# EPIGENETIC MECHANISMS OF GENE REGULATION IN HUMAN BREAST CANCER

Ashley Garrett Rivenbark

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum of Toxicology.

Chapel Hill
2007

Approved by:

William B. Coleman, Ph.D.

Channing J. Der, Ph.D.

William K. Funkhouser, M.D., Ph.D.

Wendell D. Jones, Ph.D.

William K. Kaufmann, Ph.D.

## ABSTRACT

Ashley Garrett Rivenbark: Epigenetic Mechanisms of Gene Regulation in Human Breast Cancer
(Under the direction of William B. Coleman, Ph.D.)

Breast cancer represents a significant health problem and improvements in our ability to prevent, diagnose, and treat the disease requires a greater understanding of the molecular basis of breast carcinogenesis. Epigenetic mechanisms play a major role in breast carcinogenesis, with DNA methylation accounting for most epigenetic gene silencing, affecting a number of different gene targets. However, mechanisms of DNA methylation-dependent silencing are poorly understood. To identify epigenetically-regulated genes in breast cancer, MCF-7 breast cancer cells were exposed to demethylating treatment and gene expression patterns were examined by microarray analysis. Genes with increased expression after demethylation treatment that returned to control levels after treatment withdrawal were directly assessed for DNA methylation by bisulfite sequencing. A group of 20 putative methylation-sensitive genes were identified that could be classified into three groups based upon their promoter CpG features. The majority of these methylation-sensitive genes lacked a conventional DNA methylation target (CpG island), resulting in an expanded model for epigenetic regulation of gene expression that recognizes the importance of all promoter CpGs. The breast tumor suppressor gene *CST6* (Cystatin M) is epigenetically silenced in MCF-7 breast cancer cells. *CST6* is subject to methylation-dependent regulation in multiple breast cancer cell lines, primary breast tumors, and lymph node metastases, and gene expression status correlates with promoter hypermethylation. These results suggest that

methylation dependent gene silencing of *CST6* represents an important mechanism for loss of *CST6* during breast carcinogenesis. The mechanisms that control CpG island methylation are poorly understood. *CST6* was utilized as an index gene for the identification of *cis* elements that direct promoter CpG methylation. The methylation-sensitive *CST6* promoter was assembled into luciferase reporter constructs and transfected into model breast cancer cell lines that methylate or do not methylate the *CST6* promoter. Truncation of the *CST6* promoter disassociated a putative instructional *cis* regulatory sequence located in the 5' upstream promoter region of *CST6* that functions to direct CpG methylation. The observations and results described in this dissertation significantly advance our understanding of methylation-sensitive genes and mechanisms governing DNA methylation in breast carcinogenesis.

## ACKNOWLEDGMENTS

First and foremost I would like to thank my advisor Bill Coleman, for his guidance, wisdom, support, and encouragement. He has enriched my experience in graduate school by allowing me to help him with many endeavors, including writing grants, reviews, and manuscript commentaries. My achievements and success in graduate school are mostly due to his mentorship and support. I will forever be grateful to him.

Thank you to all the people in my lab who have made coming to work exciting, always interesting, and mostly enjoyable. I want to give special thanks to Devon, who has been my 'little sister' in the lab. Teaching and learning from her has been a joy. I would like to express my gratitude to my dissertation committee members, Channing J. Der, William K. Funkhouser, Wendell D. Jones, and William K. Kaufmann. I appreciate all of their time, effort, support, and encouragement. I want to especially thank Wendell for all of his microarray analysis and statistical help.

I want to give special thanks to my parents, who have taught me to live my life according to God's will, and have encouraged and supported me throughout all endeavors. Thank you for your prayers. I love you both so much. Thank you to my brother, Justin, for supporting me in graduate school, always keeping things in perspective, and making me laugh in all situations. I want to thank Patsy, Dannie, and Ginnie for their enthusiasm and support throughout the years I have known them. Thanks to my friends who have been my support system and prayer warriors throughout graduate school (and most of my life), Stephanie, Megan, and especially Laura. Thank you Laura for being so understanding, encouraging,

enthusiastic, and someone I can count on for anything. I am so blessed to have you as a friend!

Most importantly, thank you to my husband and best friend, Jason, for your unconditional love, support, wisdom, advice, and uncanny ability to understand lab situations. I truly appreciate all you have done for me. I love you.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| 5-aza | 5-aza-2'-deoxycytidine |
| AH | Atypical hyperplasia |
| bp | Basepair |
| cm | Centimeter |
| cDNA | Complementary deoxyribonucleic acid |
| *CST6* | Cystatin M |
| CK18 | Cytokeratin 18 |
| dNTP | Deoxynucleosidetriphosphate |
| DNA | Deoxyribonucleic acid |
| DCIS | Ductal carcinoma *in situ* |
| EDTA | Ethylenedinitrilotetraacetic acid |
| EST | Expressed sequence tag |
| IHC | Immunohistochemistry |
| IDC | Invasive ductal carcinoma |
| kDa | Kilodalton |
| mRNA | Messenger ribonucleic acid |
| μg | Microgram |
| ml | Milliliter |
| mm | Millimeter |
| mM | Millimolar |
| ng | Nanogram |
| nM | Nanomolar |

| | |
|---|---|
| Neo | Neomycin |
| PCR | Polymerase chain reaction |
| PDNN | Positional dependent nearest neighbor |
| RNA | Ribonucleic acid |
| NaCl | Sodium chloride |
| S.E.M. | Standard error of the mean |
| TMA | Tissue microarray |
| TSA | Trichostatin A |
| TMI | Total methylation index |
| U | Units |

# I. INTRODUCTION

## A. Breast Cancer

### *Breast Cancer Epidemiology*

Cancer of the breast is the most common malignant neoplasm among women in the United States and the state of North Carolina. An estimated 178,000 new cases of breast cancer among women will be diagnosed in the United States in 2007, accounting for 26% of all new cancer cases among women (1). During the same period, 4870 new cases of invasive breast cancer will be diagnosed in North Carolina (1). Based on incidence rates from 2001 to 2003, approximately 13% of women will be diagnosed with breast cancer in their lifetime (2). In the 1980s, the number of new cases of breast cancer increased among women approximately 4% per year, due in part to the heightened surveillance of women in the general population using mammography, resulting in earlier breast cancer diagnosis (3). The incidence of breast cancer rates among women plateaued between 2001 and 2003, possibly due to saturation of early mammography screening and reduced use of hormone replacement therapy (1). Between 2000 and 2003, the median age of women diagnosed with breast cancer was 61 years of age, and approximately 58% of women diagnosed were between the ages of 20 and 64 (2). The majority (61%) of breast cancer cases are diagnosed when the tumor is confined to the primary site (breast) (2). In part due to early detection, the 5-year survival for breast cancer patients diagnosed between 1996 and 2002 was approximately 90%

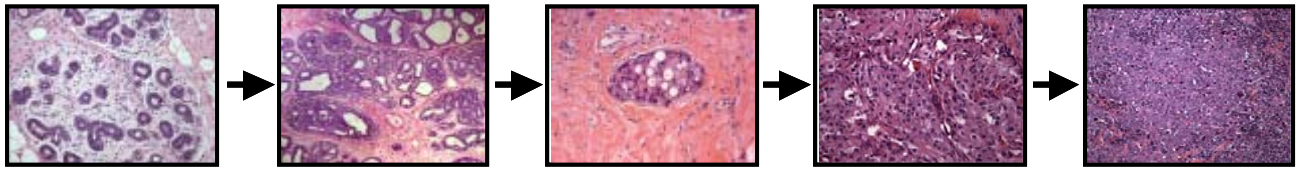
(2). However, it has been documented that breast cancer can recur after long periods of time, well after five years from the initial diagnosis (3).

In 2007, an estimated 40,000 women will die in the United States from breast cancer (1). Among females ages 20 to 59, breast cancer is the leading cause of cancer-related deaths (1). In North Carolina, 1240 breast cancer-related deaths will occur in 2007, accounting for 7% of cancer-related deaths in the state (1). Between 1990 and 2003, death rates from cancer decreased 8.5% among women, and the reduction of breast and colorectal cancer deaths combined accounted for over 60% of the decrease among women (1). The reduction in breast cancer-related deaths directly reflects improvements in early detection and therapeutic treatments.

***Natural History of Breast Cancer***

The natural history of breast cancer is characterized by a progression of preneoplastic lesions, benign neoplastic disease, and culminating in malignant disease. Figure 1 shows a highly simplified schematic representation of the natural history of breast cancer development and progression (4). Breast hyperplasia is a recognized preneoplastic lesion, and the risk for developing invasive breast cancer increases with the presence of atypia or dysplastic components. Hyperplastic lesions without atypia are less problematic, and are associated with only slightly increased risk for breast cancer development (4). Atypical hyperplasia (AH) shows some characteristics of *in situ* carcinoma, and can present as either ductal or lobular in type (4) (Figure 1). Ductal carcinoma *in situ* (DCIS) is a well-known and characterized precursor to invasive breast cancer (Figure 1). The majority of invasive breast carcinomas have a DCIS component, suggesting that DCIS is an important precursor to

**Figure 1. The natural history of breast cancer.** Representative H&E stained images corresponding to the individual stages of breast cancer development and progression are shown. The cellular changes that characterize breast tumorigenesis include preneoplastic lesions and benign lesions that confer an increased risk for development of invasive breast cancer. Invasive breast cancers will proliferate and grow destroying the surrounding breast architecture, leading to local invasion of normal tissue and eventually dissemination to distant sites, giving rise to metastatic tumors.



Normal Breast Epithelium

Atypical Breast Hyperplasia

Ductal Carcinoma *In Situ*

Invasive Ductal Carcinoma

Metastatic Disease

advanced disease (4). However, some investigators have proposed a direct transition from normal breast epithelium to malignant epithelium (5). Most invasive breast cancers (approximately 90%) are of the ductal or lobular histopathological type (6), and invasive ductal carcinoma (IDC) is the most frequently occurring breast cancer among women (4). It is now accepted that there are two major pathways of multi-step breast cancer progression, (i) well-differentiated DCIS progressing to grade I IDC, and (ii) poorly-differentiated DCIS progressing to grade III IDC (7). High grade (poorly-differentiated) DCIS is associated with necrosis, apoptosis, and cellular proliferation (6). Changes in the molecular pattern of DCIS lesions may lead to the ability to collapse the myoepithelium, escape the ductal structure, and invade the surrounding stroma forming an invasive carcinoma (6,8) (Figure 1). These IDC lesions will proliferate and grow, destroying surrounding stroma, and breast architecture. Continued disease progression can lead to tumor dissemination via lymphatic or hematogenous routes giving rise to metastatic lesions in distant organs (9) (Figure 1). Metastatic breast cancers have a tendency to metastasize to bone, lung, skin, and lymph nodes (10).

**B. Molecular Pathogenesis of Breast Cancer**

Breast cancer is a heterogeneous disease that results from the accumulation of a complex series of genetic and epigenetic events driving divergent pathways that ultimately convey varying phenotypic properties to individual neoplastic lesions. Numerous molecular markers have been examined for their predictive value in breast cancer prognostication, but histopathologic grade emerges as the most important indicator of long-term patient outcome

(6,7). However, histopathologic grade generally correlates with the expression of genes associated with increased cell proliferation (*Ki-67*, *p53*), growth (*HER-2*), and invasiveness (matrix metalloproteinases) (11,12). In contrast, low-grade breast tumors express genes associated with low cellular proliferation (*p27*) and differentiation (*ER* and *PR*) (6,13). At present, the molecular mechanisms that control tumor progression, stromal invasion, and distant metastasis are poorly understood. Nevertheless, the role of specific genes that contribute to breast tumor invasion and metastasis are beginning to be investigated and characterized.

***Breast Cancer Susceptibility Genes***

Family history constitutes a strong and independent predictor of the development of breast cancer. Women who have a family tree of relatives that have developed breast cancer exhibit a greater probability of developing breast cancer when compared to the general population. Therefore, a substantial amount of research has focused on identifying breast cancer susceptibility genes. However, only 5-10% of total breast cancer incidence is associated with genetic predisposition (4,14). Genes that confer breast cancer susceptibility include, *BRCA1*, *BRCA2*, and *p53* (15). The inheritance of a mutation in *BRCA1* and *BRCA2* genes confers a lifetime risk of breast cancer of 50-85% (16). The major functions of these protein products are DNA repair and homologous recombination. Mutations in *BRCA1* and *BRCA2* are found interspersed throughout the coding region, and the most common germline mutations found are frameshift mutations that result in the truncation of the protein product (4). Breast cancers that exhibit mutations in *BRCA1* and *BRCA2* are characterized by a large number of chromosome alterations (16). However, mutations in *BRCA1* and *BRCA2* only

account for a small percentage of familial susceptibility. In non-*BRCA1*/*BRCA2* breast cancer families, termed *BRCAx*, very little is known related to the genetic basis of inherited susceptibility (16). Histopathological studies have shown that these tumors are of lower grade and lower mitotic activity compared to breast tumors related to *BRCA1* and *BRCA2* mutation (17). Patients with Li-Fraumeni cancer-predisposition syndrome have germline mutations in the *p53* gene (18,19). Breast cancer is one of the neoplasms that affect these patients, and is characterized by early-onset, bilaterality, and association with other familial cancers (4).

***Environmental and Epigenetic Factors of Breast Cancer Susceptibility***

Although there is overwhelming evidence that breast cancer is essentially a genetically based disease, environmental and epigenetic factors play an important role in breast cancer development. However, environmental and epigenetic influences are not well understood. The major risk factors for breast cancer development include: advancing age (over 50 years of age), early age at menarche, first childbirth after the age of 35, late age at menopause, nulliparity, obesity, dietary factors (such as high-fat diets), and exposure to high-dose radiation to the chest before age 35 (20-23). Recent evidence suggests that epigenetic mechanisms play a major role in breast carcinogenesis (24). Epigenetic alterations differ from genetic alterations in that they arise more frequently, are reversible, and occur at defined regions of specific genes (25).

## C. Mechanisms of Epigenetic Regulation in Carcinogenesis

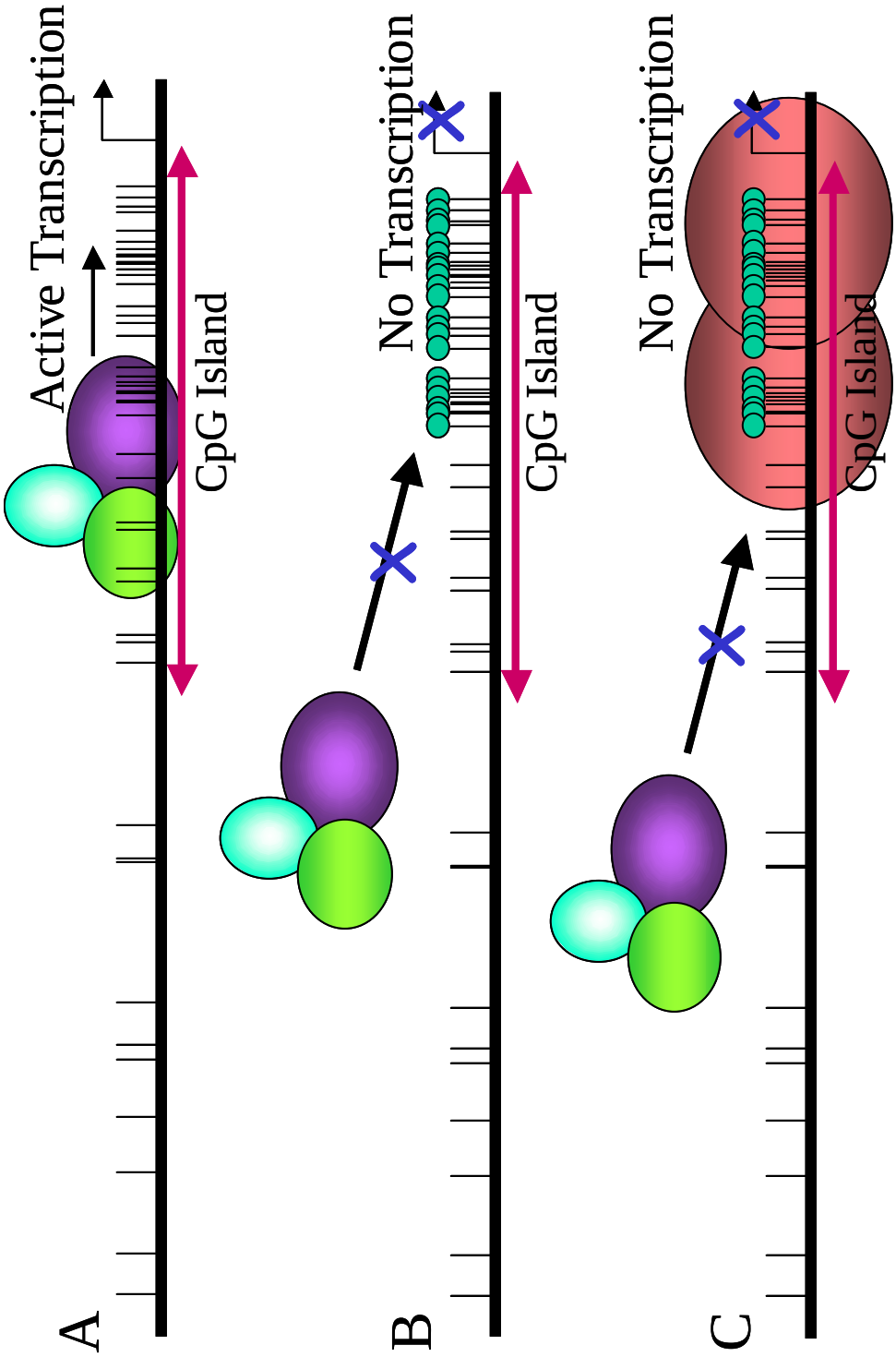
### *DNA Methylation in Cancer*

Neoplastic transformation is associated with alterations in DNA methylation, including both global hypomethylation and gene-specific hypermethylation (26-28). Hypomethylation of cancer cell genomes is associated with loss of methylation in CpG-depleted regions where most CpG dinucleotides would be expected to be methylated (29-31). The loss of methylation in these regions of the genome may be associated with aberrant or inappropriate expression of some genes that could contribute to neoplastic transformation, tumorigenesis, or cancer progression (32). In addition, genome-wide demethylation can contribute to chromosomal instability by destabilizing pericentromeric regions of certain chromosomes (33-35). Gains in DNA methylation in cancer cells typically reflect hypermethylation of CpG islands in gene promoter regions, which can lead to gene silencing (26). Methylation-dependent gene silencing is a normal mechanism for regulation of gene expression (36). However, in cancer cells methylation-dependent epigenetic gene silencing represents a mutation-independent mechanism for inactivation of tumor suppressor genes (37) (Figure 2). A significant number of cancer-related genes have been identified that are subject to methylation-dependent silencing (38), and many of these genes contribute to the hallmarks of cancer (39). These observations combine to strongly suggest that epigenetic events, and particularly those involving DNA methylation, represent fundamental aspects of cancer, and play key roles in neoplastic transformation and progression.

### *DNA Methylation in Human Breast Cancer*

It is now well recognized that epigenetic mechanisms play a major role in neoplastic transformation of breast epithelium and tumor progression (24,25). DNA methylation is a

**Figure 2. Alteration of gene expression by promoter CpG methylation.** A gene promoter CpG island located proximal to the transcription start site (indicated by the bent arrow) is depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides), including binding sites for transcription factor proteins (blue, green, and purple circles). (A) Lack of CpG island methylation allows transcription factors to bind to the gene promoter to facilitate gene transcription. (B) Promoter CpG island methylation (represented by green lollipops), inhibits transcription factor binding, resulting in inhibition of gene expression (methylation-dependent silencing). (C) Methylated DNA binding proteins (pink circles) bind to methylated CpG dinucleotides inhibiting transcription factor binding and resulting in inhibition of gene expression (methylation-dependent silencing).


Active Transcription
CpG Island
A
No Transcription
CpG Island
B
No Transcription
CpG Island
C

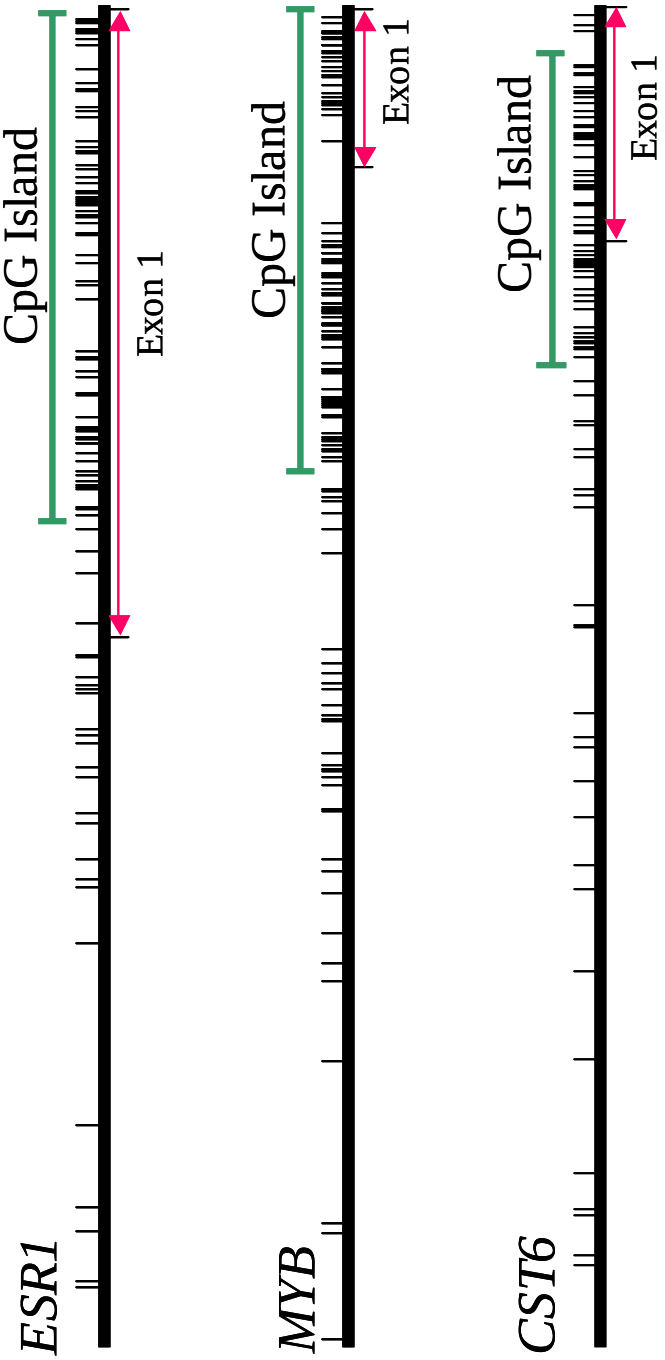
well known epigenetic mechanism, and a number of different genes have been shown to be inactivated in breast cancer through methylation-dependent gene silencing (25). Some of these genes are silenced through a direct effect of DNA methylation, while others are affected through indirect mechanisms. Genes that have been determined to be directly silenced by DNA methylation in breast cancer include cell cycle control genes (*p16$^{INK4a}$*), steroid receptor genes (*ERα*, *PR*, *RARβ2*), tumor suppressor genes (*BRCA1*), genes associated with cancer metastasis (*E-cadherin*, *TIMP-3*), and others (24,40-43). The p16$^{INK4a}$ cyclin-dependent kinase inhibitor is inactivated through methylation in several human cancers. In breast cancer, *p16$^{INK4a}$* is methylated in 20-30% of tumors and cell lines, with a concomitant loss of expression (44,45). Loss of *p16$^{INK4a}$* expression in this subset of breast cancers may contribute to unregulated cell proliferation and tumorigenesis. A significant percentage of breast cancers lack expression of the estrogen receptor (and other steroid receptors), but loss of *ER* gene expression is not associated with gene deletion or somatic mutation (46). Rather, methylation-dependent silencing of the *ER* gene is responsible for the loss of expression in these tumors (47,48). Somatic mutations of the *BRCA1* gene have not been documented in non-hereditary breast cancers (49). Therefore, an alternative mechanism for *BRCA1* inactivation involving DNA methylation was proposed (50-52). Subsequently, several studies have documented methylation-dependent epigenetic silencing of *BRCA1* in sporadic breast cancers (53-57). Loss of *E-cadherin* gene expression in breast cancer is associated with an aggressive tumor phenotype and decreased patient survival (58). Methylation-dependent loss of *E-cadherin* gene expression has been shown in 30% of primary breast cancers, and up to 60% of metastatic tumors (59). Loss of *TIMP-3* expression in breast tumors potentially results in increased proteolytic activity from matrix metalloproteinase

enzymes (60). The *TIMP-3* promoter is methylated in ~30% of primary breast cancers and breast cancer cell lines (61). Both of these methylation-related losses of gene expression are likely to contribute to tumor progression and spread.

