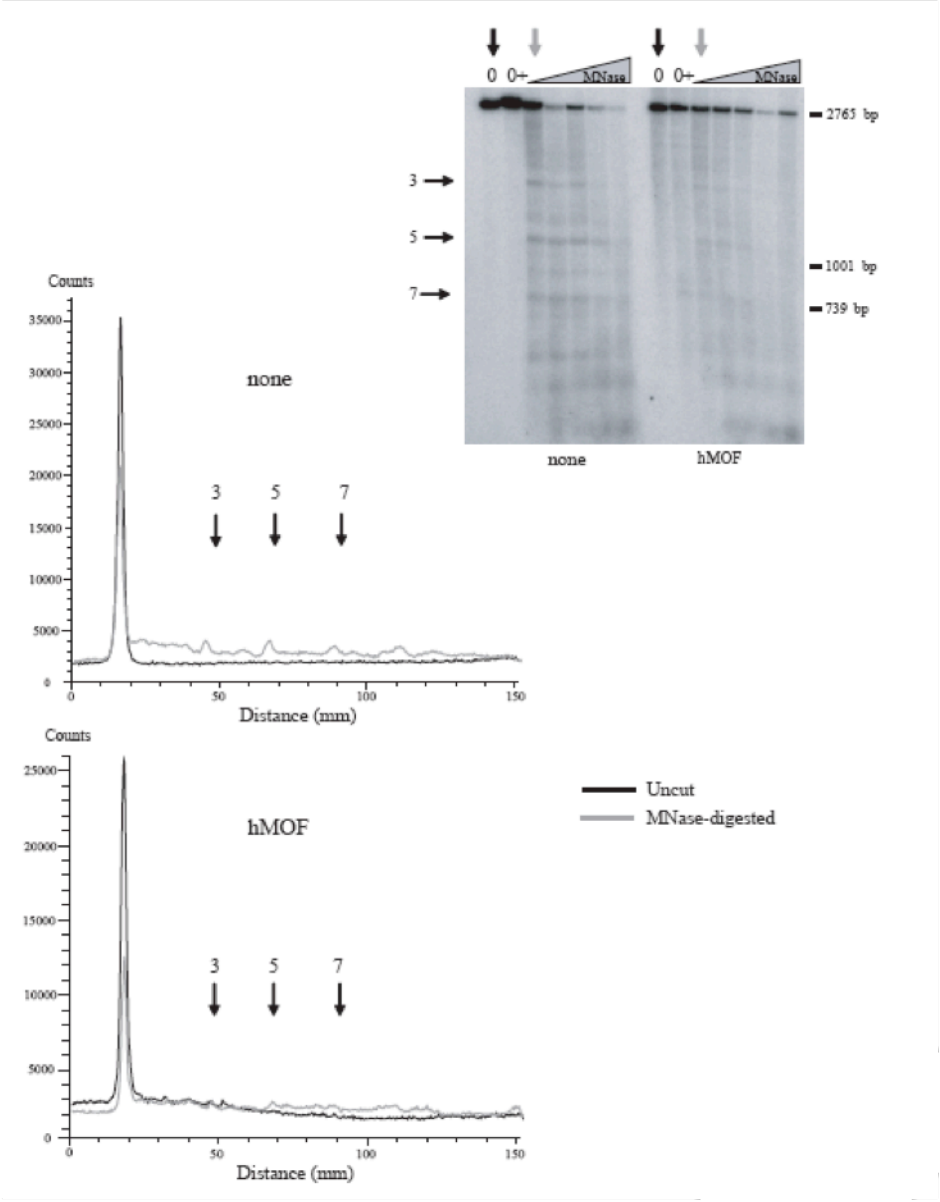
Supplemental Figure 1. Quantification of nucleosome positioning at the TMS1 locus in MCF7 and MDA-MB231 cells. The Southern blot shown in Figure 2C was used to quantify the relative frequency of the MNase-cut sites across the TMS1 locus in the indicated cell lines by pixel density analysis using the Imagequant software (Amersham Biosciences). Graphed are the relative intensity of uncut samples (black) and the lane in which there is roughly 50% digestion by MNase (grey). The lanes corresponding to the uncut and cut samples that are graphed are shown on the Southern blot (black and grey arrows, respectively). The position of cut sites 3, 5 and 7 are indicated with arrows on the blot and the graph.

Supplementary Figure 2.



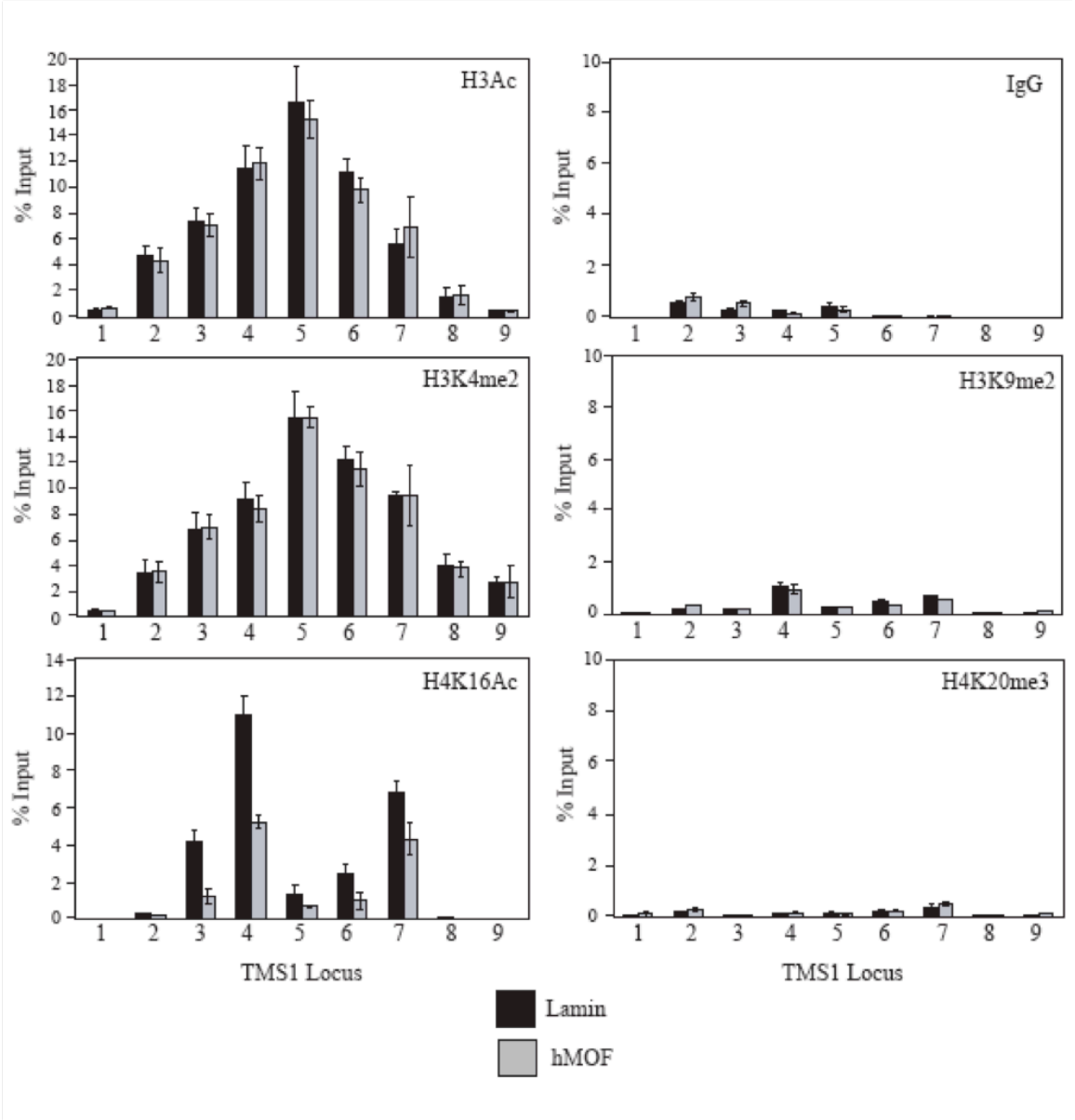
Supplemental Figure 2. Effect of hMOF downregulation on H4K16Ac and TMS1 expression. **A)** MCF7 cells were transfected with 100 nM of siRNA directed against hMOF (MOF1 and MOF2) or Lamin A/C (Lamin) (irrelevant control). Cells were harvested four days post-transfection and equal amounts of protein lysates were subjected to SDS-PAGE followed by western blotting using antibodies against hMOF, H4K16Ac, H3K9/14Ac or GAPDH (loading control). **B)** MCF7 cells infected with the pLKO.1 lentiviral vector (none) or with pLKO.1 expressing shRNA against hMOF (hMOF) were placed under puromycin selection for five days. Protein lysates were subjected to western blot analysis with hMOF, TMS1 or GAPDH antibody. **C)** IMR90 cells were infected with lentiviral constructs and subjected to western blot analysis exactly as indicated in (B). **D)** IMR90 cells infected with the pLKO.1 lentiviral vector (none) or with pLKO.1 expressing hMOF shRNA (hMOF) was subjected to ChIP with antibodies against a rabbit IgG (IgG) control or H4K16Ac (H4K16Ac). Immunoprecipitated DNA was amplified by real-time PCR using primer sets indicated in Figure 1A. Percent (%) input was determined as the amount of immunoprecipiated DNA relative to input DNA. Shown are the mean +/- standard deviation of triplicate determinations from a representative experiment.

Supplementary Figure 3.



Supplemental Figure 3. Quantitative analysis of nucleosome positioning at the TMS1 locus in MCF7 cells expressing or silenced for hMOF. The Southern blot shown in Figure 3C was used to determine the relative intensity of the MNase-digested sites across the TMS1 locus in MCF7 cells infected with pLKO.1 (none) or with pLKO.1 expressing shRNA directed against hMOF (hMOF), exactly as described in Supplemental Figure 1.

Supplementary Figure 4.



Supplemental Figure 4. Effect of hMOF downregulation on the histone modification profile at the TMS1 locus. MCF7 cells were transfected with Lamin (Lamin) or hMOF2 (hMOF) siRNA. Four days post-transfection, chromatin was isolated and immunoprecipiated using the indicated antibodies. DNA pull-downs were amplified by real-time PCR using primer sets indicated in Figure 1A. Percent (%) input was determined as the amount of immunoprecipiated DNA relative to input DNA. Each ChIP was repeated at least three times with consistent profile of enrichment across the locus. Shown are the mean +/- standard deviation of triplicate determinations from a representative experiment.

Chapter 3: The 5' and 3' boundaries of the unmethylated *TMS1* CpG island function as enhancer blockers

Kagey, JD performed all experiments in this chapter and wrote the chapter in its entirety.

This data is currently unpublished

The focus of the work in this chapter was to assay HS1 and HS3 (which flank the unmethylated CpG island at TMS1) for their ability to enhancer block in a plasmid based assay.

Introduction

The combinatorial effect of DNA methylation and the post-translational modifications of histone tails cooperate in the epigenetic regulation of genes and establish domains that are either permissive or inhibitory to transcription. Neighboring genes regulated in a similar manner are organized into chromatin domains (54). The maintenance of the boundary between domains allows is necessary component of differential gene regulation between adjacent chromatin domains. DNA sequences that act in *cis* to mediate the boundary function are known as insulators (21).

Through the study of several well-characterized insulators, we have a basic understanding of insulator function. The *H19/Ifg2* locus is an imprinted region exhibiting mono-allelic expression with *H19* expressed only from the maternal allele, and *IGF2* expressed only from the paternal allele (139). The *H19* and *Igf2* genes share a common enhancer and utilize an insulator to ensure the mono-allelic expression of both genes. This achieved through the differential binding of CTCF, a zinc finger protein capable of mediating insulator function through the binding to an imprinting control region (ICR) that if located between the two genes and a common downstream enhancer (47). CTCF's binding to the ICR is dependent on the differential DNA methylation of two CTCF binding sites within the ICR (53).

CTCF has further been demonstrated to facilitate insulator function at numerous boundaries of chromatin domains throughout the genome, including HSIV (in the *ß-globin* gene) and XL-9 (51, 140). CTCF occupancy is cell type dependent, providing

individual cells the ability to establish unique domains of epigenetic regulation. In addition to DNA methylation, the positioning of nucleosomes also has the ability to influence CTCF, through the preservation of binding sites (44). The regulation of CTCF binding and expression is essential for maintaining the overall integrity of gene expression (55). In fact up-regulation of CTCF has been correlated with a resistance to cell death within breast cancer cells through the deregulation of chromatin domains. This resistance to apoptosis may be mediated, in part, through a decrease in *Bax* transcriptional levels which directly correlates with the over expression of CTCF (141). While CTCF is not the only mammalian insulator protein to be discovered, USF-1, it is easily the most widely studied and has provided the most insight into the function and regulation of insulator regions.

At the molecular level, insulators have two distinctive functions, barrier and enhancer blocking activity. Individual insulators have been described as having either one or both of these functions. Barrier activity refers to the ability of an insulator to prevent the spread of heterochromatin from one domain to another, also known as position effect variegation (PEV). Enhancer blocking activity refers to the ability of an insulator to blocking a distal enhancer from functioning on a gene (52). Although the two functions are distinct, both can be mediated through common insulator binding proteins, such as CTCF and USF-1 (142). Barrier activity is measured experimentally through the use of a reporter gene flanked on both sides with the insulator region in question. The vector is randomly integrated into the genome and multiple clones are maintained. If the flanking sequencing is able to maintain the expression of the reporter gene in multiple clones, then the insulator sequences exhibit barrier activity (143). Enhancer blocking activity is

measured by assaying the ability of an insulator sequence to block the activity of a distal enhancer on the promoter of a reporter gene when placed between them in a reporter construct. If the sequence is capable of blocking the promoter activity of the reporter gene, the levels of luciferase expression will be repressed compared to a vector with no insulator sequence (140). Insulators have been demonstrated to exhibit one or both of these distinct functions. The insulator found at the H19/Igf2 locus exhibits both enhancer blocking activity and barrier activity, as the binding of CTCF has also been demonstrated to both prevent de novo DNA methylation from adjacent chromatin domains and prevent the enhancer from interacting with both genes on the same chromosome (144).

The mechanisms maintaining the *TMS1* CpG island in an unmethylated state in normal cells are not completely understood. We have found the unmethylated CpG island in MCF7 cells is flanked on both the 5' and 3' sides by regions of densely methylated DNA. Two DNase I hypersensitive sites (HSs) flank the CpG island and are lost upon the aberrant methylation of the CpG island (100). Distinct peaks of H4K16ac and strongly positioned nucleosomes demarcate these boundaries of DNA methylation (Chapter 2, (103)). Furthermore, HS1, which is found at the 5' boundary of DNA methylation, contains both an E-box (the binding sequence of USF-1) and HS3 contains three predicted CTCF sites using the CTCFBS prediction program (although the CTCF consensus binding sequence is rather degenerate) (145). Given these findings, we hypothesized that the DNA HSs bordering the CpG island insulators preventing the spreading of DNA methylation into the unmethylated CpG island and maintaining an active *TM1* locus.

Methods

Luciferase Enhancer Blocking Assay

The pGL3 basic vector (Promega) was used as a starting construct. The SV40 enhancer was cloned into an *MluI* site, while the HS4 (from β–globin) was cloned into a 5' *SacI* site to serve as a control insulator. This construct, along with the pXL-9 and pλ insulator constructs, which served as a positive and negative control, respectively, were a gift from the Boss lab (140). *TMS1* potential insulator sequences at HS1 and HS3 (~200 bp) were amplified centering on mapped location of the HSs (100)). Once amplified, inserts were sequenced and cloned into the vector's multiple cloning region (*NheI* and *XhoI*) and were named pHS1 and pHS3, respectively. Both sequences were amplified with a 5' *NheI* site and a 3' *XhoI* site. Once ligations were verified, plasmids were electroporated into competent DH10B cells, grown in 100 ml liquid LB culture, and isolated via the Sigma maxi-prep kit.

All enhancer-blocking plasmids were transfected along with renilla (transfection control) into either MDA-MB231 (*TMS1* methylated) or MCF7 (*TMS1* unmethylated) cells. We plated 5 x 10^4 cells into 24 well plates twenty-four hours prior to transfection. The following day cells were transfected with 2 μg of either pHS1, pHS3, pXL-9, and pλ along with 0.04 μg of renilla construct using Lipofectamine (Invitrogen) and placed at 37^o C overnight. Each construct was transfected in triplicate. Forty-eight hours after transfection cells were lysed in 100 μl passive lysis buffer and placed on shaker at room temperature for fifteen minutes. Next, 20 μl lysate was added to 100 μl luciferase assay reagent II (LAR II) and the tube was placed into the illuminometer. Data was graphed as the triplicate average of Luciferase/Renilla fluorescence.

Results

We assessed the potential of the DNase I HSs flanking the unmethylated *TMS1* CpG island to act as enhancer-blockers. Though this assay does not directly query barrier ability, enhancer-blocking ability is a closely linked phenomenon and often insulators exhibit both functions (146). Between the SV40 enhancer and the minimal promoter attached to a *luciferase* reporter a 200 bp region centered on the previously mapped HS1 and HS3 of the *TMS1* gene was cloned (100). A 200 bp region of λ DNA was used as a negative control, and a previously defined enhancer blocker, XL-9, was used as a positive control (140). The assay was conducted in both MCF7 (unmethylated for *TMS1*) and MDA-MB231 (methylated for *TMS1*) cell lines. We find that both pHS1 and pHS3 demonstrated the ability to block the SV40 enhancer acting upon the minimal promoter as indicated by a 2-3 fold reduction of *luciferase* activity compared to the known pλ (**Figure 1**). Interestingly, the construct containing HS3 (the 3' boundary of the CpG island) was a slightly more potent enhancer blocker than HS1 (the 5' boundary of the CpG island). This may result from the HS1 sequence containing a portion of the *TMS1* promoter region and thus promoting a low level of *luciferase* expression. Furthermore, we find no differences in the ability of HS1 or HS3 to serve as an enhancer blocker in either background of MCF7 or MDA-MB231 cells. This suggests that the factors involved in creating that the *trans* factors involved in mediating enhancer-blocking activity are present in MDA-MB231 cells, despite the *TMS1* HSs not being formed.

Discussion

Here we have established that both the 5' and 3' HSs that form at the boundaries of the CpG island when *TMS1* is unmethylated, have the ability to block the SV40 enhancer comparably to the known enhancer blocker XL-9 (140). In Chapter 2 we demonstrated that HS1 and HS3 not only demarcate the boundaries of DNA methylation, they also coincide with distinct peaks of H4K16ac and strongly positioned nucleosomes. It appears that these boundaries of DNA methylation do not merely represent coincidental co-localization of epigenetic factors, but that they represent functional regions of chromatin. Interestingly, HS1 and HS3 were able to enhancer block equally well in MCF7 cells (in which, *TMS1* is unmethylated) and MDA-MB231 cells (in which, *TMS1* is methylated). This suggests that the epigenetic factors responsible for mediating enhancer-blocking activity are present in MDA-MB231 cells, although the HSs do not form at *TMS1* in these cells. One possibility is that the DNA methylation of these regions in MDA-MB231 cells is inhibits the binding of proteins that would otherwise lead to the establishment of these HS regions.

Insulator activity is often mediated through the binding of non-histone proteins including CTCF and USF1. The boundaries of the *TMS1* CpG island contains potential binding sequences for both CTCF and USF-1. Since we have demonstrated both the 5' and 3' boundaries of this island exhibit enhancer blocking ability, we hypothesized that this phenomenon was mediated through CTCF or USF-1. We performed ChIP for both USF-1 and CTCF in MCF7 and MDA-MB231 cells with inconclusive results (data not shown). Since conducting these experiments, several genome-wide CTCF studies have been made publicly available (147) and show little to no enrichment of CTCF at the

boundaries of the *TMS1* CpG island. One such analysis was done in the IMR90 cell line, where it is known that the *TMS1* HSs form and are not enriched for CTCF (100, 148). This leaves the possibility that the enhancer blocking activity found at the 5' and 3' boundary of the *TMS1* CpG island is mediated through a novel insulator protein.

Given that the unmethylated *TMS1* CpG island is flanked on both sides by DNA sequences exhibiting enhancer blocking activity, and the DNA outside of the island is heavily methylated in normal cells, we proposed that HS1 and HS3 may physically interact within the nucleus facilitating the maintenance of an unmethylated domain. Recently, proximal CTCF sites have been shown to interact in close spatial proximity to create loops. To address this hypothesis we performed chromatin capture conformation (3C), which queries if non-linear DNA sequences are in close spatial proximity in the context of the chromatin structure (149). Our experiments were inconclusive as we were unable to distinguish between enrichment of interaction and background due to the short linear distance (~1kb) between the two DNA sequences in question. Furthermore, personal communication with Job Dekker has validated our hypothesis that the 3C technique is unable to resolve two sequences separated by 1kb of DNA. However, we cannot rule out this possibility.

Our finding that both HS1 and HS3 demonstrate barrier activity presents the rationale for several further experiments. This enhancer blocking activity imparts rationale to undertake the more laborious barrier activity assay using HS1 and HS3 at *TMS1* as complete insulator regions. Furthermore, the identification of the protein(s) mediating this insulator activity *in vivo* may provide important insight into the

mechanism(s) by which CpG islands are maintained in a unmethylated state how how those barriers break down in cancer.

Figures

Figure 1.

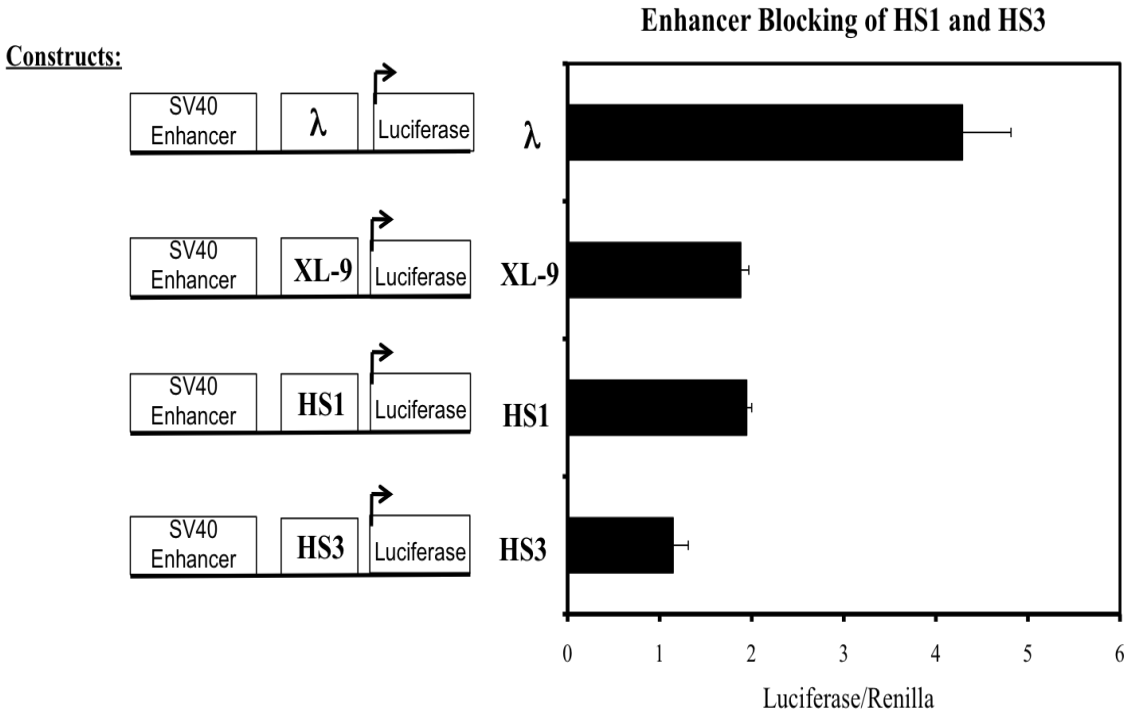


Figure 1: Enhancer blocking analysis of DNase I HSs bordering *TMS1* CpG island. Enhancer blocking constructs were transfected into MB-231 cells with renilla as a transfection control. Schematics of enhancer blocking constructs are depicted to the left of the graph. Lambda DNA was used as a negative control and XL-9 was used as a positive enhancer blocking control (140). Experiment was performed in MCF7 cells with similar results.

Chapter 4: Long-term stability of maintenance of a poised epigenetic state after transient exposure to 5-aza-2'-deoxycytidine correlates with occupancy by RNA polymerase II

Jacob D. Kagey[1,2], Priya Kapoor-Vazirani[2], Michael T. McCabe[2], Doris R. Powell[2], and Paula M. Vertino[1,2,3]

[1]Graduate Program in Genetics and Molecular Biology, Emory University [2]the Department of Radiation Oncology, Emory University School of Medicine, Atlanta, GA 30322

[3] To whom correspondence should be addressed at:
1365-C Clifton Rd, NE, Rm 4086
Atlanta, GA, 30322
Email: pvertin@emory.edu

Running Title: 5-aza-2'-deoxycytidine induces stable demethylation at TMS1

Keywords: DNA methylation, epigenetics, TMS1, chromatin, Decitabine, histone,

Kagey JD, cultured and treated cells with 5-azaCdR. Harvested DNA, RNA, and made chromatin for ChIP analysis. Performed all methylation analysis including MSP, bisulfite sequencing, COBRA, and ChIP-bisulfite sequencing. Performed QRT-PCR analysis. Wrote entire manuscript.

Kapoor-Vazirani, P performed ChIP analysis and shRNA knockdowns of *SUV4-20*

Figures 1,2,4,5,6,7, and supplemental data have been submitted as a manuscript and are under review at Molecular Cancer Therapeutics

Figure 3 is preparation for an additional manuscript

The focus of the work in this chapter was to address the role of DNA methylation and H4K20me3 in the establishment and maintenance of a silenced TMS1 locus. Furthermore, this chapter investigates the relationship between DNA methylation and individual histone modifications both at the TMS1 locus as well as CDH1 and ESR1.

Introduction

Cell-type specific gene expression patterns are established and maintained in part by epigenetic mechanisms including DNA methylation and post-translational histone modifications. In mammals, DNA methylation occurs at cytosine residues found within the context of CpG dinucleotides, which are underrepresented in the genome but can be found in clusters called CpG islands that are associated with the regulatory regions of approximately 70% of human genes (23, 24, 150). In normal cells, most CpG islands are unmethylated and transcriptionally competent. Methylation of CpG islands is associated with stable and heritable gene silencing (62). In addition to DNA methylation, post-translational modifications of the histone tails contribute to epigenetic regulation. Unmethylated CpG islands tend to be packaged into nucleosomes marked with 'active' histone modifications including, histone H3 acetylation of lysines 9 and 14 (H3ac), and histone H3 lysine 4 di- and tri -methylation (H3K4me2/3). Methylated CpG islands are generally depleted of these active marks, and are alternatively marked by a subset of 'repressive' histone modifications including histone H3 lysine 9 di- and tri-methylation (H3K9me2/3), and histone H3 lysine 27 tri-methylation (H3K27me3) (18, 151).

Genome-wide epigenetic alterations occur in cancer. There is an overall hypomethylation of the genome concurrent with the aberrant hypermethylation of a subset of CpG islands (63). This aberrant methylation of CpG islands in cancer is accompanied by a shift from a permissive to a more repressive histone modification

profile generally characterized by the loss of H3ac and H3K4me2/3 and the acquisition of H3K9me2/3 (128, 129, 152). Recent work suggests that the epigenetic silencing of some loci also involves a shift in histone H4 modifications including loss of histone H4 acetylated at K16 (H4K16Ac) and gain of histone H4 tri-methylated at lysine 20 (H4K20me3) (103). The combination of CpG island methylation and accompanying histone modifications is associated with stable gene repression and is one mechanism leading to the heritable inactivation of tumor suppressor genes during tumor progression (22, 153).

The finding that tumor suppressor genes are often inactivated by epigenetic means and that these events can play a direct role in cancer initiation and progression provides a compelling rationale for the utilization of inhibitors of DNA methyltransferases (DNMTs) and histone modifying enzymes as a therapeutic strategy (76, 151). Over the last decade, there has been considerable effort in the clinical development of DNMT and histone deacetylase (HDAC) inhibitors in cancer therapy. For example, 5-aza-2'deoxycytidine (5-azaCdR) (Decitabine) and 5-aza-cytidine (Vidaza) have become the standard of care for myelodysplastic syndrome (MDS), and have shown promise in the treatment of leukemias (86, 151). In these cases, DNMT inhibitors have demonstrated both higher response rates and increased survival when compared to more traditional cytotoxic chemotherapeutic agents (76, 154).

Despite their clinical success, there is still much to understand about the molecular mechanisms underlying the clinical activity of these agents and the durability of response. Studies in cell culture have shown that 5-azaCdR treatment induces DNA demethylation and re-activation of epigenetically silenced tumor suppressor genes. This

reactivation is accompanied by the loss of some repressive histone modifications (*e.g.* H3K9me2) and the reappearance of active histone modifications (*e.g.* H3ac and H3K4me2) (70, 84, 85). However, the chromatin structure of CpG islands does not return to a fully active configuration due to the preservation of some repressive histone modifications unaffected by DNA demethylation such as H3K27me3 and H3K9me3, leaving open the potential for re-silencing after drug removal (84). Molecular analyses from biopsy driven clinical trials indicate that global and gene-specific DNA demethylation is achievable *in vivo.* However, in cases where specific gene demethylation has been detected, remethylation is often observed within a few weeks of treatment. Furthermore, the degree to which these molecular endpoints predict clinical response remains an open question (76).

To further understand the long-term effects of transient 5-azaCdR treatment on tumor suppressor gene reactivation, we have studied the dynamics of DNA methylation, gene expression, and histone modifications at *TMS1/ASC* (*Target of Methylation-induced Silencing 1*), a CpG island-associated proapoptotic gene that is frequently silenced in conjunction with DNA hypermethylation in human breast, prostate and lung cancers (97). We find that 5-azaCdR-induced reactivation of *TMS1* is accompanied by DNA demethylation and a shift from a repressive histone profile characterized by H3K9me2 and H4K20me3 to a more active profile that includes the re-association of RNA polymerase II (Pol II) with the *TMS1* promoter. However, this shift is not complete as H4K20me3 is unaffected by 5-azaCdR. Interestingly, treatment of MDA-MB231 cells with both 5-azaCdR and shRNA against *SUV4-20* (HMT of H4K20me2/3) results in the synergistic re-activation of *TMS1* expression. Thus providing evidence for two

independent mechanisms of *TMS1* repression. Furthermore, we find that whereas a fraction of the alleles are remethylated after drug removal, there is a subpopulation of *TMS1* alleles that remained stably unmethylated for at least 27 passages in culture (~ 3 months). This subpopulation of unmethylated alleles associates with both active (H3ac, H3K4me2,) and repressive marks (H3K9me2 and H4K20me3), and remains occupied by Pol II. We further find that locus-specific differences exist in the sustainability of DNA demethylation after drug removal. Together, our data suggest that the ability to attain and to maintain Pol II occupancy is a critical factor in the long-term stability of DNA demethylation and gene expression after drug-induced reactivation.