***Targets of DNA Methylation***

DNA methylation occurs almost exclusively on cytosines within CpG dinucleotides, which are relatively rare in the genome, occurring at about 20% of the predicted frequency (25). However, regions of CpG density, termed CpG islands (62,63), occur in the promoter sequences of numerous genes, proximal to their transcription start site (64) (Figure 3). Some investigators have suggested that as many as 50% of all human genes may contain a promoter CpG island. These CpG islands are conventionally defined as >200 bp with >50% G+C and >0.6 CpG observed/CpG expected (65). Numerous studies have shown that there is a strong inverse correlation between promoter methylation status and gene expression levels (66,67). However, this inverse relationship has only been shown for methylation affecting promoter regions of genes, and not methylation that occurs in transcribed sequences (68). Studies that demonstrate extensive promoter CpG island methylation in genes that are transcriptionally silent, including imprinted genes like H19 (69), suggest that CpG island hypermethylation represents a normal mechanism for gene regulation. A significant number of CpG island containing genes have been shown to be silenced by methylation in breast cancer. The *14-3-3σ* gene is silenced in the majority of human breast cancers (94%) as a consequence of CpG island methylation (70). Likewise, the *BRCA1* gene contains a promoter CpG island that is frequently methylated in breast cancers that lack *BRCA1* expression (51,53,54). In addition, evidence for the importance of discrete methylation

**Figure 3. CpG island containing genes.** The distribution of CpG dinucleotides proximal to the transcription start site in the promoter and exon 1 of estrogen receptor 1 (*ESR1*), myeloblastosis viral oncogene (*MYB*), and cystatin M (*CST6*) are depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). CpG islands are indicated by green lines and are found in all genes. The CpG island is found in exon 1 (indicated by a pink arrow) of *ESR1*. In *MYB* and *CST6*, the CpG island is located in the proximal promoter and exon 1, spanning the transcription start site.


CpG Island
Exon 1
ESR1
CpG Island
Exon 1
MYB
CpG Island
Exon 1
CST6

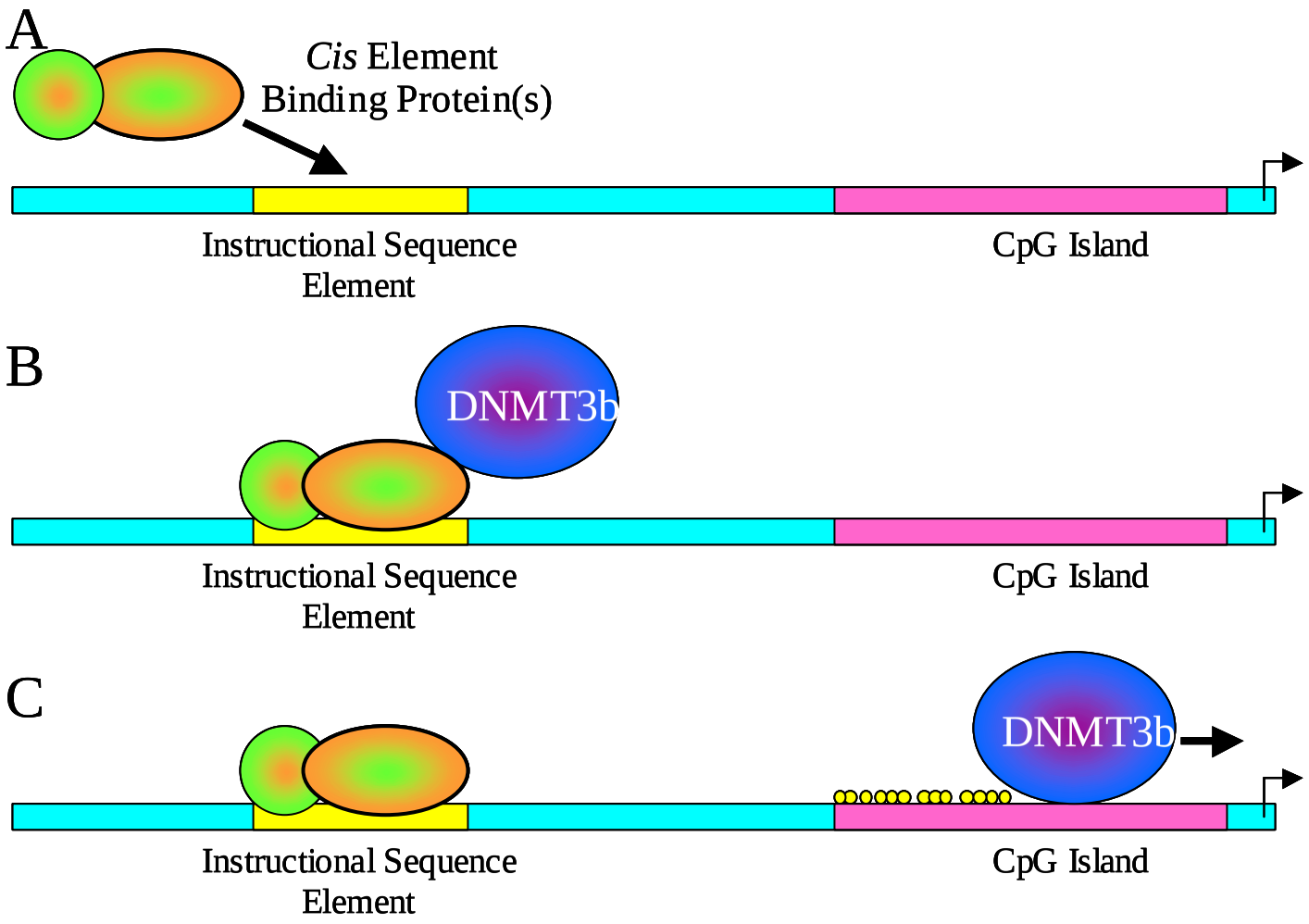
events within a larger methylation target (CpG island) has appeared in the literature. The *AP-2α* tumor suppressor gene is subject to methylation-dependent silencing through methylation of a discrete region that is contained within a larger CpG island (71). Numerous investigations have focused on methylation events that occur in CpG islands to characterize epigenetic changes in cancer (72-74). However, most of these studies acknowledge that a significant percentage (48% to 64%) of putative epigenetically-regulated genes lack these regions of CpG density (72-74), and that genes lacking CpG islands as a methylation target are frequently induced in response to demethylating drugs (75-77). Thus, it is probable that novel CpG targets for methylation are present in putative epigenetically-regulated genes that do not contain CpG islands. There is some evidence indicating that methylation events in promoters lacking CpG islands can result in down-regulation of gene expression (62). In fact, a number of studies have shown that methylation of novel CpG targets can result in epigenetic silencing of gene expression. Well-characterized examples of methylation-sensitive genes lacking CpG islands include *E-cadherin,* (25,78) *RAR-β2,* (79) *APC,* (80) and *LAMB3* (81-84). Combined, these findings from the literature suggest that targets for CpG methylation will include typical CpG islands, as well as novel methylation targets, such as specific CpG dinucleotides in critical gene regulatory regions.

### *Mechanisms of Regulation of DNA Methylation*

Both normal and cancer cells exhibit specific patterns of CpG methylation that reflect nonrandom hypermethylation of particular regions of DNA resulting in silencing of certain genes. The mechanisms that control this nonrandom distribution of CpG methylation are poorly understood. However, several lines of evidence support the notion that *cis*-acting sequence elements exist that regulate *de novo* methylation, including directive (methylation-

promoting) instructions (85,86) and protective (methylation-preventing) instructions (87) (Figure 4). It is conceivable that these directive and protective elements coexist in the promoter regions of epigenetically-regulated genes and that a balance between these forces dictate the methylation status of the promoter in specific cell types or under specific physiological conditions. A number of studies have addressed the question of whether *cis* elements direct DNA methylation of specific target genes (Figure 4). One of the most extensively studied genes is the mouse APRT gene. When an unmethylated copy of the APRT gene was transfected into APRT-deficient mouse embryonal carcinoma cells, the promoter region of the transfected gene acquired a methylation pattern that was identical to the endogenous methylation pattern of APRT confirming the presence of a *cis* element in the 5'-sequence of this gene (88,89). Deletion analysis of the mouse APRT gene localized the *cis* element to a 838 bp region in the promoter sequence (90,91). In a recent study, Feltus *et al*. performed a detailed sequence analysis of methylation-prone and methylation-resistant CpG islands to investigate the possibility that susceptibility to methylation might be conferred by *cis*-acting features of differing CpG islands (92). When general characteristics of CpG islands were evaluated (size, G+C content, CpG frequency), no significant differences were detected between methylation-prone and methylation-resistant CpG islands (92). However, using pattern recognition and supervised learning techniques to analyze sequences flanking CpG islands, a number of sequence elements were identified that predict methylation of promoter sequences with high discrimination potential (92). It is not known if these sequences function to direct or promote methylation.

**Figure 4. A model for *cis* element-mediated direction of DNA methylation.** A gene promoter is depicted schematically, including a CpG island (pink box) located proximal to the transcription start site (indicated by the bent arrow) and a putative instructional sequence element (yellow box). (A) *Cis* element binding proteins (indicated by green and orange circles) can recognize an instructional element within the upstream promoter region of a gene. (B) *Cis* element binding proteins bind to the instructional sequence element and recruit DNA methyltransferase 3b (DNMT3b, blue circle) to the gene promoter region. (C) Once recruited to the gene promoter, DNMT3b methylates (yellow circles) the target CpG island.


A
*Cis* Element
Binding Protein(s)
Instructional Sequence
Element
CpG Island
B
DNMT3b
Instructional Sequence
Element
CpG Island
C
DNMT3b
Instructional Sequence
Element
CpG Island

## D. Cystatins and Cancer: Methylation-sensitive Genes that Contribute to Breast Tumorigenesis and Progression

### *Cysteine Protease Inhibitors - Cystatins*

Cystatins function as cysteine protease inhibitors and were discovered in the 1960s with a report on a factor capable of inhibiting the clotting activity of a thiol-dependent protease in mammalian cells (93). Since that time, other groups identified cystatin proteins that control and regulate physiological processes that range from cell survival and proliferation, to differentiation, cell signaling, and immunomodulation (94,95). By the early 1980s, it was recognized that cystatins are present in lysosomes of most if not all cell types (96,97). Aberrant regulation of these important homeostatic factors contributes to a range of pathologies. Cystatins regulate the physiological activities of specific cysteine proteases (cathepsin family members) (98). There is increasing evidence that an imbalance between cysteine proteases and their inhibitors (cystatins) leads to excess protease activity due to high cathepsin levels, which contributes to tumor cell invasion (99). Consequently, imbalances in cystatins have been noted in a number of cancers (95).

### *Cystatin Super-Family*

Cysteine protease inhibitors belong to a cystatin super-family encompassing a large group of homologous proteins that inhibit papain family cysteine proteases (94,95,100). Twelve functional cystatins divide into three types based on protein structure, location in the body, and physiological role. Type 1 cystatins (cystatins A and B) are polypeptides of 98 amino acid residues and are found intracellulary, but occasionally appearing in body fluids at

detectable levels (95,100). The majority of cysteine protease inhibitors encompass type 2 cystatins including cystatin C, D, M, F, G, S, SN, and SA. Type 2 cystatins consist of 120 amino acid residues, two disulphide bridges, and an extracellular signaling peptide (101), and are found in most body fluids (95,100). Kininogens comprising type 3 cystatins are large multifunctional proteins with three type 2-like cystatin domains, of which only two are capable of inhibiting cysteine proteases (95). Kininogens are found in blood plasma (95). The tertiary structures of cystatin proteins are conserved and fold into a five-stranded beta-sheet, which wraps around a five-turn alpha-helix, termed a 'cystatin fold' (94,102). Cystatins function to protect cells from lysosomal peptidases released during normal cell death, phagocyte degranulation, and/or during cancer cell proliferation (95). Therefore, cystatins are essential in safeguarding against abnormal lysosomal cysteine protease activity that is essential for tumor invasion and metastasis.

### *CST6 (Cystatin M): A Prototype Methylation-sensitive Gene*

Cystatin M was originally identified in breast cancer cell lines isolated from a metastatic lesion and matched primary breast tumor by differential RNA display RT-PCR (103). In another investigation of EST-libraries of amniotic and fetal skin epithelial cells, cystatin M was independently cloned from cDNA (104). The biochemical properties, chromosomal localization (chromosome 11), and biological distribution of cystatin M is significantly different compared to the other cystatins (105). Cystatin M is expressed in a variety of normal human tissues including brain, lung, heart, liver, pancreas, spleen, thymus, small intestine, prostate, ovary, peripheral blood cells, and placenta (103,104). Cystatin M consists

of 121 amino acids and unlike other type 2 cystatins, is found in two different protein forms: (i) glycosylated (17 kDa), and (ii) non-glycosylated (14.4 kDa) (104).

Cystatin M is involved in regulating the activity of cathepsin B and cathepsin L, and an imbalance between these proteases and cystatin M is important in driving tumor progression (106-108). Cystatin M expression is diminished or lost in various forms of cancer including, (i) basal and squamous cell carcinomas of the skin (109), (ii) squamous cell carcinomas of the head and neck and lung regions (110), (iii) non-small cell lung cancer (111), (iv) metastatic oral cancer cell lines (105), (v) malignant glioma (112), (vi) melanoma cell lines (113), (vii) prostate cancer cell lines (113), and (viii) breast cancer (8,103,113-117). Cystatin M has been suggested to function as a breast tumor suppressor gene (116). The majority of human breast cancer cell lines derived from metastatic breast tumors lack cystatin M expression, whereas normal and premalignant cells express abundant levels of cystatin M (103,116). Exogenous expression of cystatin M in MDA-MB-435S breast cancer cells results in the suppression of cell proliferation, migration, matrix invasion, and tumor-endothelial cell adhesion *in vitro* (113). No deletions or structural rearrangements of cystatin M have been characterized, suggesting that loss of gene expression may be the result of transcriptional silencing (94,118).

Cystatin M contains a large CpG island (424 bp) including 54 CpG dinucleotides that spans the proximal promoter and exon 1, encompassing the start site for transcription. The promoter region of cystatin M contains a 8% CpG dinucleotide content 1400 bp upstream of the transcription start site, with the most CpG density (12%) occurring in the proximal 500 bp of the promoter. Several studies have shown that cystatin M is epigenetically regulated by DNA methylation-dependent silencing in breast cancer cell lines and primary invasive ductal

carcinomas (8,115). Furthermore, cystatin M was identified as a methylation-sensitive gene in glioma cell lines and primary brain tumors (112). Overall, these observations suggest strongly that methylation-dependent epigenetic silencing of cystatin M represents an important mechanism for loss of cystatin M in multiple tumor systems.

## E. Summary and Significance

The studies contained in this dissertation are relevant to breast cancer research in many important ways. While the contribution of epigenetic mechanisms to breast cancer induction and progression is well recognized, epigenetically-regulated genes in breast cancer have not been comprehensively catalogued or characterized. This dissertation characterizes a group of putative methylation-sensitive genes identified in MCF-7 breast carcinoma cells, validates that these genes are subject to methylation-dependent regulation, and identifies critical promoter methylation targets. The genes characterized include genes of unknown function, as well as genes of known (or proposed) function, among which are putative breast cancer tumor suppressor genes and genes that are associated with growth suppressive pathways. In addition, this dissertation evaluates the methylation of *CST6* in primary breast cancers and lymph node metastases, and shows that *CST6* is subject to DNA methylation-dependent epigenetic regulation *in vivo*. Thus, these studies establish a role for methylation-dependent epigenetic regulation in the loss of function of genes important for the molecular pathogenesis of breast cancer. In addition, this dissertation identifies several distinct classes of epigenetically-regulated genes and these classes can be distinguished based upon the CpG content and CpG organization of their promoters. Consequently, the establishment of a new

definition for epigenetically-regulated genes that recognizes the importance of all CpG targets has been proposed. This dissertation also addresses the unresolved question of what mechanisms govern methylation of CpG targets and identifies the existence of *cis* regulatory sequences located in the 5' upstream promoter region of *CST6* that functions to direct CpG methylation. Consequently, these results advance our understanding of mechanisms governing DNA methylation in breast carcinogenesis.

## II. EXPERIMENTAL PROCEDURES

### A. Breast Cancer Cell Line Culture

Human breast cancer cell lines were obtained from the Tissue Culture Core Facility of the UNC Lineberger Comprehensive Cancer Center at the University of North Carolina at Chapel Hill: BT-20 (ATCC#HTB-19), BT549 (HTB-122), Hs578T (HTB-126), MCF-7 (HTB-22), MDA-MB-231 (HTB-26), MDA-MB-415 (HTB-128), MDA-MB-435S (HTB-129), MDA-MB-436 (HTB-130), MDA-MB-453 (HTB-131), MDA-MB-468 (HTB-132), SK-BR-3 (HTB-30), and ZR-75-1 (CRL-1500). Normal breast epithelial cell lines, MCF12A (CRL-10782) and MCF10-2A (CRL-10781), were obtained from the ATCC (Manassas, VA). BT-20, MCF-7, and MDA-MB-231 cells were propagated in minimal essential medium (MEM) with Earle's salts, containing 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 10 μg/ml insulin (GIBCO/Invitrogen Life Technologies, Carlsbad, CA), and 10% fetal bovine serum (Hyclone, Logan, UT). Hs578T, MDA-MB-435S, MDA-MB-436, MDA-MB-453 cells were propagated in Dulbecco's modified Eagle's medium containing 4 mM L-glutamine, 10 μg/ml insulin (GIBCO/Invitrogen Life Technologies), and 10% fetal bovine serum (Hyclone). BT549 and ZR-75-1 cells were propagated in RPMI 1640 containing 2 mM L-glutamine, 1 mM sodium pyruvate, 10 μg/ml insulin (GIBCO/Invitrogen Life Technologies), and 10% fetal bovine serum (Hyclone). MDA-MB-415 and MDA-MB-468 cells were propagated in Leibovitz's L-15 medium containing 2 mM

L-glutamine, 10 μg/ml insulin (GIBCO/Invitrogen Life Technologies), 10 mg/ml glutathione (Sigma Chemical Company, St Louis, MO, USA), 10% fetal bovine serum (Hyclone). SK-BR-3 cells were propagated in McCoy's 5A medium containing 1.5 mM L-glutamine, and 10% fetal bovine serum (Hyclone). MCF12A and MCF10-2A cells were propagated in a 1:1 mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium containing 20 ng/ml human epithelial growth factor (Sigma Chemical Company, St. Louis, MO), 100 ng/ml cholera toxin (Sigma Chemical Company), 10 μg/ml insulin (GIBCO/Invitrogen Life Technologies), 500 ng/ml hydrocortisone (Sigma Chemical Company), and 5% horse serum (GIBCO/Invitrogen Life Technologies).

## B. Treatment of Human Breast Cancer Cells with Demethylating Agents

### *MCF-7 Breast Cancer Cells Treated with 5-aza-2'-deoxycytidine and Trichostatin A*

Three MCF-7 cell treatment groups were established from a single founding MCF-7 cell population: (i) control medium, (ii) medium containing 250 nM 5-aza-2'-deoxycytidine (5-aza), and (iii) medium containing 250 nM 5-aza and 50 nM trichostatin A (TSA). 5-aza and TSA were obtained from Sigma Chemical Company. Cells were plated at 5,000 cells/cm$^2$ in 150 mm polystyrene dishes (Corning Inc., Corning, NY). MCF-7 cells in the treatment groups were exposed to 5-aza or 5-aza + TSA for 3 weeks, with weekly subcultivation, followed by a 5 week recovery period in control growth medium, with weekly subcultivations during the last 3 weeks. Control MCF-7 cells were subcultivated once per week during the 8 week cell culture period. Cell cultures were fed fresh growth medium three times weekly.

***Hs578T, MDA-MB-435S, MDA-MB-436, MDA-MB-453, MCF-7, and ZR-75-1 Breast Cancer Cells Treated with 5-aza-2'-deoxycytidine***

Cell lines that lack expression of *CST6* (including Hs578T, MDA-MB-435S, MDA-MB-436, MDA-MB-453, MCF-7, and ZR-75-1) were treated with the demethylating agent 5-aza (Sigma Chemical Company), as described above. Briefly, two treatment groups were established from a single founding cell population: (i) control medium, and (ii) medium containing 250 nM 5-aza. Cells in the treatment group were exposed to 5-aza for 3 weeks, with weekly subcultivation, followed by a 5 week recovery period in control growth medium with weekly subcultivations during the last 3 weeks. Control cells were subcultivated once per week during the 8 week cell culture period. Cell cultures were fed fresh growth medium three times weekly.

**C. Human Breast and Lymph Node Tissues**

This study included 87 paraffin-embedded human tissues corresponding to primary breast tumors (n=54), lymph nodes metastases (n=22), and normal breast tissues (n=11). Twenty-one archival human tissues (primary breast tumors, lymph node metastases, and normal breast) were obtained from the University of North Carolina Lineberger Comprehensive Cancer Center and 6 archival primary breast tumors were acquired from the Louisiana State University Health Sciences Center, generously provided by Dr. Daniel Keppler (Shreveport, LA). A breast tumor microarray (Imgenex Corporation, Sorrento Valley, CA) consisting of 60 tissue cores was also utilized. In total, this study included 46 primary breast specimens diagnosed invasive ductal carcinoma (IDC), 2 breast ductal carcinoma *in situ* specimens, 1

solid papillary carcinoma, 1 medullary carcinoma, 1 signet ring cell carcinoma, 3 infiltrating lobular carcinomas, 22 lymph node metastases from IDC (n=20), atypical medullary carcinoma (n=1), and infiltrating lobular carcinoma (n=1), and 11 normal breast tissue samples. Five archival primary breast tumors were matched paired with lymph node metastases (3 independent lymph nodes corresponded to 1 primary tumor and 4 independent lymph nodes corresponded to one primary tumor). Handling of tissue specimens and protection of patient privacy followed strict policies of the institutional review board of the University of North Carolina School of Medicine.

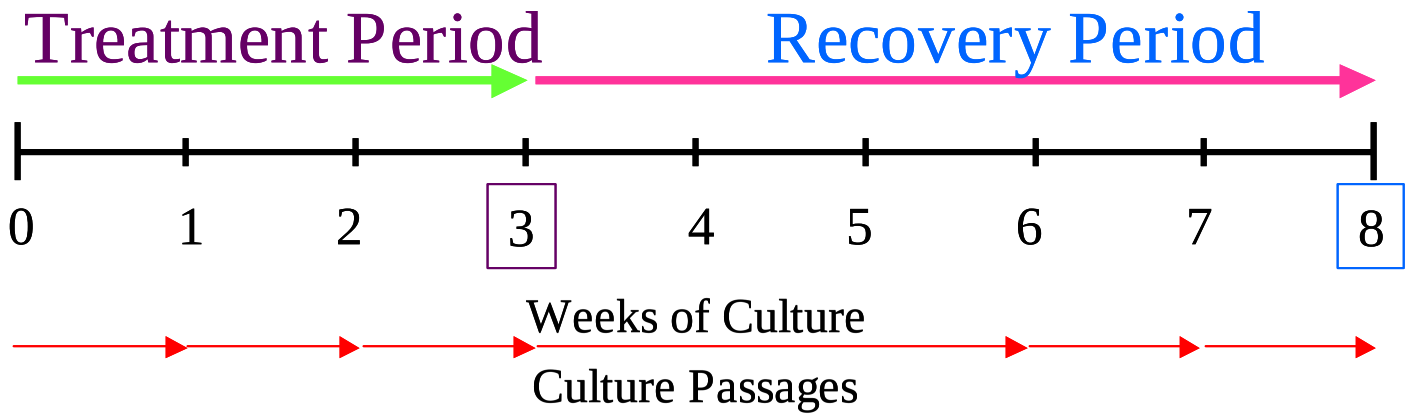
**D. RNA Isolation from Human Breast Cancer Cell Lines**

Total RNA was isolated from cultured cells using a modification of the method of Chomczynski and Sacchi (119) utilizing TRIzol Reagent (Invitrogen Corporation, Carlsbad, CA), according to the manufacturer's protocol. Control breast cancer cells were harvested weekly for RNA preparation over an 8 week culture period, whereas cells treated with 5-aza or 5-aza + TSA were harvested for RNA preparation at 3 weeks and 8 weeks. Isolated total RNA was stored at -20°C as an ethanol precipitate prior to microarray analysis or RT-PCR. Cells were counted at the end of each week using a Model Z1 Coulter Cell and Particle Counter (Beckman Coulter Inc., Fullerton, CA).

**E. Affymetrix Microarray Analysis of Gene Expression**

Large-scale gene expression analyses were performed by Expression Analysis (www.expressionanalysis.com), using the Affymetrix Human Genome GeneChip U133A oligonucleotide array (Affymetrix, Santa Clara, CA), which contains 500,000 oligonucleotides corresponding to 22,000 probe sets directed against 18,400 mRNA transcripts and 14,400 well-characterized genes. RNA samples corresponding to control MCF-7 cells (at week 3 and week 8), MCF-7 cells that were treated with 250 nM 5-aza (week 3 and week 8), and MCF-7 cells treated with 5-aza + TSA (week 3 and week 8) were utilized in this analysis. RNA samples from week 3 were derived from cells harvested after 3 weeks of exposure to 5-aza or 5-aza + TSA, whereas RNA samples from week 8 were derived from cells that were exposed to treatment for 3 weeks (to 5-aza or 5-aza + TSA) and then allowed to recover in control growth medium for 5 weeks (Figure 5). Target was prepared and hybridized according to the Affymetrix Technical Manual. Total RNA (10 μg) was converted into cDNA using Superscript II Reverse Transcriptase (Invitrogen Corporation) and a modified oligo(dT)24 primer that contains T7 promoter sequences (GenSet, Evry, France). After first strand synthesis, residual RNA was degraded by the addition of RNaseH and a double-stranded cDNA molecule was generated using DNA Polymerase I and DNA Ligase (Invitrogen Corporation). The cDNA was purified and concentrated using a standard phenol:chloroform extraction, followed by ethanol precipitation. Labeled cRNA products were generated from the purified cDNAs by incubation with T7 RNA Polymerase and biotinylated ribonucleotides, using an In Vitro Transcription kit (Enzo Diagnostics, Farmingdale, NY). cRNA products were purified on an RNeasy column (Qiagen Inc., Valencia, CA) and quantified spectrophotometrically. Purified cRNA target (20 μg) was incubated at 94°C for 35 minutes in fragmentation buffer [200 mM

**Figure 5. Demethylating treatment of MCF-7 breast cancer cells.** MCF-7 cells were exposed to 5-aza or 5-aza + TSA for three weeks, with weekly subcultivation and passage, followed by a five week recovery period in control growth medium, with weekly subcultivations and passages during the last three weeks. MCF-7 cells were harvested for RNA and DNA preparation at 1, 2, 3, 6, 7, and 8 weeks. RNA and DNA preparation at the end of the treatment period (3 weeks) and at the end of the recovery period (8 weeks) were used for microarray analysis.


Treatment Period
Recovery Period
0
1
2
3
4
5
6
7
8
Weeks of Culture
Culture Passages

Tris-acetate (pH 8.1), 500 mM potassium acetate, and 50 mM magnesium acetate], and then diluted into hybridization buffer [100 mM 2-(N-morpholino) ethanesulfonic acid, 20 mM EDTA, and 0.1% Tween 20] containing biotin-labeled OligoB2 and Eukaryotic Hybridization Controls (Affymetrix). The hybridization cocktail was denatured at 99°C for 5 minutes, incubated at 45°C for 5 minutes, and then injected onto a Human Genome U133A GeneChip cartridge. The U133A GeneChip array was incubated at 42°C for at least 16 hours in a rotating oven at 60 rpm. Subsequently, the hybridized GeneChips were washed under nonstringent conditions at 25°C in a buffer consisting of 0.9 M NaCl, 70 mM sodium phosphate (pH 7.4), 6 mM EDTA, and 0.01% Tween 20, and stringent conditions at 50°C in a buffer consisting of 100 mM 2-(N-morpholino) ethanesulfonic acid, 100 mM NaCl, and 0.01% Tween 20. The microarrays were then stained with Streptavidin Phycoerythrin and the fluorescent signal was amplified using a biotinylated antibody solution. Fluorescent images were detected in an Agilent GeneArray Scanner (Agilent Technologies Inc.). After probe-level data was extracted from the MicroArray Suite-derived CEL files, the probes were normalized using quantile probe normalization (120). Signal was computed using the Positional Dependent Nearest Neighbor (PDNN) method (121), and scaled by Expression Analysis proprietary methods to mitigate bias in fold-change underestimation. Microarray hybridizations were performed in duplicate (for each treatment group and time point) and the final values for (log) signal for all graphs were averages of the duplicates (equivalent to geometric averages of signal). The one exception is the control average, which was an average of the control at two different time points (week 3 and week 8).

## F. Semiquantitative RT-PCR Analysis of Gene Expression

Total RNA (2 μg) from control or cells treated with 5-aza or 5-aza + TSA was reverse-transcribed into cDNA using Superscript II Reverse Transcriptase (Invitrogen Corporation) and oligo(dT) as the primer, 60 minutes at 42°C, according to standard methodology. Gene-specific oligodeoxynucleotide primers were generated by the UNC Oligodeoxynucleotide Synthesis Core Facility (Chapel Hill, NC) for selected mRNAs based upon their known cDNA sequence (Genbank, www.ncbi.nih.gov). The sequences of gene-specific primers are given in Table 1. Verification of equal template concentration between samples was accomplished using primers that amplify a portion of β-actin mRNA (5'-AGAGATGGCCACGGCTGCTT-3' and 5'-ATTTGCGGTGGACGATGGAG-3'). PCR reactions were performed in a 50 μl total volume of buffer containing 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.001% gelatin, 200 μM of each dNTP (EasyStart Micro 50 PCR-mix-in-a-tube, Molecular BioProducts, San Diego, CA), 0.4 μM of each primer, and 2.5 U AmpliTaq enzyme (Perkin Elmer/Cetus, Foster City, CA). Amplifications were carried out in a Perkin Elmer 9700 Thermocycler using a step-cycle program consisting of 25-30 cycles of 94°C for denaturing (1 minute), 58°C for annealing (1 minute), and 72°C for extension (2 minutes).

## G. Quantitative Real-Time PCR

**Table 1.** Oligodeoxynucleotide Primers for RT-PCR Analysis of Gene Expression

| Gene Designation | Forward Primer | Reverse Primer | Amplicon Size |
|---|---|---|---|
| *BF* | 5’-GCCAGCAACAAAAGGAAGAG | 5’-GCAAGTATTGGGGTCAGCAT | 242 bp |
| *C8orf4* | 5’-TTTCAAACAGGTTGCACAAAA | 5’-GTTGCATGACATTTGCCAGT | 229 bp |
| *CEACAM5* | 5’-AGATTGCAGTGAGCCCAGAT | 5’-CTGCTTGATCTTGGTGGACA | 200 bp |
| *CEACAM6* | 5’-TGAGCCAGTGGTGCTAAATG | 5’-TGGAACAAGGAAACAGAACCA | 235 bp |
| *CST6* | 5’-AAGACCAGGGTCACTGGAGA | 5’-CGGGGACTTATCACATCTGC | 163 bp |
| *CYP1B1* | 5’-CCCTCATTGTGTTTCTACCG | 5’-GGCTAAGTTCTGGGACATGAA | 222 bp |
| *FLJ10134* | 5’-GGAGAACAGCTGGCTAAGGA | 5’-TTCATAGTGTGGGCATCCAA | 203 bp |
| *G1P2* | 5’-CACCTGAAGCAGCAAGTGAG | 5’-CTTTATTTCCGGCCCTTGAT | 228 bp |
| *G1P3* | 5’-CTCGCTGATGAGCTGGTCT | 5’-TGCTGGCTACTCCTCATCCT | 181 bp |
| *IFI27* | 5’-TCCTCCATAGCAGCCAAGAT | 5’-CCTGGCATGGTTCTCTTCTC | 221 bp |
| *IGFBP5* | 5’-TTCACAGACTCTGGCCTCCT | 5’-TGTGCTATCCATGTGGGCTA | 185 bp |
| *ISGF3G* | 5’-GAGCTCTTCAGAACCGCCTA | 5’-GGCTCTACACCAGGGACAGA | 226 bp |
| *KRTHB1* | 5’-TAGGCACCCCAACTCAAGTC | 5’-AAGTGGGGGATCACACAGAG | 162 bp |
| *LCN2* | 5’-ACGCTGGGCAACATTAAGAG | 5’-CGAAGTCAGCTCCTTGGTTC | 162 bp |
| *LGALS3BP* | 5’-ACCAACAGCTCGAAGAGCAC | 5’-GGTCATTGCAGAGAGGAAGG | 202 bp |
| *SAT* | 5’-ACGGGGTAAGAAGGTTCAGC | 5’-TGTCTGGCGAGTGTGAGTGT | 161 bp |
| *SCNN1A* | 5’-GCCCCCTTTGTTACTTAGGC | 5’-AAAGACACAGGGCAGAGGTG | 153 bp |
| *ZC3HDC1* | 5’-CTTATTGGCACCAGGGACAG | 5’-GTGTCAGAGCAACAGGCAGA | 191 bp |

Total RNA samples (20 μg) from Hs578T, MCF-7, MDA-MB-436, MDA-MB-453, and ZR-75-1 control and treated cells were DNAase treated (Promega, Madison, WI), purified using the Qiagen Rneasy mini-kit (Qiagen), and reversed transcribed using the High Capacity cDNA Archive Kit (Applied Biosystems, Foster City, CA) according to the manufacturer's protocol. Real-time primers and probes for *CST6* and β-actin were purchased from Applied Biosystems. Reactions were carried out using TaqMan Universal PCR Master Mix (Applied Biosystems) and the following PCR conditions: 95°C for 10 minutes, 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Gene expression levels were normalized using β-actin for each cell line and differences in *CST6* gene expression were determined using the comparative Ct method described in the ABI Prism 7700 User Bulletin #2 (Applied Biosystems).

## H. Promoter and 5'-Upstream Sequence Analysis of Putative Methylation-sensitive Genes

Genomic sequences corresponding to the promoter and 5'-upstream regions of select genes were identified using the Human Genome Browser Gateway (http://genome.ucsc.edu/cgi-bin/hgGateway) contained in the UCSC Genome Bioinformatics website (http://genome.ucsc.edu/). GenBank accession numbers were utilized to identify RefSeq records corresponding to each gene, and then the promoter and 5'-upstream sequences were identified using the Genomic Sequence Near Gene tool. For each gene of interest, 3000 bp of sequence 5'-upstream of exon 1 (containing the putative transcriptional promoter and associated elements) were identified. CpG islands were identified within

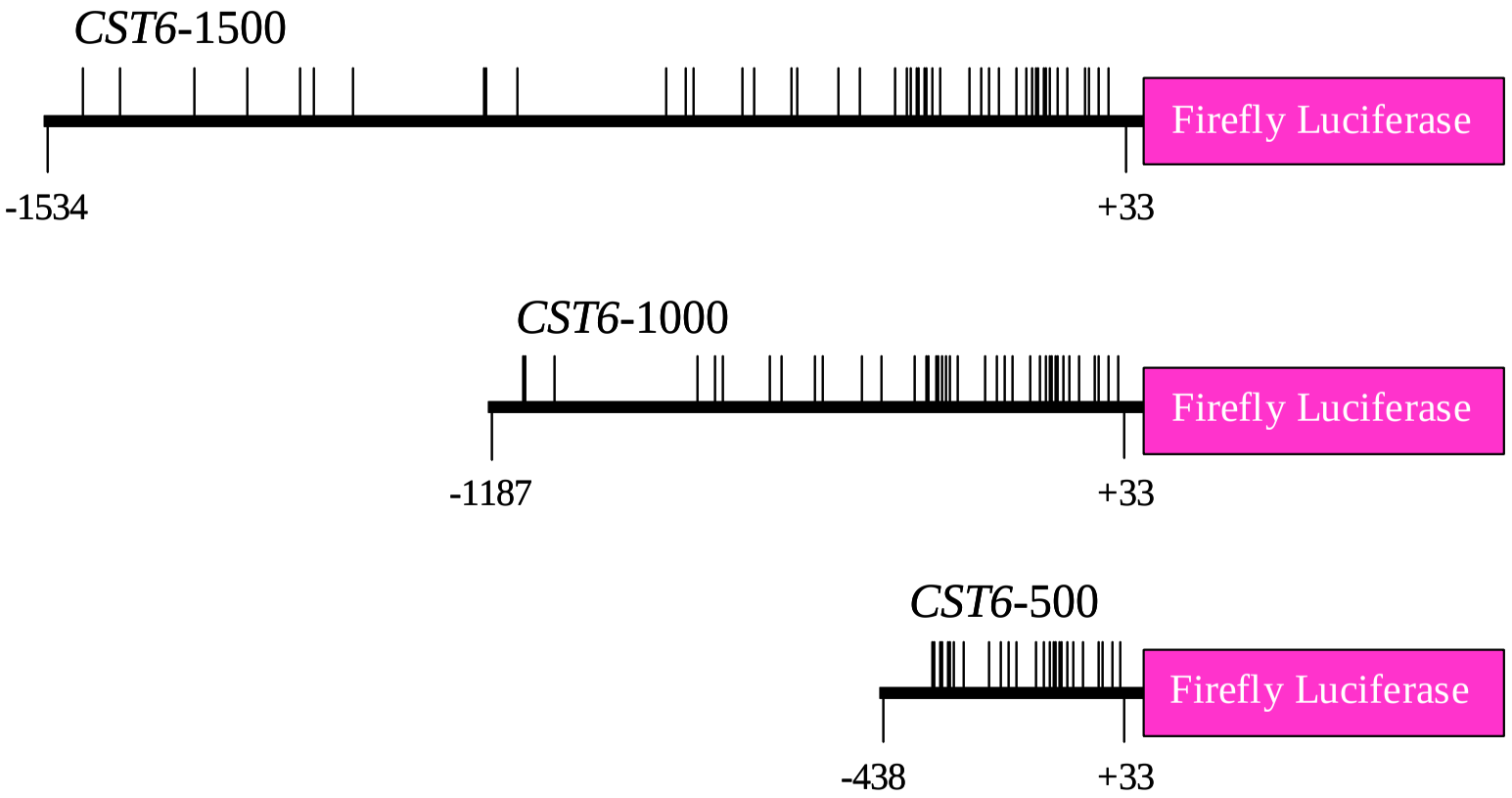
promoter and exon 1 sequences using the CpGPLOT program from the European Bioinformatics Institute website (http://www.ebi.ac.uk/emboss/CpGplot/). Typical CpG islands were defined as ≥200 bp of sequence with ≥50% C+G content and ≥0.6 CpG observed/CpG expected (22). Weak CpG islands exhibit the same features (with ≥50% G+C content and ≥0.6 CpG observed/CpG expected), but over a shorter sequence segment (>50 bp but <200 bp). Alu repetitive elements were identified using the RepeatMasker Web Server (http://repeatmasker.genome.washington.edu) from the Institute for Systems Biology at the University of Washington (Seattle, WA) and transcription factor binding sites were identified using ProSpector (http://prospector.nci.nih.gov).

## I. Construction of Reporter Gene Constructs

### *CST6 Gene Promoter Constructs*

Segments of the *CST6* promoter were amplified by PCR and inserted upstream of the firefly luciferase gene in the pGL4.17[luc2/Neo] vector (Promega, Madison, WI, USA) to generate *CST6* reporter constructs. The pGL4.17[luc2/Neo] plasmid contains a multiple cloning region, ampicillin resistance gene, *luc2* reporter gene, and a SV40 early enhancer/promoter driving a *neo* resistance gene. A series of reporter constructs were generated with different portions of the *CST6* promoter with a common 3' end, terminating at +33 bp relative to the designated transcription start site (Genebank accession number NM_001323.1). Three constructs were generated, encompassing -1534 to +33 (designated *CST6*-1500), -1187 to +33 (designated *CST6*-1000), and -438 to +33 (designated *CST6*-500) (Figure 6). The *CST6*-500 construct represents the minimal essential promoter and

**Figure 6. Design of *CST6* promoter luciferase reporter gene constructs.** A schematic representation of three *CST6* promoter reporter gene constructs is shown. Each of these reporter constructs consists of the firefly luciferase gene (pink box) driven by different portions of the *CST6* promoter: *CST6*-1500 (-1534 to +33 nucleotides), *CST6*-1000 (-1187 to +33 nucleotides), and *CST6*-500 (-438 to +33 nucleotides). The distribution of the CpG dinucleotides in the *CST6* promoter proximal to the luciferase gene is depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides).


*CST6*-1500
Firefly Luciferase
-1534
+33
*CST6*-1000
Firefly Luciferase
-1187
+33
*CST6*-500
Firefly Luciferase
-438
+33

encompasses the proximal promoter CpG island as predicted by CpGPLOT (www.ebi.uk/emboss/CpGplot/). Likewise, the *CST6*-1000 construct encompasses the proximal promoter CpG island and spans approximately 1000 bp upstream of the transcription start site, and the *CST6*-1500 construct encompasses the CpG island and spans approximately 1500 bp upstream of the transcription start site. The *CST6* gene promoter was amplified using the following forward primers designed with *BglII* restriction sites (shown underlined): *CST6*-1500, 5'-ATGCTAGAGATCTAGTTGTCAGTCCCCCTAGGTC-3', *CST6*-1000, 5'-ATGCTAGAGATCTAGGGCAGAGCTGACATGACTGA-3', *CST6*-500, 5'-ATGCTAGAGATCTAGTCCAGCACCAGACCTCTTCT-3'. A common reverse primer was used for all constructs and included a *HindIII* restriction site (shown underlined): 5'-AGTCAAGCTTAGCCTCAGAGCCGTGAGTGC-3'. Amplicons were inserted into pGEM-T Easy Vector (Promega) and propagated in JM109 bacterial cells. Ten colonies were selected per promoter construct and expanded in liquid culture. Plasmid DNA was purified using the Wizard Plus Miniprep DNA purification kit (Promega), before digestion with *BglII* and *HindIII* (New England Biolabs, Ipswich, MA) to liberate the *CST6* promoter segment. Restricted DNA samples were fractionated on 2% low temperature melting agarose gels and the cloned inserts were excised from the gel. DNA fragments were ligated into the pGL4.17[luc2/Neo] vector double digested with *BglII* and *HindIII*, cloned, and purified as described. Reporter constructs were confirmed by restriction mapping and DNA sequencing with an Applied Biosystems automated sequencer at the UNC Genome Analysis Facility (Chapel Hill, NC).

### *In Vitro CST6 Promoter Construct Methylation*

To analyze the effects of methylation on promoter activity, *CST6* reporter constructs (*CST6*-1500, *CST6*-1000, and *CST6*-500) were methylated using *SssI* methylase (M. *SssI*, New England Biolabs), which methylates all cytosine residues within CpG dinucleotides. Approximately 2 μg of each DNA construct was incubated with 4 U of *SssI* in the presence of 1600 μM *S*-adenosylmethionine at 37°C overnight. Methylated and unmethylated constructs were linearized with *BSU36I* (New England Biolabs) prior to transfection of breast cancer cells.

## J. Luciferase Reporter Assay

### *Transient Transfection of MCF-7 Breast Cancer Cells*

MCF-7 breast cancer cells were seeded at approximately $2x10^5$ per well into 6-well plates and grown to 80%-90% confluence. *CST6*-1000 and *CST6*-500 promoter constructs (2.5 μg) were transfected by *Trans*IT-LT1 Transfection Reagent (Mirus Bio, Madison, WI) according to the manufacturer's protocol. Cells were harvested 48 hours after transfection and luciferase activity was measured using the Steady-Glo Luciferase Assay System according to the manufacturer's protocol.

### *Stable Transfection of Human Breast Cancer Cell Lines*

Human breast cancer cell lines were obtained from the Tissue Culture Core Facility of the UNC Lineberger Comprehensive Cancer Center at the University of North Carolina at Chapel Hill: BT-20 (ATCC#HTB-19), MCF-7 (HTB-22), and MDA-MB-453 (HTB-131).

Normal breast epithelial cell line, MCF12A (CRL-10782) was obtained from the ATCC. All cells were propagated according to recommendations from the ATCC (see above). Human breast cells (approximately $8x10^5$) were seeded in 100 mm plates and grown to 80%-90% confluence. *CST6*-1000 and *CST6*-500 promoter constructs (2 μg) were transfected using Effectene Transfection Reagent (Qiagen) according to the manufacturer's protocol. Parallel cultures were transfected with control pGL4.17[luc2/Neo] vector, as described. Stably transfected cells were selected using 400 μg/ml G418 (Invitrogen Corporation) in BT-20 cells and MCF12A cells, 800 μg/ml G418 in MCF-7 cells, and 1200 μg/ml G418 in MDA-MB-453 cells. Luciferase activities corresponding to each reporter construct (*CST6*-1000 and *CST6*-500) were determined for each transfected cell line three times over a 5 to 7-week time point (approximately once every two weeks) using the Steady-Glo Luciferase Assay System (Promega) according to the manufacturer's protocol. Luciferase activities for each reporter construct were calculated after subtraction of background, and determined using the promoterless pGL4.17[luc2/Neo] vector.

## K. Genomic DNA Isolation

### *Human Breast Cancer Cell Lines*

Genomic DNA from $2x10^6$ cultured cells was isolated using the Puregene DNA Purification Kit (Gentra Systems, Minneapolis, PA) according to the manufacturer's protocol. Briefly, cells were lysed and incubated with proteinase K overnight at 55°C. Following incubation, RNase A solution was added and samples were incubated at 37°C for

1 hour. Subsequently, DNA was precipitated, hydrated by incubating at 65°C for 1 hour, and stored at -20°C before use.