Methods

Cell culture and 5-azaCdR treatments

MDA-MB231 cells were obtained from the American Type Culture Collection and cultured in DMEM supplemented with 10% FBS and 2 mM L-glutamine. For 5-azaCdR treatments, $5x10^4$ MDA-MB231 cells were plated in a 10 cm dish 24 hours prior to treatment with 0.5 μM 5-azaCdR. Medium containing fresh 5-azaCdR was applied every other day for six days (**Figure 1B**). Following treatment, cells were maintained in the absence of 5-azaCdR and split 1:10 every three days for 27 passages (~ 3 months). Cells were harvested and DNA, RNA, and chromatin were collected at 0, 3, 6, 9, and 27 passages post-treatment.

Lentiviral shRNA infection

Lentivirus expressing the empty pLKO.1 or pLKO.1 plasmid expressing shRNA against Suv4-*20* (TRCN0000141926) was produced in 293T cells by Fugene (Roche) transfection, essentially as described in Kapoor et al (2008). pLKO.1 plasmids were obtained from Open Biosystems. For infection, 1 x 10^6 – 2 x 10^6 MDA-MB231 cells were plated on 10 cm plates. Twenty-four hours later, cells were infected with 0.5 mL of lentiviral particles in the presence of 8 mg/mL polybrene. Cells were replenished with fresh medium containing 0.5 – 1.0 mg/mL of puromycin the following day. For infected MDA-MB231 cells also treated with 5-aza-2'-deoxycytidine (DAC), 0.5 mM of 5-azaCdR was included in the medium. Fresh medium containing puromycin (and 5-azaCdR when used) was exchanged every other day. Cells were incubated in the selection medium for 5 days and subsequently harvested for downstream assays.

Methylation Specific PCR

Genomic DNA (2 μg) was bisulfite-modified using the EZ DNA methylation kit (Zymo) and ~50 ng of modified DNA was utilized as template for Methylation-Specific PCR (MSP) as previously described (155). PCR conditions used were 5 min. at 95^o C, followed by 35 cycles of 30 s at 95^o C; 45 s at 58^o C, and 45 at 72^o C; with a final 5 min extension at 72^o C. PCR products were resolved on a 1.5% agarose gel and stained with ethidium bromide. MSP primers are described in **Supplementary Table 1**.

COmbined Bisulfite Restriction Analysis

Bisulfite-modified DNA was amplified using primers devoid of any CpGs. Amplified products were purified with the PCR Purification kit (Qiagen), digested overnight at 37^{o}C with either *FNU4*HI or *Xmn*I, precipitated, and resolved on a 2.0% agarose gel (156). Relative intensities of digested and undigested bands were quantified with ImageQuant 5.2 and percent methylation was determined as the combined intensity of the digested bands relative to that of all bands (undigested and digested). Primer sequences are in **Supplementary Table 1**.

Genomic Bisulfite Sequencing

Genomic bisulfite sequencing (GBS) was performed as previously described (118). Briefly, bisulfite-modified DNA was amplified using primers devoid of CpGs as described above. PCR products were TA-cloned (Invitrogen), transformed into chemically competent *E. Coli* and plasmid DNA isolated from 10-17 individual colonies was sequenced. Bisulfite sequencing data were analyzed using the BiQ Analyzer software (157). Primer sequences are listed in **Supplementary Table 1**.

Reverse Transcriptase PCR

RNA was isolated using the RNeasy kit (Qiagen) and was reverse-transcribed with random hexamer primers and M-MLV reverse transcriptase (118). Quantitative real-time PCR (q-PCR) was used to analyze gene-specific transcripts and their levels were normalized to *18s* rRNA as previously described (103). Primer sequences are listed in **Supplementary Table 1**.

Chromatin Immunoprecipitation and ChIP-Bisulfite Sequencing

Chromatin Immunoprecipitation (ChIP) was performed as previously reported (103). The following antibodies were used for specific immunoprecipitations, H3ac (Millipore, 06-599), H3K4me2 (Millipore, 07-030), H3K9me2 (Millipore, 07-441), H4K20me3 (Abcam, ab9053), and Pol II (Santa Cruz, SC-9001x). Immunoprecipitated DNA was analyzed by qPCR with primers for *TMS1, CDH1,* or *ESR1* (103). ChIP-bisulfite Sequencing (ChIP-bis) was adapted from the ChIP-MSP protocol initially reported by Zinn *et. al.* (158). Immunoprecipitated DNA from the ChIP procedure was subject to bisulfite conversion as described above. Modified DNA was then amplified using *TMS1*-specific bisulfite sequencing primers listed in **Supplementary Table 1**. PCR products were purified, cloned, and plasmid DNA from individual colonies was sequenced.

Micrococcal Nuclease Digestion and Indirect End-labeling

Isolation of intact nuclei and indirect end labeling was performed as described in (100) with minor modifications. Nuclei from 2 x 10^6 cells were digested with micrococcal nuclease (MNase) (10 – 200 u) at 25^{o}C for 10 min in RBS (10mM TrisHCl, 5mM $MgCl_2$, 0.5 M DTT, 0.3 mM sucrose, 0.4 mM PMSF). Reactions were stopped by incubation in 1% SDS, 20 mM EDTA, and 10 µg/ml RNaseA for 30 min at 37^{o}C, followed by incubation with 1 mg/ml Proteinase K at 50^{o}C overnight. DNA was recovered by phenol-chloroform extraction and ethanol precipitation. MNase-digested DNA (10 µg) was digested with 2 units each Hind III and Spe I at 37^{o}C overnight, separated on a 1.0%

agarose gel and transferred to a nylon membrane. Membranes were hybridized with a random-prime labeled Spe I-Xmn I probe anchored to the downstream Spe I site of TMS1. Approximately 100 pg Spe I-Hind III (2765 bp), Spe I- Bam HI (1001 bp) and Spe I-Eco RI (739 bp) fragments from the TMS1 locus were run on each gel as internal size markers. Membranes were washed to high stringency (2X SSC, 0.1% SDS at 65^{o}C), exposed to phosphor storage screens and analyzed by phosphoimage analysis using ImageQuant software (Molecular Dynamics).

Results

TMS1 is a pro-apoptotic tumor suppressor gene containing a promoter-associated CpG island that is frequently methylated and silenced in human cancers (**Figure 1**A)(97). Previously, we compared the epigenetic landscape of *TMS1* in breast cancer cell lines in which the *TMS1* gene is either unmethylated and the gene is expressed (MCF7) or methylated and silenced (MDA-MB231). We find that in MCF7 cells and other normal cells that express TMS1 (*e.g.* IMR90), the CpG island is unmethylated, with DNase I hypersensitive sites (HSs), positioned nucleosomes, and distinct peaks of H4K16ac marking the boundaries between the unmethylated CpG island domain and surrounding methylated DNA (100, 103). In this state, the unmethylated CpG island is enriched for H3K4me2 and H3ac, and depleted for H3K9me2. Conversely, MDA-MB231 cells have a densely methylated CpG island in which the CpG island-associated DNase I HSs have been lost, nucleosomes are randomly positioned, and the active histone marks (H3K4me2, H4K16ac, and H3ac) have been replaced by repressive marks including H3K9me2 and H4K20me3. In this transcriptionally repressed state, H4K20me3 is

localized to a prominent peak just upstream of transcription start, whereas H3K9me2 is enriched throughout the CpG island (103).

Long-term effects of transient 5-azaCdR treatment at the TMS1 locus

To investigate the long-term effects of a transient exposure to a DNA de-methylating agent on the chromatin architecture at the *TMS1* locus, MDA-MB231 cells were treated with 0.5 μM of 5-azaCdR every other day for 6 days and then maintained in culture for 27 passages (~80 days) in the absence of 5-azaCdR. Consistent with previous work, treatment of MDA-MB231 cells with 5-azaCdR induced the re-expression of *TMS1* mRNA (**Figure 1C**) (118). However, while *TMS1* expression was induced an average of 494-fold immediately following 5-azaCdR treatment, expression levels returned to ~3-fold over untreated cells within 3 passages after drug removal (**Figure 1B**). This low level of *TMS1* expression was then maintained for at least three months (27 passages) in the absence of 5-azaCdR.

Treatment with 5-azaCdR resulted in an average demethylation of the *TMS1* CpG island from 100% methylated to 64% methylated as determined by COBRA analysis (**Figure 1C**). After the removal of 5-azaCdR, there was an initial burst of remethylation (from 64% to 76% methylation) after 3 passages, which leveled off at ~83% and was maintained at this level for 27 passages in culture in the absence of drug (**Figure 1D**). Similar remethylation kinetics were observed when methylation was measure by MSP (**Supplementary Figure 1**).

To distinguish between the possibilities that all *TMS1* alleles exhibit a partial remethylation or that distinct subpopulations of unmethylated and methylated alleles

persist after the removal of 5-azaCdR, DNA was bisulfite modified and individual alleles were sequenced. While all alleles were densely methylated in untreated MDA-MB231 cells, treatment with 5-azaCdR induced a heterogeneous methylation pattern consisting of alleles that were densely methylated, predominantly unmethylated, and those with a mixed pattern (**Figure 1D**). Removal of 5-azaCdR resulted in the partial remethylation of the locus that, over time, resolved into two distinct subpopulations; one comprised predominantly of methylated alleles and the other comprised predominantly of unmethylated alleles (**Figure 1D**). Interestingly, where there was a more checkered methylation profile seen in early post 5-azaCdR passages, the unmethylated alleles observed 27 passages later were almost entirely devoid of methylation.

Effects of 5-azaCdR on histone modifications at TMS1

To examine the effects of 5-azaCdR on chromatin at the *TMS1* locus, we utilized ChIP to map the histone modification profile before, immediately following, and over the course of 27 passages after treatment with 5-azaCdR. In untreated MDA-MB231 cells, the *TMS1* locus was enriched for H3K9me2 and H4K20me3 (**Figure 2**). Treatment with 5-azaCdR led to the accumulation of H3ac and H3K4me2, as well as a decrease in H3K9me2 at *TMS1* compared to untreated cells. This change was further accompanied by the re-association of Pol II (**Figure 2**). In contrast, the peak of H4K20me3, observed in untreated MDA-MB231 cells was unaffected by 5-azaCdR. Additionally, H4K16Ac, which is found at the unmethylated *TMS1* locus in cells that express the gene (*e.g.* MCF7 cells,) (103), was not observed (data not shown). Thus, while 5-azaCdR treatment led to a

shift to a more transcriptionally permissive chromatin configuration, it did not fully recapitulate the conformation observed in cells that normally express *TMS1*.

Over the course of 27 passages following the removal of 5-azaCdR the enrichment of H3ac, H3K4me2, and Pol II were partially depleted at *TMS1*, but still remained three to twenty-one fold higher than that of untreated MDA-MB231 cells (**Figure 2**). H3K9me2 levels were partially restored over the same time frame, but remained at levels lower than the untreated MDA-MB231 cells. This shift in the histone profile after the cessation of treatment coincided with the partial remethylation of DNA. The peak of H4K20me3 found upstream of the *TMS1* transcription start remained unchanged throughout the time course. Thus, although there was a partial return to the pre-treatment chromatin state over time (DNA methylation and H3K9me2), a low level of active modifications and Pol II occupancy were maintained. Similar results were obtained from a second independent time course (**Supplemental Figure 2**) as well as at a second primer set (#5, (103)) in the TMS1 CpG island (data not shown).

The combined treatment of MDA-MB231 with 5-azaCdR and shRNA target against Suv4-20 synergistically re-expresses TMS1 in MDA-MB231 cell

The continued presence of H4K20me3 following 5-azaCdR treatment raised the question of whether H4K20me3 and DNA methylation represent independent layers of repression at the *TMS1* locus. Therefore, we tested the impact of simultaneous inhibition of DNA methylation and H4K20me3 on gene expression, chromatin modifications, and nucleosome positioning. shRNA constructs were designed to target SU-var 4-20 (Suv4-20), the histone methyl transferase responsible for the di and tri methylation of H4K20

(121). MDA-MB231 cells were treated with shRNA against *Suv4-20* both alone and in combination with 5-azaCdR. Down regulation of *Suv4-20* resulted in the depletion of H4K20me3 and the re-establishment of H4K16ac but had no effect on DNA demethylation, nucleosome positioning, or *TMS1* expression (**Figures 3A, 3B, 3C** and Data not shown). All other histone modifications studied (H3ac, H3K4me2, and H3K9me2) were unaltered by treatment with shRNA against *Suv4-20* (**Figure 3B**). It should be noted that these were the modifications altered at *TMS1* when treated with 5-azaCdR alone (**Figures 2, 3B**).

The combined treatment of 5-azaCdR and shRNA against *Suv4-20* led to the synergistic re-expression of *TMS1*, that was both greater than achieved by treating with 5-azaCdR alone and comparable to MCF7 cells where *TMS1* is unmethylated (**Figure 1, 3A)**. The inhibition of *Suv4-20* and 5-azaCdR also resulted in the re-positioning of nucleosomes in a pattern similar to that seen in MCF7 cells (Chapter 2, **Figure 3C**). This co-treatment did not result in additional DNA de-methylation (data not shown). These cells were maintained in culture for 6 passages after the removal of 5-azaCdR and shRNA against *Suv4-20*. By three passages after the removal of both treatments, *TMS1* exhibited a complete loss of H4K16ac along with a complete return of H4K20me3, suggesting that alterations induced from the treatment with shRNA against *Suv4-20* are transient, while the alterations accompanying DNA de-methylation are more permanent (Data not shown).

Co-existing methylated and demethylated TMS1 alleles are packaged with differentially modified histones

Previous work has suggested that treatment with 5-aza-CdR induces a "bivalent" chromatin signature at epigenetically reactivated tumor-suppressor genes (84, 159), implying that both active (*e.g.* H3K4me2) and repressive (*e.g.* H3K27me3) modifications co-exist on the same nucleosome. So far, these studies have examined histone modifications at a population level within a mixed cell sample. We similarly show above that at a population level, both active and repressive epigenetic marks were present at the *TMS1* CpG island after treatment with 5-azaCdR. However, when we analyzed individual DNA alleles two distinct subpopulations with different DNA methylation patterns were observed. To address the relationship between DNA methylation and histone modifications at the single molecule level, we employed a novel approach, ChIP-bisulfite sequencing (ChIP-bis), in which DNA was immunoprecipitated with antibodies specific to Pol II or the various histone modifications was eluted, bisulfite-modified and individual alleles were analyzed for their methylation status by bisulfite sequencing. As a control, DNA isolated from fixed and sheared total chromatin in the absence of antibody was also analyzed. Consistent with naked DNA (**Figure 2C**), DNA isolated from total chromatin directly after 5-azaCdR treatment showed a mixed DNA methylation pattern consisting of ~60% overall methylation (**Figure 4A**). ChIP-bis showed that H3Ac, H3K4me2, and Pol II were selectively associated with the unmethylated alleles in this population, as indicated by the relative enrichment of unmethylated DNA (decreased methylation density) in the immunoprecipitated DNA relative to input chromatin (**Figure 4A**). In contrast, H3K9me2 was selectively associated with the methylated alleles immediately following treatment with 5-azaCdR (84% methylation density compared to 61% in input) (**Figure 4B**). The distribution of alleles associated with H4K20me3 had a

methylation profile similar to that of input DNA (67% methylation density compared to 61%), suggesting that the nucleosomes marked by this modification associate with both methylated and unmethylated DNA, consistent with our observation that total levels of H4K20me3 levels at *TMS1* were unaltered by treatment with 5-azaCdR (**Figure 2**). Overall, these data suggest that treatment with 5-azaCdR leads to the coordinated demethylation of DNA and loss of H3K9me2, and that the chromatin associated with these demethylated alleles was selectively marked by H3Ac and H3K4me2, and bound by Pol II.

To investigate the stability of the associated marks after the removal of 5-azaCdR, we next performed ChIP-bis in cells 27 passages after drug removal**.** The DNA associated with total chromatin at this point had undergone a partial remethylation (from 61% to 79% methylation), similar to that observed for naked DNA analyzed by genomic bisulfite sequencing (compare **Figures 1D** and **4A**). After 27 passages, H3K4me2-modified histones and Pol II remained predominately associated with the unmethylated subpopulation of alleles (21% and 11% methylation density, respectively; **Figure 4B**). However, H3Ac, which selectively associated with the unmethylated alleles immediately after treatment, was now associated with both methylated and unmethylated alleles (from 12% to 68% methylation density), suggesting that the presence of H3Ac does not prevent the remethylation of *TMS1* alleles. H3K9me2 marked histones, enriched on the methylated alleles immediately after treatment, also showed no selective association with unmethylated or methylated *TMS1* alleles such that the ratio of unmethylated to methylated alleles associated with H3K9me2 was similar to that of input chromatin (76% total methylation compared with 79%). These data suggest that whereas H3K9me2 was

depleted from demethylated alleles immediately following drug removal, it re-associated with the unmethylated alleles in the absence of drug independent of DNA remethylation. H4K20me3, on the other hand, remained associated with both unmethylated and methylated alleles at a ratio similar to input both immediately after 5-azaCdR and 27 passages later. Thus, following the removal of 5-azaCdR in culture, we observe the preservation of a subset of unmethylated alleles uniquely marked by both active and repressive histone marks.

ESR1 and CDH1 are completely remethylated and silenced following transient exposure to 5-azaCdR

A comparison of the data presented above to that of other single gene studies examining the remethylation kinetics following 5-azaCdR treatment suggest that there may be gene specific variations in the remethylation kinetics following the removal of 5-azaCdR between different genes (70, 84, 160). This may reflect intrinsic differences in their underlying epigenetic regulation or could be a consequence of differences between cell lines and treatment protocols. To address this question, we monitored two additional tumor suppressor genes, *CDH1* and *ESR1* (**Figures 5A, 6A**), both of which are CpG island-associated genes that are aberrantly methylated and silent in MDA-MB231 cells (161). Like *TMS1*, the expression of *CDH1* and *ESR1* was substantially induced after 5-azaCdR treatment (**Figures 5B, 6B**). In contrast to *TMS1*, which maintained a low yet stable expression level several fold above untreated cells after drug removal, both *CDH1* and *ESR1* were completely re-silenced within three passages (~9 days) (**Figures 5B, 6B**). Analysis of a single CpG site via COBRA analysis revealed that *CDH1* and *ESR1*

exhibited similar remethylation kinetics to *TMS1* (**Figures 1C, 5C** and **6C**), in that all three genes were initially demethylated by 40-50% during treatment and then ultimately remethylated to 82-85%. However, a more detailed methylation analysis by genomic bisulfite sequencing revealed distinctions in the patterns of their remethylation. After 27 passages in the absence of 5-azaCdR, both *CDH1* and *ESR1* exhibited more uniform patterns of methylation, in that nearly all alleles were exhibited some level of remethylation (**Figures 6D, 7D**). Unlike *TMS1,* neither locus retained a subset of predominantly unmethylated alleles after 5-azaCdR treatment.

We also examined alterations in the histone profiles at the *CDH1* and *ESR1* CpG islands throughout the 5-azaCdR time course. As observed at *TMS1*, 5-azaCdR-induced reactivation of *CDH1* and *ESR1* was associated with the accrual of H3ac, H3K4me2, the loss of H3K9me2 and the re-establishment of Pol II occupancy (**Figures 5E** and **6E**). H4K20me3 was not present at significant levels at either gene in untreated MDA-MB231 cells (data not shown) (103) and was not further analyzed. Although 5-azaCdR treatment induced a similar set of epigenetic changes at all three genes, there were differences in the ability to maintain the induced histone profiles after drug removal. At *CDH1*, low levels of H3K4me2 and H3ac were maintained throughout the time course, while these marks were completely depleted from *ESR1* by 27 passages after 5-azaCdR removal (**Figures 5E** and **6E**). Somewhat surprisingly, Pol II was still present at both *CDH1* and *ESR1* 6 passages after drug removal, despite the lack of detectable gene expression. Ultimately, however, both *CDH1* and *ESR1* exhibited a complete loss of Pol II occupancy between 6 and 27 passages after drug removal (**Figures 5E**, **6E**). These differences in maintained histone profile may reflect the subtle differences in the extent

of remethylation of individual alleles observed in genomic bisulfite sequencing (**Figures 5D**, **6D**). Similar results were obtained from a second independent time course (**Supplementary Figures 4, 5**). Thus, although a similar subset of epigenetic alterations are induced by treatment with 5-azaCdR at *TMS1, CDH1* and *ESR1,* the inability to maintain Pol II occupancy may ultimately allow for the remethylation of DNA and the return to stable repression following the removal of drug.

Discussion

Previously, we have shown that the *TMS1* CpG island is aberrantly methylated and silenced in MDA-MB231 cells. This silencing is accompanied with packing into histones that display the repressive histone modifications H3K9me2 and H4K20me3 (**Figure 7**). Here we have shown that the treatment of MDA-MB231 cells with 5-azaCdR results in a spectrum of epigenetic alterations at the *TMS1* locus including DNA demethylation, gene re-expression, the acquisition of H3ac, H3K4me2, and the depletion of H3K9me2. This allows for the re-engagement of RNA Pol II on the TMS1 promoter. Using a novel approach, ChIP-bis, we show that these acquired active marks (H3ac, H3K4me2, and Pol II) preferentially associate with de-methylated alleles and that H3K9me2 was enriched for those alleles that remained methylated (modeled in **Figure 7).** H4K20me3 was unaltered by 5-azaCdR treatment and therefore was found on both subpopulations of alleles. When cells were co-treated with 5-azaCdR and shRNA against *Suv4-20* we find a synergistic re-expression of *TMS1* that is accompanied by the loss of H4K20me3 along with the re-association of H4K16ac and the re-positioning of nucleosomes. However, the alterations induced by removing *Suv4-20* were entirely

transient. Thus we only focused on *TMS1* regulation following 5-azaCdR treatment. Maintenance of these cells in the absence of 5-azaCdR allowed us to investigate the durability of these induced alterations. After 27 passages in the absence of drug, a subpopulation of unmethylated alleles remained associated with the active marks H3K4me2, H3ac, and Pol II, as well as the repressive mark H4K20me3. Interestingly, H3K9me2, which was preferentially associated with methylated alleles immediately following treatment, was able to re-associate with the unmethylated alleles over time. Furthermore, H3ac, which exclusively associated with the unmethylated alleles at the time of treatment, was now bound by both subpopulations of alleles. This suggests that any *TMS1* alleles that remethylated after drug removal lost both H3K4me2 and Pol II occupancy while maintaining H3ac. Taken together these data demonstrate that following the removal of 5-azaCdR, the *TMS1* CpG island is maintained in a unique epigenetic domain, consisting of two distinct subpopulations of alleles neither of which fully resembles either the repressed state in untreated MDA-MB231 cells, or that induced immediately after 5-azaCdR treatment.

Consistent with previous studies (103, 128, 158), we found that H3K9me2, H3K4me2 and DNA methylation status are tightly linked at the *TMS1* locus (as well as at *CDH1* and *ESR1*) (70, 84, 85). Our single molecule approach confirms that H3K9me2 is selectively depleted from the unmethylated alleles (*i.e.* it remains selectively associated with methylated alleles), suggesting that DNA methylation is necessary to maintain H3K9me2. Recent work demonstrates that the histone methyltransferase G9a, thought to catalyze most H3K9me2 in euchromatin, interacts with components of the DNMT1 complex (38, 72, 73). This points to a model in which DNA methylation and H3K9me2

may be coordinately maintained during DNA replication. Given that the incorporation of 5-azaCdR into DNA precipitates the degradation of DNMT1 (162), it is possible that the depletion of H3K9me2 is an indirect consequence of loss of DNMT1. After removal of 5-azaCdR, H3K9me2 showed an increased association with unmethylated *TMS1* alleles. These data suggest that H3K9me2 can re-associate with the unmethylated subpopulation of alleles in the absence of *TMS1* remethylation and may precede their re-methylation, as has been suggested by work in DNMT knockout cells (163). This re-association of H3K9me2 with the unmethylated alleles may account for the partial transcriptional re-silencing of *TMS1* after removal of 5-azaCdR.

After the removal of 5-azaCdR, we find H3ac to be associated with both methylated and unmethylated *TMS1* alleles. Studies on *hTERT* gene similarly found H3ac to associate with both methylated and unmethylated alleles (158). Our data suggest that drug-induced DNA demethylation is necessary to re-establish H3ac at *TMS1*, but once established, the presence of H3ac does not prevent the remethylation of DNA. Nor does DNA remethylation drive the de-acetylation of H3. Thus, although the combination of DNMT and HDAC inhibitors results in the synergistic re-activation of epigenetically silenced genes (89), they may have little impact on DNA remethylation and ultimately gene re-silencing after the cessation of treatment. Indeed, treatment with an HDAC inhibitor does not prevent the remethylation of the *p16* gene after 5-azaCdR-induced DNA demethylation (164).

Unlike many densely methylated genes reactivated by 5-azaCdR treatment which have a propensity to undergo remethylation and resilencing after drug removal (85, 164), we found that a subset (~20%) of *TMS1* molecules are maintained in an unmethylated

state at the *TMS1* locus for more than 80 days in culture in the absence of drug. These alleles were selectively occupied by RNA Pol II and marked by H3K4me2, and likely account for the residual low-level gene expression observed. In contrast, *ESR1* and *CDH1* were completely resilenced and all alleles remethylated over the same time frame. Studies on the *MLH1* gene in colon cancer cells have similarly shown the maintenance of a stable subpopulation of unmethylated alleles (~7%) for at least 44 days after 5-aza-CdR-induced reactivation (70). This study showed that 5-aza-CdR induced DNA demethylation was associated with the eviction of a single nucleosome near the transcription start site of *MLH1* and that the maintenance of this pattern of nucleosome depletion was both heritable and selective for the stably unmethylated subset of alleles (70). Although we did not examine nucleosome occupancy directly in this study, we have previously shown that the *TMS1* CpG island is spanned by positioned nucleosomes flanking a single nucleosome gap at the transcription start site in its unmethylated and actively transcribed state (103). We propose that this 5-azaCdR-induced nucleosome eviction allows for the re-engagement of the Pol II complex with the demethylated promoter, perhaps filling the gap left behind by the evicted nucleosome, and that it is the maintenance of Pol II occupancy and/or the associated H3K4me2 deposition that prevents DNA remethylation. Consistent with this idea, recent studies examining histone methylation at 5-azaCdR-induced genes across the genome showed that reactivation of methylated genes was associated with accumulation of H3K4me2 distributed to either side of a characteristic "dip'' centered over transcription start (159). *CDH1* and *ESR1*, which exhibited total allelic remethylation after drug removal, also failed to maintain significant Pol II occupancy. The idea that Pol II occupancy may protect CpG islands

from *de novo* methylation is further supported by recent work indicating that Pol II occupancy at CpG islands is better correlated with resistance to aberrant DNA methylation in cancer cells than is gene expression (165).

The fact that many genes appear to re-silence to some extent after the removal of 5-azaCdR may be a direct result of the underlying chromatin structure. Whereas some repressive marks like H3K9me2 are reversed by 5-azaCdR-induced demethylation, others (H3K9me3, H3K27me3) are not affected, and can be maintained in the absence of DNA methylation (84, 164). It has been suggested that the persistence of these marks may precipitate re-silencing after drug removal. The presence of H3K27me3 predisposes certain genes to *de novo* methylation in cancer although the molecular mechanisms precipitating this event remain unclear (64-67). We show here that H4K20me3 is similarly unaffected by alterations in DNA methylation and remains associated with both unmethylated and methylated alleles after the initial treatment with 5-azaCdR and throughout the course of DNA remethylation. These data suggest that there is a mechanism to maintain H4K20me3 at *TMS1* that is independent of DNA methylation. Furthermore, we find that the presence of H4K20me3 at *TMS1* is inhibitory to H4K16ac occupancy and optimal *TMS1* expression. This provides support for that H4K20me3 highlights an additional repressive pathway at *TMS1*, one that is independent of DNA methylation. At present the mechanisms targeting H4K20me3 to specific loci are unclear. Enzymatically, methylation of H4K20 proceeds in a step-wise fashion with PR-SET7 catalyzing the mono-methylation at this position, which is then acted on by SUV4-20h1/2 to catalyze di- and tri-methylation (166-169). As has been suggested for H3K27me3, the residual presence of H4K20me3 may facilitate the partial re-silencing at *TMS1*. Whether

the continued long-term inhibition of histone methyltransferases might facilitate the maintenance of tumor suppressor gene activity after demethylation-induced reactivation remains an open question, and warrants consideration of combination therapies and dosing schedules that incorporate DNA methylation inhibitors and histone methyltransferase inhibitors.

The DNA methyltransferase inhibitors are now in widespread use for the treatment of MDS, and are currently in clinical trials for AML and other solid tumor types (86). Molecular analyses of bone marrow biopsies from patients treated with these agents have demonstrated that global and gene-specific DNA demethylation is achievable *in vivo* (81, 82). The degree of demethylation varies between patients, and whether this is an important indicator of clinical response remains controversial and may depend upon the compartment being analyzed (*i.e*. repetitive element methylation versus methylated tumor suppressor genes) or the surrogate marker measured (DNA methylation versus gene expression) (76, 83). In most cases, a gradual return to pre-treatment methylation levels was observed within a few weeks or by the start of the next treatment cycle (81, 83, 170). We show here that, at least *in vitro*, the kinetics of DNA remethylation and gene silencing do not necessarily parallel each other, and vary at different loci. Whereas *CDH1* and *ESR1* were resilenced within a week after drug removal and before the complete remethylation of H3K9me2 and DNA, a subset of *TMS1* alleles remained stably unmethylated and occupied by Pol II. The remethylation potential of a particular gene may be determined in part by the histone code present at that gene prior to demethylation. A thorough understanding of the underlying causes of these differences would be

valuable in both developing new combined approaches incorporating inhibitors of other histone modifying enzymes and improving existing epigenetic therapy regimens.

Figures

Figure 1.