***Human Primary Breast Tumors, Lymph Nodes, and Normal Breast Tissue***

Paraffin-embedded tissue specimens were scraped or microdissected from slides using a clean razor blade, deparaffinized, and genomic DNA was isolated using a QIAamp DNA Micro kit (Qiagen, Inc., Valencia, VA) according to manufacturer's instructions. Briefly, tissue samples were incubated overnight at 56°C with proteinase K. Subsequently, carrier RNA (1 μg/μl) was added and DNA samples were applied to columns, washed, and eluted with 35 μl of distilled water.

**L. Bisulfite Modification of Genomic DNA, Cloning, and Sequencing**

***Human Breast Cancer Cell Lines***

Bisulfite modification of genomic DNA was performed by a procedure adapted from Grunau *et al.* (122), generously provided by Dr. Randy Jirtle (Duke University, Durham, NC). Genomic DNA (3 μg) was digested with 1 U of *Xho I* (New England Biolabs) overnight in 12 μl total volume and heat inactivated at 65°C for 20 minutes; 5 μl of digest was subjected to bisulfite modification. Briefly, approximately 1.5 μg of DNA in 45 μl of distilled water was denatured by adding 5 μl 3 M NaOH and incubating for 20 minutes at 42°C, followed by addition of 450 μl of sodium bisulfite solution (saturated sodium bisulfite, 10 mM hydroquinone, pH 5.0) and incubation at 55°C for 4 hours. Bisulfite modified DNA (500 μl) was purified using the Wizard DNA Clean-Up kit (Promega), reconstituted with 50

µl of Tris-Cl (pH 8.0) and desulfonated by addition of 5.5 µl 3 M NaOH and incubation at 37°C for 20 minutes. The solution was precipitated by adding 40 µl 7.5 M ammonium acetate and 300 µl 100% ethanol at -20°C for at least 30 minutes. The DNA pellet was washed with 70% ethanol, dried briefly, and resuspended in 20 µl 1 mM Tris-Cl (pH 8.0). Bisulfite converted DNA was amplified using primers directed to specific segments within the promoter regions and exon 1 of selected genes (Table 2). PCR amplification was accomplished using a step-cycle program consisting of 40 cycles of 94°C for denaturing (1 minute), 55°C for annealing (1.5 minutes), and 72°C for extension (2 minutes). PCR products were fractionated on 2% agarose gels containing 40 mM Tris-acetate/1.0 mM EDTA (pH 8.0) and visualized by ethidium bromide staining. A portion of the PCR products (1 to 5 µl) was cloned into pGEM-T Easy Vector (Promega, Madison, WI). Three to 12 colonies were selected per gene segment and expanded in liquid culture. Plasmid DNA was purified using the Wizard Plus Miniprep DNA purification kit (Promega, Madison, WI), prior to digestion with *NcoI* and *NdeI* (New England Biolabs, Beverly, MA) to confirm the presence and size of the cloned insert. Validated clones were sequenced using the universal M13R3 primer with an Applied Biosystems automated sequencer at the UNC Genome Analysis Facility (Chapel Hill, NC). The bisulfite conversion efficiency was calculated for each sequenced clone based upon the ratio of converted Cs (non-CpG) to total number of Cs (non-CpG) in a given gene segment. Only clones determined to have a conversion efficiency of >95% were included in the present study. The results of methylation analyses were expressed as total methylation index (TMI). This measure of methylation can be applied to single CpG dinucleotides, select groups of CpG dinucleotides, or to continuous groups of CpG dinucleotides in a given gene segment. TMI was calculated for each cell line and clone

**Table 2.** Oligodeoxynucleotide Primers for Bisulfite Sequencing Analysis of Selected Genes

| Gene | Gene Segment | Forward Primer | Reverse Primer | Amplicon Size |
|---|---|---|---|---|
| *CST6* | Segment A (-636 to -206) | 5'-GGTTGGAATGTTGTAGTGGT | 5'-CCCCAACAACAAATACCAA | 413 bp |
| | Segment B (-228 to +10) | 5'-TTGTATTGGTATTTGTTGTTGG | 5'-TTACGACCGCGCAACTTTTA | 238 bp |
| | Segment C (-228 to + 242) | 5'-GGTTGGAATGTTGTAGTGGT | 5'-CTACCCATATTATAACTAACC | 470 bp |
| | Segment D (-228 to +118) | 5'-GGTTTTTTGGGTTTTTTGAATTT | 5'-CTACCCATATTATAACTAACC | 346 bp |
| *C8orf4* | Segment A (-1090 to -763) | 5'-GAATTAAAATATAAGGAGAGTTTT | 5'-AACATTACCCAAACATAAAACAA | 328 bp |
| | Segment B (-786 to -278) | 5'-TTGTTTTATGTTTGGGTAATGTT | 5'-CCCCAACAACAAATACCAA | 529 bp |
| | Exon 1 (+23 to +453) | 5'-TTTGGAAAGTTTGGGAGTTG | 5'-ATCTCACAAACTCACTTAATC | 430 bp |
| *IFI27* | Segment A (-1287 to -883) | 5'-GGTGGTGGTAGTATTTTATAA | 5'-AACACAAATCCTAAAATATAC | 404 bp |
| | Segment B (-905 to -463) | 5'-GTATATTTTAGGATTTGGTGTT | 5'-CCTCACAAAATTTACCAAC | 442 bp |
| | Segment C (-399 to -16) | 5'-GTGGGTTTTTAGTTGAGATT | 5'-CCAAACCTTTTAAACAACTC | 385 bp |

by dividing the number of methylated CpGs observed by the total CpGs analyzed and expressed as percent methylation. For instance, in an analysis of a gene segment containing 55 CpG dinucleotides and three clones sequenced, TMI would be calculated based upon 165 possible CpG methylation events (3 x 55).

***Human Primary Breast Tumors, Lymph Nodes, and Normal Breast Tissue***

Genomic DNA was modified with sodium bisulfite using the EZ DNA Methylation-Gold Kit (ZYMO Research Co., Orange, CA) according to the manufacturer's protocol. Briefly, approximately 1.5 μg denatured genomic DNA was treated with conversion reagent, incubated at 98°C for 10 minutes, and 53°C for 30 minutes, followed by a step-cycle program consisting of 8 cycles of 53°C for 6 minutes and 37°C for 30 minutes. Subsequently, samples were applied to columns, washed, desulfonated, washed and then eluted with 20 μl of elution buffer. In general, 2 μl of modified DNA was used in subsequent PCR reactions as described above. A portion of each PCR product (1 to 5 μl) was cloned into pGEM-T Easy Vector (Promega), expanded in liquid culture, and plasmid DNA was purified as described above. The results of methylation analyses were expressed as total methylation index (TMI) as described above. Tumors with a *CST6* promoter TMI >11% were considered hypermethylated.

***CST6 Reporter Gene Constructs***

Bisulfite modification of genomic DNA was performed by a procedure adapted from Grunau *et al.* (122). Genomic DNA (3 μg) was digested with 1 U of *EcoRI* (New England Biolabs, Beverly, MA) overnight in 12 μl total volume and heat inactivated at 65°C for 20

min; 5 µl of digest was subjected to bisulfite modification as described above. Bisulfite converted DNA was PCR amplified using primers directed to stably transfected *CST6* promoter constructs. The primers were designed to encompass a region of the *CST6* promoter and luciferase reporter gene within the pGL4.17[luc2/Neo] vector (5'-TTGTATTGGTATTTGTTGTTGG-3' and 5'-CTTCATAACTTTATACAACTAC-3'). A portion of the PCR product was cloned, purified, and sequenced according to procedures described above.

**M. Immunohistochemical Analysis of Human Primary Breast Tumors, Lymph Node Metastases, and Normal Breast Tissue**

Formalin-fixed, paraffin-embedded human breast tissues and lymph nodes were sectioned (5 µm thick) and mounted on glass microscope slides. Immunohistochemical staining was performed according to standard methods. Briefly, tissue sections were incubated on a slide warmer at 60°C for 15 minutes, deparaffinized in xylene, incubated with 3% $H_2O_2$ in methanol to block endogenous peroxidase activity, and rehydrated through a series of ethanol washes. Antigen retrieval was accomplished by steaming in 1x citrate buffer (Dako Inc., Carpinteria, CA) for 30 minutes. After incubation with serum-free protein block (Dako Inc.) for 10 minutes, tissues were incubated for 2 hours at room temperature with polyclonal rabbit anti-cystatin M antibodies diluted 1:1000 (116) generously supplied by Dr. Daniel Keppler (Louisiana State University Health Sciences Center, Shreveport, LA). Subsequently, tissues were washed and layered with a two-step secondary set-up including a anti-rabbit biotintylated link and streptavidin-conjugated HRP solution (Dako Inc.) for 10 minutes each,

incubated with HRP substrate containing 3,3'diaminobenzidine (Dako Inc.) for a total of 5 minutes, followed by counterstaining with hematoxylin. Control immunostaining reactions were performed at room temperature with mouse monoclonal anti-cytokeratin 18 (Santa Cruz Biotechnology Inc., Santa Cruz, CA) antibodies diluted 1:1000. Negative control staining followed the same procedure except sections were incubated with either rabbit preimmune serum or 1x wash buffer instead of anti-cystatin M antibody. Normal breast tissue was used as a positive control for the anti-cystatin M antibody.

**N. Statistical Analysis**

Values included in the text represent the mean ± S.E.M. for CpG content (observed CpG dinucleotides/total dinucleotides x 100). The values for the mean and S.E.M. were calculated using the statistical function of KaleidaGraph Version 3.5 (Synergy Software, Essex Junction, VT). Statistical significance was determined using an unpaired t-test (KaleidaGraph).

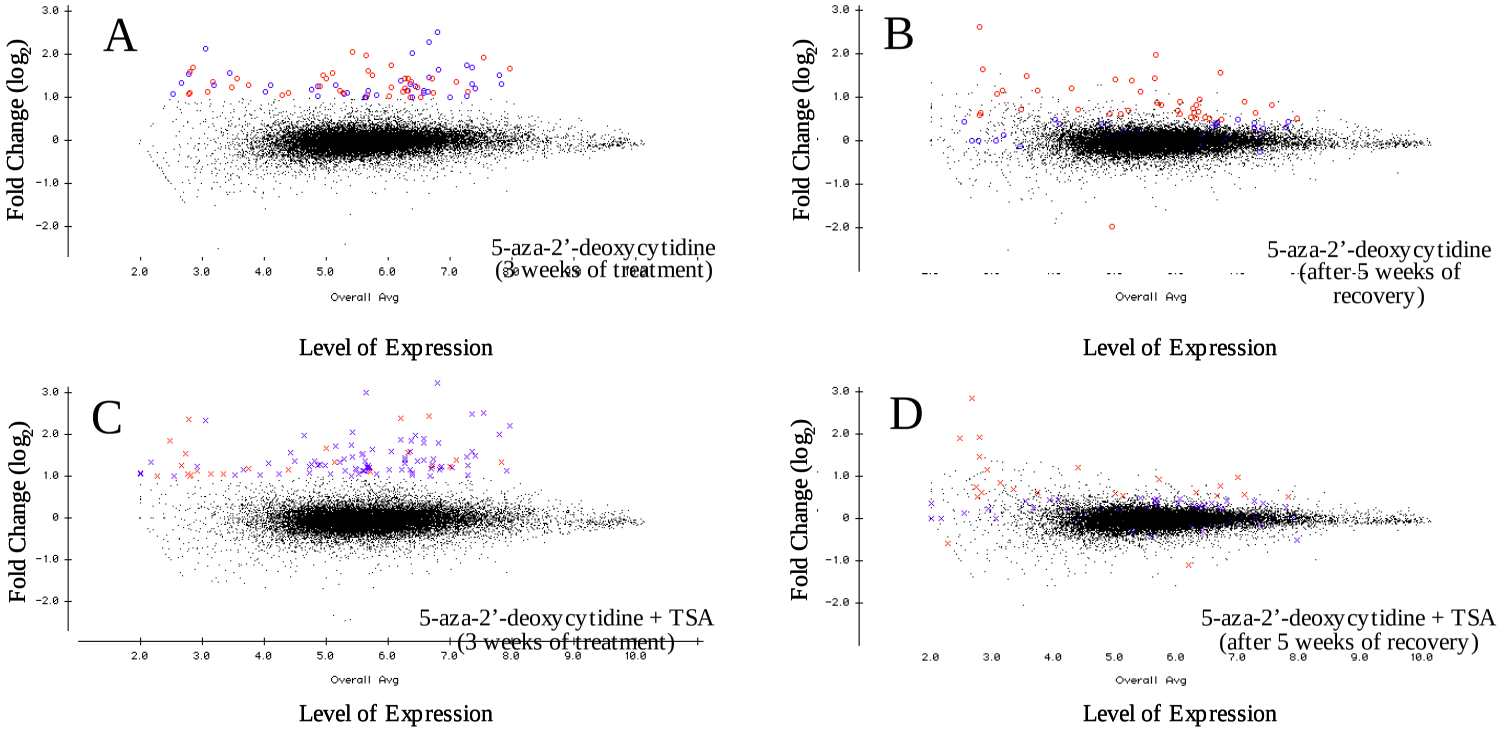
# III. RESULTS

## A. DNA Methylation-Dependent Epigenetic Regulation of Gene Expression in MCF-7 Breast Cancer Cells

### *Identification of Putative Epigenetically-regulated Genes in MCF-7 Breast Cancer Cells Through Microarray Analysis of Gene Expression*

Microarray analysis of gene expression was performed using RNA samples from control MCF-7 cells harvested at the week 3 and week 8 time points of the cell culture period. The gene expression profiles of these control cultures were found to be remarkably consistent when the levels of expression of individual transcripts were compared between the two time points. Analysis of the week 3 versus week 8 control expression data produced a positive correlation coefficient close to 1 ($r = 0.96$, $r^2 = 0.92$). Based upon this result, the expression data for the two control time points were averaged and utilized for normalization of the expression data obtained with RNA samples from treated cells. Treatment of MCF-7 cells with 250 nM 5-aza for 3 weeks resulted in an approximate $\geq$2-fold increased expression of 79 genes (Figure 7A). Likewise, treatment of MCF-7 cells with 250 nM 5-aza + 50 nM TSA for 3 weeks produced an approximate $\geq$2-fold increased expression of 107 genes (Figure 7C). To reduce the numbers of genes for analysis, and to enrich for genes that are putatively epigenetically regulated, we analyzed the microarray data to identify genes that were modified by treatment (increased expression levels) but then returned to control expression

**Figure 7. Identification of putative epigenetically-regulated genes in MCF-7 breast cancer cells after exposure to demethylating treatment.** M versus A plot of microarray data. The values on the Y-axis reflect $\log_2$-scale fold-change (log ratios) for treatment samples relative to control values. The values on the x-axis reflect the average signal intensity for individual control probe sets (transcripts). For panels A and B, genes with $\geq$2-fold (log ratio $\geq$1) increased expression in MCF-7 cells after 3 weeks of 5-aza treatment are shown in red, and genes that returned to control values after a 5 week recovery period (following withdrawal of 5-aza) are shown in blue (Panel A, week 3; Panel B, week 8). This analysis identified 37 genes with increased expression in response to 5-aza treatment in week 3 that returned to control levels by week 8. For panels C and D, genes showing $\geq$2-fold increased expression in MCF-7 cells after 3 weeks of 5-aza + TSA treatment are shown in red, and genes that returned to control values after a 5 week recovery period (following withdrawal of 5-aza + TSA) are shown in blue (Panel C, week 3; Panel D, week 8). This analysis identified 70 genes with increased expression in response to 5-aza + TSA treatment in week 3 that returned to control levels by week 8.


A
Fold Change (log2)
Overall Avg
5-aza-2'-deoxycytidine
(3 weeks of treatment)
Level of Expression
B
Fold Change (log2)
Overall Avg
5-aza-2'-deoxycytidine
(after 5 weeks of
recovery)
Level of Expression
C
Fold Change (log2)
Overall Avg
5-aza-2'-deoxycytidine + TSA
(3 weeks of treatment)
Level of Expression
D
Fold Change (log2)
Overall Avg
5-aza-2'-deoxycytidine + TSA
(after 5 weeks of recovery)
Level of Expression

levels following the withdrawal of treatment. This analysis identified 37 genes in 5-aza-treated MCF-7 cells and 70 genes in 5-aza + TSA treated MCF-7 cells that increased ≥2-fold at 3 weeks and returned to control level after 8 weeks (Figure 7B and 7D). Comparison of these gene lists identified 20 genes in common between the 5-aza and 5-aza + TSA treatment groups (Table 3). Most of these genes (16/20, 80%) have not been shown previously to be subject to methylation-dependent silencing in cancer cells. However, there is evidence for epigenetic regulation of *C8orf4*, *CYP1B1*, *PSG6*, and *SAT* (72,123,124). Suzuki and colleagues identified *C8orf4* and *PSG6* among genes that are up-regulated in human RKO colorectal carcinoma cells in response to 5-aza + TSA treatment (72). *CYP1B1* has been shown to be methylated in primary breast cancers (124), and *SAT* is subject to silencing through X-chromosome inactivation (123). Genes that responded to demethylating treatment with either 5-aza (n=17) or 5-aza + TSA (n=50), but not both, are given in Table 4. Most of these putative epigenetically-regulated genes were apparently induced with both demethylating treatments, but failed to indicate a greater than 2-fold difference in expression for one of the treatments. For example, *CYP1A1* was estimated to increase 2-fold in response to 5-aza treatment, but only 1.9-fold in response to 5-aza + TSA. Likewise, *SYNGR3* was estimated to increase 2.3-fold in response to 5-aza + TSA, but only 1.9-fold in response to 5-aza (Table 4).

***Validation of Treatment-related Changes in Gene Expression by RT-PCR***

RT-PCR was employed to validate the changes in gene expression identified by microarray analysis that occur in MCF-7 cells with 5-aza or 5-aza + TSA treatment. RT-PCR analysis of RNA samples prepared from control MCF-7 cells at 1, 2, 3, 6, 7, and 8

**Table 3.** Putative Epigenetically-regulated Genes Identified in MCF-7 Breast Cancer Cells After Exposure to Demethylating Treatment

| Gene Designation | Genbank Accession Number | Relative Expression Level[a] | Promoter and Exon 1 Sequence Features: Typical CpG Island[b] Promoter | Typical CpG Island[b] Exon 1 | Weak CpG Island[c] Promoter | Weak CpG Island[c] Exon 1 | CpG Content Entire Promoter[d] | CpG Content Proximal Promoter[e] | CpG Content Promoter +Exon |
|---|---|---|---|---|---|---|---|---|---|
| ***Genes with Typical CpG Features*** | | | | | | | | | |
| *CRIP1* | NM_001311.1 | 3.1 | Yes | No | Yes | No | 8.1% | 13.2% | 8.4% |
| *CST6* | NM_001323.1 | 2.4 | Yes | Yes | Yes | No | 4.9% | 11.6% | 6.6% |
| *CYP1B1* | NM_000104.2 | 2.1 | Yes | Yes | Yes[g] | No | 10.9% | 16.8% | 11.2% |
| *FLJ10134* | NM_018004 | 2.9 | Yes | Yes | Yes | No | 4.1% | 14.8% | 4.6% |
| *G1P2* | NM_005101.1 | 3.6 | Yes | No | Yes[g] | No | 5.5% | 6.4% | 5.5% |
| *G1P3* | NM_022873.1 | 3.2 | Yes | No | No | No | 2.7% | 2.4% | 2.9% |
| *KRTHB1* | NM_002281.1 | 3.1 | No | Yes | No | No | 1.4% | 2.0% | 3.5% |
| *SAT* | NM_002970.1 | 2.5 | Yes | Yes | No | Yes | 4.9% | 10.4% | 5.5% |
| *ZC3HDC1* | NM_022750 | 2.4 | No | Yes | Yes[g] | Yes[g] | 2.8% | 4.4% | 8.2% |
| ***Genes with Intermediate CpG Features*** | | | | | | | | | |
| *BF* | NM_001710.1 | 2.8 | No | No | Yes | Yes | 2.7% | 2.0% | 2.7% |
| *C8orf4* | NM_020130 | 4.8 | No | No | No | Yes | 1.9% | 2.0% | 2.1% |
| *CEACAM5* | NM_004363.1 | 3.2 | No | No | Yes | No | 3.5% | 3.0% | 3.5% |
| *CEACAM6* | NM_002483.3 | 5.2 | No | No | Yes[g] | No | 2.8% | 2.4% | 2.7% |
| *IGFBP5* | NM_000599.2 | 2.1 | No | No | No | Yes[g] | 2.1% | 4.0% | 4.4% |
| *ISGF3G* | NM_006084.1 | 2.8 | No | No | Yes[g] | No | 3.5% | 2.4% | 3.5% |
| *LCN2* | NM_005564.1 | 4.4 | No | No | Yes | No | 4.1% | 2.8% | 4.0% |
| *PSG6* | NM_002782 | 2.3 | No | No | Yes[g] | No | 3.3% | 1.2% | 3.2% |

**Table 3.** Continued

| Gene Designation | Genbank Accession Number | Relative Expression Level[a] | Promoter and Exon 1 Sequence Features: Typical CpG Island[b] Promoter | Typical CpG Island[b] Exon 1 | Weak CpG Island[c] Promoter | Weak CpG Island[c] Exon 1 | CpG Content Entire Promoter[d] | CpG Content Proximal Promoter[e] | CpG Content Promoter +Exon |
|---|---|---|---|---|---|---|---|---|---|
| ***Genes with Atypical CpG Features*** | | | | | | | | | |
| *IFI27* | NM_005532.1 | 7.6 | No | No | No | No | 2.3% | 2.0% | 2.3% |
| *LGALS3BP* | NM_005567.2 | 2.8 | No | No | No | No | 3.5% | 2.4% | 3.7% |
| *SCNN1A* | NM_001038.1 | 2.3 | No | No | No | No | 2.9% | 3.6% | 3.1% |

[a]Relative expression levels are expressed as average fold control levels of expression at the end of 3 weeks of demethylating treatment (5-aza and 5-aza + TSA).
[b]Typical CpG islands were defined using the conventionally accepted criteria for these promoter elements, consisting of a region of $\geq$200 bp with $\geq$50% C+G, and $\geq$0.6 CpG observed/CpG expected (65).
[c]Weak CpG islands were defined as a region of >50 bp but <200 bp with $\geq$50% C+G and $\geq$0.6 CpG observed/CpG expected.
[d]%CpG dinucleotides in the putative gene promoter, defined as 3000 bp upstream of exon 1 for the purpose of this analysis (observed CpG dinucleotides/total dinucleotides x 100).
[e]%CpG dinucleotides in the proximal promoter region, defined as 0 to -500 bp upstream of exon 1 (observed CpG dinucleotides/total dinucleotides x 100).
[f]%CpG dinucleotides in the putative gene promoter (3000 bp) and exon 1 (observed CpG dinucleotides/total dinucleotides x 100).
[g]These genes contained multiple distinct weak CpG islands.

**Table 4.** Putative Epigenetically-regulated Genes Identified in MCF-7 Cells After Demethylating Treatment with Either 5-aza or 5-aza + TSA[a]

| Gene Designation | Gene Name | GenBank Accession Number | Relative Expression Level[b] |
|---|---|---|---|
| **Genes Responding to 5-aza Treatment** (n=17)[c] | | | |
| ***Genes with Typical CpG Features***[d] | | | |
| *CENTB2* | Centaurin, beta 2 | NM_012287.3 | 2.2 |
| *CYP1A1* | Cytochrome P450, family 1, subfamily A, polypeptide 1 | NM_000499.2 | 2.0 |
| *ID1* | Inhibitor of DNA binding 1 | NM_002165.2 | 2.0 |
| *MAP3K8* | Mitogen-activated protein kinase kinase kinase 8 | NM_005204.2 | 2.4 |
| ***Genes with Intermediate CpG Features***[d] | | | |
| *COL4A6* | Collagen type IV alpha 6 | NM_001847 | 2.2 |
| *GDF-15* | Growth differentiation factor 15 | NM_004864.1 | 2.5 |
| *LCP2* | Lymphocyte cytosolic protein 2 | NM_005565.3 | 3.0 |
| *LXN* | Latexin protein | NM_020169.2 | 2.0 |
| *NOX1* | NADPH oxidase 1 | NM_007052.3 | 2.8 |
| *SLICK* | Potassium channel, subfamily T, member 2 | NM_198503.2 | 2.5 |
| *WISP2* | WNT1 inducible signaling pathway protein 2 | NM_003881.2 | 2.2 |
| *ZFHX1B* | Zinc finger homeobox 1b | NM_014795.2 | 2.1 |
| ***Genes with Atypical CpG Features***[d] | | | |
| *CTAGE-1* | CTAGE-1 protein | NM_022663.1 | 2.2 |
| *GH2* | Growth hormone 2 | NM_022558 | 2.5 |
| *GNRH1* | Gonadotropin-releasing hormone 1 | NM_000825.2 | 3.0 |
| *RARRES3* | Retinoic acid receptor responder | NM_004585.2 | 2.2 |
| **Genes Responding to 5-aza + TSA Treatment** (n=50)[c] | | | |
| ***Genes with Typical CpG Features***[d] | | | |
| *ABCG2* | ATP-binding cassette, subfamily G, member 2 | NM_004827.2 | 3.6 |
| *AQP3* | Aquaporin 3 | NM_004925.3 | 2.7 |
| *BUB1* | Budding uninhibited by benzimidazoles 1 | NM_004336 | 2.4 |
| *FLJ90013* | Cytomegalovirus partial fusion receptor | NM_153365.1 | 2.1 |
| *GREM1* | Gremlin 1, cysteine knot superfamily | NM_013372.4 | 2.1 |
| *HAPLN1* | Homo sapiens hyaluronan and proteogylcan link protein 1 | NM_001884.2 | 2.2 |
| *HLA-B* | Major histocompatibility complex, class 1, B | NM_000885.3 | 2.2 |
| *HLA-C* | Major histocompatibility complex, class 1, C | NM_002117 | 2.4 |
| *IFITM1* | Interferon-inducible protein 9-27 | NM_003641 | 2.0 |
| *IGFBP3* | Insulin-like growth factor binding protein 3 | NM_000598.4 | 3.1 |
| *INHA* | Inhibin alpha | NM_002191.2 | 2.6 |
| *ITGA4* | Integrin, alpha 4 (antigen CD49D) | NM_000885.4 | 2.1 |
| *KRTHB6* | Keratin, hair, basic, 6 (monilethrix) | NM_002284.2 | 2.5 |
| *L1CAM* | L1 cell adhesion molecule | NM_000425.2 | 2.6 |
| *LOXL2* | Lysyl oxidase-like 2 | NM_002318.1 | 2.5 |
| *PLSCR1* | Phospholipid scramblase 1 | NM_021105.1 | 2.5 |
| *PSMB9* | Proteosome (prosome, macropain), subunit beta, type 9 | NM_002800.3 | 2.3 |
| *RAFTLIN* | Raft-linking protein | NM_015150.1 | 3.1 |
| *RGS16* | Regulator of G-protein signaling 16 | NM_002928.2 | 2.3 |
| *RIG-I* | DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide | NM_014314.2 | 2.1 |
| *STAT1* | Signal transducer and activator of transcription 1 | NM_007315.2 | 2.0 |
| *SYNGR3* | Synaptogyrin 3 | NM_004209.4 | 2.3 |
| *TAP1* | Transporter 1, ATP-binding cassette, subfamily B | NM_00593.5 | 2.8 |
| *TUBB* | Tubulin, beta polypeptide | NM_178014.2 | 2.2 |

| | | | |
|---|---|---|---|
| *TXNRD1* | Thioredoxin reductase 1 | NM_003330.2 | 2.4 |
| *UBE2L6* | Ubiquitin-conjugating enzyme E2L 6 | NM_004223.3 | 2.3 |
| | | | |
| ***Genes with Intermediate CpG Features***[d] | | | |
| *BST2* | Bone marrow stromal cell antigen 2 | NM_004335.2 | 2.7 |
| *C3* | Complement component 3 | NM_000064.1 | 2.3 |
| *CGB* | Chorionic gonadotropin, beta polypeptide | NM_000737.2 | 3.8 |
| *DIO2* | Deiodinase, iodothyronine, type II | NM_013989.2 | 4.2 |
| *FLJ20035* | Hypothetical protein FLJ20035 | NM_017631.3 | 3.0 |
| *GAGE4* | G antigen 4 | NM_001474 | 2.2 |
| *IFIT1* | Interferon-induced protein with tetratricopeptide repeats 1 | NM_001548.1 | 3.4 |
| *LY6D* | Lymphocyte antigen 6 complex, locus D | NM_003695.2 | 2.2 |
| *PDE6C* | Phosphodiesterase 6C | NM_006204.2 | 2.0 |
| *PLAC8* | Placenta-specific 8 | NM_016619 | 2.6 |
| *S100P* | S100 calcium binding protein P | NM_05980.2 | 2.2 |
| *SCGB1A1* | Secretoglobin, family 1A, member 1 (uteroglobin) | NM_003357.3 | 2.6 |
| *SP110* | SP110 nuclear body protein | NM_004509.1 | 2.2 |
| | | | |
| ***Genes with Atypical CpG Features***[d] | | | |
| *CGA* | Glycoprotein hormone, alpha polypeptide | NM_000735.2 | 8.1 |
| *CYP11A1* | Cytochrome P450, family 11, subfamily A, polypeptide 1 | NM_000781.1 | 4.0 |
| *FLG* | Filaggrin | NM_002016.1 | 2.4 |
| *GJA1* | Gap junction protein, alpha 1 | NM_00165.2 | 2.6 |
| *ITGB6* | Integrin, beta 6 | NM_000888.3 | 2.5 |
| *KYNU* | Kynureninase (L-kynurenine hydrolase) | NM_003937.1 | 2.1 |
| *OAS1* | 2',5'-Oligoadenylate synthetase 1 | NM_016816 | 2.3 |
| *S100A8* | S100 calcium binding protein A8 (calgranulin A) | NM_002964.3 | 2.9 |
| *S100A9* | S100 calcium binding protein A9 (calgranulin B) | NM_002965.2 | 5.8 |
| *S100A12* | S100 calcium binding protein A12 (calgranulin C) | NM_005621.1 | 2.1 |

[a]Genes listed in this table were found to display increased expression in MCF-7 cells in response to treatment with either 5-aza and 5-aza + TSA, but not both. Genes that displayed increased expression in response to both treatments are listed in Table 3.

[b]Relative expression levels are expressed as average fold control levels of expression at the end of 3 weeks of demethylating treatment (5-aza or 5-aza + TSA).

[c]Genes responding to 5-aza treatment included FLJ12055 (Genbank accession AK022117), and genes responding to 5-aza + TSA included DKFZp761G18121 (Genbank accession BC018100). These genes were omitted from this analysis as no promoter sequence information was available.

[d]Genes with typical CpG features contain typical CpG islands (defined as a region of $\geq$200 bp with $\geq$50% C+G and $\geq$0.6 CpG observed/CpG expected) (65). Genes with intermediate CpG features contain weak CpG islands (defined as a region of >50 bp but <200 bp with $\geq$50% C+G and $\geq$0.6 CpG observed/CpG expected). Genes with atypical CpG features do not contain CpG islands (typical or weak).

weeks of cell culture revealed no significant variations in gene expression level for *C8orf4* or *ZC3HDC1* across all time points (data not shown). *C8orf4* was not expressed at any time point and *ZC3HDC1* was expressed at low (but detectable) levels at all time points. β-actin RNA was also expressed evenly across all time points in control MCF-7 cells (data not shown). These results indicate that MCF-7 cells propagated in control growth medium produce consistent patterns of gene expression over time in cell culture. Subsequent analyses focused on the differential expression of putative epigenetically-regulated genes (n=20) in response to 5-aza and 5-aza + TSA treatment. In total, 15/20 (75%) of these genes were shown by RT-PCR to be increased in response to 5-aza + TSA (including *BF*, *C8orf4*, *CEACAM5*, *CEACAM6*, *CST6*, *CYP1B1*, *FLJ10134*, *G1P2*, *G1P3*, *IFI27*, *ISGF3G*, *KRTHB1*, *LCN2, SCNN1A*, and *ZC3HDC1*), and 18/20 (90%) of these genes were shown by RT-PCR to be increased in response to 5-aza treatment alone (those listed above and *IGFBP5*, *LGALS3BP*, and *SAT*). The remaining two genes were not examined (*CRIP1* and *PSG6*). Figure 8 shows representative RT-PCR reactions for seven genes (*C8orf4*, *CEACAM5*, *CEACAM6*, *IFI27*, *ISGF3G*, *SCNN1A*, and *ZC3HDC1*). *C8orf4*, *IFI27*, and *ZC3HDC1* were expressed at low or undetectable levels in control MCF-7 cells, but demonstrated significantly increased expression 3 weeks following treatment with either 5-aza or 5-aza + TSA (Figure 8B-D). Likewise, *BF, CST6*, *CYP1B1*, *FLJ10134*, *G1P2*, *G1P3*, *KRTHB1*, *IGFBP5*, *LCN2*, *LGALS3BP*, and *SAT* were expressed at very low levels in control MCF-7 cells followed by an increase in expression with 5-aza and/or 5-aza + TSA treatment (data not shown). *CEACAM5*, *CEACAM6*, *ISGF3G*, and *SCNN1A* were expressed at moderate levels in control MCF-7 cells, and each of these genes showed significantly increased levels of expression 3 weeks after treatment (Figure 8A, E-G).

**Figure 8. Expression of putative epigenetically-regulated genes in response to demethylating treatment in MCF-7 breast cancer cells.** Representative agarose gels of RT-PCR products are shown. In each panel, lane 1 corresponds to a no cDNA template control, lane 2 corresponds to cDNA from untreated (control) MCF-7 cells, and lanes 3-4 correspond to MCF-7 cells after 3 weeks of treatment with 5-aza or 5-aza + TSA, respectively. Panel A, *CEACAM6*; Panel B, *C8orf4*; Panel C, *IFI27*; Panel D, *ZC3HDC1*; Panel E, *CEACAM5*; Panel F, *SCNN1A*; Panel G, *ISGF3G*.

A
-
untreated
5-aza (week 3)
5-aza + TSA (week 3)
CEACAM6
B
-
untreated
5-aza (week 3)
5-aza + TSA (week 3)
C8orf4
C
-
untreated
5-aza (week 3)
5-aza + TSA (week 3)
IFI27
D
-
untreated
5-aza (week 3)
5-aza + TSA (week 3)
ZC3HDC1
E
-
untreated
5-aza (week 3)
5-aza + TSA (week 3)
CEACAM5
F
-
untreated
5-aza (week 3)
5-aza + TSA (week 3)
SCNN1A
G
-
untreated
5-aza (week 3)
5-aza + TSA (week 3)
ISGF3G

***Promoter Sequence Features of Putative Epigenetically-regulated Genes***

An analysis of the promoters and 5'-upstream sequences (3000 bp) for each of the 20 genes identified in MCF-7 cells that responded to both 5-aza and 5-aza + TSA treatment was performed to identify common sequence features that may be associated with methylation-dependent epigenetic regulation, with emphasis on CpG dinucleotide frequency and distribution. This analysis revealed a tremendous variation in promoter CpG content and organization among these putative epigenetically-regulated genes (Table 3). Based upon a comparative analysis of the CpG features of their promoter and proximal sequences (exon 1), we grouped the putative epigenetically-regulated genes identified in this study into three distinct classes, including: (i) genes with typical CpG features (typical CpG islands within the promoter or exon 1), (ii) genes with intermediate CpG features (weak CpG islands within the promoter or exon 1), and (iii) genes with atypical CpG features (no CpG islands).

Using the commonly accepted criteria for a typical CpG island (65), 9/20 (45%) genes were found to contain a CpG island in either their promoter and/or exon 1 (Table 3). This subset of genes exhibits the typical features expected for an epigenetically-regulated gene. Among these genes, 4/9 (44%) contain distinct (typical) CpG islands in both the promoter region and exon 1 (Table 3). The CpG islands found in the promoter and/or exon 1 sequences of *CST6*, *CYP1B1*, *KRTHB1*, *SAT*, and *ZC3HDC1* withstood a more rigorous CpG island analysis ($\geq$200 bp with $\geq$60% G+C and $\geq$0.7 CpG observed/CpG expected), which approximates a new standard suggested by Takai and Jones (63). In 7/9 (78%), genes with typical CpG features, distinct weak CpG islands were detected in the promoter and/or exon 1 (Table 3). As expected, all of the CpG island-containing genes demonstrated significant promoter CpG content, with the highest concentration of CpG dinucleotides in the first 500

bp upstream of the transcription start site in most cases (Table 3). However, CpG islands were detected in several genes with relatively low CpG content (including *G1P3* and *KRTHB1*). In some cases, the CpG content of exon 1 exceeds that of the proximal promoter (like in the case of *ZC3HDC1*), reflecting the presence of a typical CpG island (Table 3). Five genes contain >10% CpG content in the first 500 bp upstream of the transcription start site, some with much more extensive regions of CpG density (Table 3). Other genes contain more focused regions of CpG density that are confined to the portion of the promoter sequence that is proximal to the transcriptional start site. These include *CST6* (17.6% CpG in proximal 250 bp), *G1P2* (10.4% CpG in proximal 250 bp), *SAT* (18.4% CpG in proximal 250 bp), and *FLJ10134* (20.8% CpG in proximal 250 bp). While 8/9 (89%) genes with typical CpG islands contained Alu repeats (1-6 repeats; average = 3 repeats per promoter), the CpG-rich regions (CpG islands) did not correspond to Alu repetitive elements.

Weak CpG islands were detected in the promoter and/or exon 1 sequences of 8/20 (40%) genes that lacked typical CpG islands (Table 3). We have described this subset of genes as displaying intermediate CpG features based upon the observation that they lack typical CpG islands, but contain smaller regions of CpG density (weak CpG islands). These weak CpG islands occur most often in gene promoter sequences (6/8, 75%), rarely in exon 1 alone (2/8, 25%), or in both the promoter and exon 1 (1/8, 13%) (Table 3). Genes with weak CpG islands display lower promoter and 5'-upstream sequence CpG content than genes containing typical CpG islands (3.0 $\pm$ 0.3% versus 5.0 $\pm$ 1.0%, N.S.), but this difference was most pronounced when the first 500 bp upstream from the transcription start site was examined (2.5 $\pm$ 0.3% versus 9.1 $\pm$ 1.8%, $P$=0.0066). Alu repeats were detected in 5/8 (63%) genes

with intermediate CpG features, but with fewer repeats than genes with typical CpG features (1-4 repeats; average = 2 per promoter).

Three genes (*IFI27*, *LGALS3BP*, and *SCNN1A*) contain no CpG islands (typical or weak) in their promoter or exon 1 sequences (Table 3). These genes are CpG-deficient, with no regions of CpG density and no clustering of CpG dinucleotides. Based upon these CpG characteristics, we have described this subset of genes as exhibiting atypical CpG features. Similar to genes with intermediate CpG features, genes with atypical CpG features display significantly lower promoter and 5'-upstream sequence CpG content when compared to genes with typical CpG features. The average CpG content of the atypical features genes was 2.9% when 3000 bp of sequence was examined and 2.7% when the first 500 bp proximal to the transcriptional start site was evaluated (Table 3). Single Alu repeats were detected in the promoters of each of these genes.

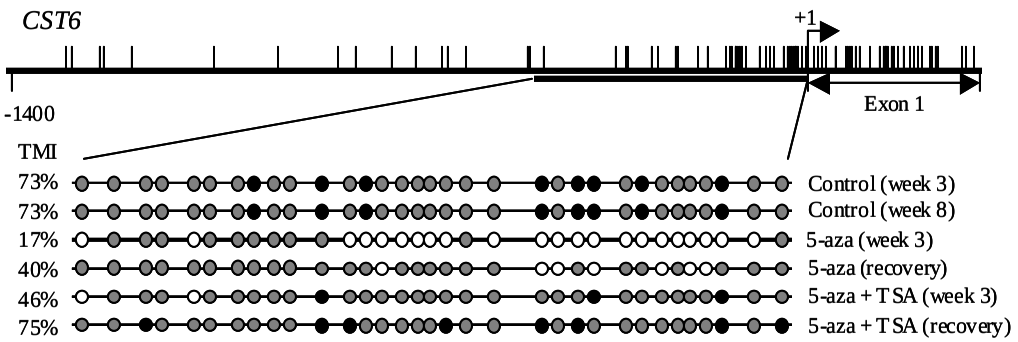
A similar analysis of promoter and 5'-upstream CpG sequence features was performed for genes that respond to 5-aza (n=17) or 5-aza + TSA treatment (n=50), but not both. *FLJ12055* and *DKFZp761G1812* were omitted from this analysis due to a lack of known promoter sequence. Genes with CpG sequence characteristics corresponding to each of the proposed classes of putative epigenetically-regulated genes were identified in these groups of genes. Among those responding to 5-aza alone (n=16), 4/16 (25%) genes exhibit typical CpG features, 8/16 (50%) exhibit intermediate CpG features, and 4/16 (25%) display atypical CpG features (Table 4). Likewise, among genes responding to 5-aza + TSA (n=49), 26/49 (53%) exhibit typical CpG features, 13/49 (27%) exhibit intermediate CpG features, and 10/49 (20%) display atypical CpG features (Table 4).

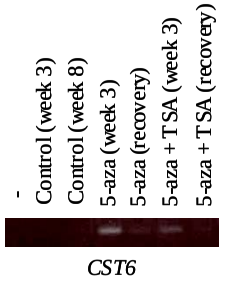
***Bisulfite Sequencing Demonstrates that Putative Epigenetically-regulated Genes are Subject to Methylation-dependent Regulation***

Methylation analysis by bisulfite sequencing (122) was performed on the promoter regions of prototype genes from each proposed class of putative epigenetically-regulated gene: *CST6* (typical CpG features), *C8orf4* (intermediate CpG features), and *IFI27* (atypical CpG features). This analysis produced evidence for the direct regulation of genes in each proposed class by CpG methylation in MCF-7 breast cancer cells, and began to identify CpG methylation events that are critical for gene silencing.

*Methylation Analysis of the CST6 Promoter in MCF-7 Breast Cancer Cells.* To directly address whether *CST6* is methylated in MCF-7 cells, we analyzed 33 CpG dinucleotides from a segment of the promoter region (+10 to -636) that contains a typical CpG island (Figure 9). *CST6* is not expressed in untreated MCF-7 cells, but expression is significantly increased with 5-aza and 5-aza + TSA treatment (Figure 9). Furthermore, the treatment-related increase in gene expression is reversible, and treatment withdrawal results in significant reduction of *CST6* mRNA levels (Figure 9). In untreated MCF-7 cells, 8/33 (24%) CpGs were 100% methylated, 25/33 (76%) CpGs were intermediately methylated, and 0/33 (0%) CpGs were unmethylated, producing a TMI for the promoter of 73%. Treatment of MCF-7 cells with 5-aza or 5-aza + TSA resulted in demethylation of 30/33 (91%) and 20/33 (61%) CpGs respectively, resulting in TMI values of 17% and 46%, and *CST6* was expressed at detectable levels (Figure 9). Treatment of MCF-7 cells with 5-aza resulted in a relatively higher level of expression for *CST6*, which appears to correlate with the degree of promoter demethylation in this region (Figure 9). Withdrawal of 5-aza or 5-aza + TSA treatment

**Figure 9. Correlative analysis of promoter methylation and gene expression for *CST6* in MCF-7 cells.** The distribution of CpG dinucleotides proximal to the transcription start site in the promoter (0 to -1400 nucleotides) and exon 1 (0 to +294 nucleotides) of *CST6* are depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). Methylation analysis was performed on a region of the promoter spanning from +10 to -636 (indicated by a solid horizontal line), which contains 33 CpG dinucleotides and is part of a large CpG island. A summary of results for the methylation analysis is shown for control MCF-7 cells and cells treated with 5-aza and 5-aza + TSA for the week 3 and week 8 time points. Each circle represents 3-5 replicates of bisulfite sequencing. Black circles correspond to fully (100%) methylated CpGs, gray circles correspond to CpGs with intermediate methylation, and open circles correspond to unmethylated CpGs. TMI values for the promoter region (33 CpGs) are given for control MCF-7 cells and cells treated with 5-aza and 5-aza + TSA for the week 3 and week 8 time points. Representative RT-PCR reactions are shown demonstrating the level of *CST6* expression in control and treated MCF-7 cells at each time point. The correspondence between *CST6* promoter methylation status and gene expression for all treatments is shown in the inset table.



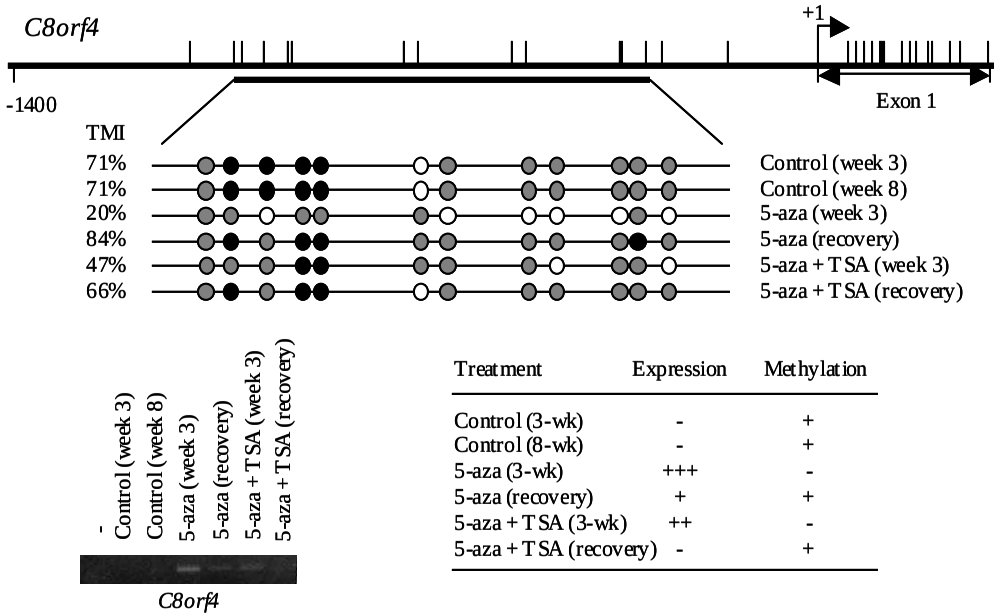




| Treatment | Expression | Methylation |
|---|---|---|
| Control (3-wk) | - | + |
| Control (8-wk) | - | + |
| 5-aza (3-wk) | +++ | - |
| 5-aza (recovery) | + | + |
| 5-aza + TSA (3-wk) | ++ | - |
| 5-aza + TSA (recovery) | - | + |

resulted in silencing of *CST6* gene expression concurrent with remethylation of the majority of CpG dinucleotides producing TMI values of 40% and 75%, respectively (Figure 9).