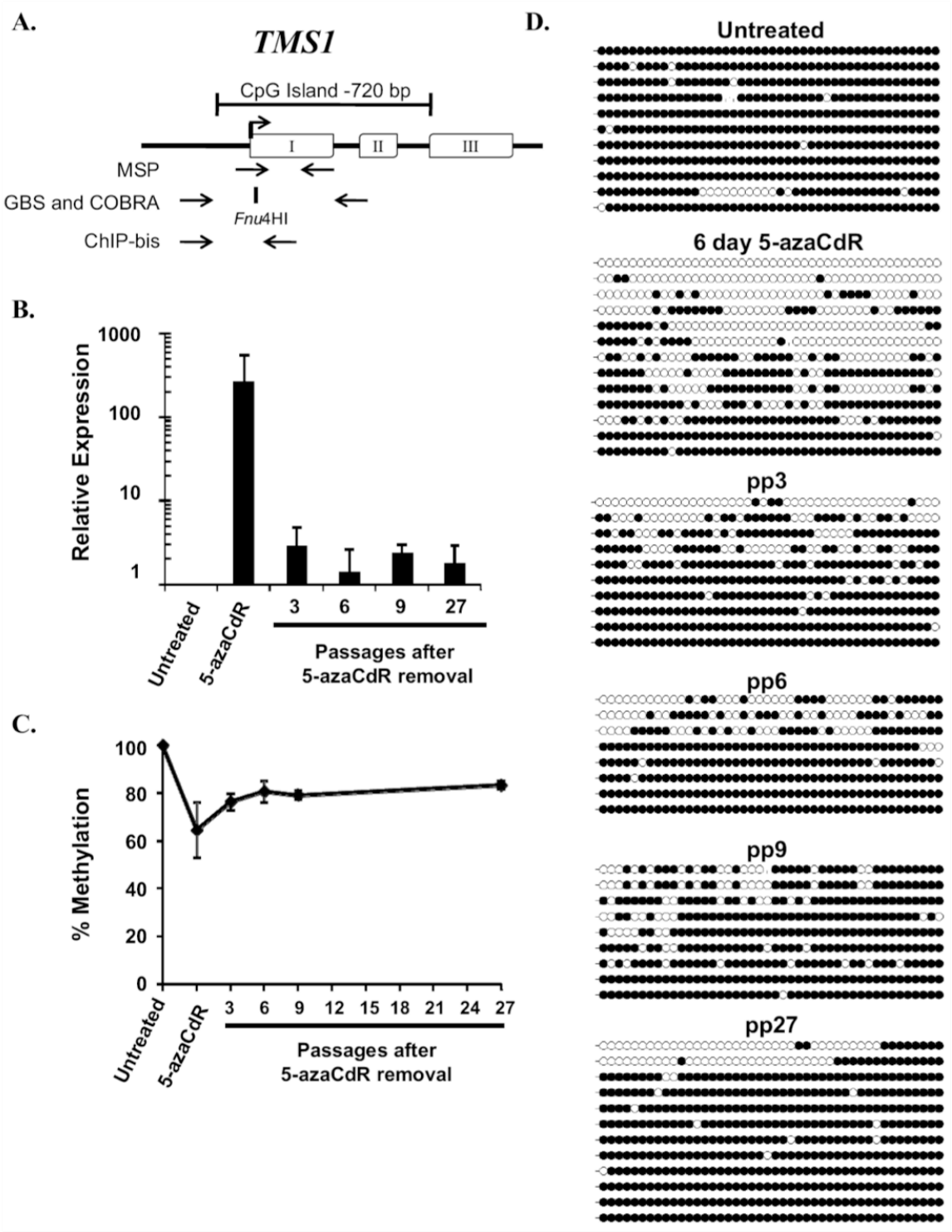


Figure 1: *TMS1* expression and DNA methylation following the removal of 5-azaCdR. A. Diagram of the *TMS1* gene. Exons, open boxes; transcription start site, arrow. The CpG island is indicated by the bracket. Primers utilized for MSP, GBS, and COBRA analyses are indicated below the gene, as is the position of the *Fnu*4HI restriction site used in COBRA analysis. **B.** *TMS1* mRNA abundance was measured immediately after treatment (5-azaCdR) or at the indicated timepoints after drug removal (pp, passages post 5-azaCdR) by reverse transcriptase qPCR and normalized to 18S rRNA. Shown is the fold change in expression (mean ± standard deviation) relative to untreated cells from three independent time-course experiments assayed in triplicate. **C.** COBRA analysis of DNA methylation following the removal of 5-azaCdR at the *TMS1* locus. Data represents the mean percent methylation (± standard deviation) from three independent time course experiments **D.** DNA methylation was analyzed by bisulfite sequencing at the indicated time points. Each line represents a single colony isolate (8-13 isolated per sample). Open circles, unmethylated CpG; filled circles, methylated CpG.

Figure 2.

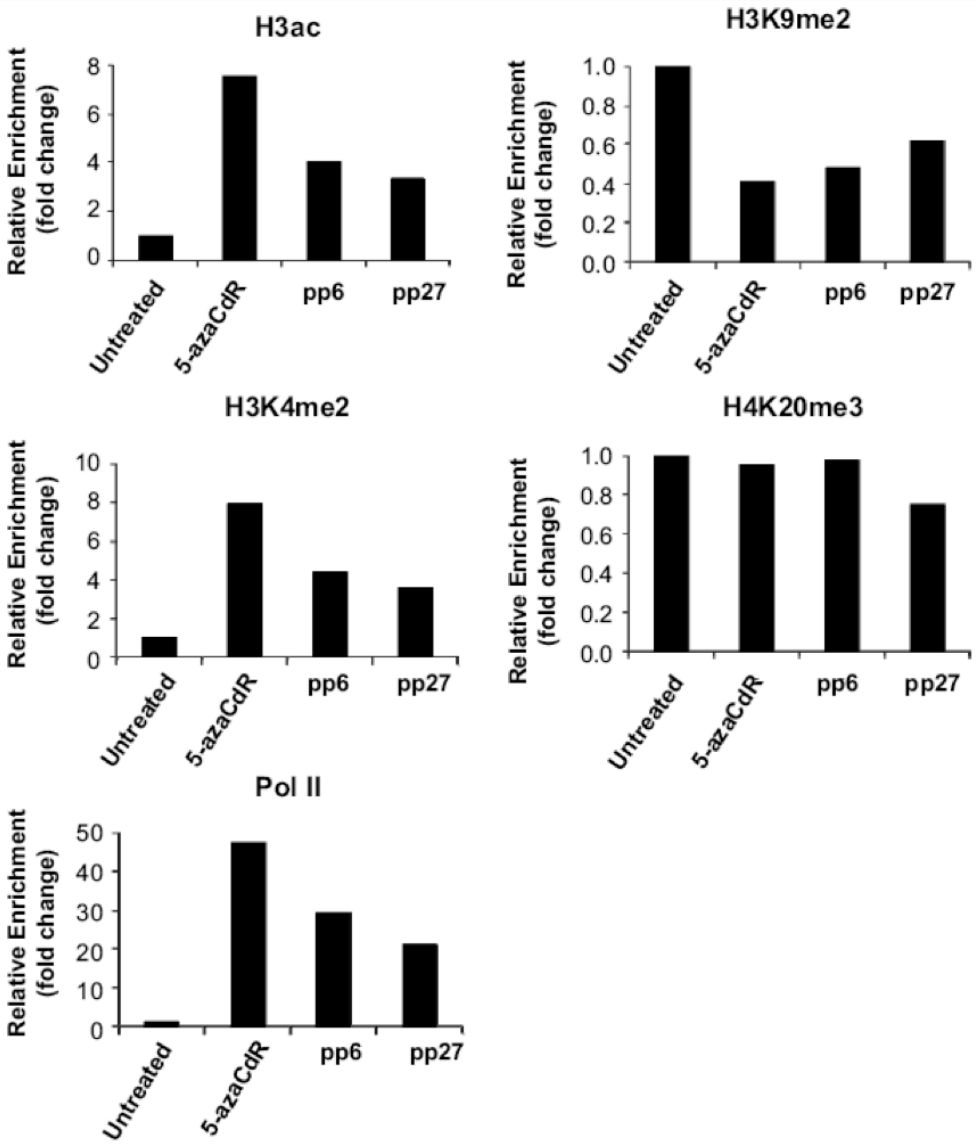


Figure 2: Histone modifications and RNA Pol II occupancy at *TMS1* after the removal of 5-azaCdR. MDA-MB231 cells were left untreated or treated with 0.5uM 5-azaCdR for six days. Chromatin was isolated immediately after treatment (5-azaCdR) or at the indicated time after drug removal (pp, passages post 5-azaCdR). Histone modifications and RNA Pol II occupancy were analyzed by ChIP followed by qPCR. Percent enrichment was determined in triplicate by comparison of immunoprecipitated DNA relative to input DNA at each time point using primer set 3 of the TMS1 locus (103). Plotted is the fold change in enrichment relative to untreated MDA-MB231 cells from a single time course experiment. Similar results were obtained from a second independent time course (**Supplementary Figure 2**)

Figure 3.

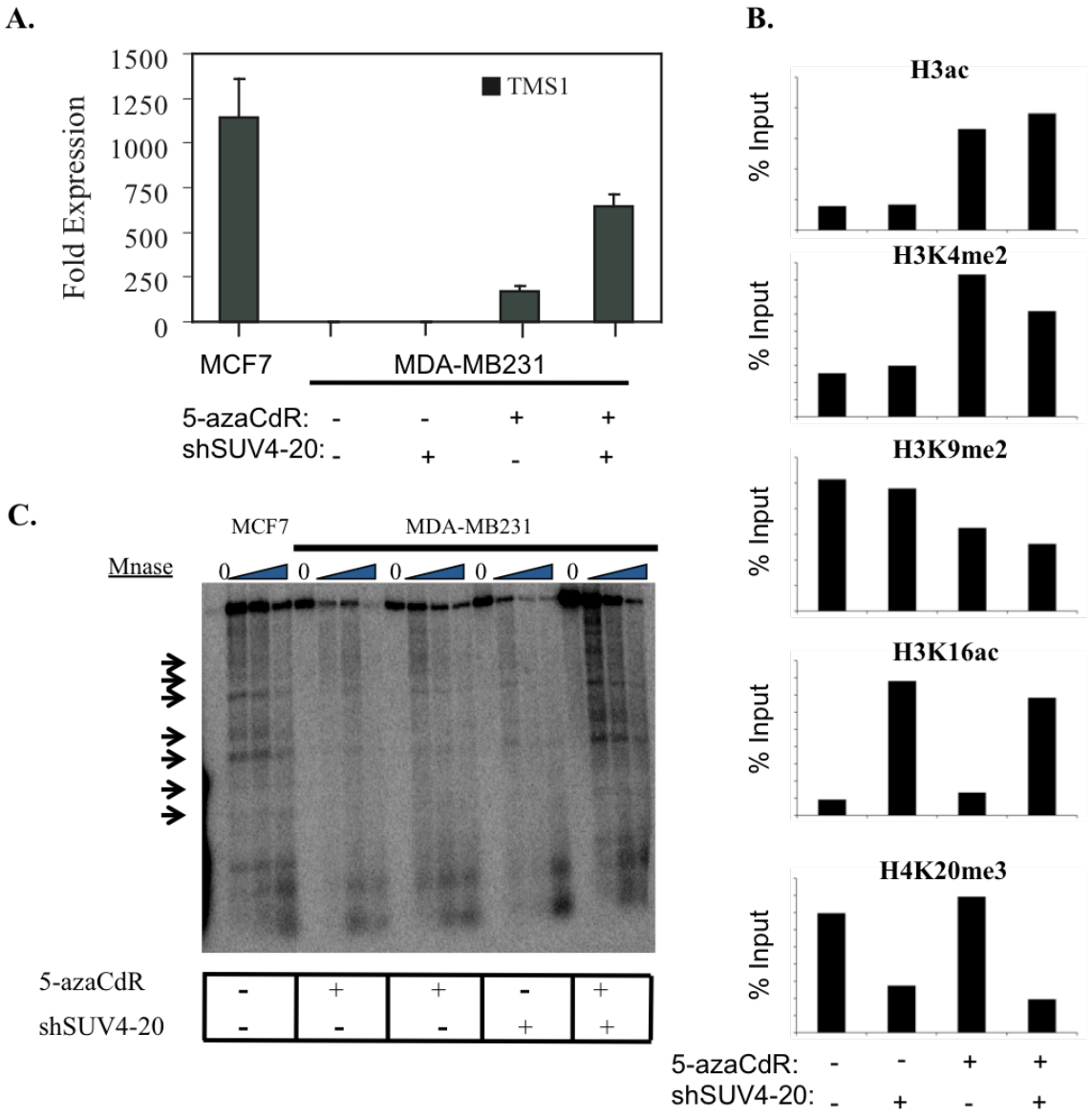

A.
B.
C.
Fold Expression
1500
1250
1000
750
500
250
0
TMS1
MCF7
MDA-MB231
5-azaCdR: - - + +
shSUV4-20: - + - +
H3ac
H3K4me2
H3K9me2
H3K16ac
H4K20me3
% Input
Mnase
0 0 0 0 0
5-azaCdR
shSUV4-20

Figure 3: Treatment with 5-azaCdR and shRNA against *Suv4-20* synergistically reactivates *TMS1* expression. A. *TMS1* mRNA abundance was measured immediately after treatment 5-azaCdR or shRNA targeting *Suv4-20* by reverse transcriptase qPCR and normalized to 18S rRNA. MCF7 is shown as an unmethylated control. Shown are the average of three independent experiments. **B.** Histone modifications were analyzed by ChIP followed by qPCR. Percent enrichment was determined by comparison of immunoprecipitated DNA relative to input DNA at each time point using primer set 3 of the TMS1 locus shown. A representative experiment, similar resulted were obtained from three independent experiments. **C.** Positioned nucleosomes in MDA-MB231 cells untreated, treated with 5-azaCdR, treated with shRNA against *Suv4-20*, or co-treated with both were mapped by indirect end labeling. Preferential Mnase cut sites are marked with arrows. MCF7 cells shown as an unmethylated and expressing control.

Figure 4.

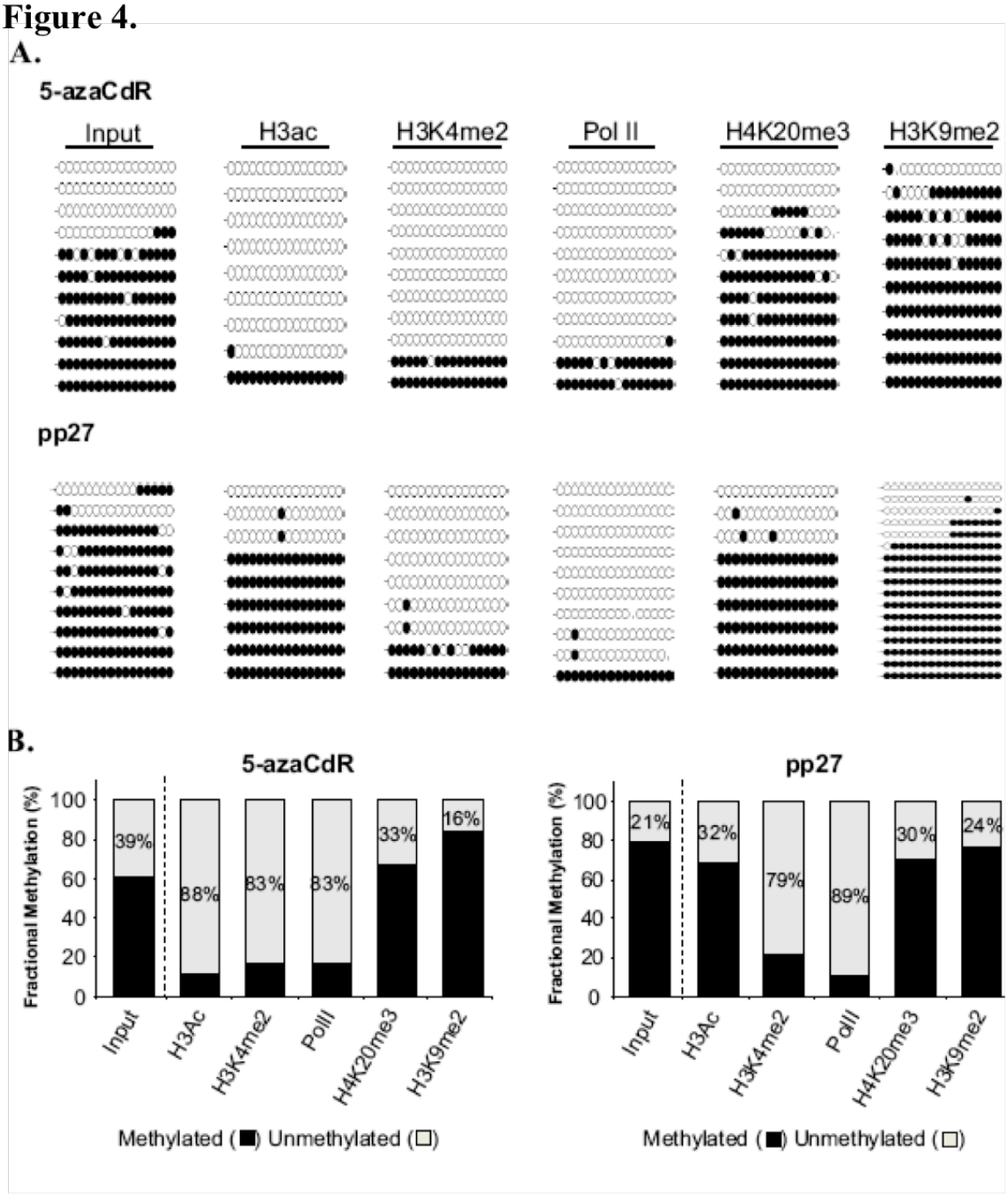

A.
5-azaCdR
Input
H3ac
H3K4me2
Pol II
H4K20me3
H3K9me2
pp27
B.
5-azaCdR
pp27
Fractional Methylation (%)
0
20
40
60
80
100
39%
88%
83%
83%
33%
16%
21%
32%
79%
89%
30%
24%
Input
H3Ac
H3K4me2
PolII
H4K20me3
H3K9me2
Methylated (■) Unmethylated (□)

Figure 4: Single molecule association between DNA methylation and chromatin modifications at the *TMS1* locus. MDA-MB231 cells were left untreated or treated with 0.5uM 5-azaCdR for six days (5-azaCdR) **A.** Chromatin was isolated immediately after (5-azaCdR) or at the indicated time after drug removal and immunoprecipitated with the indicated antibodies. Precipitated DNA was eluted, bisulfite modified, and amplified with the bisulfite sequencing primers indicated in **Figure 1A**. For each immunoprecipitation, 9-17 individual clones were sequenced. Open circles, unmethylated CpGs; filled circles, methylated CpGs. Vertical hash marks represents missing data. **B.** Overall methylation density was determined as the total number of methylated CpGs relative to total number of CpGs in all alleles analyzed.

Figure 5.

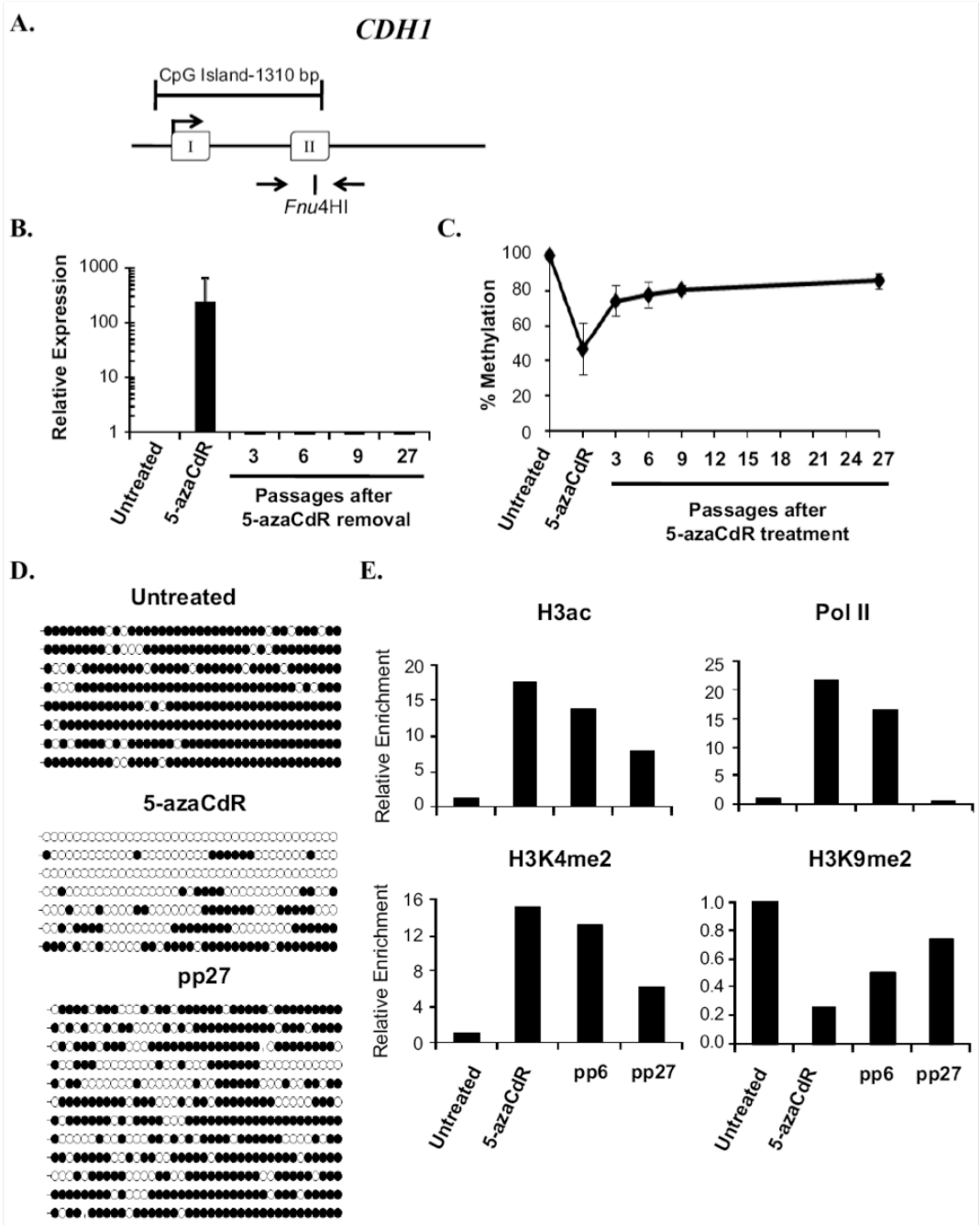

A.
CDH1
CpG Island-1310 bp
I
II
Fnu4HI
B.
Relative Expression
1000
100
10
1
Untreated
5-azaCdR
3
6
9
27
Passages after
5-azaCdR removal
C.
% Methylation
100
80
60
40
20
0
Untreated
5-azaCdR
3 6 9 12 15 18 21 24 27
Passages after
5-azaCdR treatment
D.
Untreated
5-azaCdR
pp27
E.
H3ac
Pol II
Relative Enrichment
H3K4me2
H3K9me2
Untreated
5-azaCdR
pp6
pp27

Figure 5: Analysis of *CDH1* during and following treatment with 5-azaCdR. A. Diagram of *CDH1* CpG island. Open boxes, exons; arrow, transcription start site. The position of primers used for methylation analyses and the *Fnu*4H1 restriction enzyme site used for COBRA analysis are indicated. **B.** *CDH1* mRNA expression was determined by reverse transcriptase qPCR and was normalized to 18S rRNA. Shown is the fold change in expression (mean ± standard deviation) relative to untreated cells from three independent experiments assayed in triplicate. **C.** COBRA analysis of DNA methylation following the removal of 5-azaCdR at the *CDH1* locus. Data represents the mean percent methylation (± standard deviation) from three independent time course experiments. **D.** DNA methylation was further analyzed by bisulfite sequencing at the indicated time points. Each line represents a single colony isolate (8-11 isolated per sample). Open circles, unmethylated CpG; filled circles, methylated CpG. pp, passages post 5-azaCdR treatment. **E.** Histone modifications and RNA Pol II occupancy were determined by ChIP as described in the legend to Figure 2. Plotted is the fold change in enrichment relative to untreated MDA-MB231 cells from a representative time course experiment. Similar results were obtained from a second independent time course (**Supplementary Figure 3**).

Figure 6.

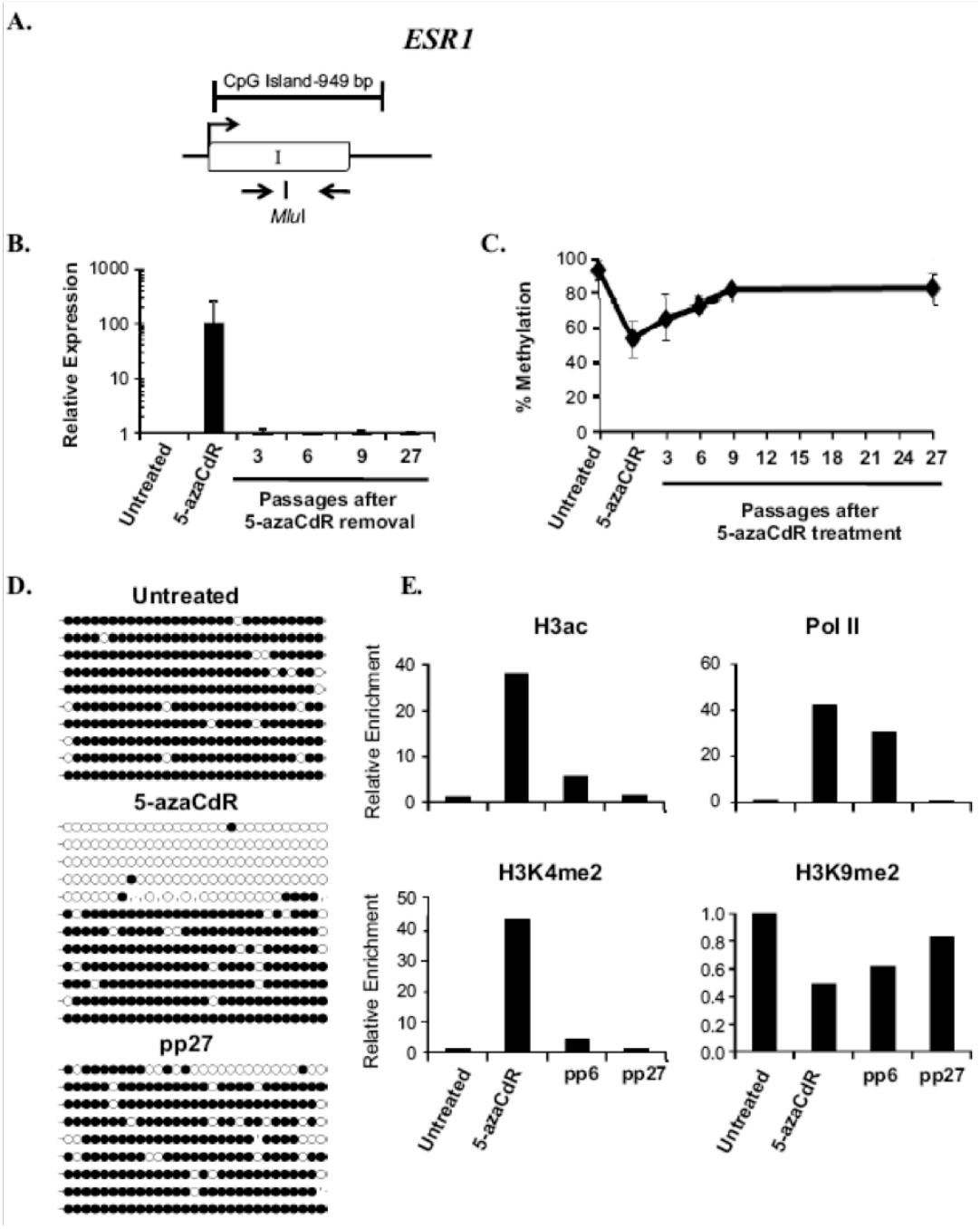

A.
ESR1
CpG Island-949 bp
I
MluI
B.
Relative Expression
1000
100
10
1
Untreated
5-azaCdR
3
6
9
27
Passages after
5-azaCdR removal
C.
% Methylation
100
80
60
40
20
0
Untreated
5-azaCdR
3 6 9 12 15 18 21 24 27
Passages after
5-azaCdR treatment
D.
Untreated
5-azaCdR
pp27
E.
H3ac
Pol II
H3K4me2
H3K9me2
Relative Enrichment
Untreated
5-azaCdR
pp6
pp27

Figure 6: Analysis of *ESR1* during and following treatment with 5-azaCdR. A Diagram of *ESR1* CpG island. Open boxes, exons; arrow, transcription start site. The position of primers used for methylation analyses and the *Mlu1* restriction enzyme site used for COBRA analysis are indicated. **B.** mRNA expression was determined using primers specific to *ESR1*. Shown is the fold change in expression (mean ± standard deviation) relative to untreated cells after internal normalization to 18S rRNA from three independent experiments assayed in triplicate. **C.** COBRA analysis of DNA methylation at time points following the removal of 5-azaCdR at the *ESR1* locus as described in the legend to Figure 1 **D.** DNA methylation was analyzed by bisulfite sequencing at the indicated time points. Each line represents a single colony isolate (9-12 isolated per sample). Open circles, unmethylated CpG; filled circles, methylated CpG. pp, passages post 5-azaCdR treatment. **E.** Histone modifications and RNA Pol II occupancy were determined by ChIP as described in the legend to Figure 2. Plotted is the fold change in enrichment relative to untreated MDA-MB231 cells from a representative time course experiment. Similar results were obtained from a second independent time course (**Supplementary Figure 4**).

Figure 7.