*Methylation Analysis of the C8orf4 Promoter in MCF-7 Breast Cancer Cells.* The promoter of *C8orf4* is CpG-deficient, but does contain a weak CpG island in exon 1. *C8orf4* is not expressed at detectable levels in MCF-7 cells. We analyzed an 812 bp segment (-278 to -1090) of the *C8orf4* promoter containing 12 CpG dinucleotides (Figure 10). In untreated MCF-7 cells, 4/12 (33%) of these CpGs were 100% methylated, 7/12 (58%) CpGs were methylated at an intermediate level, and 1/12 (8%) CpGs were unmethylated, with the greatest concentration of methylated CpGs in a 103 bp region (-926 to -1029) containing 5 CpG dinucleotides (100% methylated; n=5), and producing a TMI value of 71% (Figure 10). Treatment of MCF-7 cells with 5-aza or 5-aza + TSA resulted in demethylation of 10/12 (83%) and 8/12 (67%) CpG dinucleotides resulting in TMI values of 20% and 47% respectively, and coordinate expression of *C8orf4* (Figure 10). Treatment of MCF-7 cells with 5-aza resulted in a relatively higher level of expression for *C8orf4*, which appeared to correlate with the degree of promoter demethylation in this region (Figure 10). However, treatment withdrawal resulted in silencing or significantly lower expression of *C8orf4* and coordinate remethylation of the majority of these CpGs (Figure 10). We also examined the methylation status of *C8orf4* exon 1 (between +23 and +453) which contains 14 CpGs, forming a weak CpG island. In control MCF-7 cells, 7/14 (50%) CpGs are 100% methylated and 7/14 (50%) CpGs were methylated at an intermediate level, producing a TMI for the exon 1 region of 73%. Following treatment with 5-aza or 5-aza + TSA, 11/14 (79%) CpGs become demethylated resulting in TMI values of 31% and 29% respectively, and concurrent reexpression of the gene. Withdrawal of the treatment resulted in silencing of gene

**Figure 10. Correlative analysis of promoter methylation and gene expression for *C8orf4* in MCF-7 cells.** The distribution of CpG dinucleotides proximal to the transcription start site in the promoter (0 to -1400 nucleotides) and exon 1 (0 to +529 nucleotides) of *C8orf4* are depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). *C8orf4* contains no typical CpG islands in the promoter, but does contain a weak CpG island in exon 1 (see results). Methylation analysis was performed on a region of the promoter spanning from -278 to -1090 (indicated by a solid horizontal line), which contains 12 CpG dinucleotides. A summary of results for the methylation analysis is shown for control MCF-7 cells and cells treated with 5-aza and 5-aza + TSA for the week 3 and week 8 time points. Each circle represents 5 replicates of bisulfite sequencing. Black circles correspond to fully (100%) methylated CpGs, gray circles correspond to CpGs with intermediate methylation, and open circles correspond to unmethylated CpGs. TMI values for the promoter region (12 CpGs) are given for control MCF-7 cells and cells treated with 5-aza and 5-aza + TSA for the week 3 and week 8 time points. The methylation status of *C8orf4* exon 1 (between +23 and +453) which contains 14 CpGs was also analyzed (see results). Representative RT-PCR reactions are shown demonstrating the level of *C8orf4* expression in control and treated MCF-7 cells at each time point. The correspondence between *C8orf4* promoter methylation status and gene expression for all treatments is shown in the inset table.




| Treatment | Expression | Methylation |
|---|---|---|
| Control (3-wk) | - | + |
| Control (8-wk) | - | + |
| 5-aza (3-wk) | +++ | - |
| 5-aza (recovery) | + | + |
| 5-aza + TSA (3-wk) | ++ | - |
| 5-aza + TSA (recovery) | - | + |

expression, but without significant remethylation of this weak CpG island (data not shown). These results suggest that CpG methylation events occurring within the promoter region rather than exon 1 may be most important for the silencing of *C8orf4* in MCF-7 cells.

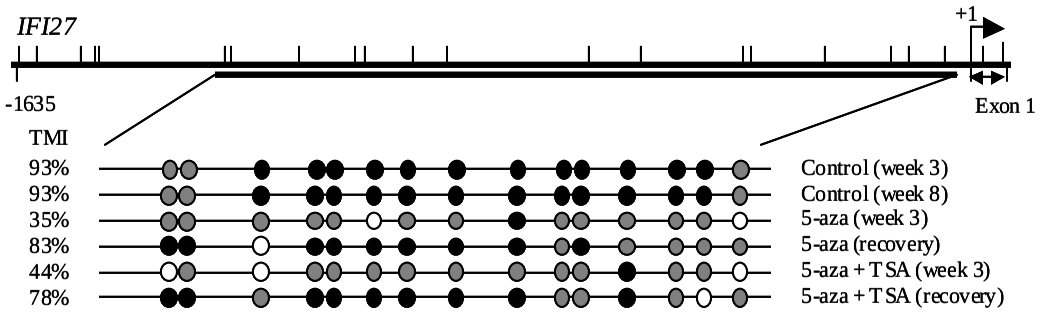
*Methylation Analysis of the IFI27 Promoter in MCF-7 Breast Cancer Cells.* We examined the methylation status of 15 CpGs within a 1271 bp segment (-16 to -1287) of the *IFI27* promoter, which lacks typical or weak CpG islands. Untreated MCF-7 cells do not express *IFI27* and 12/15 (80%) CpG dinucleotides within the promoter region are 100% methylated and 3/15 (20%) are methylated at an intermediate level, producing a TMI value of 93% (Figure 11). Treatment of MCF-7 cells with 5-aza or 5-aza + TSA resulted in expression of *IFI27* and demethylation of the majority of CpG dinucleotides: 12/15 (80%) CpGs were demethylated following 5-aza treatment and 9/15 (60%) CpGs were demethylated after 5-aza + TSA treatment, resulting in TMIs of 35% and 44% respectively (Figure 11). Withdrawal of 5-aza treatment leads to remethylation of the majority of CpG dinucleotides (13/15, 87%) with a TMI value of 83%. Likewise, withdrawal of 5-aza + TSA treatment resulted in remethylation of 12/15 (80%) CpGs exhibiting a TMI value of 78% with concurrent loss of gene expression (Figure 11).

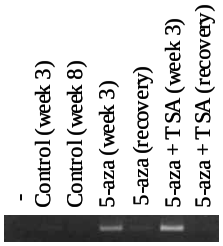
## B. DNA Methylation-Dependent Silencing of *CST6* in Human Breast Cancer Cell Lines

### *CST6 is Differentially Expressed Among Breast Cancer Cell Lines*

*CST6* gene expression was examined by semiquanitative RT-PCR in MDA-MB-231, MDA-MB-415, MDA-MB-435S, MDA-MB-436, MDA-MB-453, MDA-MB-468, BT-549, BT-20, Hs578T, MCF-7, SK-BR-3, and ZR-75-1 breast cancer cell lines and two normal

**Figure 11. Correlative analysis of promoter methylation and gene expression for *IFI27* in MCF-7 cells.** The distribution of CpG dinucleotides proximal to the transcription start site in the promoter (0 to -1635 nucleotides) and exon 1 (0 to +65 nucleotides) of *IFI27* are depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). *IFI27* contains no CpG islands (typical or weak) or other regions of CpG density. Methylation analysis was performed on a region of the promoter spanning from -16 to -1287 (indicated by a solid horizontal line), which contains 15 CpG dinucleotides. A summary of results for the methylation analysis is shown for control MCF-7 cells and cells treated with 5-aza and 5-aza + TSA for the week 3 and week 8 time points. Each circle represents 4-5 replicates of bisulfite sequencing. Black circles correspond to fully (100%) methylated CpGs, gray circles correspond to CpGs with intermediate methylation, and open circles correspond to unmethylated CpGs. TMI values for the promoter region (15 CpGs) are given for control MCF-7 cells and cells treated with 5-aza and 5-aza + TSA for the week 3 and week 8 time points. Representative RT-PCR reactions are shown demonstrating the level of *IFI27* expression in control and treated MCF-7 cells at each time point. The correspondence between *IFI27* promoter methylation status and gene expression for all treatments is shown in the inset table.




*IFI27*

| Treatment | Expression | Methylation |
|---|---|---|
| Control (3-wk) | - | + |
| Control (8-wk) | - | + |
| 5-aza (3-wk) | +++ | - |
| 5-aza (recovery) | + | + |
| 5-aza + TSA (3-wk) | ++ | - |
| 5-aza + TSA (recovery) | - | + |

mammary epithelial cell lines, MCF10-2A and MCF12A. Five of the 12 breast cancer cell lines (42%) express detectable levels of *CST6* mRNA: MDA-MB-468 and SK-BR-3 cells express low levels of *CST6*, MDA-MB-231 and MDA-MB-415 cells express moderate levels of *CST6*, and BT-20 cells strongly express *CST6*. However, no *CST6* mRNA was found in the remaining 7 cell lines (58%) (Figure 12). Both MCF10-2A and MCF12A cells transcribe *CST6* at low levels, and β-actin mRNA was expressed evenly across all cell lines examined (Figure 12). These results are consistent with the recently published studies on MCF-7, SK-BR-3, and MDA-MB-231 cells (8). This analysis identified subsets of breast cancer cell lines that differentially express *CST6* mRNA, providing the cellular reagents for examination of methylation-dependent epigenetic regulation of *CST6* in breast cancer cells.

***5-aza Treatment Induces CST6 Expression in Breast Cancer Cell Lines***

*CST6*-negative cell lines Hs578T, MCF-7, MDA-MB-435S, MDA-MB-436, MDA-MB-453, and ZR-75-1 were treated with 5-aza, to determine if gene silencing was the likely result of DNA methylation. Four of 6 (67%) of these cell lines expressed significantly increased levels of *CST6* mRNA in response to 5-aza treatment (Figure 12B). The significant induction of *CST6* mRNA in response to 5-aza treatment in MCF-7 cells observed in this study is consistent with similar published studies (8). In contrast, 5-aza exposure of MDA-MB-453 and MDA-MB-436 cells resulted in a modest, but detectable increase in *CST6* mRNA (Figure 12B). The 5-aza-induced increase in *CST6* expression was completely reversible. Following a period of recovery after treatment withdrawal, *CST6* mRNA diminished to control levels (corresponding to untreated cells) in all cell lines examined (Figure 12B). The 5-aza treatment-related increases in *CST6* expression, combined with the

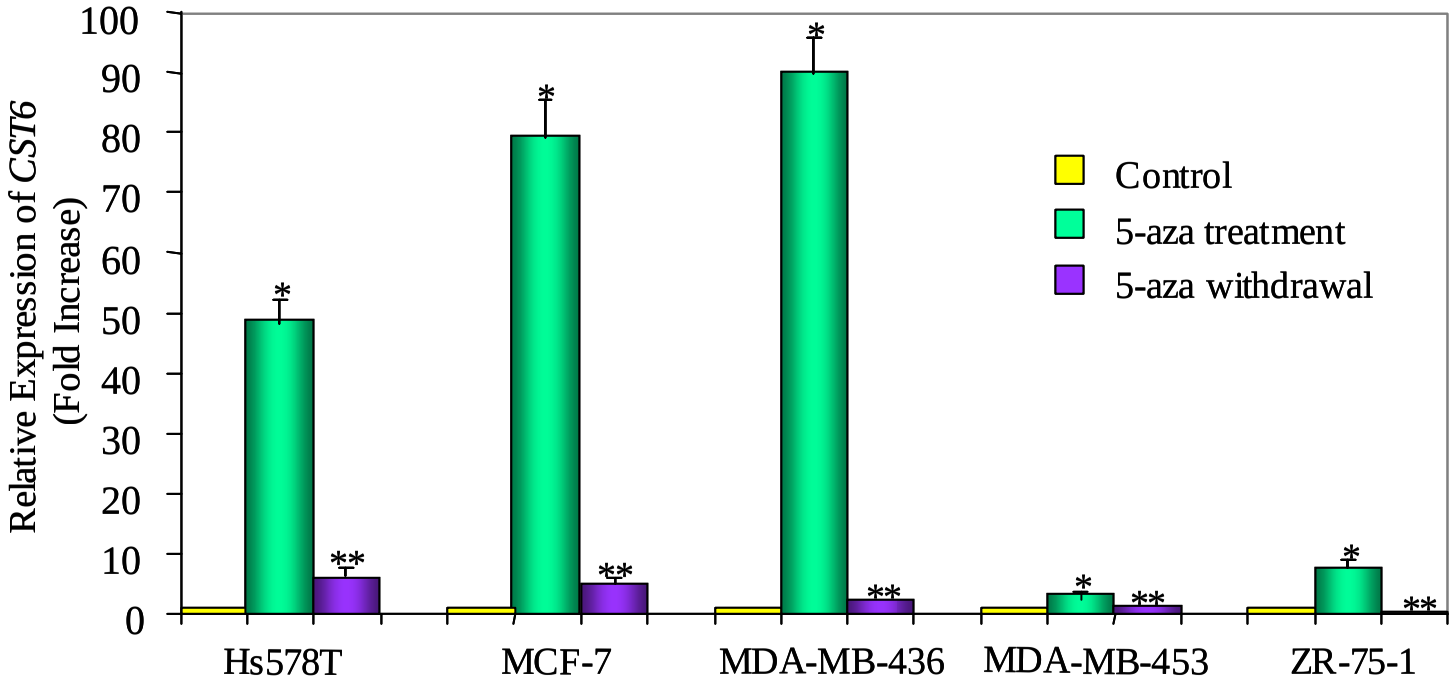
**Figure 12. *CST6* expression in human breast cancer cell lines and normal mammary epithelial cells.** Representative agarose gels of RT-PCR products are shown. (A) Differential expression of *CST6* among breast cancer cell lines and two normal mammary epithelial cell lines (MCF10-2A and MCF12A). The *CST6* RT-PCR product is 163 bp in size. β-actin was utilized as a sample control. (B) Lane 1 corresponds to cDNA from indicated untreated breast cancer cell lines; lanes 2 and 3 correspond to cells after 5-aza treatment or treatment withdrawal, respectively. β-actin RNA was evenly expressed in all samples (data not shown).

A
CST6
β-actin
DNA Ladder
-
MDA-MB-231
MDA-MB-415
MDA-MB-435S
MDA-MB-436
MDA-MB-453
MDA-MB-468
BT-549
DNA Ladder
BT-20
Hs578T
MCF-7
MCF10-2A
MCF12A
SK-BR-3
ZR-75-1
B
untreated
5-aza treated
5-aza withdrawal
Hs578T
MCF-7
MDA-MB-435S
MDA-MB-436
MDA-MB-453
ZR-75-1

loss of *CST6* expression following treatment withdrawal, suggests that *CST6* may be subject to methylation-dependent silencing in these breast cancer cell lines.

Expression of *CST6* was analyzed using real-time PCR in *CST6*-negative cell lines Hs578T, MCF-7, MDA-MB-436, MDA-MB-453, and ZR-75-1 to quantitate the 5-aza treatment-related increases in gene expression. Consistent with the RT-PCR results, untreated cell lines express extremely low levels of *CST6* mRNA, and exposure to 5-aza resulted in significant increases in *CST6* mRNA (Figure 13). Exposure of Hs578T cells to 5-aza led to a 49-fold increase in *CST6* expression ($P$=0.0036 compared to control) and withdrawal of 5-aza resulted in a significant reduction ($P$=0.0021 compared to 5-aza treated) of gene expression to a level that approaches that of control (untreated) cells (Figure 13). Likewise, 5-aza treatment of MCF-7 and MDA-MB-436 cells produced 80-fold ($P$=0.0058) and 90-fold ($P$=0.0042) increases in *CST6* expression, respectively, and withdrawal of the 5-aza treatment lead to significant reductions in *CST6* expression in both cell lines ($P$=0.0062 and $P$=0.0043, respectively) (Figure 13). Of note, the induction of *CST6* mRNA in MDA-MB-436 cells in response to 5-aza treatment was consistently demonstrable with all methods, but the magnitude of increased gene expression detected by real-time RT-PCR was greater than that detected using RT-PCR, possibly due to the increased sensitivity of the real-time method. Exposure of MDA-MB-453 and ZR-75-1 cells to 5-aza produced modest increases in *CST6* mRNA (3-fold and 8-fold, respectively) that were statistically significant ($P$=0.0017 and $P$=0.0479), and withdrawal of treatment resulted in significant decreases in *CST6* expression to levels that were comparable to untreated cells ($P$=0.0070 and $P$=0.0393, respectively) (Figure 13). These data show that 5-aza treatment of *CST6*-negative cell lines

**Figure 13. Quantitative real-time PCR analysis of *CST6* in breast cancer cell lines.** β-actin was used to normalize gene expression levels for each cell line and differences in *CST6* expression were determined using the comparative Ct method. *CST6* gene expression for cell lines treated with 5-aza (green bars) and after treatment withdrawal (purple bars) are expressed as relative fold-change compared to control values (set at 1.0). Error bars reflect S.E.M. Values for gene expression that do not show error bars reflect data where the S.E.M could not be depicted graphically. *, Denotes statistical significance at $P<0.05$ compared to control values. **, Denotes statistical significance at $P<0.04$ compared to 5-aza treatment values.


Relative Expression of CST6
(Fold Increase)
100
90
80
70
60
50
40
30
20
10
0
Control
5-aza treatment
5-aza withdrawal
Hs578T
MCF-7
MDA-MB-436
MDA-MB-453
ZR-75-1

results in statistically significant increases in *CST6* mRNA, and suggests strongly that *CST6* is subject to methylation-dependent silencing in a variety of breast cancer cell lines.

***Methylation-dependent Silencing of CST6 in Select Breast Cancer Cell Lines***

To facilitate a correlative analysis of *CST6* gene expression and *CST6* CpG island methylation status, we analyzed 55 CpG dinucleotides from a segment of the proximal promoter region and exon 1 (+242 to -228) in normal breast epithelial cells (MCF12A), two breast cancer cell lines (BT-20 and SK-BR-3) that express *CST6*, and five breast cancer cell lines (Hs578T, MCF-7, MDA-MB-435S, MDA-MB-436, and MDA-MB-453) that lack expression of *CST6*. Multiple clones (n=3-5) corresponding to the *CST6* promoter and exon 1 from each cell line were analyzed by sodium bisulfite sequencing and individual CpGs were scored for methylation status. In MCF12A cells, 35/55 (64%) CpGs were not methylated, 18/55 (33%) CpGs were methylated at an intermediate level, and 2/55 (3%) CpGs were 100% methylated, producing a TMI for the promoter/exon 1 of 16% (Figure 14). BT-20 cells exhibit sparse methylation of *CST6*: 53/55 (96%) CpGs were unmethylated in all clones analyzed, resulting in a TMI of 1%. The *CST6* gene in SK-BR-3 cells was significantly more methylated than MCF12A or BT-20 cells (especially within exon 1): 39/55 (71%) CpGs show some level of methylation, producing a TMI of 45% (Figure 14). Overall, the *CST6*-positive cell lines (MCF12A, BT-20, and SK-BR-3) exhibit low levels of methylation within the proximal promoter/exon1 of *CST6* (mean TMI = 21 ± 13%) (Table 5). In contrast, breast cancer cell lines that do not express *CST6* exhibit hypermethylation of the *CST6* promoter/exon 1 region, with TMI values ranging from 72% to 98% (Figure 14). MDA-MB-435S and MDA-MB-453 cells were 100% methylated at 52/55 (95%) and 51/55

**Figure 14. Methylation analysis of the *CST6* proximal promoter and exon 1 in breast cancer cell lines and normal mammary epithelial cells that differentially express *CST6*.** A summary of the methylation analysis of the *CST6* promoter (23 CpGs) and exon 1 (32 CpGs) is shown. The black arrow indicates the start of transcription between CpGs 23 and 24. Black circles correspond to fully (100%) methylated CpGs, gray circles correspond to CpGs with intermediate methylation, and open circles correspond to unmethylated CpGs. TMI values for the entire promoter/exon 1 region (55 CpGs) are given for each breast cancer cell line. MCF12A, SK-BR-3, and BT-20 cells express *CST6*, while the remaining cell lines lack *CST6* expression.