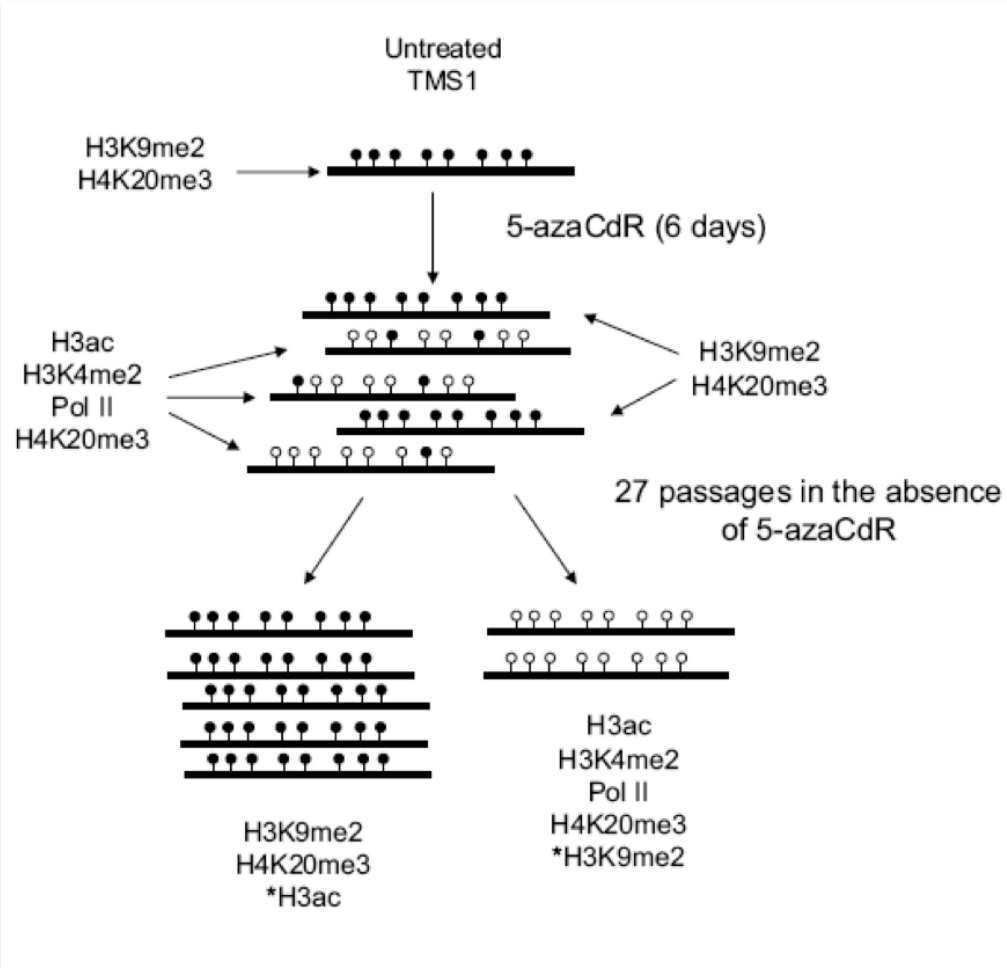
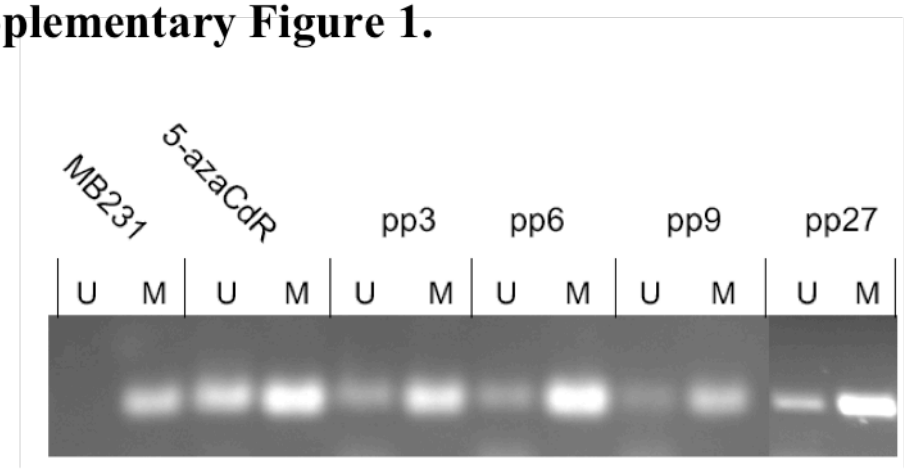

Untreated
TMS1
H3K9me2
H4K20me3
5-azaCdR (6 days)
H3ac
H3K4me2
Pol II
H4K20me3
H3K9me2
H4K20me3
27 passages in the absence
of 5-azaCdR
H3K9me2
H4K20me3
*H3ac
H3ac
H3K4me2
Pol II
H4K20me3
*H3K9me2

Figure 7: Model of the epigenetic regulation at *TMS1* during 5-azaCdR time course. *TMS1* is densely methylated in MDA-MB231 cells (filled circles) and marked by H3K9me2 and H4K20me3. Six days of treatment with 5-azaCdR resulted in the demethylation (open circles) of a subpopulation of *TMS1* alleles. These alleles are wrapped into nucleosomes that are depleted of H3K9me2 and gain H3ac, and H3K4me2, but retain H4K20me3. These alleles are selectively re-engaged by RNA Pol II. After 3 months in the absence of drug, a subpopulation of *TMS1* alleles remained stably unmethylated and bound by H3ac, H3K4me2, and Pol II as well as by the repressive marks H4K20me3 and gained H3K9me2. The subset of alleles that remethylated in this time period appears to retain H3ac.

Supplementary Table I

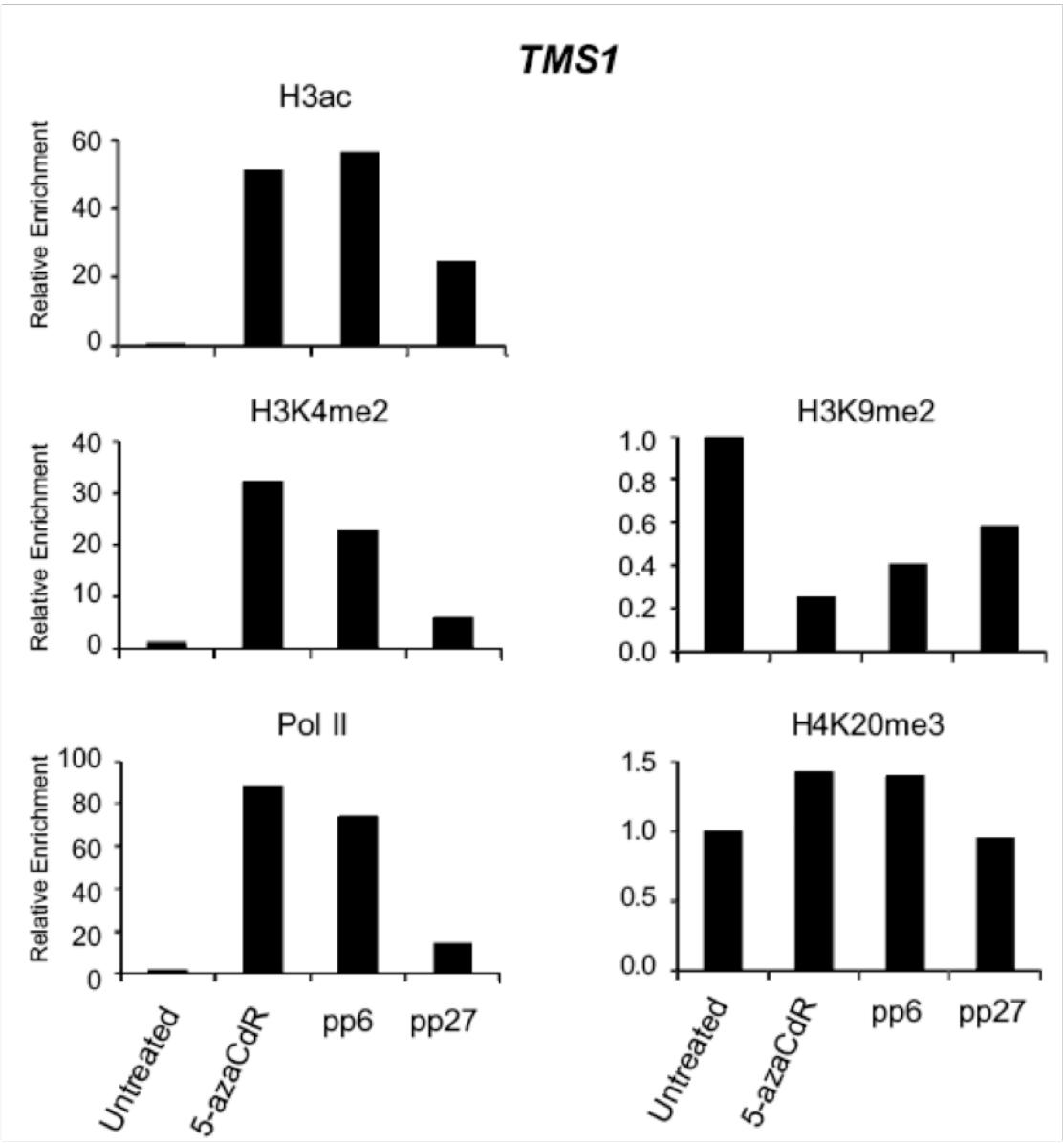
Primer	Sequence (5' to 3')	Reference
TMS1 cDNA	S - TGGGCCTGCAGGAGATG AS -ATTTGGTGGGATTGCCAG	(Stimson and Vertino, 2002)
TMS1 MSP	**Methylated** S - TTGTAGCGGGGTGAGCGGC AS - TTCGTCCATAAACAACAACGCG **Unmethylated** S - GGTTGTAGTGGGGTGAGTGGT AS -	(Levine et al., 2003)
TMS1 GBS/COBRA	S – TTGGTGTAAGTTTAGAGATAAGT AS - ACCATCTCCTACAAACCCATA	(Stimson and Vertino, 2002)
TMS1 ChIP	S – GCTTCTAGCTGTCATGATTCC AS - CTTGCACCAGCGGGTACAGA	(Kapoor-Vazirani et al., 2008)
TMS1 ChIP-bis	S – GGTTTGGGGTTTTAATTTAGAG AS – CCCCATAACTCCAAAATCC	This manuscript
CDH1 cDNA	S – GAACGCATTGCCACATACAC AS – AGCACCTTCCATGACAGACC	(Graff et al., 1997)
CDH1 GBS/COBRA	S – GAGGGAAGGAGAGGGGTATT AS – CCCTCACCTCTACCCAAAAC	This manuscript
CDH1 ChIP	S – GACGATCTTCGAGGGAAGG AS - CACTCCTATTTACGACCTTTC	(Kapoor-Vazirani et al., 2008)
ESR1 cDNA	S – GAACGCATTGCCACATACAC AS – AGCACCTTCCATGACAGACC	(Lapidus et al., 1998)
ESR1 GBS/COBRA	S – AGGTGTATTTGGATAGTAGTAAG AS – CAAATAATAAAACACCTACTAACC	This manuscript
ESR1 ChIP	S – GCATCAGATCCAAGGGAACG AS - TTGCTGCTGTCCAGGTACAC	(Kapoor-Vazirani et al., 2008)

Supplementary Figure 1.



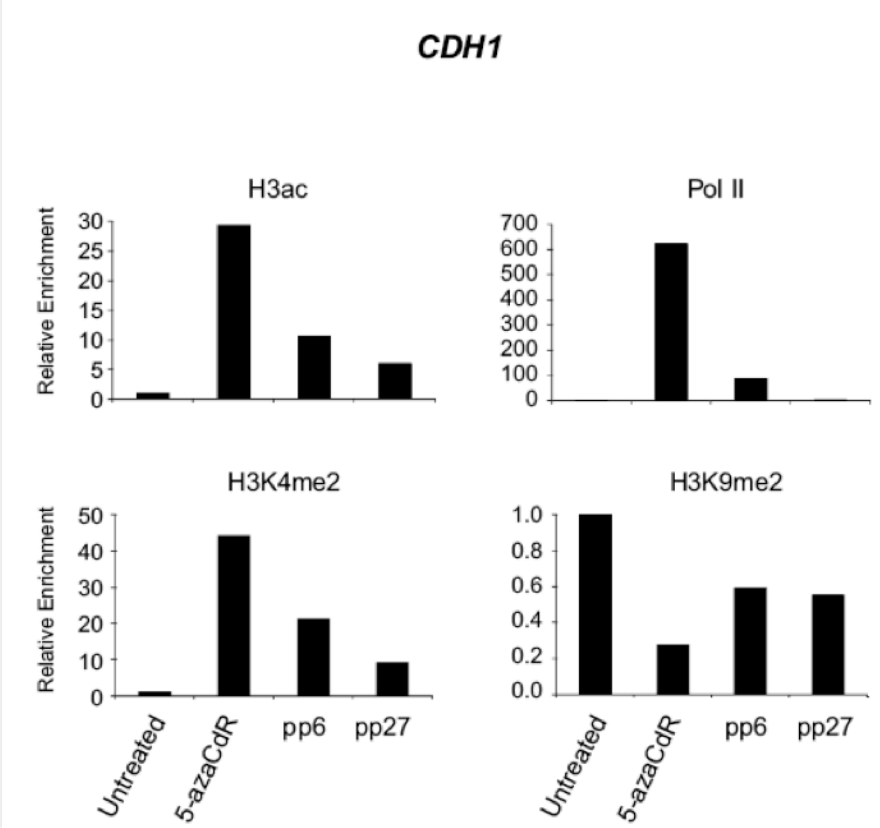
Supplementary Figure 1: **MSP analysis of *TMS1* during 5-azaCdR time course.** DNA methylation at the *TMS1* CpG island was analyzed by methylation –specific PCR. MDAMB231 cells were left untreated (MB231) or treated with 0.5uM 5-azaCdR for six days. DNA was isolated immediately after treatment (5-azaCdR) or at the indicated time after drug removal (pp, passages post 5-azaCdR). DNA was bisulfite modified and amplified with primers specifically recognizing either unmethylated (U) or methylated (M) DNA after bisulfite conversion. Primers utilized are listed in Supplementary Table I.

Supplementary Figure 2.



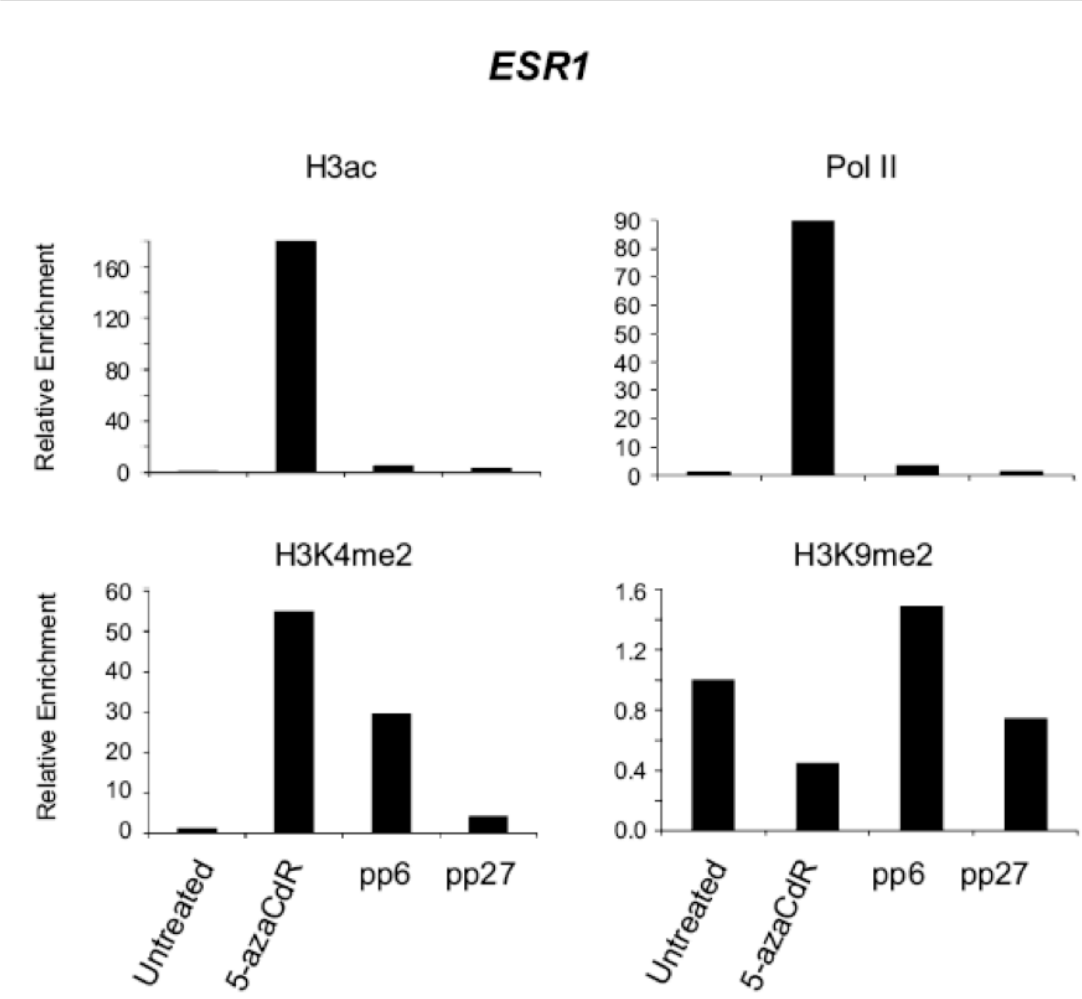
Supplementary Figure 2: **Histone modifications and RNA Pol II occupancy at *TMS1* after the removal of 5-azaCdR.** MDA-MB231 cells were left untreated or treated with 0.5uM 5-azaCdR for six days. Chromatin was isolated immediately after treatment (5-azaCdR) or at the indicated time after drug removal (pp, passages post 5-azaCdR). Histone modifications and RNA Pol II occupancy were analyzed by ChIP followed by qPCR. Percent enrichment was determined in triplicate by comparison of immunoprecipitated DNA relative to input DNA at each time point using primer set 3 of the TMS1 locus (Kapoor-Vazirani *et al.*, 2008). Plotted is the fold change in enrichment relative to untreated MDA-MB231 cells from a representative time course experiment.

Supplementary Figure 3.



Supplementary Figure 3: **Histone modifications and RNA Pol II occupancy at *CDH1* after the removal of 5-azaCdR.** MDA-MB231 cells were left untreated or treated with 0.5uM 5-azaCdR for six days. Chromatin was isolated immediately after treatment (5-azaCdR) or at the indicated time after drug removal (pp, passages post 5-azaCdR). Histone modifications and RNA Pol II occupancy were analyzed by ChIP followed by qPCR. Percent enrichment was determined in triplicate by comparison of immunoprecipitated DNA relative to input DNA at each time point using primers specific to the *CDH1* (Supplementary Table I). Plotted is the fold change in enrichment relative to untreated MDA-MB231 cells from a second time course experiment.

Supplementary Figure 4.



Supplementary Figure 4: **Histone modifications and RNA Pol II occupancy at *ESR1* after the removal of 5-azaCdR.** MDA-MB231 cells were left untreated or treated with 0.5uM 5-azaCdR for six days. Chromatin was isolated immediately after treatment (5-azaCdR) or at the indicated time after drug removal (pp, passages post 5-azaCdR). Histone modifications and RNA Pol II occupancy were analyzed by ChIP followed by qPCR. Percent enrichment was determined in triplicate by comparison of immunoprecipitated DNA relative to input DNA at each time point using primers specific to the ESR1 locus (Supplementary Table I). Plotted is the fold change in enrichment relative to untreated MDA-MB231 cells from a second time course experiment.

Chapter 5: Determinants of genome-wide remethylation kinetics after transient exposure to 5-azaCdR

Kagey, JD performed all cell culture and drug treatment experiments, molecular analyses, and assisted in data processing. Wrote the entire chapter.

McCabe, MT performed data processing and bioinformatics

This data is currently in preparation

The focuses of the work in this chapter was to build upon the findings of the previous chapter to define differences in DNA remethylation kinetics among individual CpG sites throughout the genome and begin to define factors responsible for creating differences in remethylation kinetics.

Introduction

Global alterations in the epigenome accompany tumorigenesis including genome-wide hypomethylation and the aberrant hypermethylation of a subset of CpG islands. This hypermethylation has been shown to be a mechanism of tumor suppressor gene (TSGs) silencing in cancer cells (8). Furthermore, the hypermethylation of CpG islands is accompanied by a shift in the histone profile at these loci from an active to a more repressive one (13, 24). Although epigenetically silenced, these genes retain the wild type DNA sequence, providing a unique opportunity for the re-establishment of normal gene expression, the ultimate goal of epigenetic therapy. Currently DNMT inhibitors, including 5-aza-2'-deoxycyitidine (5-azaCdR), are the standard of care for the treatment of myelodysplastic syndrome (MDS) and are in clinical trials for other cancer types (86, 151). Although these drugs have demonstrated substantial success in the clinic, the comprehensive understanding of the molecular nature underlying their efficacy remains incompletely understood.

Although there have been numerous investigations into both DNA methylation and gene expression alterations induced by treatment with 5-azaCdR, only a handful of previous studies have investigated the ability of specific loci to sustain 5-azaCdR induced DNA demethylation following drug removal. Several genes were found to undergo a complete remethylation and resilencing after 5-azaCdR removal including *p16, FMR1,*

CDH1, and *ESR1* ((85, 171), Chapter 4). In other cases, a fraction of unmethylated alleles were stably maintained after the removal of 5-azaCdR at other loci including *TMS1* and *MLH1* ((70), Chapter 4). This suggested that the long-term impact of transient 5-azaCdR treatment differs between loci. Although genes appear to exhibit differences in their abilities to stably maintain 5-azaCdR induced alteration, these studies were conducted in different cell lines while utilizing different regiments of 5-azaCdR treatment presenting the possibility that this accounts for the differences observed. Two studies, one by Bender *et. al.* and the other performed by our group (highlighted in Chapter 4), investigated the propensity for remethylation at multiple loci during the same 5-azaCdR time course (171). Both studies reported differences in extent of post-5-azaCdR remethylation levels, which were mediated through epigenetic factors including the proximity of the CpG to the transcription start site, underlying histone modifications, or the continued occupancy of Pol II.

DNA isolated from MDS patients treated with 5-azaCdR has demonstrated the complete *in vivo* remethylation of LINE elements following removal of treatment (172). With the differences in remethylation kinetics observed at a handful of individual loci in cell culture, it is possible that different genes (or classes of genes) also differ in their remethylation kinetics within patients. Given that 5-azaCdR is the standard of care for MDS, the understanding of both the extent and long-term sustainability of 5-azaCdR induced demethylation remains vital for improving treatment regimen.

In this study we sought to determine the susceptibility of CpGs to 5-azaCdR induce demethylation and subsequent remethylation on a genome-wide scale. MDA-MB231 cells were treated with 5-azaCdR for 6 days and then kept them in culture in the

absence of drug for 27 passages (~3 months). We monitored genome-wide demethylation and subsequent remethylation using the Illumina Infinium methylation platform and found individual CpGs differ in the ability to initially demethylate with 5-azaCdR and then to sustain that demethylation. Furthermore, we found that the patterns of remethylation differ between CpG probes that are normally methylated compared with those that undergo cancer specific hypermethylation in breast cancer cells. Among those CpGs that exhibited cancer specific hypermethylation, we found their remethylation kinetics were linked with the occupancy of SUZ12 (a member of the PRC2 complex, responsible for the tri-methylation of H3K27) in embryonic stem cells. These data provide rationale for the use of genome-wide methylation patterns in individual tumors as a predictive marker of 5-azaCdR successes and perhaps in the personalization of regimens of epigenetic therapy.

Materials and Methods

Cell culture and 5-azaCdR treatments.

MDA-MB231 cells were obtained from ATCC and cultured as discussed in chapter 4. $5x10^5$ cells were plated and treated the following day with 0.5 μM of 5-azaCdR every other day for six days. After treatment DNA was harvest with a DNA extraction kit (Qiagen). Cells were then maintained in culture without 5-azaCdR for 27 passages, with cells being passed every three days 1:10. Three independent 5-azaCdR time courses were performed.

Methylation Specific PCR and COmbined Bisulfite Restriction Analysis

Isolated genomic DNA (2 μg) was converted with sodium bisulfite as previously described (118). Following ethanol precipitation, bisulfite modified DNA was resuspended in 40 μl dH_20. For Methylation Specific PCR (MSP) 1μl of bisulfite-modified template was used. MSP done as in Chapter 4 and as initially reported by Herman *et. al.* (155). PCR conditions used were 95°C 5 minutes, 80°C hold while 1 unit of Taq polymerase was added to each sample, followed by 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C seconds, this was followed by a final extension at 72°C for 5 minutes. Gene-specific primers designed to amplify methylated or unmethylated DNA following sodium bisulfite conversion are found in **Table 1**.

Combined bisulfite and restriction analysis (COBRA) was done as described in chapter 4 and initially reported (156). Primers devoid of any CpG were designed for *FGFR3, CDK3,* and *ARHGEF4* and their accompanying restriction enzymes are found in **Table 1.** PCR conditions were the same as in Chapter 4. Digested PCR products were analyzed on a 2% agarose gel and relative band intensities were determined using ImageQuant 5.2 and percent methylation was determined as the intensity of the digested bands divided by the intensity of all bands (digested and undigested).

Illumina Infinium genome-wide methylation analysis

DNA from the following cell lines were sent to the Emory University core for genome wide methylation analysis in biological triplicates: untreated MDA-MB231, 5-azaCdR treated, 3, 6, 9, and 27 passages after removal of 5-azaCdR. HMEC (human mammary epithelial cells) were also included to determine normal breast epithelial

methylation patterns. Briefly, the Infinium platform quantitatively measures DNA methylation levels of nearly 27,000 CpGs, through the use of probes for a methylated CpG or unmethylated CpG. Methylation levels were calculated as the ratio of the fluorescence of the methylated probe to the unmethylated probe, following bisulfite conversion. This process is diagramed in **Figure 1**.

Data Processing

The data from the three biologically independent 5-azaCdR time courses were averaged at each time point for each individual CpG. Those probes with an average β value < 0.3 in HMEC cells but >0.75 in untreated MDA-MB231 cells were designated cancer-specific methylation. Those probes with a β value of > 0.7 in HMEC cells and > 0.75 in untreated MDA-MB231 were designated normal methylation.

To analyze demethylation and remethylation of MDA-MB231 cells treated with 5-azaCdR the data set was filtered to only include samples with a β value of 0.75 or greater in the untreated MDA-MB231 cells. In addition CpGs were defined as either enriched or depleted for SUZ12 enrichment in embryonic stem cells. The SUZ12 genome wide data was from a data set published by Pasani, *et. al.* (173). These CpGs were next defined as low CpG island, high CpG island, or no CpG island. A probe was designated as a low CpG island if that probe fell within 500 bp of a CpG island as desribed by Gardiner-Garden *et. al.* (>50% CpG content over a 200 bp or greater stretch of DNA)(174). A probe was designated as a high CpG island if that probe fell within 500 bp of the more stringent classification of CpG islands as defined by Takai and Jones (>55% CpG content over a 500 bp or greater stretch of DNA) (175). Finally, a CpG probe was

designated as no CpG island if that probe did not fall within 500 bp of a CpG island as defined by either criteria.

Data analysis

Probes were filtered for those with an average β value of 0.75 or great in untreated MDA-MB231 cells followed by pair-wise average linkage clustering (176). HMEC cells (also filtered to remove probes < 0.75 β value in untreated MDA-MB231 cells) were included in the clustering as a reference of normal breast epithelial methylation patterns.

To identify differences in remethylation kinetics among different CpGs, we utilized Self Organizing Maps (SOMs). SOMs cluster data, using an algorithm reported by Tamayo *et. al.* to group individual CpG probe based on similar remethylation kinetics (177). The probes are randomly assigned to a pattern. This is followed by up to 50,000 iterations of individual remethylation curves in an effort to improve the clustering. Empirically, we decided up 16 as the optimal number of SOMs, providing both tight clustering within each group and a small enough number of SOMs that we could investigate them individually or in small groups.

Results

The CpG islands of six methylated genes exhibit differences in their demethylation and remethylation kinetics

As described in Chapter 4, three CpG islands differed in their ability to maintain an unmethylated domain after drug removal. This presented the hypothesis that genes

may have varying patterns of remethylation kinetics after 5-azaCdR removal and these differences may be predicated, in part, by other epigenetic factors. To gain an insight into the underlying mechanisms responsible for *de novo* DNA methylation, we examined 5-azaCdR-induced demethylation and remethylation kinetics as six genes as a pilot study. MDA-MB231 cells were treated with 5-azaCdR for 6 days (described in Chapter 4). After treatment, cells were maintained in culture for 9 passages in the absence of drug to measure stability. MCF7 cells were included as an unmethylated control. We found distinct differences in the demethylation and subsequent remethylation ability among these six genes. Two genes, *TMS1* and *CDH1* underwent an initial DNA de-methylation followed by a partial but incomplete remethylation 9 passages after the removal o 5-azaCdR (**Figure 2A**). Not all genes exhibited this identical remethylation kinetics, as *Lif, Progestrone Receptor (PR),* and *Estrogen Receptor (ESR1)* all demonstrated DNA demethylation followed by a complete remethylation within 9 passages of drug removal. Interestingly, *AdamTS5,* did not demonstrate any demethylation, which may be the result of rapid remethylation before the time point was taken. The heat map in **Figure 2B** is a visual interpretation of the MSP gels (63). Overall, there appeared to be measureable differences between genes in their ability to first demethylate and then maintain that demethylation after the removal of 5-azaCdR. This preliminary data provided rational for the expansion of our analysis genome-wide.

Characterization of CpG probes in MDA-MB231 cells in comparison to normal HMEC methylation patterns

For genome-wide demethylation and remethylation analysis we used the Illumina Infinium platform, which quantitatively measures methylation levels of 27,578 individual CpG sites spanning 14,495 different genes. To gain insight into the different behaviors of individual CpG loci we compared the methylation status of CpGs in Human Mammary Epithelial Cells (HMEC) which constitute the normal breast epithelial methylation pattern with those uniquely methylated in MDA-MB231 cells, constituting cancer specific methylation. To accomplish this analysis cutoffs were established. Probes with a β value of < 0.3 in HMEC cells were classified as normally unmethylated and probes with a β value > 0.7 were classified as normally methylated. This constituted 15,950 individual CpG probes. All other probes had a β value between 0.3 and 0.7 in either HMEC cells or MDA-MB231 cells were removed from initial characterization since they could not be definitively identified as either methylated or unmethylated. Based on these criteria, the vast majority of probes were unmethylated in both HMEC and MDA-MB231 cells (**Figure 3**). The overwhelming majority of these probes (n= 11,518) associated with CpG islands (98.5%). Conversely, 2750 probes were methylated in both HMEC and MDA-MB231 cells, and will be referred to as normal methylation. Only 57.2% of these probes were CpG island associated. We also observed 95 probes that were methylated in HMEC cells but unmethylated in MDA-MB231 cells indicating DNA hypomethylation in cancer; a phenomenon that has been reported for nearly thirty year (57). The relative small number of hypomethylated probes is likely a result of the bias of the Infiunium

platform to include CpG island probes (~91% of all probes are CpG island associated) coupled with the findings that DNA hypomethylation during carcinogenesis is typically found in bulk CpG methylation, not found in CpG island. The last group annotated included those probes unmethylated in HMEC cells and methylated in MDA-MB231 cells, which will be referred to as cancer specific methylation. This group consisted of 1,587 derived almost exclusively from CpG islands associated probes (95.9%) (178). This further supports that a subset of CpG islands were hypermethylated during carcinogenesis (13). A substantial amount of the normal methylation was found in non-CpG island associated probes while those CpGs unmethylated in HMEC cells that either remained unmethylated or underwent cancer specific methylation tended to be CpG island associated.