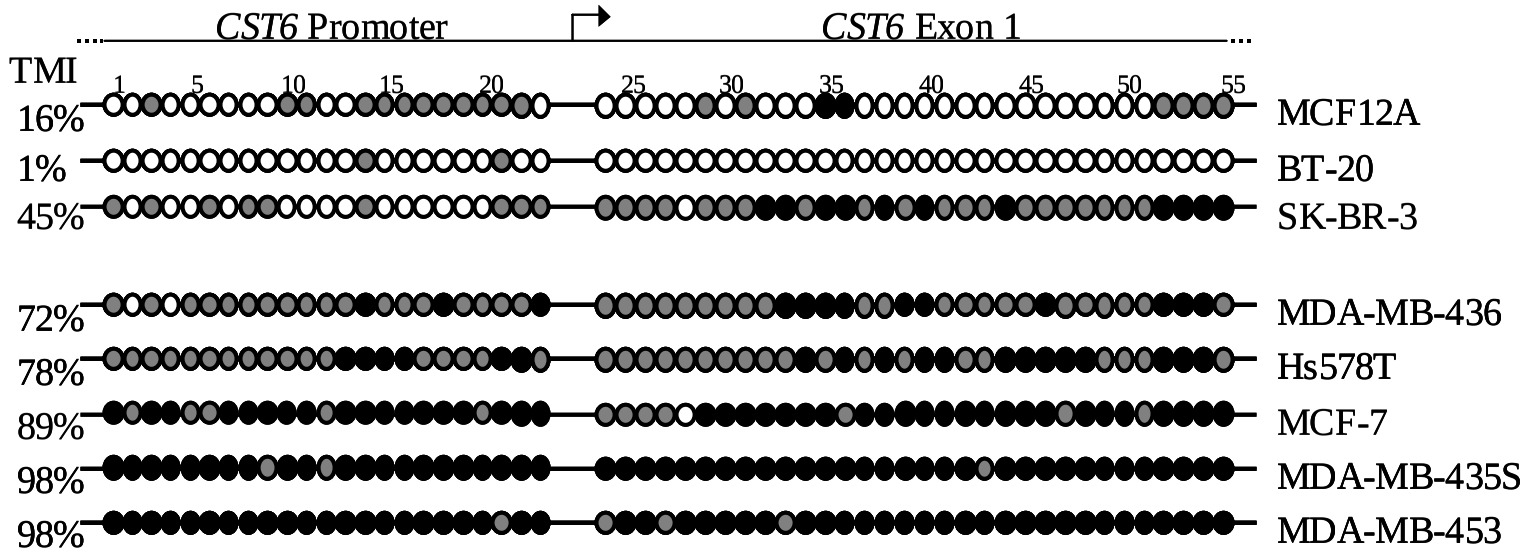

CST6 Promoter
CST6 Exon 1
TMI
1
5
10
15
20
25
30
35
40
45
50
55
16%
MCF12A
1%
BT-20
45%
SK-BR-3
72%
MDA-MB-436
78%
Hs578T
89%
MCF-7
98%
MDA-MB-435S
98%
MDA-MB-453

**Table 5.** Segmental Methylation Analysis of the *CST6* Promoter/Exon 1 in *CST6*-postive and -negative Breast Cancer Cell Lines

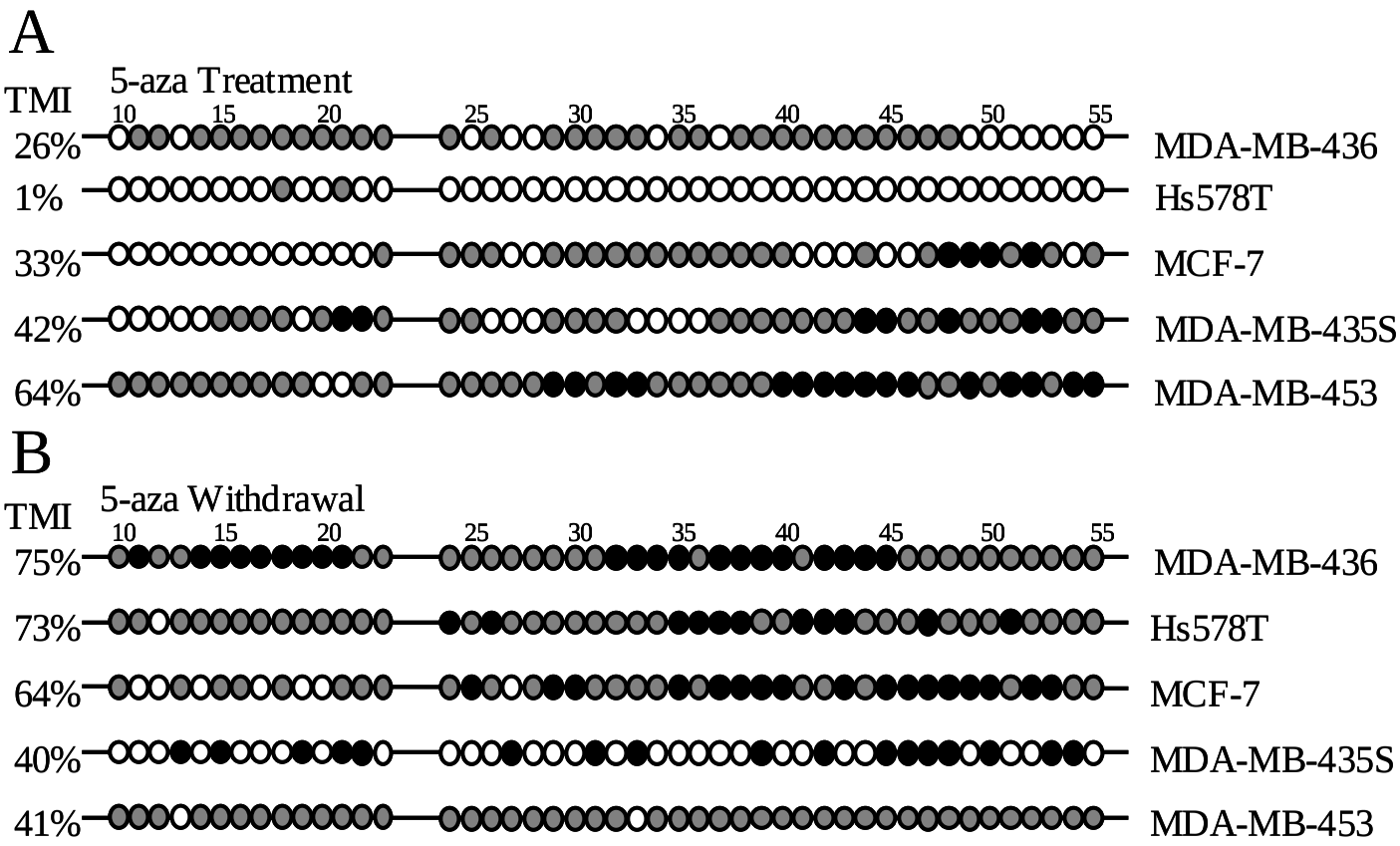
| Cell Line | CpGs 1-55[a] Control | CpGs 10-55[b] Control | CpGs 10-55[b] 5-aza Treatment | CpGs 10-55[b] 5-aza Withdrawal | CpGs 1-23[c] Control | CpGs 24-55[d] Control | CpGs 1-31[e] Control |
|---|---|---|---|---|---|---|---|
| *CST6-Positive* | | | | | | | |
| BT-20 | 1% | 1% | - | - | 3% | 0% | 2% |
| MCF12A | 16% | 18% | - | - | 19% | 14% | 15% |
| SK-BR-3 | 5% | 50% | - | - | 16% | 67% | 19% |
| *CST6-Negative* | | | | | | | |
| MDA-MB-435S | 98% | 99% | 42% | 40% | 96% | 99% | 97% |
| MDA-MB-453 | 98% | 97% | 64% | 41% | 99% | 97% | 97% |
| MCF-7 | 89% | 90% | 33% | 64% | 90% | 89% | 83% |
| Hs578T | 78% | 86% | 1% | 73% | 72% | 80% | 72% |
| MDA-MB-436 | 72% | 76% | 26% | 75% | 65% | 77% | 66% |
| *CST6-Positive* | 21 ± 13 | 23 ± 14 | - | - | 13 ± 5 | 27 ± 20 | 12 ± 5 |
| *CST6-Negative* | 87 ± 5 | 90 ± 4 | 33 ± 10 | 59 ± 8 | 84 ± 7 | 88 ± 4 | 83 ± 6 |
| t-test | $P$=0.0227[f] | $P$=0.0349 | $P$=0.0033[g] | N.S.[h] | $P$<0.0001[f] | N.S.[f] | $P$<0.0001[f] |

[a]CpG dinucleotides from a segment of the promoter region and exon 1 (+242 to -228).
[b]CpG dinucleotides from a segment of the proximal promoter region and exon 1 (+118 to -228). The values indicated under 5-aza treatment are from cell lines at the conclusion of a 3-week treatment period. The values indicated under 5-aza withdrawal are from cell lines that had recovered from 5-aza treatment for 5 weeks.
[c]CpG dinucleotides from the proximal promoter region (0 to -228).
[d]CpG dinucleotides from exon 1 (0 to +242).
[e]CpG dinucleotides from the proximal promoter region and 8 CpGs from exon 1 (+50 to -228).
[f]*P* values provided reflect a statistical comparison of average TMI values corresponding to *CST6*-positive (n=3) and *CST6*-negative (n=5) cell lines for the designated gene segments.
[g]*P* value for the comparison of average TMI values corresponding to control and 5-aza treated *CST6*-negative cell lines for the designated gene segment.
[h]*P* value for the comparison of average TMI values corresponding to 5-aza treated *CST6*-negative cell lines and the same cell lines 5-aza withdrawal.

(93%) CpGs, respectively (Figure 14). In MCF-7 cells, 42/55 (76%) CpGs were 100% methylated, 12/55 (22%) CpGs were methylated at an intermediate level, and 1/55 (2%) CpGs was unmethylated (Figure 14). Hs578T and MDA-MB-436 cells were 100% methylated at 19/55 (35%) and 13/55 (24%) CpGs, respectively, with the remaining CpGs methylated at an intermediate level. The average TMI for the *CST6*-negative cell lines was 87 ± 5% (n=5), reflecting *CST6* promoter/exon 1 hypermethylation among these cells (Table 5). When the methylation status of *CST6* promoter/exon 1 was compared between groups of cell lines that differentially express *CST6*, a significant association ($P$=0.0227) between *CST6* promoter/exon 1 region methylation (CpGs 1-55) and *CST6* gene expression was found (Table 5). However, hypermethylation of the proximal promoter of *CST6* (CpGs 1-23 and 1-31, $P$<0.001) was more strongly associated with loss of *CST6* expression status than methylation involving exon 1 (CpGs 24-55, N.S.) (Table 5). These observations suggest that hypermethylation of the *CST6* CpG island contributes to the silencing of *CST6* expression in breast cancer cell lines, and that hypermethylation of the proximal promoter is most important for down-regulation of *CST6* gene expression.

In order to identify CpG dinucleotides that are critical in *CST6* silencing, 5 *CST6*-negative cell lines (Hs578T, MCF-7, MDA-MB-435S, MDA-MB-436, and MDA-MB-453) were treated with 5-aza, and 46 CpG dinucleotides from the proximal promoter region and exon 1 (+118 to -228, CpGs 10-55) were evaluated by sodium bisulfite sequencing (Figure 15). The *CST6* promoter/exon 1 became significantly demethylated in response to 5-aza treatment in each cell line examined ($P$<0.0001), resulting in lower values for TMI (33 ± 10%, range 1-64%) (Figure 15A). Following withdrawal of 5-aza treatment, remethylation of the *CST6* promoter/exon 1 occurred in MDA-MB-436, Hs578T, and MCF-7 cells (Figure

**Figure 15. Methylation analysis of the *CST6* proximal promoter and exon 1 in *CST6*-negative breast cancer cell lines that have been exposed to demethylating treatment.** A summary of the methylation analysis of the *CST6* promoter (14 CpGs) and exon 1 (32 CpGs) is shown. Black circles correspond to fully (100%) methylated CpGs, gray circles correspond to CpGs with intermediate methylation, and open circles correspond to unmethylated CpGs. TMI values for the promoter/exon 1 region (46 CpGs) are given for each breast cancer cell line. (A) CpG methylation analysis of the *CST6* promoter/exon 1 region after treatment with 5-aza. (B) CpG methylation analysis of the *CST6* promoter/exon 1 region after withdrawal of 5-aza treatment.

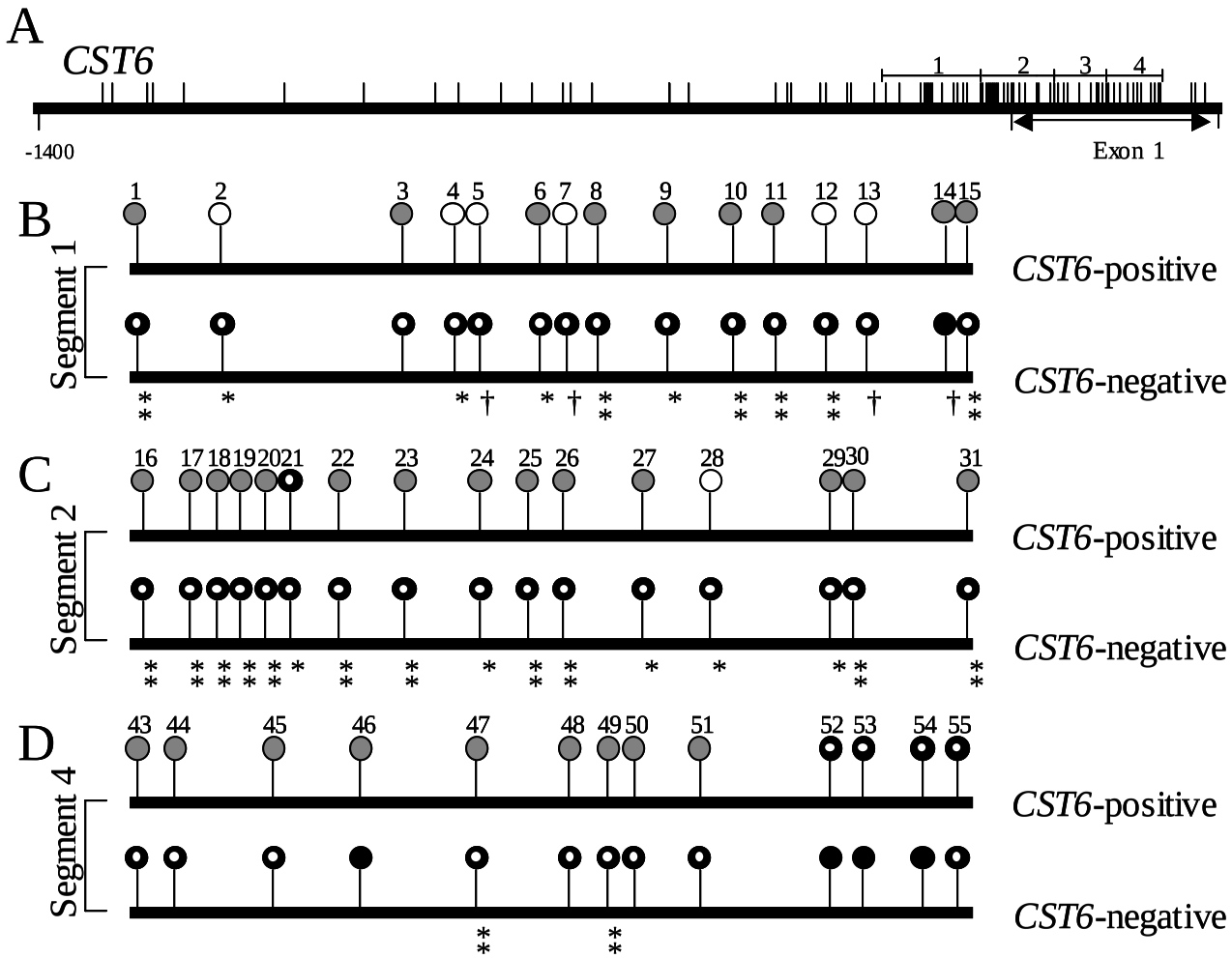

A
5-aza Treatment
TMI
10
15
20
25
30
35
40
45
50
55
26%
MDA-MB-436
1%
Hs578T
33%
MCF-7
42%
MDA-MB-435S
64%
MDA-MB-453
B
5-aza Withdrawal
TMI
10
15
20
25
30
35
40
45
50
55
75%
MDA-MB-436
73%
Hs578T
64%
MCF-7
40%
MDA-MB-435S
41%
MDA-MB-453

15B). The extent of methylation after 5-aza withdrawal was indistinguishable from that of controls for MDA-MB-436 and Hs578T cells, reflecting a complete remethylation of the promoter/exon 1 ($P<0.0001$) (Figure 14 and 15B). Likewise, significant remethylation of *CST6* occurred in MCF-7 cells after treatment withdrawal, with methylation levels approaching that of control (TMI of 64% versus 89%) (Figure 14 and 15B). In contrast, there was no apparent change in *CST6* methylation in MDA-MB-435S cells after treatment withdrawal based upon calculated TMIs (42% versus 40%), but there was a significant qualitative change in the methylation pattern (Figure 15B). MDA-MB-435S cells exhibit a loss of *CST6* expression when the demethylating treatment was withdrawn, suggesting that some or all of the 12 CpGs that were differentially remethylated (CpGs 13, 15, 19, 27, 31, 33, 39, 42, 46, 47, 50, 54, Figure 15) may be critical for *CST6* silencing. *CST6* expression after 5-aza withdrawal was diminished in MDA-MB-453 cells in the absence of a dramatic increase in *CST6* methylation, although loss of expression was accompanied by a redistribution of methylation across the *CST6* promoter and exon 1 (Figure 15). There is a significant association between *CST6* methylation status and expression among *CST6*-negative cell lines treated with 5-aza for both the promoter/exon 1 region (CpGs 10-55, control versus 5-aza, $P=0.0033$), as well as exon 1 alone (CpGs 24-55, control versus 5-aza, $P=0.0127$) (Table 5).

A comparative analysis of CpG dinucleotides in the promoter and exon 1 regions of *CST6*-positive and -negative cell lines was performed to identify methylation events involving individual CpGs or regions of CpG density that are important for the silencing of *CST6*. Average TMI values for individual CpG dinucleotides were calculated for *CST6*-positive (BT-20, MCF12A, and SK-BR-3) and *CST6*-negative cell lines (Hs578T, MCF-7,

MDA-MB-435S, MDA-MB-436, and MDA-MB-453) (Figure 16). The *CST6*-positive cell lines exhibit a low level of methylation (TMI $\leq$ 33%) for the 15 CpG dinucleotides contained in segment 1 (corresponding to -50 to -200). Of note, 6 CpGs (CpGs 2, 4, 5, 7, 12, and 13) were not methylated in these cell lines, suggesting a possible role for these CpGs and/or the sequences containing these CpGs in the positive regulation of *CST6* expression (Figure 16B). In contrast, a high level of methylation for CpGs in segment 1 (TMI $\geq$ 73%) was found in *CST6*-negative cell lines, including CpG 14, which was fully (100%) methylated in all cell lines (Figure 16B). The extent of methylation of individual CpG dinucleotides correlated with *CST6* gene expression for 14/15 CpGs in segment 1 ($P$=0.0197 to $P$<0.0001). Segment 2 spans the transcriptional start site of *CST6* (+50 to -50), and consists of CpG dinucleotides 16-31 (Figure 16C). *CST6*-positive cell lines have a relatively low level of methylation at individual CpG dinucleotides (TMI range: 0% to 44%), whereas *CST6*-negative cell lines contain high levels of methylation in this region of the *CST6* promoter/exon 1 region (TMI range: 66% to 93%) (Figure 16C). The extent of methylation of individual CpG dinucleotides was significantly associated with *CST6* gene expression for all CpGs in segment 2 ($P$=0.0343 to $P$=0.0046). Segments 3 and 4 encompass CpG dinucleotides 32-55 of *CST6* exon 1 (+242 to +50). CpG dinucleotides in this region were moderately methylated in *CST6*-positive cell lines with TMI values ranging from 11% to 67% (Figure 16D). Among *CST6*-negative cell lines, the calculated TMI values ranged from 80% to 100%, reflecting exon 1 hypermethylation in both segments. With few exceptions there was no significant correlation between *CST6* gene expression and methylation status of individual CpG dinucleotides in segments 3 and 4.

**Figure 16. Methylation analysis for individual CpG dinucleotides in *CST6*-positive and *CST6*-negative breast cancer cell lines.** (A) Distribution of CpG dinucleotides proximal to the transcriptional start site in the promoter (0 to -1400 nucleotides) and exon 1 (0 to +294 nucleotides) of *CST6* are depicted schematically. Vertical lines represent the relative position of individual CpG dinucleotides and the segmented horizontal lines (designated 1-4) indicate the location of individual CpG dinucleotides depicted in (B) (segment 1), (C) (segment 2), and (D) (segment 4). The results for CpGs in segment 3 are not shown. Representative CpG dinucleotides are indicated by lollipops corresponding to various segments of the promoter or exon 1 (+242 to -228). TMI values represent averages of the three *CST6*-positive and five *CST6*-negative breast cell lines: black circles correspond to 100% methylated CpGs, partially-filled circles correspond to >40% methylated CpGs, gray circles correspond to <40% methylated CpGs, and white circles correspond to unmethylated CpGs. †, Denotes statistical significance at $P<0.001$; **, denotes statistical significance at $P<0.01$; and *, denotes statistical significance at $P<0.03$, when comparing individual CpG methylation to *CST6* gene expression status. (B) CpGs 1-15 from segment 1 (-50 to -200) within the promoter region. (C) CpGs 16-31 from segment 2 (+50 to -50) spanning the transcriptional start site. (D) CpGs 43-55 from segment 4 (+242 to +138) within exon 1.


A
CST6
1
2
3
4
-1400
Exon 1
B
Segment 1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
CST6-positive
CST6-negative
C
Segment 2
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
CST6-positive
CST6-negative
D
Segment 4
43 44 45 46 47 48 49 50 51 52 53 54 55
CST6-positive
CST6-negative

## C. Methylation-dependent Silencing of *CST6* in Primary Human Breast Tumors and Metastatic Lesions

### *Immunohistochemical Analysis of Cystatin M in Primary Breast Tumors and Lymph Node Metastases*

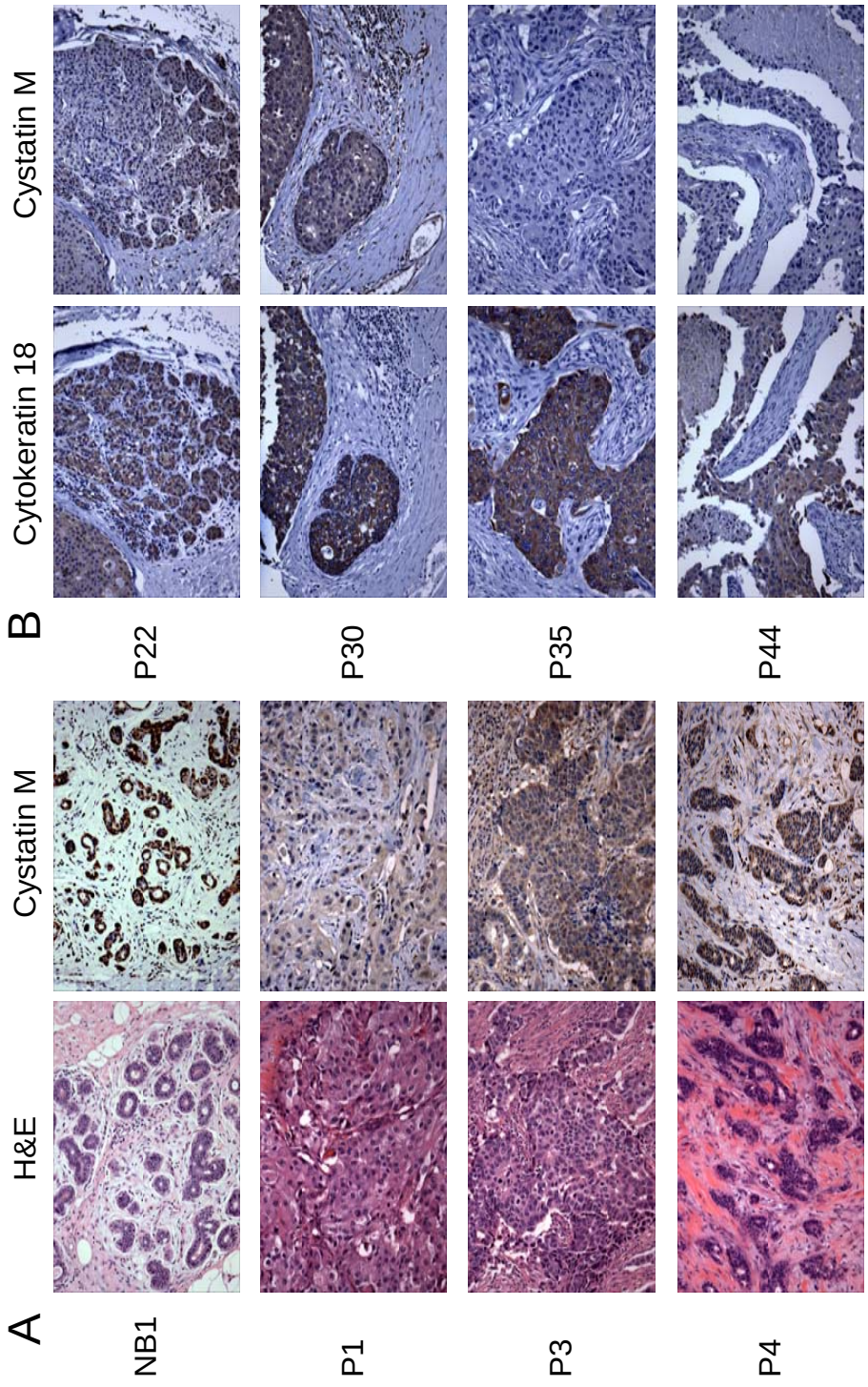
The expression of cystatin M was examined at the protein level using paraffin-embedded tissues and immunohistochemical (IHC) staining. Table 6 shows primary breast tumor designations along with tumor type, pTNM, pathological stage, and cystatin M protein expression status. A breast tissue microarray (TMA) containing 60 tissue cores was immunostained for cystatin M and cytokeratin 18 (CK18). Seventeen tissue cores from the TMA could not be scored due to an absence of CK18 immunostaining. The remaining 43/60 (72%) cores (including 31 primary tumors) showed strong staining for CK18. Therefore, these 43 tissue samples from the TMA were analyzed for cystatin M expression. Immunodetection of cystatin M in select normal human breast tissues (total n=5) and primary breast tumors (total n=45) are shown in Figure 17 and summarized in Table 6. Epithelial and myoepithelial cells of 5/5 (100%) normal breast tissues showed strong immunostaining for cystatin M (Figure 17A, NB1). Likewise, 20/45 (44%) primary tumors (Figure 17, P3, P4, P22, and P30) were positive for cystatin M expression. In contrast, 25/45 (56%) primary breast tumors were found to be negative for cystatin M (Figure 17, P1, P35 and P44), including 21/38 (55%) IDC, 1/2 (50%) ductal carcinoma *in situ* (DCIS), 1/1 (100%) solid papillary carcinoma, 1/3 (33%) infiltrating lobular carcinoma, and 1/1 (100%) signet ring cell carcinoma.