CpG sites follow distinct remethylation profiles following a transient exposure to 5-azaCdR

To analyze DNA remethylation kinetics of individual the Infinium dataset was filtered to include only those probes that had an average β value > 0.75 in the untreated MDA-MB231 cells (n=6,308 probes). In this filtered group of probes the average β value was 0.87. Treatment with 5-azaCdR decreased the average β value of these probes to 0.64 (**Figure 4**). Time point analyzed after the removal of 5-azaCdR showed a global remethylation of CpGs with an increase in β value from 0.64 to 0.75 after 27 passages in the absence of 5-azaCdR. The corresponding HMEC probes had an average β value of 0.52. The heat map in **Figure 4** was created through clustering of individual probes by

Pearson average linkage and depicts gross differences in remethylation kinetics across the genome.

CpG probes falling into distinct remethylation categories following a transient exposure to 5-azaCdR

The basic demethylation and remethylation kinetics among CpG sites observed in **Figure 4** suggested that there exist different patterns of remethylation. Individual CpG demethylation-remethylation kinetics were group by similarity utilizing Self Organizing Maps (SOMs). The use of SOMs is a technique utilized to visualize data that has been processed and sorted by a neural network. For this particular investigation we chose sixteen individual groups as the optimal number to highlight the differences in CpG demethylation and remethylation kinetics. We found that sixteen categories provided us with a substantial number of CpGs in each category, and relatively small variations in patterns within each group. We find that the CpG probes methylated in the untreated MDA-MB231 cells could be grouped into three general yet distinct patterns of remethylation kinetics (**Figure 5A**). The first trend involved CpG sites that exhibited little initial demethylation followed by a partial but incomplete remethylation. This trend was seen within patterns 0, 1, 2, 4, and 5. It is possible the low level of demethylation observed may actually represent rapid remethylation occurring before the 5-azaCdR time point was collected. The second trend encompassed those CpGs that underwent extensive demethylation followed by almost complete remethylation following drug removal (patterns 6, 8, 9, 12, 13, and 14). Finally, the third general group was composed of CpGs that underwent significant demethylation followed by little to no remethylation (patterns

3, 7, 10, 11, and 15). It should be noted that more subtle differences in remethylation kinetics were observed between patterns within each general group.

To provide molecular validation, three genes were chosen for methylation analysis by COBRA for comparison to the appropriate remethylation pattern assigned by the SOMs. Genes were chosen on the criteria of having multiple probes methylated in the untreated cells and all of these probes classified into the same remethylation pattern. We selected *CDK3* (multiple probes assigned to pattern 8), *ARHGEF4* (multiple probes assigned to pattern 12), and *FGFR3* (multiple probes assigned to pattern 15). The COBRA analysis of all three genes exhibited similar remethylation kinetics in comparison to the pattern in which they were classified (**Figure 5B**). Taken together, this demonstrated we are able to identify variations in remethylation kinetics within the 6,308 probes methylated in MDA-MB231.

Additional epigenetic factors contribute to the remethylation kinetics of individual CpG sites

To address the factors dictating why individual CpG sites displayed certain demethylation and remethylation kinetics, we compared individual patterns for enrichment or depletion of factors other than remethylation kinetics. We first compared normal methylation and cancer-specific methylation. Normal methylation was defined as those probes methylated in both HMEC and MDA-MB231 cells, while cancer-specific methylation was defined as those probes unmethylated in HMEC cells and methylated in MDA-MB231 cells. The fraction of normal methylation compared to cancer-specific methylation of each individual remethylation pattern is displayed by pie charts

underneath each pattern (**Figure 6A**). We then calculated the expected number of cancer specific methylation CpGs for each pattern. The total number of cancer-specific methylation probes compared to all probes methylated in MDA-MB231 cells was used to calculate the expected number of cancer-specifically methylated probes for each individual remethylation pattern. Chi squared analysis revealed that patterns 0, 1, 5, and 6 were significantly enriched for cancer-specific methylation ($p < 0.0001$) (**Figure 6B**). Pattern 15 was also significantly enriched for cancer specific methylation, just not as significantly ($p = 0.0019$). Conversely, patterns 8, 9, 10, 12, 13, and 14 were significantly depleted for cancer specific methylation and enrichment for normal methylation ($p < 0.0001$). Taken together these data suggest that whether or not a CpG is normally methylated in HMEC cells or is specifically methylated in cancer is able to serve as a predictor of that particular probe's demethylation and remethylation kinetics. This may represent underlying biological differences in the *de novo* methylation of CpGs either during normal development or aberrantly during carcinogenesis.

Given that probes specifically methylated in MDA-MB231 cells significantly clustered into two basic groups of demethylation-remethylation patterns, we hypothesized that the presence of underlying epigenetic factors might contribute to this behavior. One such factor we investigated was the relationship between remethylation kinetics and Polycomb Repressive Complex (PRC) occupancy, which is responsible for the tri-methylation of H3K27. Previous work has shown that the PRC complex marks many CpG islands prone to hypermethylation in human cancers (63). To determine the relationship between SUZ12 (a member of the PRC2 complex) occupancy and remethylation profiles we compared the individual profiles to SUZ12 occupancy in

human embryonic stem cells using recent genome-wide data from Pasini *et. al.* (173). The use of SUZ12 occupancy in ES cells has been demonstrated to correlate with CpG islands targeted for hypermethylation in cancer (63). Here we have used the same dataset to correlated SUZ12 occupancy in ES cells with specific remethylation kinetics in MDA-MB231 cells following the removal of 5-azaCdR. The ratio of CpG sites overlapping with regions enriched for SUZ12 was determined and graphed as pie charts underneath each individual remethylation pattern (**Figure 7A**). Since SUZ12 enriched probes only constituted a portion of our data set (985 of 6,308 probes (15.6%)) we calculated the expected number of SUZ12 positive probes for each individual pattern as 15.6% of probes in that particular remethylation pattern. The ratio of observed over expected SUZ12 positive probes was graphed in **Figure 7B**. We found that patterns 0, 1, and 5 were all significantly enriched for SUZ12 occupancy ($p < 0.0001$), pattern 6 was also enriched, just not to same level of significance $p = 0.0164$. These patterns were also all significantly enriched for probes specifically methylated in cancer. Also, we find that patterns significantly enriched for normal methylation are also significantly depleted for SUZ12 (patterns 8, 12, 13, and 14 ($p < 0.0001$)). Interestingly, pattern 15, which was enriched for cancer specific methylation, was also significantly depleted for SUZ12 suggesting that not all cancer specific methylation can be accounted for by SUZ 12 occupancy in human embryonic stem cells.

Discussion

5-azaCdR is currently the standard of care for MDS and involved in clinical trials for leukemia and other cancer types. While the transient effects of 5-azaCdR have been

defined in cell culture models and in patients, the understanding of what happens to DNA methylation after 5-azaCdR cessation remains poorly understood. The analysis of remethylation kinetics following the removal of 5-azaCdR has been limited to a handful of studies examining independent loci or more globally within DNA repetitive elements (Chapter 4, (70, 84, 85, 171, 179). Here we have taken a genome-wide approach to analyze differences in 5-azaCdR induced demethylation and remethylation kinetics following drug removal. Our data indicate that have different kinetics of remethylation after drug removal. These probes can be grouped into different patterns of demethylation-remethylation kinetics as defined through the use of SOMs. A low level of initial demethylation followed by a partial but incomplete remethylation highlighted one group of patterns (0, 1, 2, 4, and 5). Another group or patterns (6, 8, 9, 12, and 13) demonstrated a strong demethylation followed by a complete remethylation. Lastly, we observed a group of remethylation pattern that underwent initial demethylation followed by little to no demethylation, including patterns 3, 7, 11, and 15.

It seems unlikely that the differences between patterns of remethylation kinetics were random. We sought to identify factors contributing to the remethylation kinetics found of the different patterns. We found that those patterns exhibiting strong demethylation followed by complete remethylation (patterns 8, 9, 12, and 13) were all significantly enriched for CpGs found in the normal breast epithelial methylation pattern. While those CpGs that constituted the cancer-specific methylation profile fall into two distinct groups of remethylation kinetics. One group of patterns (0, 1, 4, and 5) demonstrated low levels of demethylation followed by a near complete remethylation as well as a strong enrichment of SUZ12 binding in embryonic stem cells. The other group,

also enriched for cancer specific methylation, was highlighted by strong demethylation followed by little to no remethylation (pattern 15) and demonstrated a significant depletion of SUZ12 occupancy.

The mechanisms underlying the clinical effectiveness low levels of 5-azaCdR treatment in cancer patients remains poorly understood. Examination of patient's methylation status after 5-azaCdR treatment has demonstrated little correlation between the extent of demethylation and patient response. Furthermore, while cell culture models have consistently demonstrated demethylation upon 5-azaCdR treatment, patient demethylation has been controversial, with results varying depending on the particular CpG methylation measured (81, 82). Here we have supported this notion with the finding that CpGs exhibit differences in the ability to demethylate from 5-azaCdR treatment. By only measuring a few specific loci, different conclusions on the extent demethylation may be reached, providing rationale for the use of genome wide methylation analysis when analyzing patient's DNA following 5-azaCdR treatment.

The clinical use of 5-azaCdR has experienced two distinct waves of success. The first coming in the 1970s when it was shown that high doses of 5-azaCdR resulted in a dose-dependent response in cancer patients. Unfortunately, these high doses of 5-azaCdR correlated with substantial side effects (79). Recently, clinical trials have utilized lower doses of 5-azaCdR given over a longer period of time and demonstrated anti-tumor activity without the inhibitory side effects. Currently, it is not fully understood why normal cells appear unaffected by these low levels of 5-azaCdR. A recent review by J.P. Issa presented two hypotheses underlying this phenomenon (76). One hypothesis included the notion that cancer cells depend the cancer-specific methylation of particular

genes for survival, and disruption of this methylation would preferentially affect cancer cells. The second hypothesis is based on the fact 5-azaCdR is an S-phase drug requiring active cell division, might preferentially target to cancer cells (76). Here we have provided molecular support for the cancer cell specific effects of 5-azaCdR *in vivo*. Those CpG that encompass the normal breast epithelial methylation profile exhibit a common pattern of demethylation and remethylation. These probes tend to undergo significant demethylation, followed by a nearly complete remethylation. This has lead us to propose that those CpGs that are normally methylated, although initially demethylated by 5-azaCdR, would be able to recover, fully restoring the normal pattern of DNA methylation in non-cancerous cells.

Given that 5-azaCdR is an S-phase drug, incorporated into DNA during synthesis, it is difficult to imagine that areas of the genome would demethylate to different extents. Nevertheless, we find a group of CpGs that exhibit little demethylation immediately after exposure to 5-azaCdR (patterns 0 and 1, in particular). The CpGs may be resistant to demethylation by 5-azCdR. Alternatively, these loci may represent CpGs undergoing a rapid and active remethylation. Our 5-azaCdR time course lasted for 6 days with 3 treatments of 0.5 μM 5-azaCdR administered every other day. Cells analyzed at the time of treatment had received their last dose of 5-azaCdR 48 hours prior. This time may provide the cells the opportunity for active DNA remethylation of certain CpGs. These CpGs also tend to be enriched for SUZ12 binding in embryonic stem cells. H3K27me3, the modification mediated by PRC2 (of which SUZ12 is a member), is present at a subset of CpG islands aberrantly methylated in cancer (69, 84, 180). Furthermore, the H3K27me3 mark is unaffected by treatment with 5-azaCdR, and it has been suggested

that the persistence of this mark may precipitate re-silencing (84). The finding that members of PRC2 are responsible for the tri-methylation of H3K27, associate with the DNA methyl transferases (66) presents the possibility that DNA methyltransferases are targeted to these loci and may facilitate their rapid re-methylation. This suggests that the inhibition of EZH2 and subsequently H3K27me3 may act synergistically with 5-azaCdR to have a greater impact on the long-term demethylation of cancer specific methylated CpGs.

Perhaps the most interesting group of CpGs were those that exhibited substantial demethylation followed by little to no remethylation after 27 passages in the absence of 5-azaCdR (Patterns 10, 11, 14, and 15). Patterns 10, 11, and 14 were enriched for those normally methylated probes. It is possible that the mechanisms responsible for establishing these methylation events in HMEC cells are no longer present and their demethylation from 5-azaCdR is more permanent. Conversely, pattern 15 was enriched for cancer specific methylation but was significantly depleted for SUZ12 occupancy. These CpGs may not be associated with those repressive histone modifications unaltered by 5-azaCdR (H3K27me3, H3K9me3, H4K20me3) (Chapter 4, (84, 85), therefore once demethylated there are no residual repressive modifications capable of facilitating the resilencing of these loci. One gene represented by this pattern of remethylation kinetics is *FGFR3*, which is a gene that is targeted for methylation not only in MDA-MB231 cells, but also several tumor types including lung cancer (181). Further studies into why these genes are capable of remaining demethylated following the removal of 5-azaCdR would be useful in understanding the underlying epigenetic factors that contribute to *de novo* methylation of these probes that are unable to facilitate remethylation.

There is still much to ascertain from this data set. Here we have shown that CpGs sites throughout the genome exhibit differential demethylation and remethylation potentials when challenged with the DNMT inhibitor, 5-azaCdR. These differences can be correlated, in part, with HMEC methylation status (normal vs. cancer specific methylation) and SUZ12 occupancy (H3K27me3) in embryonic stem cells. However, these two epigenetic factors do not fully explain all differences observed and future studies into other histone modifications, proximity to DNA repeat elements, chromosomal locations, and proximity to non-histone protein binding (such as CTCF) may provide further insight into the inherent differences in remethylation kinetics.

Figures

Figure 1.

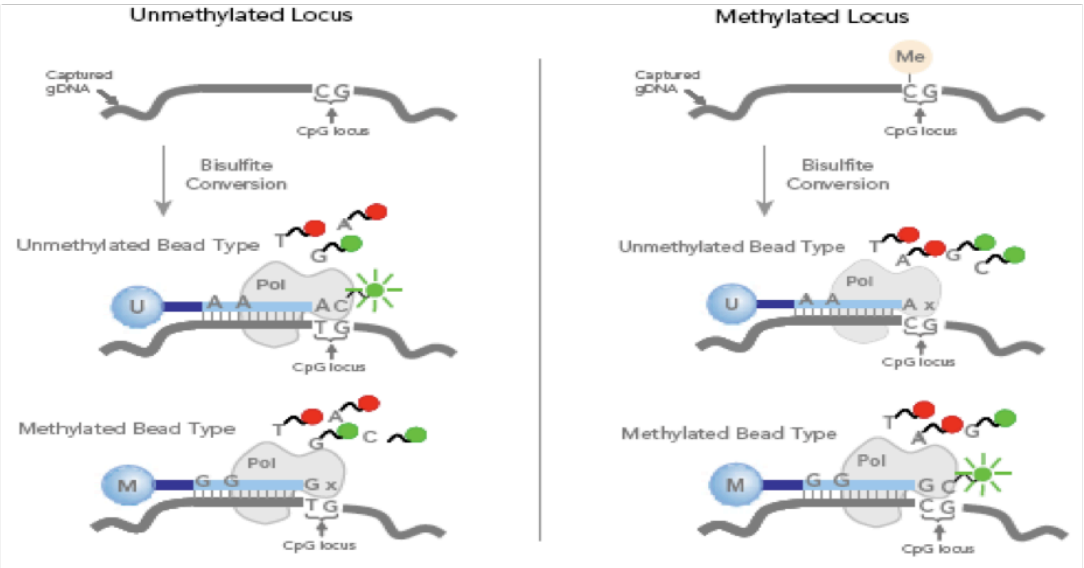


Figure 1: Schematic of Illumina's Infinium methylation analysis platform. Following bisulfite conversion DNA is hybridized to an array with probe representing both the unmethylated and methylated CpG. The array assays nearly 27,000 different CpGs as a quantitative ration of methylated binding (fluorescence) to unmethylated binding (fluorescence). Figure adapted from http://www.illumina.com/.

Figure 2.

HMEC	MDA-MB-231	High CpGi	Low CpGi	No CpGi	Totals
M	M	885	688	1177	2750
M	U	31	19	45	95
U	M	1450	73	64	1587
U	U	11090	262	166	11518

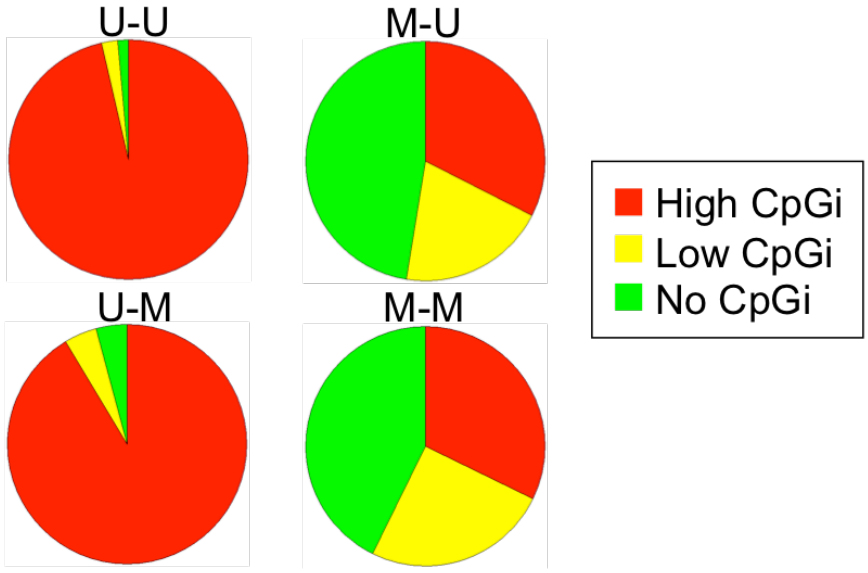


Figure 2: Characterization of CpGs as either normal methylation or cancer-specific methylation. CpG probes with a β value < 0.3 were designated as unmethylated (U) in HMEC or MDA-M231 cells. CpG probes with a β value > 0.7 were designated as methylated (M). High (>55% CpG content, and > 500 bp), low (>50% CpG content and > 200 bp), and no CpG island designations. Probes were categorized as normal methylation, methylated in both HMEC and MDA-MB231 and methylated in HMEC and umethylated in MDA-MB231; or cancer-specific methylation unmethylated in HMEC and methylated in MDA-MB231.

Figure 3.

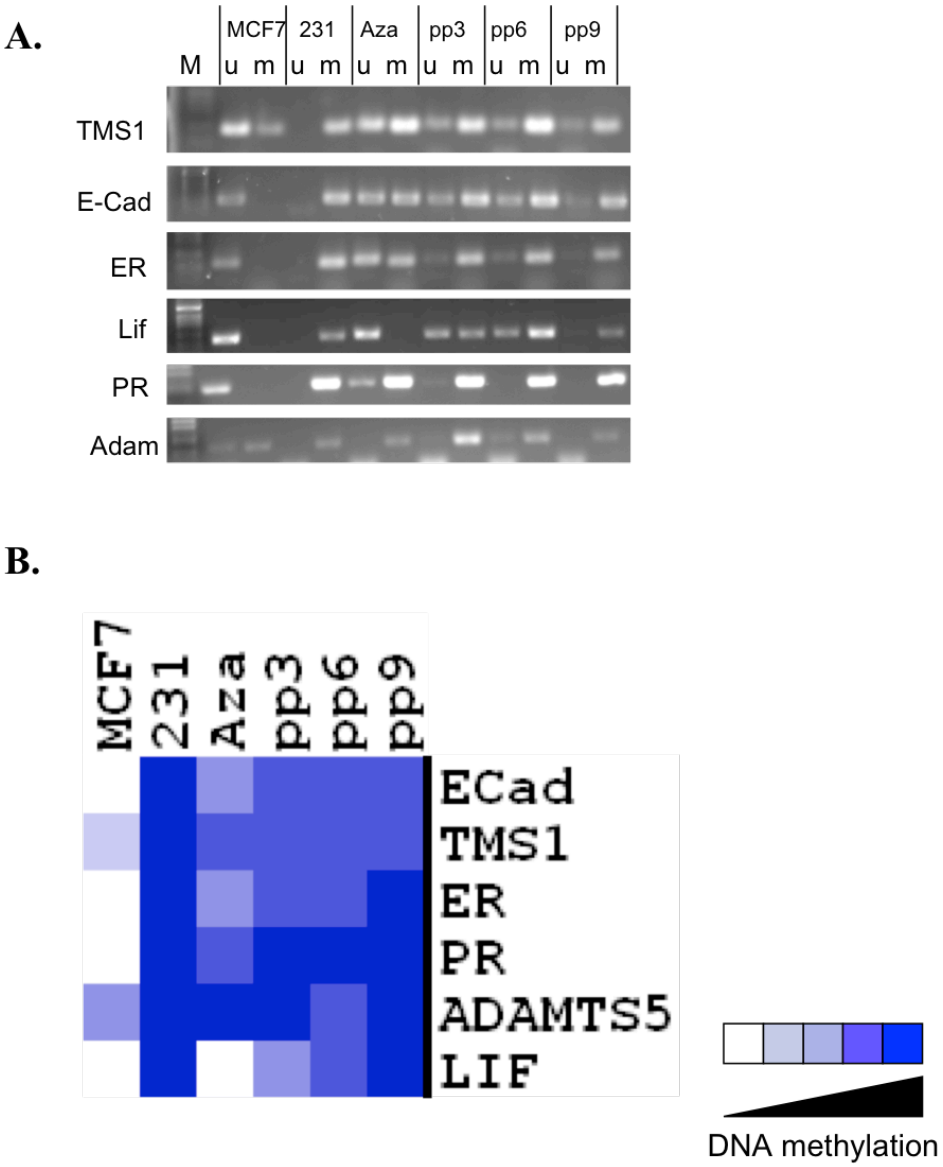


Figure 3: Difference exists between genes in their ability to demethylate due to 5-azaCdR and then remethylate after drug removal at the individual loci. A. MSP analysis of 6 CpG island containing genes immediately after treatment and periodically for 9 passages after treatment. U, unmethylated M, methylated, and pp, passage post 5-azaCdR treatment. **B.** Visual depiction of MSP data from part A. done as previously published (63). MSPs were performed in triplicate, a representative gel is shown.

Figure 4.

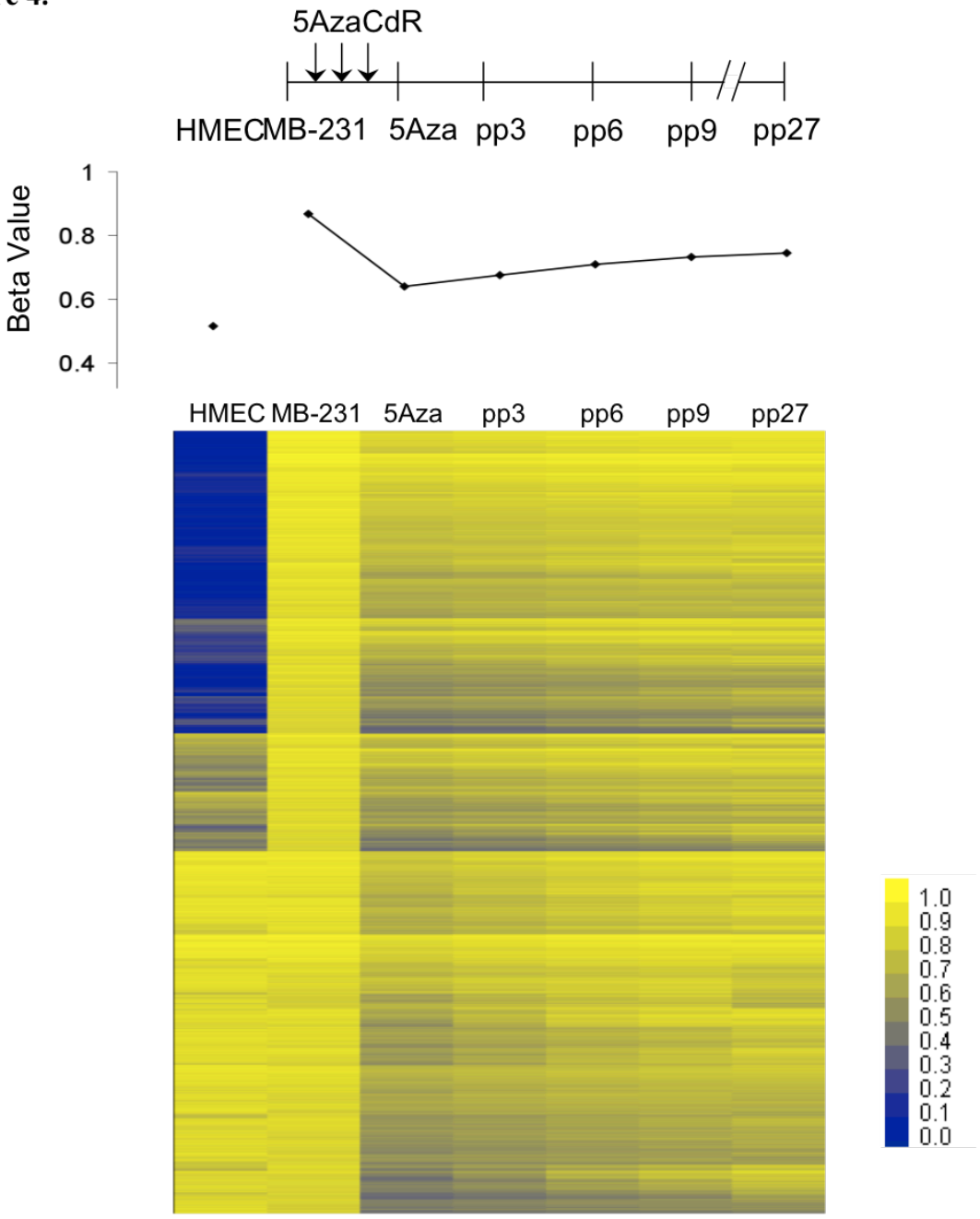


Figure 4: Genome wide clustering of demethylating and remethylated CpGs following Illumina Infinium methylation analysis. Data was filtered to include CpGs with a β value >0.75 in untreated MDA-MB231 cells (6,308 probes). Graph represents average β value of all samples at each time point. Heat map was generated by pair wise linkage analysis. HMEC included for normal breast epithelial methylation pattern. pp, passages post 5-azaCdR treatment.

Figure 5.

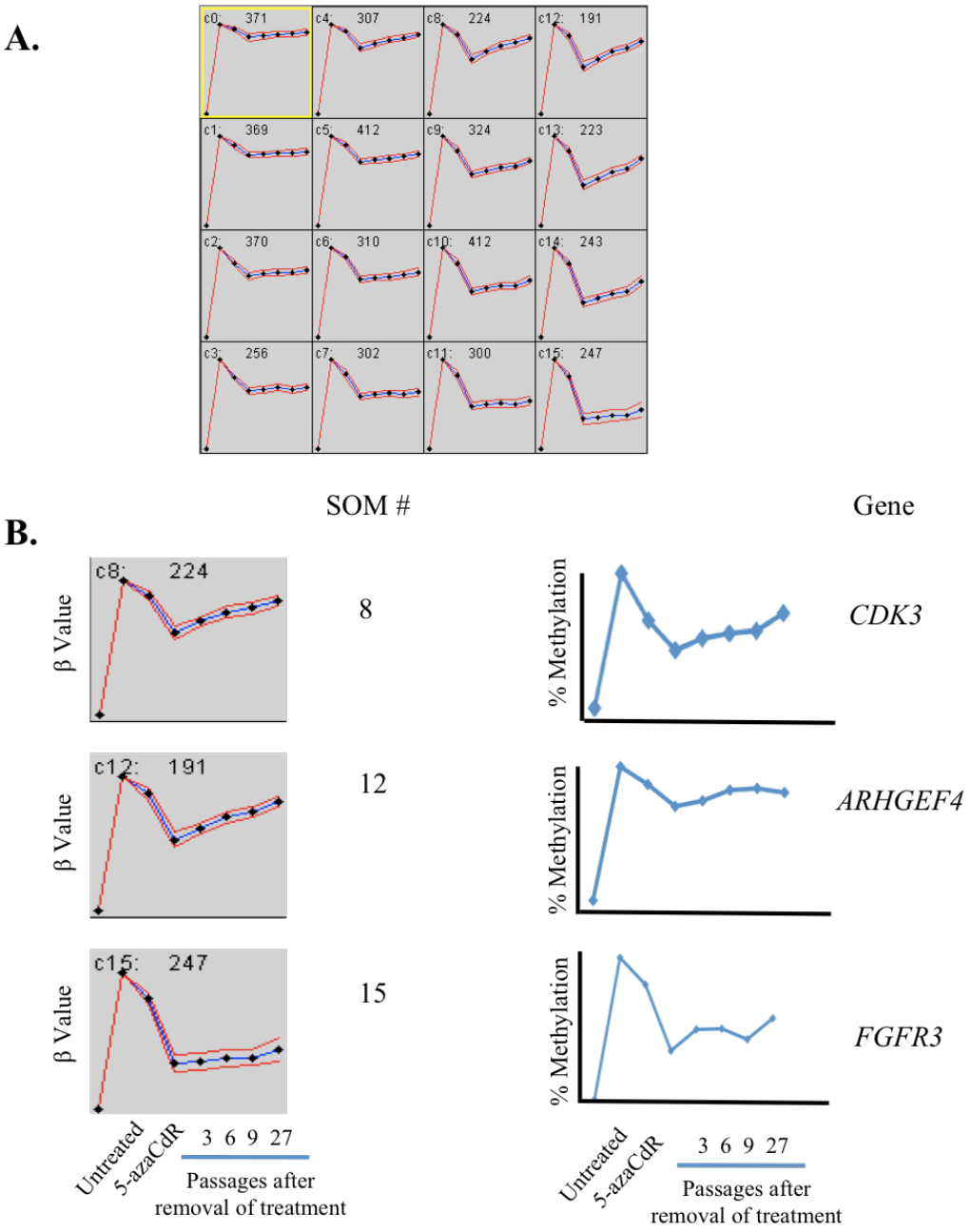
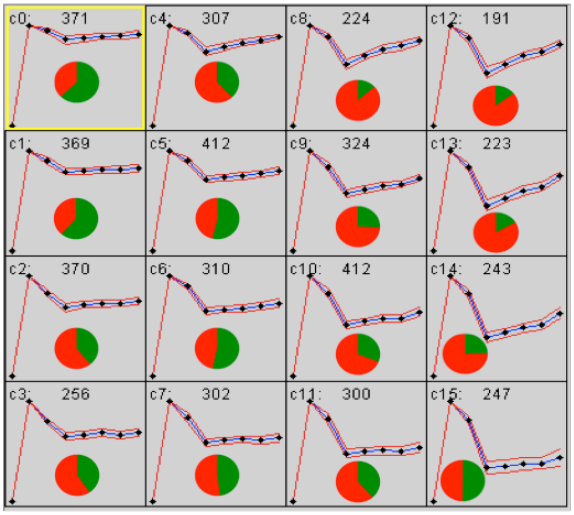


Figure 5: CpGs differ can be classified into distinct groups of remethylation kinetics after the removal of 5-azaCdR. A. All CpG remethylation curves were placed into sixteen patterns through the se of Self-Organizing Maps. The patterns are numbered 0 -15 and the number of CpGs in each map is denoted. The first and second point in each map is 0, 1 respectively to provide a uniform Y-axis. **B.** Molecular validation of remethylation kinetics by COBRA analysis. Three SOMs (left) were selected and a representative gene from each map was chosen for COBRA analysis. COBRA data (right) also has a 0,1 for Y-axis scaling and the data represents the average of three independent experiments.

Figure 6.

A.



- normal HMEC methylation

- cancer-specific methylation

B.

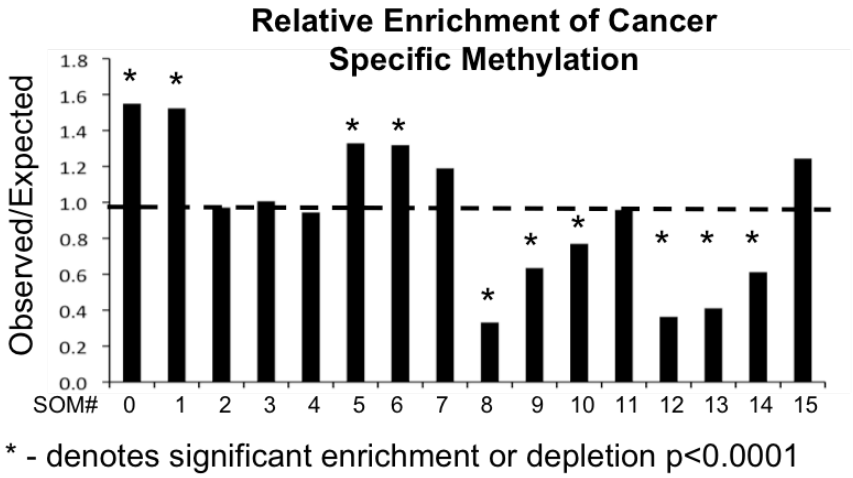


Figure 6: Relationship between remethylation kinetics and methylation class A. Each pattern from Figure 5 were further analyzed for CpGs that are either normally methylated (HMEC) or cancer specific methylated (MDA-MB231 only). Pie chart on each map represents the fraction of normal methylation and cancer specific methylation among CpGs in that pattern. **B.** Statistical analysis of significance of enrichment. Chi squared analysis was done for each pattern based on the predicted number of cancer specific probes and the actual number. * Denotes a p value < 0.0001.

Figure 7.

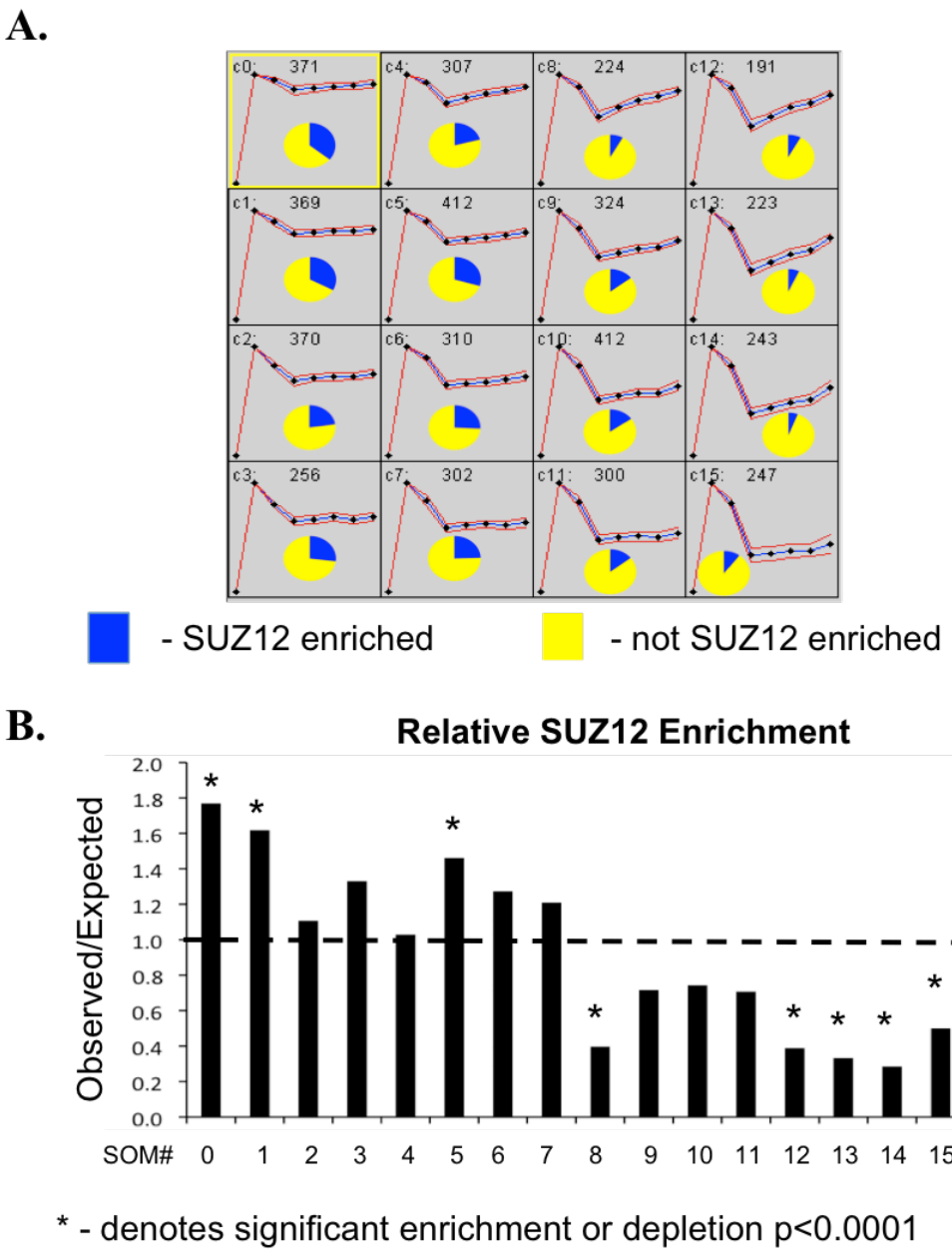


Figure 7: A sub-set of remethylation patterns are significantly enriched or depleted SUZ12 occupancy in embryonic stem cells. A. Each of the patterns from Figure 5 was further analyzed for SU12 occupancy in embryonic stem cells. Pie chart on each map represents the fraction of enriched to non-enriched sites within each pattern. **B.** Statistical analysis of significance of enrichment. Chi squared analysis was done for each pattern based on the predicted number probes occupied by SUZ12 and the actual number. * Denotes a p value < 0.0001.

Table 1

MSP

Gene	U Primer	M Primer
TMS1	F - GGTTGTAGTGGGGTGAGTGGT R - CAAAACATCCATAAACAACAACACA	F - TTGTAGCGGGGTGAGCGGC R - TTCGTCCATAAACAACAACGCG
CDH1	F - AGTTTGGTTGTTTGTTGGAAGGT R - AAACAACAACCCAAAATCTAACCA	F - TTCGGTTGTTCGTTGGAAGGC R - CGACAACCCGAAATCTAACCG
ESR1	F - ATGAGTTGGAGTTTTTGAATTGTTT R - ATAAACCTACACATTAACAACAACCA	F - CGAGTTGGAGTTTTTGAATCGTTC R - CTACGCGTTAACGACGACCG
ADAMTS5	F - GAGTTGTGAGGTTAGTGTTGTGT R - ATAAATCCACACAAATCCTACACA	F - GTTGCGAGGTTAGTGTTGCGC R - AATCCGCACGAATCCTACACG
LIF	F - GGAGTTTGTAAGTTGTTTGTTGTTT R - TCCTCCTCACTCCAAACACTCA	F - AGTTTGTAAGTTGTTCGTCGTTC R - CTCCTCGCTCCGAACACTCG

COBRA

Gene	Primer	Cut Site
CDK3	F – TATTGGGAGTTTAGTTTTTTGG R - CTACTATTTCCTACTAACTACC	*BssHII*
FGFR3	F - GATTTTTAAGGGTGGGTGTG R - AACCAAAACCTCCCTCCAC	*MluI*
AGHGEF4	F - AGAGTTTGGGAGAGTGTTGG R - CCAAAATCCCCTAAAATCCCC	*Xmn1*

Chapter 6: Discussion

Discussion

Summary

Here we have utilized the *TMS1* gene to study the relationships between DNA methylation, gene expression and chromosome structure. Previous studies from our group found the unmethylated *TMS1* CpG island was highlighted by 3 DNAse I HSs, enrichment for H3ac, and gene expression. Two of the HSs (HS1 and HS3) coordinated with the boundaries of DNA methylation. When the CpG island was aberrantly methylated in cancer cells, *TMS1* demonstrated a loss of the three island-associated HSs, enrichment of H3ac, and gene expression (100). At the start of this thesis, the above summarizes our understanding of the relationship between DNA methylation and basic chromatin structure at the *TMS1* locus.

Our initial aim was to further our understanding of the relationship between CpG island methylation and other epigenetic factors at *TMS1*. In addition to the previous findings that gene expression, H3ac, and three island-associated DNase I HSs all corresponded to the unmethylated CpG island, we found that the unmethylated *TMS1* CpG island was also enriched for Pol II, H3K4me2, and H4K16ac (**Figure 1**). Uniquely, H4K16ac localized to two peaks correlating with both the boundaries of DNA methylation and HS1 and HS3. H3ac, H3K4me2, and Pol II were all enriched across the entire CpG island. Furthermore, we found positioned nucleosomes spanning the entire unmethylated CpG island, save for a 400 bp nucleosome-free region overlapping the transcription start site. Although nucleosomes were positioned across the entirety of the unmethylated island, there were differences in the strength of positioning, with the most strongly positioned nucleosomes established at HS1 and HS3, the peaks of H4K16ac, and

the boundaries of DNA methylation. Conversely, when *TMS1* was methylated we found a loss of the active histone marks (H3ac, H4K16ac, H3K4me2, and Pol II) that coincided with the loss of gene expression and nucleosome positioning. In this state *TMS1* was enriched with H3K9me2 and a distinct peak of H4K20me3, immediately upstream of transcription start site (**Figure 1**) (Chapter 2, (103)).

Once we had thoroughly defined the *TMS1* epigenetic landscape in both the unmethylated and methylated states we next sought to experimentally assign function to the different epigenetic features associated with the CpG island both when the gene was active and when the gene was silent. Given the abundance of epigenetic factors co-localizing to the boundaries of DNA methylation we hypothesized that HS1 and HS3 may be serving as barrier elements, protecting the CpG island from aberrant methylation. We tested the hypothesis and found HS1 and HS3 were both capable of functioning as robust enhancer blockers, a phenomena related to barrier activity, providing initial support for their role in preventing the spread of methylation into the *TMS1* CpG island (Chapter 3).

We next utilized the DNMT inhibitor 5-azaCdR to study of role of DNA methylation plays in the epigenetic repression of *TMS1* (Chapter 4). A transient exposure to 5-azaCdR induced both DNA demethylation and gene re-expression at *TMS1*. These alterations were accompanied by the partial loss of H3K9me2 and the accrual of H3K4me2, H3ac, and Pol II. 5-azaCdR treatment did not completely recapitulate the unmethylated *TMS1* domain however, as H4K20me3 was unaffected and remained at a prominent peak immediately upstream of the transcription start site.

After the six-day treatment with 5-azaCdR, cells were maintained in culture to investigate the long-term stability of these induced alterations. After 5-azaCdR removal, *TMS1* did not fully remethylate after 5-azaCdR removal and maintained a subpopulation of demethylated alleles for 27 passages in culture (~3 months). Furthermore, we found the stably demethylated alleles preferentially associated with H3ac, H3K4me2, and Pol II as well as the repressive marks H3K9me2 and H4K20me3, creating a uniquely maintained epigenetic state that did not fully resemble either the unmethylated or methylated domain. This finding led to our investigation of the other epigenetically silenced TSGs (*ESR1* and *CDH1*) to determine if the maintained subpopulation of unmethylated alleles was common to all genes after 5-azaCdR treatment. We found that while all three genes (*TMS1, ESR1*, and *CDH1)* exhibited alterations in a similar subset of histone modifications, once 5-azaCdR was removed both *ESR1* and *CDH1* remethylated and resilenced within 27 passages in culture. The differences in the ability to maintain unmethylated alleles in absence of drug may reflect differences in the underlying epigenetic factors unaltered by changes in DNA methylation (H4K20me3, H3K9me3, and H3K27me3).

Given that H4K20me3 was the only repressive mark at *TMS1* unaltered through 5-azaCdR induced demethylation, we next co-treated with 5-azaCdR and shRNA constructs against *SuVar4-20* (sh*SuVar4-20*), the HMT for H4K20 di and tri methylation. We hypothesized co-treatment was would further induce *TMS1* re-expression through the synergistic removal of all known repressive marks and following the removal of treatments, *TMS1* alleles may be more robustly maintained than seen with 5-azaCdR alone (if some is good more is better) (Chapter 4). We first found that, indeed, co-

treatment with 5-azaCdR and sh*SuVar4-20* resulted in a more robust re-expression of *TMS1,* comparable to unmethylated cell lines, and significantly greater than observed from 5-azaCdR alone. H4K16ac re-association and the re-positioning of nucleosomes, as seen in unmethylated cells, accompanied the decreased levels of DNA methylation and H4K20me3 at TMS1. Surprisingly, we found that after removal of 5-azaCdR and sh*SuVar4-20,* the remethylation kinetics of *TMS1* was identical to after the removal of 5-azaCdR alone.

To bring our findings on 5-azaCdR induced demethylation and subsequent remethylation kinetics of *TMS1* into a more genome-wide spectrum we analyzed the remethylation kinetics of nearly 24,000 individual CpG sites (Chapter 5). We found individual CpG probes differed in the extent of both their initial DNA demethylation and subsequent remethylation after 5-azaCdR removal. One group of CpGs exhibited both a strong demethylation followed by remethylation to near pre-treatment levels. This group was significantly enriched for CpGs that were part of the HMEC methylation profile (normal methylation). Another subset of CpGs exhibited a robust DNA demethylation followed by little to no remethylation. This group was enriched for CpG sites specifically methylated in cancer. Finally, a distinct group of CpGs also methylated specifically in cancer demonstrated low levels of demethylation followed by a partial, but incomplete remethylation. This group was not only significantly enriched for cancer hypermethylation, but also significantly enriched for binding by PRC2 in stem cells. PRC2 is known to mediate H3K27me3, and this association may account for the low levels of DNA methylation. Another interpretation of is that the rapid and active remethylation occurred before the 5-azaCdR time point was taken. This preliminary study

provides evidence for the dramatic differences in remethylation kinetics follow 5-azaCdR treatment. The particular findings of this thesis will be discussed below in greater detail.

The relationship between positioned nucleosomes and TMS1 transcription

At *TMS1* nucleosomes were strongly positioned throughout the unmethylated CpG island while the methylation of the island correlated with the random positioning of nucleosomes. This finding was not unprecedented as Patel *et. al.* previously demonstrated a loss of nucleosome positioning accompanying promoter hypermethylation (45). *In vivo,* a positioned nucleosome can be the result of a number of different factors including the underlying DNA sequence, neighboring positioned nucleosomes, and the binding of non-histone proteins. The understanding of why nucleosomes are positioned may provide insight into the overall epigenetic regulation of *TMS1*.

In the yeast genome, approximately one half of the positioned nucleosomes could be accounted for based on solely on DNA sequence alone (41). This study, by Segal *et. al.* (41), published software to predicting nucleosome positioning based on a given DNA sequence. Although the predictive software was based on nucleosome positioning sequences observed in Yeast and Chicken, these sequences are conserved through evolution. We queried the sequence for *TMS1* to help determine to origin of positioning in the unmethylated island. We find that positioned nucleosomes are predicted, based on DNA sequence, across the *TMS1* locus except for the CpG island, where there was an approximate 1 kb void of predicted positioning (**Figure 2**). This was not entirely surprising, since the consensus sequences for nucleosome positioning are enriched for A/T base pairs allowing the DNA to easily bend and more easily accommodate the

histone octomer, and the CpG island is enriched for CpG dinucleotides. However, we observed strong nucleosome positioning throughout this region when the CpG island was unmethylated. Given that the CpG island was void of any predicted nucleosome positioning sequences, one possibility is that the positioning of nucleosomes at the unmethylated *TMS1* CpG island is based on interactions between nucleosomes and other epigenetic factors regulating the gene, and not from nucleosome positioning sequences within the DNA.

To determine the factors underlying the translational positioning of nucleosomes at the unmethylated CpG island we considered the following data: 1.) Unmethylated MCF7 cells exhibit a distinct pattern of nucleosome positioning across the CpG island while the methylated MDA-MB231 cells exhibit a random positioning of nucleosomes (Chapter 2). 2.) MCF7 cells transfected with siRNA targeting against *hMOF* (the HAT for H4K16) resulted in decreased nucleosome positioning concurrent with *TMS1* expression, while no other histone modifications were altered (Chapter 2). 3.) Treating MDA-MB231 cells with either 5-azaCdR or transfected with shRNA targeting *Suv4-20* (the HMT for H4K20me 2/3) alone did not result in the repositioning of nucleosomes (Chapter 4). 4.) However, the co-treatment of MDA-MB231 cells with 5-azaCdR and shRNA targeting *Suv4-20* did lead to the repositioning of nucleosomes concurrent with robust *TMS1* expression (Chapter 4). This may indicate that nucleosome positioning at *TMS1* is a consequence of robust expression. Taken together these findings suggest that the translational positioning of nucleosomes in the unmethylated CpG island at *TMS1* is directly related to the level of *TMS1* expression. The only two conditions in which we observe nucleosome positioning are in normally unmethylated cells (*e.g.* MCF7 cells)

and MDA-MB231 cells co-treated with 5-azaCdR and shRNA against *SUV4-20*. These are the two experimental conditions in which *TMS1* expression was the greatest. MDA-MB231 cells treated with 5-azaCdR alone resulted in *TMS1* re-expression, but we did not observe positioned nucleosomes. The lack of observed nucleosome positioning upon 5-azaCdR treatment may be due to the low level of *TMS1* expression in this condition coupled with the detection limits of in-direct end labeling.

The role of chromatin architecture in the protection of the unmethylated TMS1 CpG island from aberrant methylation

The finding that the peaks of H4K16ac, strongly positioned nucleosomes, and DNase I HSs all co-localized at the boundaries of DNA methylation of the unmethylated *TMS1* CpG island, we hypothesized that these regions (which I will refer to as HS1 and HS3) function as barrier elements, potentially protecting the unmethylated *TMS1* locus from position effect variegation and subsequent gene silencing. This hypothesis is further supported through the finding that several well-defined barrier elements were initially reported as DNase I HSs, such as HSIV within regulatory region of *β-globin*, located at the boundary of DNA methylation (146). As preliminary study, we assayed HS1 and HS3 for their ability to block a distal enhancer from the promoter of the luciferase reporter gene in a plasmid based system. Although enhancer-blocking activity is distinct from insulator activity, the two often coexist at transitional regions between chromatin domains and may be mediated by the same factors (55, 146). We found both HS1 and HS3 to exhibit strong enhancer blocking activity, comparable to a known enhancer-blocking element from the *XL-9* locus (Chapter 3, (140)). Interestingly, the ability to

enhancer block was similar in the MCF7 cells (*TMS1* unmethylated) and MDA-MB231 cells (*TMS1* methylated), suggesting all factors responsible for enhancer blocking function are also present in MDA-MB231 cells even though HS1 and HS3 do not form in these cells at the endogenous *TMS1* locus. Given these data, we developed the following model for how the unmethylated *TMS1* locus is protected from aberrant methylation and silencing in normal cells. Our model contends that the two strongly positioned nucleosomes marked by H4K16ac at HS1 and HS3 associate with each other in the context of chromatin structure, creating a protective loop in which the unmethylated CpG island is insulated from methylated DNA and position effect variegation in both directions. The creation of HS1 and HS3 maybe dependent on the binding of unknown non-histone proteins, and these proteins may mediate the barrier activity (**Figure 3**).

In an attempt to address this hypothesis, we conducted Chromatin Confirmation Capture (3C), which assays whether two non-linear regions of DNA are held in close spatial proximity within the context of the chromatin structure, creating a loop. After extensive attempts to assay HS1 and HS3 for their association in both MCF7 (*TMS1* unmethylated) and MDA-MB231 (*TMS1* methylated) cells, we concluded that the linear distance between HS1 and HS3 (1 kb) was too small to assay in this manner (personal communication Dekker, J). It is also possible this model is over simplified, given recent studies that have found that clusters of genes that are all regulated in a similar manner, localize in nuclear domains creating transcriptional clusters (55). *TMS1* could be part of a more complicated looping structure with other genes or even with a distal enhancer region in addition to our model in which the unmethylated CpG island is maintained in a protective chromatin loop.

The relationship between DNA methylation and individual histone modifications

We utilized the DNA demethylating agent 5-azaCdR to examine what other epigenetic alterations occurred, in conjunction with DNA demethylation. Our objective with 5-azaCdR treatments was to determine the relationship between DNA methylation and the repressive histone modifications known to mark *TMS1* (H3K9me2 and H4K20me3). To achieve maximal demethylation in culture we utilized a low dose of 5-azaCdR for an extended period of time. This treatment regimen has been shown both in cell culture and in the clinic to maximize DNA demethylation (182). We found that after 6-days of treatment MDA-MB231 cells were still viable, however when we increased the treatment to 8-days, cells often died (data not shown). Treatment with 5-azaCdR resulted in the loss of H3K9me2 and the accrual of H3ac, H3K4me2, and Pol II at *TMS1*. These modifications are commonly altered with 5-azaCdR treatments as demonstrated by several previous studies, providing striking evidence for their direct linkage to DNA methylation status (84, 85, 159, 171). These studies described the associated histone modifications of 5-azaCdR treated loci at the population level. While at the population level individual loci appear to have both repressive and active histone modifications, it remained a possibility that 5-azaCdR treated cells resolved into distinct sub-populations of alleles, uniquely marked by histone modifications. To address individual histone modifications for their associative methylation status both immediately after treatment with 5-azaCdR and three months after the removal of treatment, we employed a novel approach, ChIP-bis. ChIP-bis allowed us to study the relationship between individual

histone modifications with DNA methylation in a more detailed manner than has previously been achieved.

Treatment with 5-azaCdR resulted in H3ac re-association at the *TMS1* locus exclusively with the demethylated alleles. However, following 27 passages without 5-azaCdR H3ac associated with both the unmethylated and methylated alleles. This demonstrates a lack of de-acetylation on those alleles that are remethylated following 5-azaCdR removal. These data suggest that the continued presence of H3ac is not inhibitory to DNA remethylation nor is DNA remethylation inhibitory to the maintenance of H3ac marked nucleosomes. At *CDH1*, which is homogenously remethylated after the removal of 5-azaCdR, we observed a complete loss of expression and Pol II occupancy, however H3ac remained at low levels. This finding was a bit surprising since H3ac has traditionally been thought of as associating exclusively within actively transcribed genes. Here we have observed that H3ac is more durable to remethylation than H3K4me2 and Pol II and that H3ac is maintained despite DNA remethylation on alleles that are not transcribing as suggested by the absence of Pol II.

These findings may have important clinical implications. Cell culture studies have found that co-treatment with 5-azaCdR and HDACi resulted in synergistic reactivation of silenced tumor suppressor genes through the combined DNA demethylation and histone hyperacetylation (89). However, co-treatment with 5-azaCdR and the HDACi, 4-phenylbutyric acid, does not retard the resilencing of the *p16* gene observed after the removal of 5-azaCdR alone in T24 bladder cancer cells (164). This demonstrates that while the use of HDACi may act synergistically with 5-azaCdR to reactivate TSGs, the inhibition of HDAC activity is unlikely to aide in long-term maintenance of gene activity.

Practically, it may the more effective to use HDACi during the initial treatment and the switch to an alternative therapy targeting a different histone modification after the cessation of 5-azaCdR treatment.

The DNA methylation associated with H3K9me2 during the 5-azaCdR time course was also noteworthy. Initially, following treatment with 5-azaCdR H3K9me2 was lost from all demethylated alleles. From our current data we cannot determine if this was an active demethylation of the H3K9 residue or whether the H3K9me2 marked histone was replaced during transcription or replication. Given that 5-azaCdR is incorporated into the DNA during replication it seems reasonable that unmodified histones be deposited during this replication-dependent demethylation. Another possibility is the deposition of a histone unmethylated for H3K9 following 5-azaCdR induced transcription. In fact, the transcriptional deposition of the histone variant H3.3 has been demonstrated to be a mechanism of removing H3K9me from induced genes (183). This could provide another opportunity for addition of an unmodified histone at those demethylated and actively transcribing alleles.

After the removal of 5-azaCdR there was a shift in the distribution of H3K9me2 marked histones from the complete enrichment on methylated alleles (immediately after treatment) to being enriched on both methylated and unmethylated alleles (27 passages after treatment). The association of H3K9me2 with the subpopulation of unmethylated may explain the decrease in *TMS1* expression observed after the removal of 5-azaCdR. This hypothesis is further supported from a study by Backman *et. al.* (163). This study utilized HCT116 cells where *p16* is heritably silenced via DNA methylation. Several DNMTs were then knocked out in these cells resulting in demethylation and re-

expression of *p16.* However, over time in culture the *p16* locus was occupied by H3K9me2 and became heritably silenced (163). This heritable silencing was independent of DNA methylation and may demonstrate that H3K9me2 has the ability to epigenetically silence genes on unmethylated alleles.

H3K9me2 presence on both methylated and unmethylated alleles following drug removal also suggests a model in which H3K9me2 re-occupancy precedes DNA remethylation. The HMT G9a facilitates the H3K9me2 mark in human cells. Recent work has demonstrated G9a interacts with DNMT, providing a mechanism for the targeting of DNA methylation to loci marked by H3K9me2 or the co-deposition of H3K9me histones during replication (38). We propose that those *TMS1* alleles that were initially demethylated and then bound by G9a, which facilitates H3K9me2, targets the DNMT proteins to the *TMS1* locus to facilitate remethylation. We cannot rule out the possibility that over a longer time course the continued presence of H3K9me2 may eventually lead to the remethylation of all unmethylated alleles.

Previous studies that have reported on the alterations of histone modifications following treatment with 5-azaCdR in culture have only assayed histone on a population level. This type of analysis has led to the deduction that certain genes that are aberrantly silenced in cancer exhibit bivalent chromatin domain with both active and repressive modifications in cancer cells (184). Furthermore, bivalent chromatin domains were again reported as the result of 5-azaCdR treatment (84). While it is possible that each molecule of DNA at these loci is associated with histones marked by both active and repressive histone modifications, it is also possible that there are sub-populations of alleles that are differentially marked by histone modifications and the bivalent appearance is a

consequence of a mixed population of alleles. Through the use of ChIP-bis we have demonstrated that although on the population level *TMS1* is found in a bivalent domain following 5-azaCdR treatment, the regulation of histone modifications is a bit more complex and represents aspects of both a bivalent chromatin domain and independent sub-populations that are differentially marked by histone modifications.

Positioned nucleosomes correlate with the ability to maintain unmethylated alleles following a transient exposure to 5-azaCdR

Currently, it remains unknown why certain *TMS1* alleles are targeted for remethylation while others are maintained stably unmethylated for 27 passages after drug removal. Bender *et. al.* demonstrated that *p16* was completely remethylated and resilenced within 9 passages of drug removal (179). Unfortunately, the *p16* locus is homozygously deleted in MDA-MB231 cells; so we were unable utilize the *p16* remethylation kinetics as a comparison in our time course (185). We instead monitored two other well-known TSGs, methylated in MDA-MB231 cells, *CDH1* and *ESR1*, and found that both of these genes were homogenously remethylated following the removal of 5-azaCdR. The loss of Pol II at both of these loci 27 passages after the removal of 5-azaCdR suggests a loss of poise for transcription of both genes and this loss may facilitate DNA remethylation of all alleles. In fact, the only other published report of a gene maintaining a subpopulation of unmethylated alleles after 5-azaCdR removal was the *MLH1* locus (70). One common feature between *TMS1* and *MLH1* is the transcriptional dependent positioning of nucleosomes within the CpG island. It is possible that the re-establishment of Pol II and other transcription factors during 5-azaCdR

treatment leads to the positioning of nucleosomes on those demethylated alleles at these loci. The positioning of nucleosomes may facilitate the reestablishment of barrier elements that, once 5-azaCdR is gone, are capable of protecting the locus from remethylation. Both genes do undergo a partial remethylation after 5-azaCdR removal, which may represent those demethylated alleles that were unable to fully reestablish functional insulators during treatment. Those genes that are completely remethylated after 5-azaCdR removal (*CDH1, ESR1*, and *p16*) may not reposition nucleosomes and re-establish insulators; so that once the drug is gone they are subjected complete remethylation.

Two independent epigenetic repressive pathways are found at TMS1

5-azaCdR treatment did not fully reconstitute the euchromatic characteristics observed when *TMS1* is unmethylated (*e.g.* MCF7 cells). The peak of H4K20me3 upstream of the *TMS1* transcription start site was unaltered by treatment with 5-azaCdR (Chapter 4); to our knowledge this is the first demonstration of the separation between H4K20me3 occupancy and DNA methylation. Other tri-methylated marks such as H3K27me3 and H3K9me3 are also unaltered by 5-azaCdR treatment (84). This suggests there is a subset of repressive histone modifications, which may vary from gene to gene, that are unaffected by changes in DNA methylation.

Given that 5-azaCdR alone was unable to decrease the levels of H4K20me3 at *TMS1* concurrent with the lack of H4K16ac we next investigated the relationship between H4K20me3 and H4K16ac. To address the role of H4K20me3 at *TMS1* we used an shRNA construct targeting *Suv4-20*, the HMT responsible for di- and tri- methylating

H4K20 (121). Knocking down *Suv4-20* in MDA-MB231 cells resulted in the decrease of H4K20me3 concurrent with the reappearance of H4K16ac at the *TMS1* CpG island (Chapter 4). Furthermore, the knockdown of *Suv4-20* alone does not alter DNA methylation or H3K9me2 and is unable induce *TMS1* re-expression, demonstrating that DNA methylation is epistatic to the repressive effects of H4K20me3. However, when MDA-MB231 cells were co-treated with 5-azaCdR and shRNA against *Suv4-20* there was a synergistic reactivation of *TMS1* expression. Together these data suggest that two independent epigenetic pathways negatively regulate *TMS1*.

DNA methylation and H3K9me2 highlight one of the repressive pathways. Alterations in DNA methylation via 5-azaCdR result in the loss of H3K9me2 concurrent with the re-association of H3ac, H3K4me2, and Pol II to the *TMS1* locus (**Figure 4**). It is possible that each of these active marks (H3ac, H3K4me2, and Pol II) require an unmethylated locus to initially associate with *TMS1*, and that the presence of DNA methylation or H3K9me2 is inhibiting their association. It should be noted that these marks are sufficient for a minimal level of *TMS1* re-expression. However, once H3K9me2 and/or DNA methylation returns, the majority of this re-expression is lost.

H4K20me3 highlights the second repressive pathway. Altering this pathway results in the return of H4K16ac, but is not sufficient to re-express *TMS1*. However, when MDA-MB231 cells are co-treated *TMS1* is synergistically reactivated, nearing expression levels of unmethylated cell lines. From other unpublished work in our lab, we postulate that H4K20me3 actively prohibits H4K16ac from associating with *TMS1*. The role of H4K16ac is to facilitate Pol II elongation, thus increasing the transcription levels only when Pol II is already present. Providing further evidence for the separation of these

two pathways was our investigation of cells culture after treatment with both 5-azaCdR and shRNA against *Suv4-20*. Although these cells exhibit a synergistic reactivation upon treatment, once both drugs are removed these cells identically resemble those cell treated with only 5-azaCdR. Once *SUV4-20* expression returns we observed an immediate loss of H4K16ac concurrent with the return of H4K20me3. This rapid return of H4K20me3 further suggests there is a recruitment factor at *TMS1,* unaltered by either treatment that is capable of re-targeting H4K20me3 and the de-acetylation of H4K16ac upon the return of *Suv4-20*.

Genome-wide remethylation kinetics

We expanded our analysis of the 5-azaCdR time course to other loci to determine if all genes are regulated in a manner similar to *TMS1.* Given that several previous studies have investigated DNA remethylation following 5-azaCdR have demonstrated differences in both the extent of remethylation coupled with our studying finding subtle difference in the remethylation kinetics of three additional genes provided rationale for a genome-wide investigation of remethylation kinetics. To this end we analyzed the same 5-azaCdR time course by the Illumina Infinium methylation analysis and analyzed the demethylation and subsequent remethylation kinetics of over 6,000 individual CpG sites methylated in untreated MDA-MB231 cells. In accordance with the investigations of individual loci, we observed differences in the remethylation kinetics of CpGs across the genome (Chapter 5). We then began to focus on the underlying epigenetic factors associated with these differences.

One underlying theme that precipitated from our genome-wide analysis of 5-azaCdR demethylation and subsequent remethylation was the distinction between the establishment and maintenance of epigenetic domains characterized by DNA methylation. Recent work from Wheeler *et. al.* utilized the L5 repeat element to study the establishment and maintenance of heterochromatic domains in *Schizosaccharomyces pombe*. Removal of this L5 sequence by cre-mediated recombination did not abolish the heterochromatic domain, indicating that the epigenetic mechanisms utilized for establishment and maintenance of repressed regions of the genome are distinct events ((186), personal communication Scott, K).

The heritability (maintenance) of CpG methylation is facilitated through DNMT1, the maintenance methyltransferase, responsible for methylating hemi-methylated DNA following replication (25). The methylation patterns of MDA-MB231 cells were nearly identical between the biological replicates, suggesting that once established, DNA methylation is stable, independent of whether or not the methyl mark was the result of the normal methylation pattern or the cancer-specific methylation pattern. It seems that while all methylated domains are maintained in a similar manner, they are established in unique ways. CpGs differed in their ability to remethylate (re-establish) following a transient demethylation by 5-azaCdR. Using self-organizing maps to group the remethylation of individual remethylation kinetics of all CpGs three general patterns were observed. One group was composed of CpGs that underwent substantial demethylation immediately after 5-azaCdR treatment followed by a near complete remethylation after three months in the absence of drug. This group was significantly enriched for probes that were also methylation in HMEC cells, normal methylation. The exact factors initially dictating

tissue-specific methylation patterns are not known, although DNMT3a/3b likely deposits the *de novo* methylation during development (187). It is likely that these factors are still present in MDA-MB231 cells and the disruption of normal methylation by 5-azaCdR does not inhibit its full re-establishment after drug removal (modeled in **Figure 5**).

A second group of CpGs exhibited lower levels of demethylation immediately after 5-azaCdR treatment, followed by a partial but incomplete remethylation following 27 passages in the absence of drug. This group was significantly enriched for those probes methylated in MDA-MB231 cells, but not HMEC cells (cancer-specific methylation). These probes were significantly enriched for SUZ12 binding in embryonic stem cells (Chapter 5, (173)). The seemingly low level of initial demethylation may instead reflect rapid remethylation. Vire, *et. al.* demonstrated that EZH2 interacts with DNMTs and may target genes marked by H3K27me3 for *de novo* DNA methylation (66). Furthermore, H3K27me3 is a repressive histone mark that has been demonstrated to be unaffected by treatment with 5-azaCdR (84). We propose that the continued occupancy of H3K27me3 during 5-azaCdR treatment precipitates this rapid remethylation. The idea of rapid *de novo* remethylation following 5-azaCdR is not a novel concept. Work from Bender *et. al.* demonstrated the CpGs found in the body of genes were targeted for rapid and *de novo* remethylation after the removal of 5-azaCdR (171). Here we propose a mechanism behind the targeting, which is the continued presence of H3K27me3. Interestingly, this group does not exhibit a complete remethylation after 27 passages in the absence of drug (**Figure 5**). Given that these probes are aberrantly methylated to begin with it is possible that the incomplete remethylation of these CpGs reflects the incomplete reestablishment of this type of DNA methylation. On the individual gene

level we observed this phenomenon with *ESR1* and *CDH1* (Chapter 4). Both of these genes are enriched for H3K27me3 and both demonstrate a partial but incomplete remethylation (Chapter 4, (129)). Although these genes are not completely remethylated, they do exhibit remethylation in all alleles, the complete loss of Pol II gene silencing after the removal of 5-azaCdR. Overall, it seems that the presence of H3K37me3 correlates with the reestablishment of heritable gene silencing at these loci, which may not require complete remethylation. One caveat of this conclusion is that the SUZ12 occupancy was determined via a publically available data set created from human embryonic stem (ES) cells. The use of SUZ12 binding in ES cells to predict remethylation kinetics in breast cancer cells leaves the possibility that sites enriched in ES cells may not also be enriched in MDA-MB231 cells and vice versa. However, studies form our lab, and others, have shown a significant correlation between SUZ12 occupancy in ES cells and a hypersensitivity to hypermethylation during carcinogenesis (63). Given that these CpG islands in cancer cells are accurately predicted to be methylation prone based on SUZ12 binding in ES cells, we argue that the same SUZ12 occupancy may be used to predict remethylation kinetics following a transient exposure to 5-azaCdR.

A third group of CpGs demonstrated substantial demethylation followed by little to no remethylation after drug removal. This group was also enriched for probes specifically methylated in cancer but was not enriched for SUZ12 in embryonic stem cells, suggesting a different class of genes. The fact that these CpGs did not undergo remethylation suggests that the mechanisms leading to the initial establishment of methylation in cancer cells may are no longer present (**Figure 5**). It is possible that over time, the mechanisms leading to the *de novo* methylation of these CpGs were not selected

for, since once the domains are established they are maintained. Another possibility is that the initial insult leading to the aberrant methylation of these CpGs was an isolated event.

Clinical implications

Our findings suggest that in general CpGs methylated in normal cells reestablish this methylation pattern after treatment with 5-azaCdR while a portion of the CpGs specifically methylated in cancer cells are resistant to remethylation. This finding has several important clinical implications for the use and scheduling of 5-azaCdR in patients. First treating patients with multiple cycles of 5-azaCdR with breaks between each cycle might allow time for the normal methylation patterns to re-establish within the normal cells while cancer-specific methylation may remain stably demethylated. It is possible those cancer-specifically methylated CpGs that exhibited a partial but incomplete remethylation could be stably demethylated in a step-wise fashion given multiple cycles of treatment over time. Over time, the continued demethylation of regions specifically methylated in cancer concurrent with the re-establishment of normal methylation patterns would serve to re-activate silenced TSGs leading to the arrest growth or apoptosis of tumor cells while having minimal effects on normal cell methylation, due to the selective and complete remethylation of the normal methylation pattern.

Another clinical implication of these data is the use of tumor specific methylation as a predictive biomarker prior to 5-azaCdR treatment. By categorizing individual patient's pattern of cancer-specific methylation, it might be possible optimize the

treatment schedule. If a particular patient had an abundance of cancer specific methylation bound by H3K27me3, then perhaps multiple cycles of 5-azaCdR treatments would have the greatest efficacy on that particular cancer, leading to the step-wise demethylation over time. On the other hand, if a patient displayed an abundance of cancer-specific methylation from the CpGs not associated with SUZ12, perhaps a one-time treatment with a higher dose of 5-azaCdR would be the most effective. Most likely, a patient would have cancer specific methylation from both groups, and the analysis of pre-treatment methylation levels would help to monitor progression during treatment.

Finally, the analysis of a patient's particular cancer-specific methylation patterns may allow for the optimization of drugs to use in combination with 5-azaCdR. Currently, HDACi are in clinical trials both alone and in combination with 5-azaCdR (188). Although HDACi have been shown to act synergistically with 5-azaCdR to re-activate silent genes, their ability to prolong the demethylation effects of 5-azaCdR appears unlikely, given our finding that H3ac is insufficient to prevent DNA remethylation at *TMS1* (89, 92). Previous work has demonstrated that co-treatment with 5-azaCdR and the HDACi 4-phenylbutyric acid has no impact on the resilencing kinetics of *p16* observed after 5-azaCdR treatment alone (164). There may be key histone modifications`, that when altered would facilitate the pro-longed demethylation of the cancer specific methylation found in Group 2. An obvious target might be a member of the polycomb complex`, removing H3K27me3`, which is enriched for CpGs that are resistant to the initial DNA methylation (or exhibit rapid remethylation) upon treatment with 5-azaCdR. Furthermore`, our data analyses at the TMS1 locus also present H4K20me3 as another possible target. It is possible that concurrent inhibition of DNA

methylation and H4K20me3 might result in stably demethylated CpGs. H3K27me3 is one of several repressive histone modifications enriched in epigenetically silent TSGs (63), and targeting other histone modifications may also be advantageous for the sustained long-term demethylation of cancer-specific methylation.

Conclusions and future directions

Overall this thesis has focused on *TMS1* as a model gene to study epigenetic silencing in cancer. Here we have furthered our understanding of the relationship between DNA methylation at the *TMS1* CpG island and histone modifications, gene expression, positioned nucleosomes, and overall chromatin structure (**Figure 1**). We have also uncovered a functional contribution of several of these factors including nucleosome positioning, H4K20me3, and DNA methylation. We find that positioned nucleosomes is directly linked to *TMS1* gene expression. Two strongly positioned nucleosomes also demarcate the boundaries of DNA methylation and represent enhancer-blocking regions of chromatin. H4K20me3 was found to repress *TMS1* expression through the inhibition of H4K16ac, which is responsible for Pol II elongation (personal communication Kapoor-Vazirani, P). Finally, DNA methylation also represses *TMS1* expression through the inhibition of both Pol II and H3K4me2 at *TMS1*. In addition to our findings, several areas of *TMS1* epigenetic regulation still remain incompletely understood.

Our preliminary study demonstrated that both HS1 and HS3 were capable enhancer blockers. Future experiments would include the identification of non-histone proteins bound to these HSs facilitating the enhancer blocking function, and to perform assays to understand if the enhancer blocking ability of HS1 and HS3 is predictive of

insulator activity. Insulator activity is typically measured through the insertion of two potential insulator sequences flanking a reporter gene. If the reporter gene expression is maintained in a number of inserted clones, the sequences demonstrate insulator function. If HS1 and HS3 were demonstrated to have insulator function, *TMS1* would serve as a model of insulator function *in vivo*. Another unknown aspect of the epigenetic regulation surrounding *TMS1* is discovery of the repressive mechanism(s) that re-target H4K20me3 back to *TMS1* following knockdown of *Suv4-20*. Unearthing this mechanism may provide evidence of another level of epigenetic repression at *TMS1* unaffected by changes in either DNA methylation or the removal of H4K20me3.

We then expanded our understanding the relationship between DNA methylation and 5-azaCdR at *TMS1* genome-wide to identify differences CpGs exhibit in their ability to demethylated and subsequently remethylate. We find striking differences in the remethylation kinetics of individual CpGs based on whether or not that CpG is part of the normal methylation profile or the cancer specific profile. We can further sub-classify the cancer specific methylation based on SUZ12 occupancy in embryonic stem cells. We propose two distinct mechanisms that underlie the remethylation of individual CpG sites following a transient treatment with 5-azaCdR. One group of CpGs, which are not enriched for SUZ12 and are methylation in HMEC cells undergo a rapid a near complete remethylation. It is possible that the mechanisms utilized by the HMEC cells to establish the normal methylation profile remain in MDA-MB231 cells and once 5-azaCdR is removed those CpGs are capable of a rapid and complete remethylation. Those CpGs, which are specifically methylated in cancer, had distinct patterns of remethylation kinetics, suggesting a different mechanism of remethylation. A number of these probes

were also marked by SUZ12 in ES cells, suggesting their possible occupancy by H3K27me3 in MDA-MB231 cells. Given that components of the PcG complex, which is responsible for H3K27me3, have been shown to directly interact with DNMTs, we propose that those CpGs that are associated with H3K27me3 are targeted for an active remethylation. This would explain why the majority of probes enriched for SUZ12 in ES cells demonstrated little to no demethylation, which can also be interpreted as active and rapid remethylation (**Figure 5**). Taking this one step further, there are also probes specifically methylated in cancer, which are not marked by SUZ12 in ES cells, and these probes are enriched for significant demethylation followed by little to know remethylation, suggesting that no remethylation mechanisms are in place for these particular CpGs. Further analysis of this dataset is needed to identify additional features of remethylation ability. The ability to predict the initial demethylation and remethylation of a CpG based on a series of underlying epigenetic factors provides rationale for the use of tumor methylation patterns as a predictive markers of response to epigenetic therapies. This study also provides rationale for the development of novel epigenetic drugs targeting factors enriched for CpGs specifically methylated in cancer, including H3K27me3. Finally, the predictive nature of tumor specific methylation patterns may help to optimize 5-azaCdR treatment dosage and scheduling. The addition of gaps between 5-azaCdR treatment cycles may allow for the normal methylation patterns to reestablish while a portion of the cancer specific methylation remains partially demethylated. Our study provides a foundation upon which future basic and clinical experiments can be based to improve the development and use of epigenetic therapies in the treatment of cancer.

Figures

Figure 1.

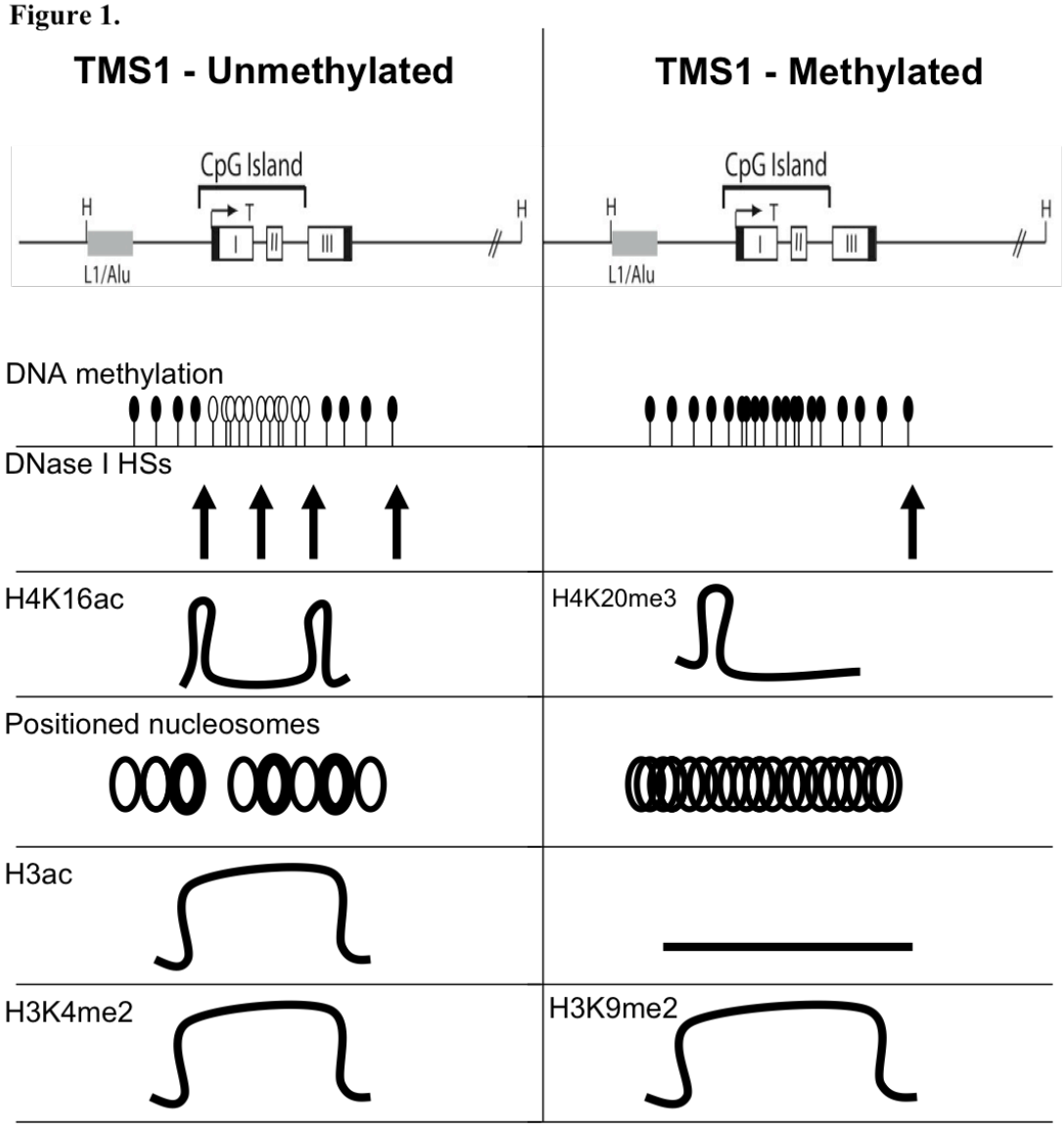


Figure 1. Summary of the epigenetic factors associated with the *TMS1* CpG island in the methylated and unmethylated state.

Figure 2.

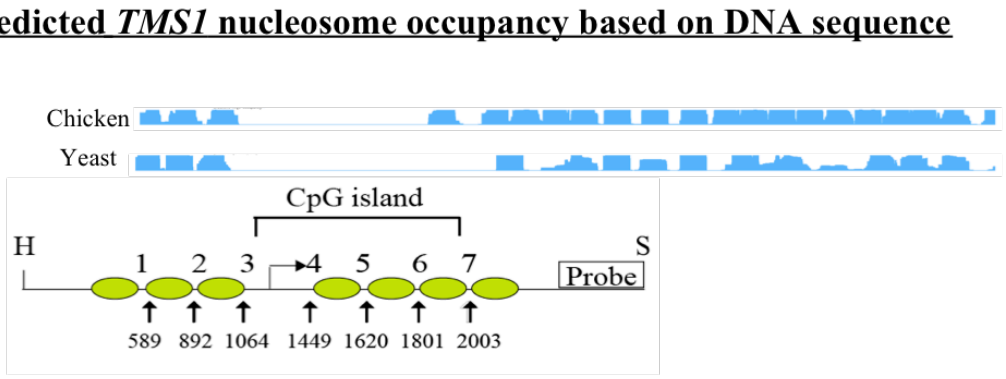


Figure 2. The *TMS1* CpG island exhibits a void of predicted nucleosome occupancy based on DNA sequence. A 5.5 kb region of the *TMS1* locus was analyzed for predicted nucleosome occupancy based on genome-wide predictions derived from Yeast and Chicken by Segal *et. al* (41). *TMS1* schematic indicates location of mapped positioned nucleosomes *in vivo*.

Figure 3.

Model of chromatin structure maintaining *TMS1* in an unmethylated state

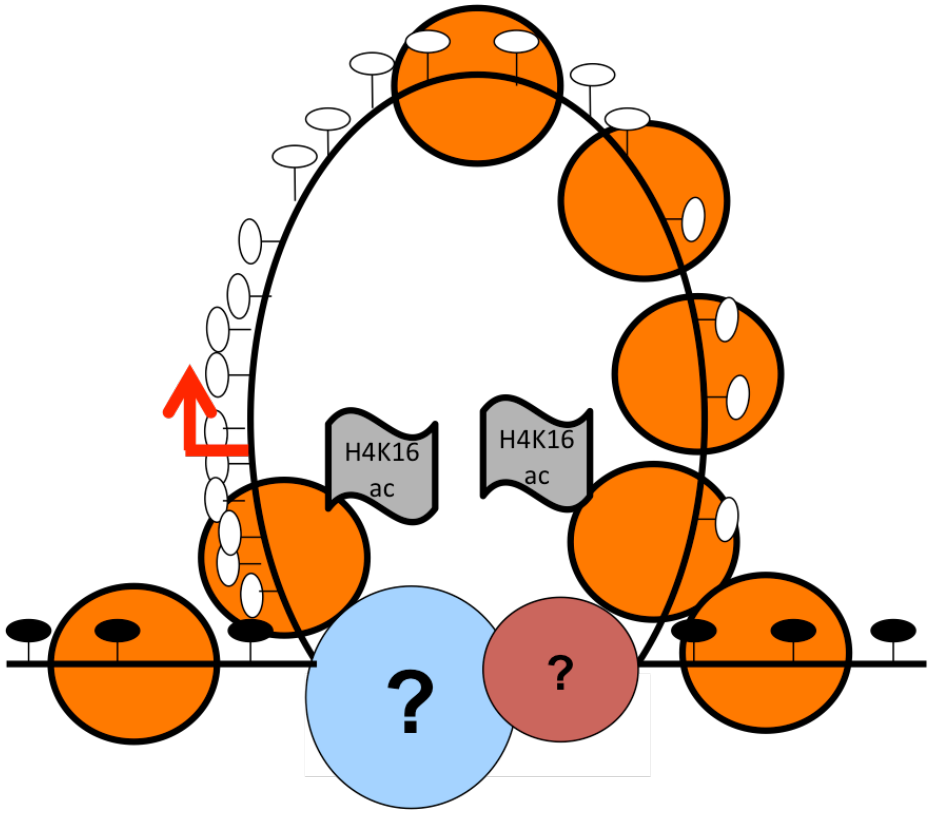


Figure 3. Model of chromatin structure protecting the unmethylated *TMS1* locus from position effect variegation. Orange circles represent nucleosomes, filled lollipops represent methylated DNA, and open lollipops represent unmethylated DNA, transcription start site represented by red arrow, and unknown proteins represented by question marks.

Figure 4.

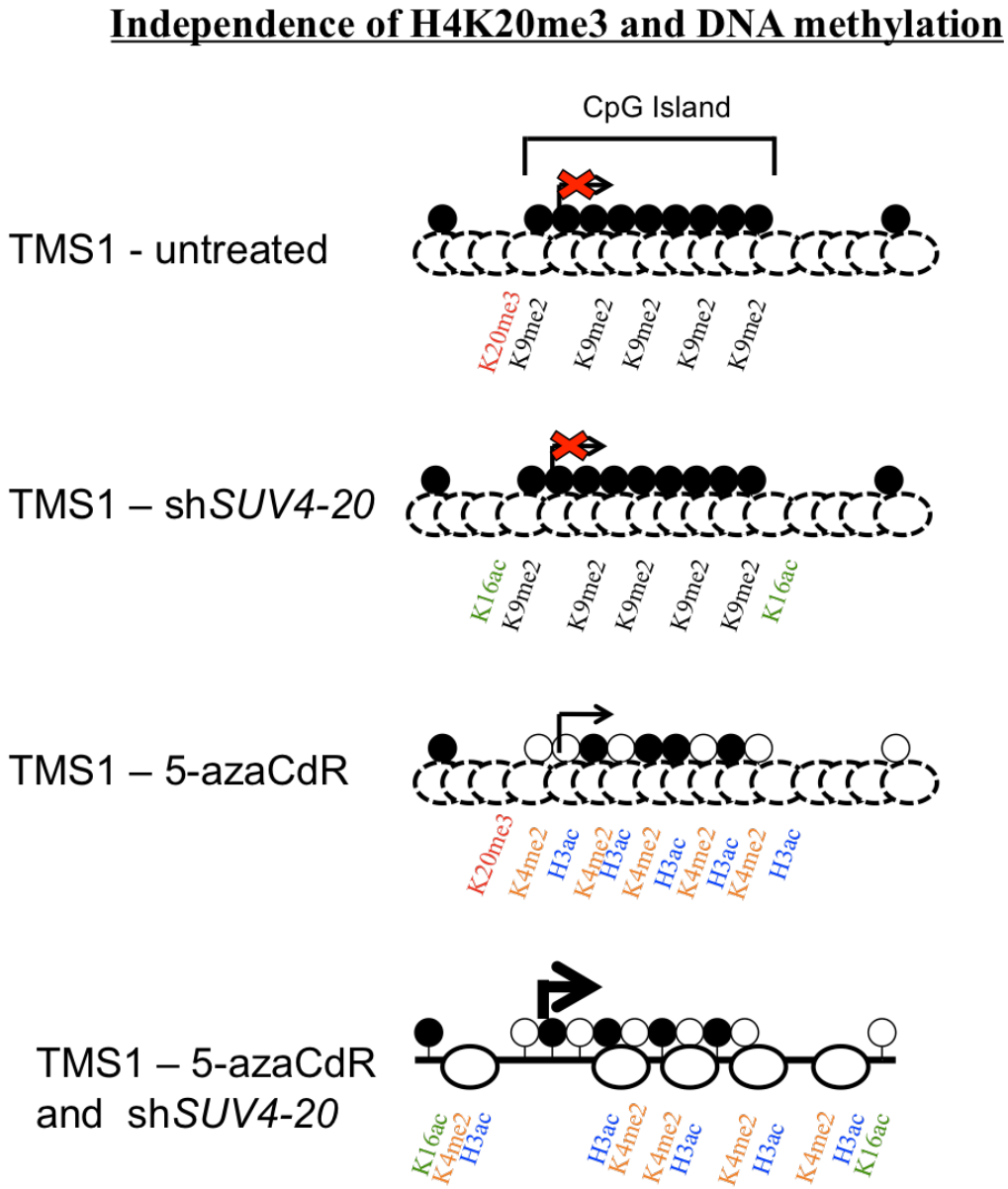


Figure 4. *TMS1* is regulated by two independent epigenetic repressive mechanisms. Schematic diagrams of *TMS1* regulation in MDA-MB231 cells either untreated, treated with either 5-azaCdR or shRNA targeted against SU*V4-20*, or treatment with both.

Figure 5.

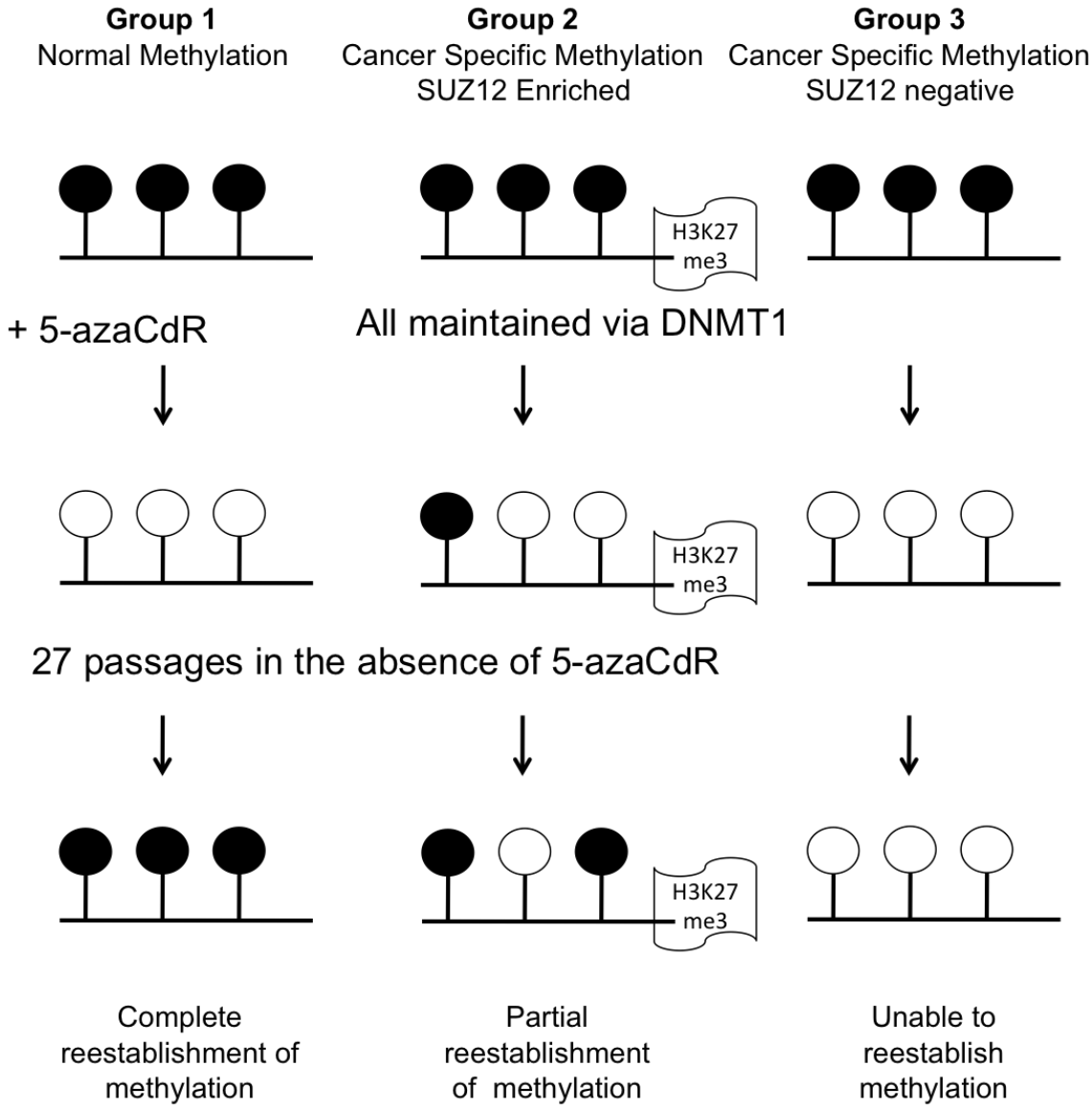


Figure 5. CpGs fall into three distinct groups of remethylation kinetics following the treatment and removal of 5-azaCdR.

References

1. Boveri T. Concerning the Origin of Malignant Tumors by Theodore Boveri. Translated and annotated by Henry Harris. Journal of Cell Science 2007; 121: 1-84.
2. Rous P. A Sarcoma of the fowl transmissible by an agent separable from the tumor cells. J Exp Med 1911; 13: 397-411.
3. Avery OT, MacLeod CM, McCarty M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III. 1944. Mol Med 1995; 1: 344-65.
4. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953; 171: 737-8.
5. Franklin RE, Gosling RG. Molecular configuration in sodium thymonucleate. Nature 1953; 171: 740-1.
6. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 1971; 68: 820-3.
7. Stehelin D. The transforming gene of avian tumor viruses. Pathol Biol (Paris) 1976; 24: 513-5.
8. Greger V, Passarge E, Hopping W, Messmer E, Horsthemke B. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum Genet 1989; 83: 155-8.
9. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000; 100: 57-70.
10. Murphree AL, Benedict WF. Retinoblastoma: clues to human oncogenesis. Science 1984; 223: 1028-33.
11. Smigal C SR, Jemal A. Breast Cancer Facts and Figures 2005-2006. American Cancer Society 2005.
12. Tanaka K, Oshimura M, Kikuchi R, Seki M, Hayashi T, Miyaki M. Suppression of tumorigenicity in human colon carcinoma cells by introduction of normal chromosome 5 or 18. Nature 1991; 349: 340-2.
13. Esteller M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene 2002; 21: 5427-40.
14. Hedenfalk I, Duggan D, Chen Y, et al. Gene-expression profiles in hereditary breast cancer. N Engl J Med 2001; 344: 539-48.
15. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003; 33 Suppl: 245-54.
16. Brandeis M, Ariel M, Cedar H. Dynamics of DNA methylation during development. Bioessays 1993; 15: 709-13.
17. Waddington CH. The genetic control of wing development in Drosophila. Journal of Genetics 1940; 41: 75-113.
18. Jenuwein T, Allis CD. Translating the histone code. Science 2001; 293: 1074-80.
19. Yasuhara JC, Wakimoto BT. Oxymoron no more: the expanding world of heterochromatic genes. Trends Genet 2006; 22: 330-8.
20. Talbert PB, Henikoff S. Spreading of silent chromatin: inaction at a distance. Nat Rev Genet 2006; 7: 793-803.

21. Gaszner M, Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet 2006; 7: 703-13.
22. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 2003; 349: 2042-54.
23. Bird AP. CpG-rich islands and the function of DNA methylation. Nature 1986; 321: 209-13.
24. McCabe MT, Brandes JC, Vertino PM. Cancer DNA methylation: molecular mechanisms and clinical implications. Clin Cancer Res 2009; 15: 3927-37.
25. Turek-Plewa J, Jagodzinski PP. The role of mammalian DNA methyltransferases in the regulation of gene expression. Cell Mol Biol Lett 2005; 10: 631-47.
26. Reik W, Dean W. DNA methylation and mammalian epigenetics. Electrophoresis 2001; 22: 2838-43.
27. Ooi SK, Bestor TH. The colorful history of active DNA demethylation. Cell 2008; 133: 1145-8.
28. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 1997; 389: 251-60.
29. Karrer KM, VanNuland TA. Nucleosome positioning is independent of histone H1 in vivo. J Biol Chem 1999; 274: 33020-4.
30. Allfrey VG, Faulkner R, Mirsky AE. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. Proc Natl Acad Sci U S A 1964; 51: 786-94.
31. Turner BM. Cellular memory and the histone code. Cell 2002; 111: 285-91.
32. Biel M, Wascholowski V, Giannis A. Epigenetics--an epicenter of gene regulation: histones and histone-modifying enzymes. Angew Chem Int Ed Engl 2005; 44: 3186-216.
33. Marmorstein R, Trievel RC. Histone modifying enzymes: structures, mechanisms, and specificities. Biochim Biophys Acta 2009; 1789: 58-68.
34. Shi Y, Lan F, Matson C, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 2004; 119: 941-53.
35. Annunziato AT. Split decision: what happens to nucleosomes during DNA replication? J Biol Chem 2005; 280: 12065-8.
36. Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. Differential chromatin marking of introns and expressed exons by H3K36me3. Nat Genet 2009; 41: 376-81.
37. Bernstein BE, Mikkelsen TS, Xie X, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006; 125: 315-26.
38. Esteve PO, Chin HG, Smallwood A, et al. Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes Dev 2006; 20: 3089-103.
39. Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL. Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science 2006; 311: 844-7.
40. Yuan GC, Liu YJ, Dion MF, et al. Genome-scale identification of nucleosome positions in S. cerevisiae. Science 2005; 309: 626-30.
41. Segal E, Fondufe-Mittendorf Y, Chen L, et al. A genomic code for nucleosome positioning. Nature 2006; 442: 772-8.

42. Chen C, Yang TP. Nucleosomes are translationally positioned on the active allele and rotationally positioned on the inactive allele of the HPRT promoter. Mol Cell Biol 2001; 21: 7682-95.
43. Pennings S, Allan J, Davey CS. DNA methylation, nucleosome formation and positioning. Brief Funct Genomic Proteomic 2005; 3: 351-61.
44. Davey C, Fraser R, Smolle M, Simmen MW, Allan J. Nucleosome positioning signals in the DNA sequence of the human and mouse H19 imprinting control regions. J Mol Biol 2003; 325: 873-87.
45. Patel SA, Graunke DM, Pieper RO. Aberrant silencing of the CpG island-containing human O6-methylguanine DNA methyltransferase gene is associated with the loss of nucleosome-like positioning. Mol Cell Biol 1997; 17: 5813-22.
46. Choi JK, Bae JB, Lyu J, Kim TY, Kim YJ. Nucleosome deposition and DNA methylation at coding region boundaries. Genome Biol 2009; 10: R89.
47. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 1999; 98: 387-96.
48. Lewis A, Murrell A. Genomic imprinting: CTCF protects the boundaries. Curr Biol 2004; 14: R284-6.
49. Geyer PK, Corces VG. DNA position-specific repression of transcription by a Drosophila zinc finger protein. Genes Dev 1992; 6: 1865-73.
50. Bushey AM, Ramos E, Corces VG. Three subclasses of a Drosophila insulator show distinct and cell type-specific genomic distributions. Genes Dev 2009; 23: 1338-50.
51. Chung JH, Whiteley M, Felsenfeld G. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell 1993; 74: 505-14.
52. Burgess-Beusse B, Farrell C, Gaszner M, et al. The insulation of genes from external enhancers and silencing chromatin. Proc Natl Acad Sci U S A 2002; 99 Suppl 4: 16433-7.
53. Mukhopadhyay R, Yu W, Whitehead J, et al. The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. Genome Res 2004; 14: 1594-602.
54. Donze D, Kamakaka RT. Braking the silence: how heterochromatic gene repression is stopped in its tracks. Bioessays 2002; 24: 344-9.
55. Bushey AM, Dorman ER, Corces VG. Chromatin insulators: regulatory mechanisms and epigenetic inheritance. Mol Cell 2008; 32: 1-9.
56. Jair KW, Bachman KE, Suzuki H, et al. De novo CpG island methylation in human cancer cells. Cancer Res 2006; 66: 682-92.
57. Gama-Sosa MA, Slagel VA, Trewyn RW, et al. The 5-methylcytosine content of DNA from human tumors. Nucleic Acids Res 1983; 11: 6883-94.
58. Lengauer C, Kinzler KW, Vogelstein B. DNA methylation and genetic instability in colorectal cancer cells. Proc Natl Acad Sci U S A 1997; 94: 2545-50.
59. Baylin SB, Hoppener JW, de Bustros A, Steenbergh PH, Lips CJ, Nelkin BD. DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. Cancer Res 1986; 46: 2917-22.
60. Feltus FA, Lee EK, Costello JF, Plass C, Vertino PM. DNA motifs associated with aberrant CpG island methylation. Genomics 2006; 87: 572-9.

61. Conway KE, McConnell BB, Bowring CE, Donald CD, Warren ST, Vertino PM. TMS1, a novel proapoptotic caspase recruitment domain protein, is a target of methylation-induced gene silencing in human breast cancers. Cancer Res 2000; 60: 6236-42.
62. Baylin SB. DNA methylation and gene silencing in cancer. Nat Clin Pract Oncol 2005; 2 Suppl 1: S4-11.
63. McCabe MT, Lee EK, Vertino PM. A multifactorial signature of DNA sequence and polycomb binding predicts aberrant CpG island methylation. Cancer Res 2009; 69: 282-91.
64. Ohm JE, McGarvey KM, Yu X, et al. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 2007; 39: 237-42.
65. Schlesinger Y, Straussman R, Keshet I, et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 2007; 39: 232-6.
66. Vire E, Brenner C, Deplus R, et al. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 2006; 439: 871-4.
67. Widschwendter M, Fiegl H, Egle D, et al. Epigenetic stem cell signature in cancer. Nat Genet 2007; 39: 157-8.
68. Fraga MF, Ballestar E, Villar-Garea A, et al. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat Genet 2005; 37: 391-400.
69. Kondo Y, Shen L, Suzuki S, et al. Alterations of DNA methylation and histone modifications contribute to gene silencing in hepatocellular carcinomas. Hepatol Res 2007; 37: 974-83.
70. Lin JC, Jeong S, Liang G, et al. Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. Cancer Cell 2007; 12: 432-44.
71. Tazi J, Bird A. Alternative chromatin structure at CpG islands. Cell 1990; 60: 909-20.
72. Dong KB, Maksakova IA, Mohn F, et al. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. EMBO J 2008; 27: 2691-701.
73. Kim JK, Esteve PO, Jacobsen SE, Pradhan S. UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. Nucleic Acids Res 2009; 37: 493-505.
74. Ng HH, Bird A. DNA methylation and chromatin modification. Curr Opin Genet Dev 1999; 9: 158-63.
75. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007; 128: 683-92.
76. Issa JP, Kantarjian HM. Targeting DNA methylation. Clin Cancer Res 2009; 15: 3938-46.
77. Flotho C, Claus R, Batz C, et al. The DNA methyltransferase inhibitors azacitidine, decitabine and zebularine exert differential effects on cancer gene expression in acute myeloid leukemia cells. Leukemia 2009; 23: 1019-28.
78. Cihak A. Biological effects of 5-azacytidine in eukaryotes. Oncology 1974; 30: 405-22.

79. Momparler RL, Rivard GE, Gyger M. Clinical trial on 5-aza-2'-deoxycytidine in patients with acute leukemia. Pharmacol Ther 1985; 30: 277-86.
80. Jones PA, Taylor SM. Cellular differentiation, cytidine analogs and DNA methylation. Cell 1980; 20: 85-93.
81. Mund C, Hackanson B, Stresemann C, Lubbert M, Lyko F. Characterization of DNA demethylation effects induced by 5-Aza-2'-deoxycytidine in patients with myelodysplastic syndrome. Cancer Res 2005; 65: 7086-90.
82. Gore SD, Baylin S, Sugar E, et al. Combined DNA methyltransferase and histone deacetylase inhibition in the treatment of myeloid neoplasms. Cancer Res 2006; 66: 6361-9.
83. Fandy TE, Herman JG, Kerns P, et al. Early epigenetic changes and DNA damage do not predict clinical response in an overlapping schedule of 5-azacytidine and entinostat in patients with myeloid malignancies. Blood 2009; 114: 2764-73.
84. McGarvey KM, Fahrner JA, Greene E, Martens J, Jenuwein T, Baylin SB. Silenced tumor suppressor genes reactivated by DNA demethylation do not return to a fully euchromatic chromatin state. Cancer Res 2006; 66: 3541-9.
85. Coffee B, Zhang F, Ceman S, Warren ST, Reines D. Histone modifications depict an aberrantly heterochromatinized FMR1 gene in fragile x syndrome. Am J Hum Genet 2002; 71: 923-32.
86. Oki Y, Aoki E, Issa JP. Decitabine--bedside to bench. Crit Rev Oncol Hematol 2007; 61: 140-52.
87. Schrump DS. Cytotoxicity mediated by histone deacetylase inhibitors in cancer cells: mechanisms and potential clinical implications. Clin Cancer Res 2009; 15: 3947-57.
88. Spange S, Wagner T, Heinzel T, Kramer OH. Acetylation of non-histone proteins modulates cellular signalling at multiple levels. Int J Biochem Cell Biol 2009; 41: 185-98.
89. Cameron EE, Bachman KE, Myohanen S, Herman JG, Baylin SB. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. Nat Genet 1999; 21: 103-7.
90. Marks PA, Richon VM, Miller T, Kelly WK. Histone deacetylase inhibitors. Adv Cancer Res 2004; 91: 137-68.
91. Bueso-Ramos CE, Cortes J, Talpaz M, et al. Imatinib mesylate therapy reduces bone marrow fibrosis in patients with chronic myelogenous leukemia. Cancer 2004; 101: 332-6.
92. Shaker S, Bernstein M, Momparler LF, Momparler RL. Preclinical evaluation of antineoplastic activity of inhibitors of DNA methylation (5-aza-2'-deoxycytidine) and histone deacetylation (trichostatin A, depsipeptide) in combination against myeloid leukemic cells. Leuk Res 2003; 27: 437-44.
93. Vertino PM, Yen RW, Gao J, Baylin SB. De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5-)-methyltransferase. Mol Cell Biol 1996; 16: 4555-65.
94. Feltus FA, Lee EK, Costello JF, Plass C, Vertino PM. Predicting aberrant CpG island methylation. Proc Natl Acad Sci U S A 2003; 100: 12253-8.

95. McConnell BB, Vertino PM. Activation of a caspase-9-mediated apoptotic pathway by subcellular redistribution of the novel caspase recruitment domain protein TMS1. Cancer Res 2000; 60: 6243-7.
96. Virmani A, Rathi A, Sugio K, et al. Aberrant methylation of TMS1 in small cell, non small cell lung cancer and breast cancer. Int J Cancer 2003; 106: 198-204.
97. McConnell BB, Vertino PM. TMS1/ASC: the cancer connection. Apoptosis 2004; 9: 5-18.
98. Parsons MJ, Vertino PM. Dual role of TMS1/ASC in death receptor signaling. Oncogene 2006.
99. Parsons MJ, Patel P, Brat DJ, Colbert L, Vertino PM. Silencing of TMS1/ASC promotes resistance to anoikis in breast epithelial cells. Cancer Res 2009; 69: 1706-11.
100. Stimson KM, Vertino PM. Methylation-mediated silencing of TMS1/ASC is accompanied by histone hypoacetylation and CpG island-localized changes in chromatin architecture. J Biol Chem 2002; 277: 4951-8.
101. Lucas ME, Crider KS, Powell DR, Kapoor-Vazirani P, Vertino PM. Methylation-sensitive regulation of TMS1/ASC by the Ets factor, GA-binding protein-alpha. J Biol Chem 2009; 284: 14698-709.
102. Taylor SM. p53 and deregulation of DNA methylation in cancer. Cellscience Review 2006; 2: 1742-8130.
103. Kapoor-Vazirani P, Kagey JD, Powell DR, Vertino PM. Role of hMOF-dependent histone H4 lysine 16 acetylation in the maintenance of TMS1/ASC gene activity. Cancer Res 2008; 68: 6810-21.
104. Antequera F, Boyes J, Bird A. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. Cell 1990; 62: 503-14.
105. Jones PA, Gonzalgo ML, Tsutsumi M, Bender CM. DNA methylation in bladder cancer. Eur Urol 1998; 33 Suppl 4: 7-8.
106. Laird PW. Cancer epigenetics. Hum Mol Genet 2005; 14 Spec No 1: R65-76.
107. Davie JR, Spencer VA. Control of histone modifications. J Cell Biochem 1999; Suppl 32-33: 141-8.
108. Bernstein BE, Kamal M, Lindblad-Toh K, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell 2005; 120: 169-81.
109. Weber M, Hellmann I, Stadler MB, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007; 39: 457-66.
110. Heard E, Rougeulle C, Arnaud D, Avner P, Allis CD, Spector DL. Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. Cell 2001; 107: 727-38.
111. Plath K, Fang J, Mlynarczyk-Evans SK, et al. Role of histone H3 lysine 27 methylation in X inactivation. Science 2003; 300: 131-5.
112. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet 2007; 8: 286-98.
113. Ohtsuka T, Ryu H, Minamishima YA, et al. ASC is a Bax adaptor and regulates the p53-Bax mitochondrial apoptosis pathway. Nat Cell Biol 2004; 6: 121-8.
114. Masumoto J, Taniguchi S, Ayukawa K, et al. ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells. J Biol Chem 1999; 274: 33835-8.

115. Moriai R, Tsuji N, Kobayashi D, et al. A proapoptotic caspase recruitment domain protein gene, TMS1, is hypermethylated in human breast and gastric cancers. Anticancer Res 2002; 22: 4163-8.
116. Yokoyama T, Sagara J, Guan X, et al. Methylation of ASC/TMS1, a proapoptotic gene responsible for activating procaspase-1, in human colorectal cancer. Cancer Lett 2003; 202: 101-8.
117. Stone AR, Bobo W, Brat DJ, Devi NS, Van Meir EG, Vertino PM. Aberrant methylation and down-regulation of TMS1/ASC in human glioblastoma. Am J Pathol 2004; 165: 1151-61.
118. Levine JJ, Stimson-Crider KM, Vertino PM. Effects of methylation on expression of TMS1/ASC in human breast cancer cells. Oncogene 2003; 22: 3475-88.
119. Kirmizis A, Bartley SM, Kuzmichev A, et al. Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev 2004; 18: 1592-605.
120. Dou Y, Milne TA, Tackett AJ, et al. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell 2005; 121: 873-85.
121. Schotta G, Lachner M, Sarma K, et al. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev 2004; 18: 1251-62.
122. Smith ER, Cayrou C, Huang R, Lane WS, Cote J, Lucchesi JC. A human protein complex homologous to the Drosophila MSL complex is responsible for the majority of histone H4 acetylation at lysine 16. Mol Cell Biol 2005; 25: 9175-88.
123. Taipale M, Rea S, Richter K, et al. hMOF histone acetyltransferase is required for histone H4 lysine 16 acetylation in mammalian cells. Mol Cell Biol 2005; 25: 6798-810.
124. Gupta A, Guerin-Peyrou TG, Sharma GG, et al. The mammalian ortholog of Drosophila MOF that acetylates histone H4 lysine 16 is essential for embryogenesis and oncogenesis. Mol Cell Biol 2008; 28: 397-409.
125. Pfister S, Rea S, Taipale M, et al. The histone acetyltransferase hMOF is frequently downregulated in primary breast carcinoma and medulloblastoma and constitutes a biomarker for clinical outcome in medulloblastoma. Int J Cancer 2008; 122: 1207-13.
126. Graff JR, Herman JG, Myohanen S, Baylin SB, Vertino PM. Mapping patterns of CpG island methylation in normal and neoplastic cells implicates both upstream and downstream regions in de novo methylation. J Biol Chem 1997; 272: 22322-9.
127. Ottaviano YL, Issa JP, Parl FF, Smith HS, Baylin SB, Davidson NE. Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells. Cancer Res 1994; 54: 2552-5.
128. Nguyen CT, Weisenberger DJ, Velicescu M, et al. Histone H3-lysine 9 methylation is associated with aberrant gene silencing in cancer cells and is rapidly reversed by 5-aza-2'-deoxycytidine. Cancer Res 2002; 62: 6456-61.
129. Kondo Y, Shen L, Issa JP. Critical role of histone methylation in tumor suppressor gene silencing in colorectal cancer. Mol Cell Biol 2003; 23: 206-15.
130. Johnson CA, O'Neill LP, Mitchell A, Turner BM. Distinctive patterns of histone H4 acetylation are associated with defined sequence elements within both heterochromatic and euchromatic regions of the human genome. Nucleic Acids Res 1998; 26: 994-1001.

131. Grunstein M. Yeast heterochromatin: regulation of its assembly and inheritance by histones. Cell 1998; 93: 325-8.
132. Suka N, Luo K, Grunstein M. Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. Nat Genet 2002; 32: 378-83.
133. Matangkasombut O, Buratowski S. Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. Mol Cell 2003; 11: 353-63.
134. Corona DF, Clapier CR, Becker PB, Tamkun JW. Modulation of ISWI function by site-specific histone acetylation. EMBO Rep 2002; 3: 242-7.
135. Pruitt K, Zinn RL, Ohm JE, et al. Inhibition of SIRT1 reactivates silenced cancer genes without loss of promoter DNA hypermethylation. PLoS Genet 2006; 2: e40.
136. Macleod D, Charlton J, Mullins J, Bird AP. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev 1994; 8: 2282-92.
137. Gal-Yam EN, Jeong S, Tanay A, Egger G, Lee AS, Jones PA. Constitutive nucleosome depletion and ordered factor assembly at the GRP78 promoter revealed by single molecule footprinting. PLoS Genet 2006; 2: e160.
138. Gupta A, Sharma GG, Young CS, et al. Involvement of human MOF in ATM function. Mol Cell Biol 2005; 25: 5292-305.
139. Kanduri M, Kanduri C, Mariano P, et al. Multiple nucleosome positioning sites regulate the CTCF-mediated insulator function of the H19 imprinting control region. Mol Cell Biol 2002; 22: 3339-44.
140. Gomez JA, Majumder P, Nagarajan UM, Boss JM. X box-like sequences in the MHC class II region maintain regulatory function. J Immunol 2005; 175: 1030-40.
141. Docquier F, Farrar D, D'Arcy V, et al. Heightened expression of CTCF in breast cancer cells is associated with resistance to apoptosis. Cancer Res 2005; 65: 5112-22.
142. West AG, Huang S, Gaszner M, Litt MD, Felsenfeld G. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. Mol Cell 2004; 16: 453-63.
143. Recillas-Targa F, De La Rosa-Velazquez IA, Soto-Reyes E, Benitez-Bribiesca L. Epigenetic boundaries of tumour suppressor gene promoters: the CTCF connection and its role in carcinogenesis. J Cell Mol Med 2006; 10: 554-68.
144. Kurukuti S, Tiwari VK, Tavoosidana G, et al. CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. Proc Natl Acad Sci U S A 2006; 103: 10684-9.
145. Bao L, Zhou M, Cui Y. CTCFBSDB: a CTCF-binding site database for characterization of vertebrate genomic insulators. Nucleic Acids Res 2008; 36: D83-7.
146. Recillas-Targa F, Pikaart MJ, Burgess-Beusse B, et al. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. Proc Natl Acad Sci U S A 2002; 99: 6883-8.
147. Kuhn RM, Karolchik D, Zweig AS, et al. The UCSC Genome Browser Database: update 2009. Nucleic Acids Res 2009; 37: D755-61.
148. Kim TH, Abdullaev ZK, Smith AD, et al. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 2007; 128: 1231-45.
149. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science 2002; 295: 1306-11.

150. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001; 409: 860-921.
151. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature 2004; 429: 457-63.
152. Fahrner JA, Eguchi S, Herman JG, Baylin SB. Dependence of histone modifications and gene expression on DNA hypermethylation in cancer. Cancer Res 2002; 62: 7213-8.
153. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet 2002; 3: 415-28.
154. Silverman LR, Demakos EP, Peterson BL, et al. Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B. J Clin Oncol 2002; 20: 2429-40.
155. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. Proc Natl Acad Sci U S A 1996; 93: 9821-6.
156. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res 1997; 25: 2532-4.
157. Bock C, Reither S, Mikeska T, Paulsen M, Walter J, Lengauer T. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics 2005; 21: 4067-8.
158. Zinn RL, Pruitt K, Eguchi S, Baylin SB, Herman JG. hTERT is expressed in cancer cell lines despite promoter DNA methylation by preservation of unmethylated DNA and active chromatin around the transcription start site. Cancer Res 2007; 67: 194-201.
159. McGarvey KM, Van Neste L, Cope L, et al. Defining a chromatin pattern that characterizes DNA-hypermethylated genes in colon cancer cells. Cancer Res 2008; 68: 5753-9.
160. Coffee B, Zhang F, Warren ST, Reines D. Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. Nat Genet 1999; 22: 98-101.
161. Nass SJ, Herman JG, Gabrielson E, et al. Aberrant methylation of the estrogen receptor and E-cadherin 5' CpG islands increases with malignant progression in human breast cancer. Cancer Res 2000; 60: 4346-8.
162. Ghoshal K, Datta J, Majumder S, et al. 5-Aza-deoxycytidine induces selective degradation of DNA methyltransferase 1 by a proteasomal pathway that requires the KEN box, bromo-adjacent homology domain, and nuclear localization signal. Mol Cell Biol 2005; 25: 4727-41.
163. Bachman KE, Park BH, Rhee I, et al. Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. Cancer Cell 2003; 3: 89-95.
164. Egger G, Aparicio AM, Escobar SG, Jones PA. Inhibition of histone deacetylation does not block resilencing of p16 after 5-aza-2'-deoxycytidine treatment. Cancer Res 2007; 67: 346-53.
165. Takeshima H, Yamashita S, Shimazu T, Niwa T, Ushijima T. The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands. Genome Res 2009.
166. Yang H, Mizzen CA. The multiple facets of histone H4-lysine 20 methylation. Biochem Cell Biol 2009; 87: 151-61.

167. Pannetier M, Julien E, Schotta G, et al. PR-SET7 and SUV4-20H regulate H4 lysine-20 methylation at imprinting control regions in the mouse. EMBO Rep 2008; 9: 998-1005.
168. Oda H, Okamoto I, Murphy N, et al. Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. Mol Cell Biol 2009; 29: 2278-95.
169. Schotta G, Sengupta R, Kubicek S, et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. Genes Dev 2008; 22: 2048-61.
170. Braiteh F, Soriano AO, Garcia-Manero G, et al. Phase I study of epigenetic modulation with 5-azacytidine and valproic acid in patients with advanced cancers. Clin Cancer Res 2008; 14: 6296-301.
171. Bender CM, Gonzalgo ML, Gonzales FA, Nguyen CT, Robertson KD, Jones PA. Roles of cell division and gene transcription in the methylation of CpG islands. Mol Cell Biol 1999; 19: 6690-8.
172. Jabbour E, Issa JP, Garcia-Manero G, Kantarjian H. Evolution of decitabine development: accomplishments, ongoing investigations, and future strategies. Cancer 2008; 112: 2341-51.
173. Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol 2007; 27: 3769-79.
174. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol 1987; 196: 261-82.
175. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A 2002; 99: 3740-5.
176. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95: 14863-8.
177. Tamayo P, Slonim D, Mesirov J, et al. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc Natl Acad Sci U S A 1999; 96: 2907-12.
178. Jones PA, Takai D. The role of DNA methylation in mammalian epigenetics. Science 2001; 293: 1068-70.
179. Bender CM, Pao MM, Jones PA. Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines. Cancer Res 1998; 58: 95-101.
180. Cao Q, Yu J, Dhanasekaran SM, et al. Repression of E-cadherin by the polycomb group protein EZH2 in cancer. Oncogene 2008; 27: 7274-84.
181. Cortese R, Hartmann O, Berlin K, Eckhardt F. Correlative gene expression and DNA methylation profiling in lung development nominate new biomarkers in lung cancer. Int J Biochem Cell Biol 2008; 40: 1494-508.
182. Cheng JC, Weisenberger DJ, Gonzales FA, et al. Continuous zebularine treatment effectively sustains demethylation in human bladder cancer cells. Mol Cell Biol 2004; 24: 1270-8.
183. Daury L, Chailleux C, Bonvallet J, Trouche D. Histone H3.3 deposition at E2F-regulated genes is linked to transcription. EMBO Rep 2006; 7: 66-71.

184. Rodriguez J, Munoz M, Vives L, Frangou CG, Groudine M, Peinado MA. Bivalent domains enforce transcriptional memory of DNA methylated genes in cancer cells. Proc Natl Acad Sci U S A 2008; 105: 19809-14.
185. Kamb A, Gruis NA, Weaver-Feldhaus J, et al. A cell cycle regulator potentially involved in genesis of many tumor types. Science 1994; 264: 436-40.
186. Wheeler BS, Blau JA, Willard HF, Scott KC. The impact of local genome sequence on defining heterochromatin domains. PLoS Genet 2009; 5: e1000453.
187. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev 2002; 16: 6-21.
188. Ellis L, Atadja PW, Johnstone RW. Epigenetics in cancer: targeting chromatin modifications. Mol Cancer Ther 2009; 8: 1409-20.