**Table 6.** Characteristics of Human Primary Breast Tumors and Normal Breast Tissues

| Tissue Designation[1] | Tumor Type[2] | pTNM | Stage | Cystatin M Expression |
|---|---|---|---|---|
| P1 | IDC | T2N1Mx | IIB | No |
| P2 | IDC | T1cN3aMx | IIIC | Yes |
| P3 | IDC | T2N1Mx | IIB | Yes |
| P4 | IDC | T2N1Mx | IIB | Yes |
| P5 | IDC | T1cN2aMx | IIIA | No |
| P6 | IDC | T1NxMx | UNK | No |
| P7 | IDC | TxNxMx[3] | IIA | Yes |
| P8 | IDC | T2N1Mx | IIB | No |
| P9 | IDC | T1bN0Mx | I | No |
| P10 | IDC | T1cN1Mx | IIA | No |
| P11 | IDC | T1N1Mx | IIA | No |
| P12 | IDC | T2N1Mx | IIB | No |
| P13 | IDC | T4N3M1 | IV | Yes |
| P14 | IDC | T1N0Mx | I | Yes |
| P15 | solid papillary carcinoma | T3N1aM0 | IIIA | No |
| P16 | IDC | T4bN2aM0 | IIIB | No |
| P17 | IDC | T3N0M0 | IIB | No |
| P18 | IDC | T3N1aM0 | IIIA | No |
| P19 | IDC | T3N3aN0 | IIIC | No |
| P20 | infiltrating lobular carcinoma | T2N1aM0 | IIB | Yes |
| P21 | IDC | T3N1aM0 | IIIA | Yes |
| P22 | IDC | T2N1aM0 | IIB | Yes |
| P23 | infiltrating lobular carcinoma | T3N3aM0 | IIIC | Yes |
| P24 | infiltrating lobular carcinoma | T3N3aM0 | IIIC | No |
| P25 | IDC | T2N1aM0 | IIB | No |
| P26 | IDC | T2N0M0 | IIA | No |
| P27 | DCIS | TisN0M0 | 0 | No |
| P28 | IDC | T2N2aM0 | IIIA | Yes |
| P29 | IDC | T2N3bM0 | IIIC | Yes |
| P30 | DCIS | TisN0M0 | 0 | Yes |
| P31 | IDC | T2N2M0 | IIIA | Yes |
| P32 | IDC | T2N0M0 | IIA | No |
| P33 | IDC | T2N0M0 | IIA | Yes |
| P34 | signet ring cell carcinoma | T3N0M0 | IIB | No |
| P35 | IDC | T3N2aM0 | IIIA | No |
| P36 | IDC | T3N2aM0 | IIIA | No |
| P37 | IDC | T2N0M0 | IIA | Yes |
| P38 | IDC | T2N0M0 | IIA | Yes |
| P39 | IDC | T2N2aM0 | IIIA | Yes |
| P40 | IDC | T2N0M0 | IIB | Yes |
| P41 | IDC | T3N1aM0 | IIIA | No |
| P42 | IDC | T2N1aM0 | IIB | Yes |
| P43 | IDC | T2N0M0 | IIA | No |
| P44 | IDC | T3N2aM0 | IIIA | No |
| P45 | IDC | T3N1aM0 | IIIA | No |
| NB1 | Normal | NA | NA | Yes |
| NB2 | Normal | NA | NA | Yes |
| NB3 | Normal | NA | NA | Yes |
| NB4 | Normal | NA | NA | Yes |
| NB5 | Normal | NA | NA | Yes |

[1]Primary breast tumors are indicated as Px and normal breast tissues are designated NBx.
[2]IDC refers to invasive ductal carcinoma and DCIS refers in ductal carcinoma *in situ*.
[3]TxNxMx refers to a pTNM that is unknown.

**Figure 17. Immunohistochemical analysis of cystatin M expression in human primary breast tumors.** (A) Panels show H&E and cystatin M immunostaining in the same tumors. Normal breast (NB1) and primary tumors P3 and P4 show positive staining for cystatin M. Tumor P1 shows reduced cystatin M staining compared to NB1. (B) Panels show cytokeratin 18 (CK18) and cystatin M immunostaining in the same tumors. All tumors show strong staining for CK18. Tumors P22 and P30 exhibit positive cystatin M immunostaining. Tumors P35 and P44 show reduced cystatin M staining. (Original objective lens magnification 10x).


A
H&E
Cystatin M
NB1
P1
P3
P4
B
Cytokeratin 18
Cystatin M
P22
P30
P35
P44

To examine cystatin M protein expression in metastatic lesions, 20 lymph nodes were obtained for immunostaining (12 of the lymph nodes were derived from 5 primary breast tumors). Table 7 contains information related to the lymph nodes analyzed, including tumor designations, along with tumor type, pTNM, pathological stage from matched primary tumors, and cystatin M protein expression status. Figure 18 shows representative cystatin M IHC staining in these metastatic lesions. The majority (17/20, 85%) of metastatic lesions in lymph nodes were negative for cystatin M expression (Figure 18, P2N1, P4N1, LNM2, LNM5, LNM8, and LNM11). In contrast, 3/20 (15%) metastatic lesions were positive for cystatin M expression (Figure 18, P1N1). Overall, these numbers reflect the presence of cystatin M-negative lymph node metastases in 13/16 (81%) patients. These observations are consistent with the suggestion that loss of cystatin M expression is a common feature of metastatic breast tumors. However, the presence of cystatin M-positive breast tumor in some lymph node specimens suggests that loss of cystatin M expression is not absolutely required for tumor invasion and metastasis.

To examine the possibility that loss of cystatin M reflects a tumor progression-related event, 5 primary breast tumors and matched lymph node metastases were immunostained. Figure 19 shows representative examples of these primary tumor/lymph node pairs. One of five (20%) matched pairs were negative for cystatin M expression in both the primary breast tumor and lymph node metastasis, indicating an early loss of cystatin M protein during breast tumorigenesis, with persistence in the metastatic lesion (Figure 19, P5 and P5N1). Additionally, 1/5 (20%) matched pairs showed positive cystatin M staining in both the primary breast and lymph node tissues, suggesting that tumor metastasis in this patient was mediated through a cystatin M-independent pathway (Figure 19, P3 and P3N1). The

**Table 7.** Characteristics of Human Lymph Node Metastases

| Tissue Designation[1] | Tumor Type | pTNM[2] | Stage[2] | Cystatin M Expression |
|---|---|---|---|---|
| P1N1 | lymph node metastasis | T2N1Mx | IIB | Yes |
| P2N1-N3[3] | lymph node metastasis | T1cN3aMx | IIIC | No |
| P3N1 | lymph node metastasis | T2N1Mx | IIB | Yes |
| P4N1 | lymph node metastasis | T2N1Mx | IIB | No |
| P5N1 | lymph node metastasis | T1cN2aMx | IIA | No |
| LNM1-4[4] | lymph node metastasis | TxNxMx | UNK | No |
| LNM5 | lymph node metastasis | TxNxMx | UNK | No |
| LNM6 | lymph node metastasis | T3N3aM0 | IIIC | Yes |
| LNM7 | lymph node metastasis | T3N3aM0 | IIIC | No |
| LNM8 | lymph node metastasis | T3N3aM0 | IIIC | No |
| LNM9 | lymph node metastasis | T3N3M0 | IIIC | No |
| LNM10 | lymph node metastasis | T3N2aM0 | IIIA | No |
| LNM11 | lymph node metastasis | T1cN1aM0 | IIA | No |
| LMN12 | lymph node metastasis | T2N3aM0 | IIIC | No |
| LMN13 | lymph node metastasis | T1bN3aM0 | IIIC | No |

[1]Metastaic lesions corresponding to a matched primary breast tumor are indicated as PxNx and lesions that have no matched primary are designated LNMx.
[2]pTNM and pathological stage designations for lymph node metastases are derived from primary breast tumor designations. TxMxNx refers to a pTNM that is unknown.
[3]P2N1-N3 designates 3 independent lymph nodes corresponding to one primary tumor.
[4]LNM1-4 designates 4 independent lymph nodes corresponding to one primary tumor.

**Figure 18. Immunohistochemical analysis of cystatin M expression in lymph node metastases.** (A) Panels show H&E and cystatin M immunostaining in the same lymph nodes. Lymph node P1N1 shows positive staining for cystatin M. Lymph nodes P2N1, P4N1, LNM2, and LNM5 show reduced cystatin M immunostaining. (B) Panels show cytokeratin 18 (CK18) and cystatin M immunostaining in the same lymph nodes. All metastatic lesions show strong staining for CK18 and exhibit reduced cystatin M immunostaining. (Original objective lens magnification 10x).

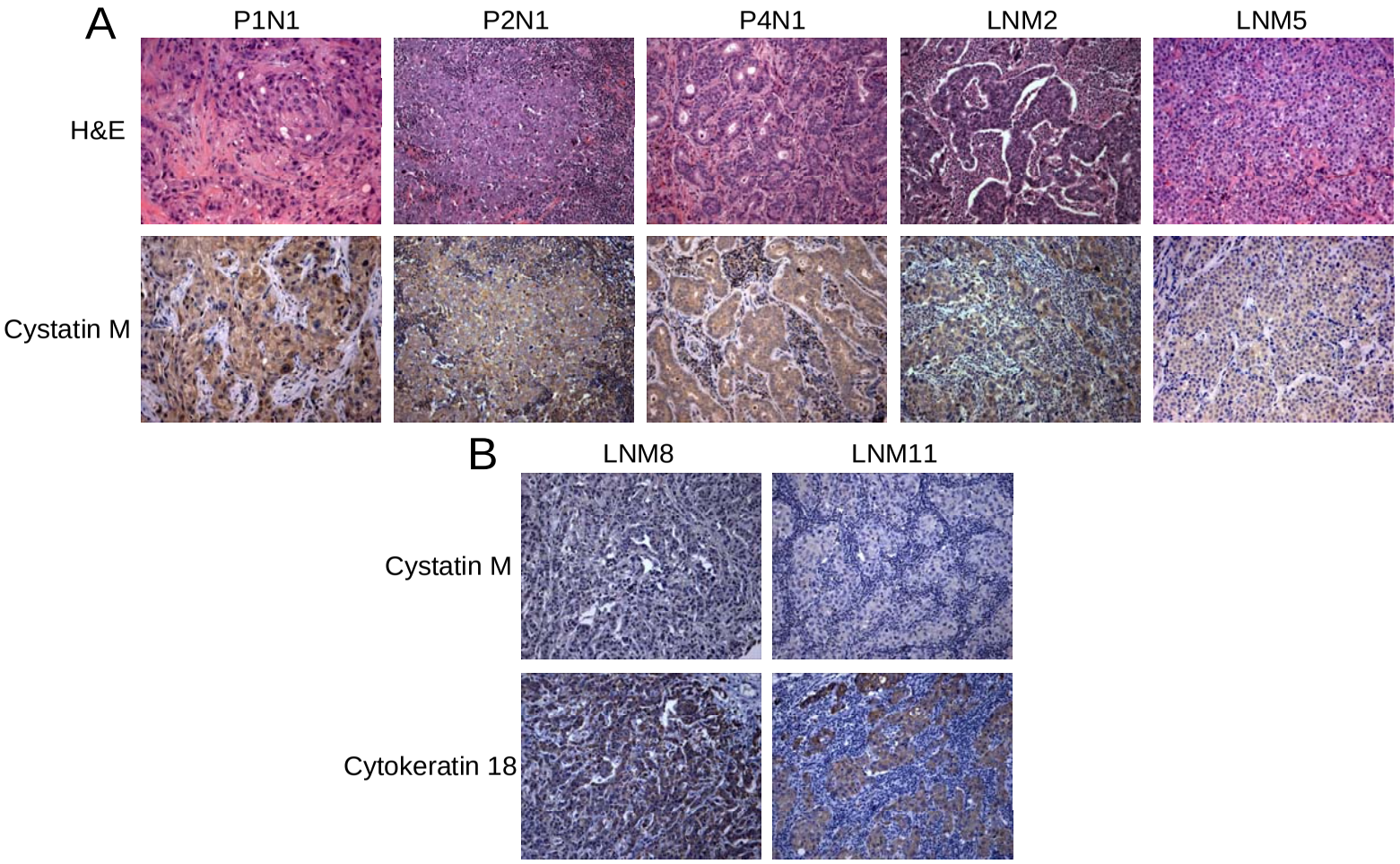
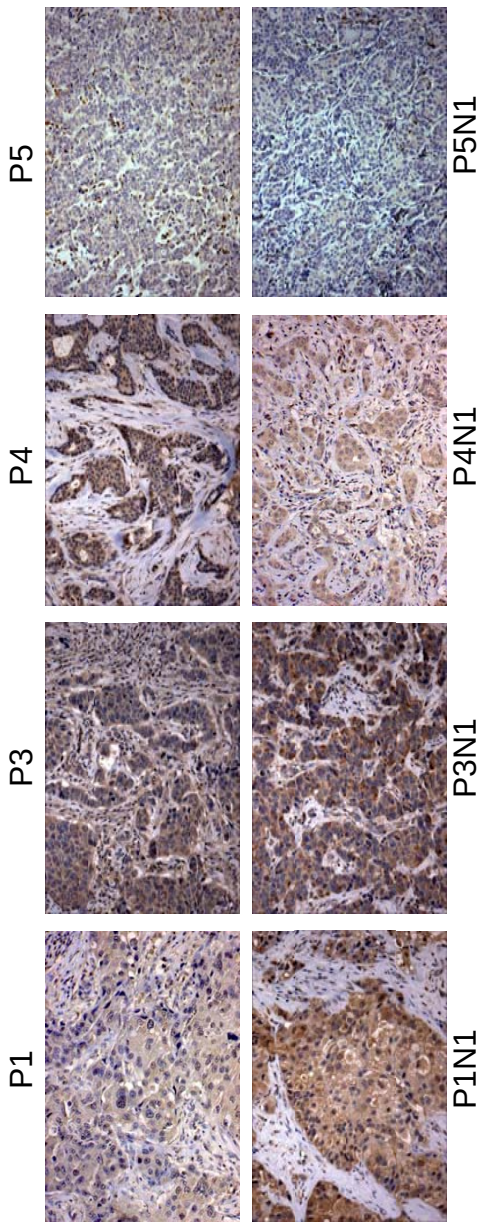

A
P1N1
P2N1
P4N1
LNM2
LNM5
H&E
Cystatin M
B
LNM8
LNM11
Cystatin M
Cytokeratin 18

**Figure 19. Immunohistochemical analysis of cystatin M expression in matched primary breast tumors and lymph node metastases.** Representative examples of matched pairs of primary breast tumors (top panel) and lymph node metastasis (bottom panel) are shown. Primary breast tumor P1 shows reduced cystatin M immunostaining compared to its matched lymph node P1N1. Primary breast tumor P3 and lymph node metastasis P3N1 both show positive staining for cystatin M. Primary breast tumor P4 shows positive staining for cystatin M compared to its matched lymph node metastasis P4N1. Primary breast tumor P5 and lymph node metastasis P5N1 are negative for cystatin M expression. (Original objective lens magnification 10x).

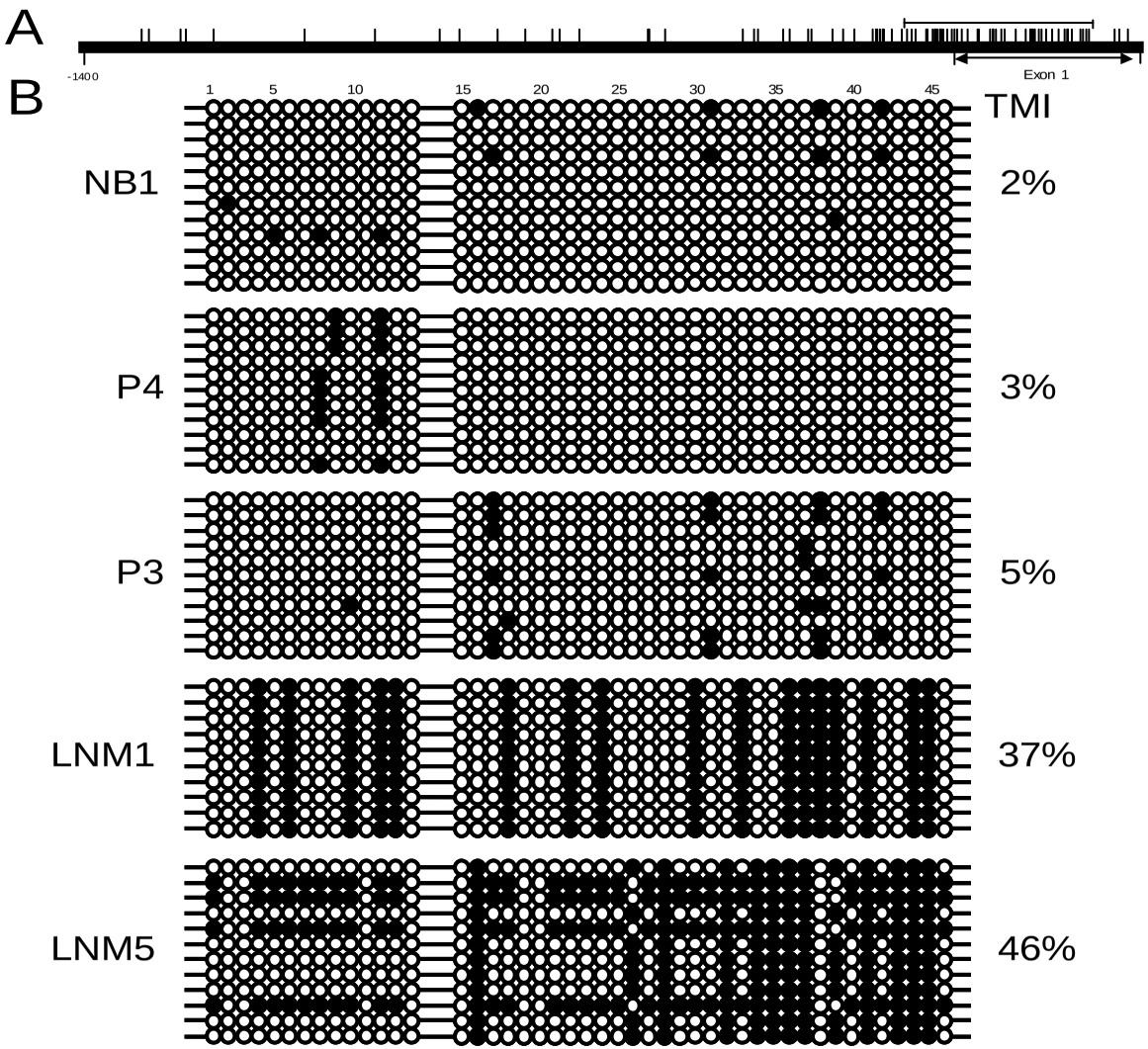

P1
P3
P4
P5
P1N1
P3N1
P4N1
P5N1
Primary Breast
Lymph Node

remaining 3 patients had cystatin M-positive primary tumors, but 2/3 (67%) of the matched lymph node metastases lacked cystatin M expression (Figure 19, P4 and P4N1). This observation is consistent with a progression related loss of cystatin M expression during the evolution of the metastatic clone. The remaining primary tumor/lymph node pair lacks cystatin M expression in the primary tumor, while the lymph node metastasis exhibits stronger staining for the cystatin M protein (Figure 19, P1 and P1N1). This result is unclear, but may reflect heterogeneity of cystatin M expression in this tumor. Thus, a cystatin M-positive tumor cell population may have given rise to this metastatic lesion through a cystatin M-independent pathway.

***Methylation-Dependent Silencing of CST6 in Primary Breast Tumors and Lymph Node Metastases***

To explore the possibility that loss of cystatin M expression is related to epigenetic silencing of *CST6*, we analyzed primary breast tumors and lymph node metastases for *CST6* promoter methylation. A segment of the proximal promoter and exon 1 (-118 to +242, Figure 20A) containing 46 CpG dinucleotides was analyzed in normal breast tissue, 11 primary breast tumors (5 express cystatin M and 6 lack cystatin M), and 12 lymph node metastases (2 express cystatin M and 10 lack cystatin M). Multiple clones (n=5-12) corresponding to the *CST6* promoter and exon 1 from each primary tumor or lymph node metastasis were analyzed by sodium bisulfite sequencing and individual CpGs were scored for methylation status. Representative examples are shown in Figure 20B. In normal breast tissue, 33/46 (72%) CpGs were consistently unmethylated, 13/46 (28%) CpGs were methylated at an intermediate level, and 0/46 (0%) were 100% methylated, producing a TMI

**Figure 20. Methylation analysis of the *CST6* proximal promoter and exon 1 in representative primary breast tumors and lymph node metastases.** (A) The distribution of CpG dinucleotides proximal to the transcription start site in the promoter (0 to -1400 nucleotides) and exon 1 (0 to +294 nucleotides) of *CST6* are depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). Methylation analysis was performed on a region of the promoter spanning from -118 to +242 (indicated by a solid horizontal line), which contains 46 CpG dinucleotides and encompasses a large CpG island. (B) All clones analyzed for methylation of the *CST6* promoter and exon 1 (46 CpGs) are shown for representative primary breast tumor and lymph node metastases examples. Black circles correspond to methylated CpGs and open circles correspond to unmethylated CpGs. TMI values for the entire promoter/exon 1 region (46 CpGs) are given for each primary breast tumor and lymph node metastasis. NB1, P4, and P3 express cystatin M, while LNM1 and LNM5 lack cystatin M protein expression.


A
-1400
Exon 1
B
1
5
10
15
20
25
30
35
40
45
TMI
NB1
2%
P4
3%
P3
5%
LNM1
37%
LNM5
46%

of 2% (Figure 20B, NB1). The majority (3/5, 60%) of primary tumors that stain positive for cystatin M lack appreciable levels of methylation, with TMI values of 0 to 5% (Figure 20B, P4 and P3). Primary breast tumors P2 and P9, exhibit methylation and express cystatin M. In P2, 23/46 (50%) CpGs were unmethylated, 22/46 (48%) CpGs were methylated at an intermediate level, and 1/46 (2%) was 100% methylated, producing a TMI of 38%. In P9, the majority (33/46, 72%) of CpGs were unmethylated, although 13/46 (28%) were 100% methylated resulting in a TMI of 28%. The continued expression of cystatin M in P2 and P9, despite *CST6* promoter hypermethylation, suggests that there may be mechanisms to transcriptionally bypass promoter methylation. Most, (5/6, 83%) primary breast tumors that are negative for cystatin M expression exhibit very low levels of *CST6* methylation (TMI = 0 to 3%). This finding suggests that there may be other epigenetic or mutational mechanisms responsible for the silencing of cystatin M in these primary breast tumors. In contrast, one tumor (P1) was negative for cystatin M protein expression and 45/46 (98%) CpGs were methylated at an intermediate level, resulting in a TMI of 28%. Overall, a subset of primary tumors (3/11, 27%) exhibits *CST6* promoter hypermethylation, and in one case this methylation was associated with loss of cystatin M expression.

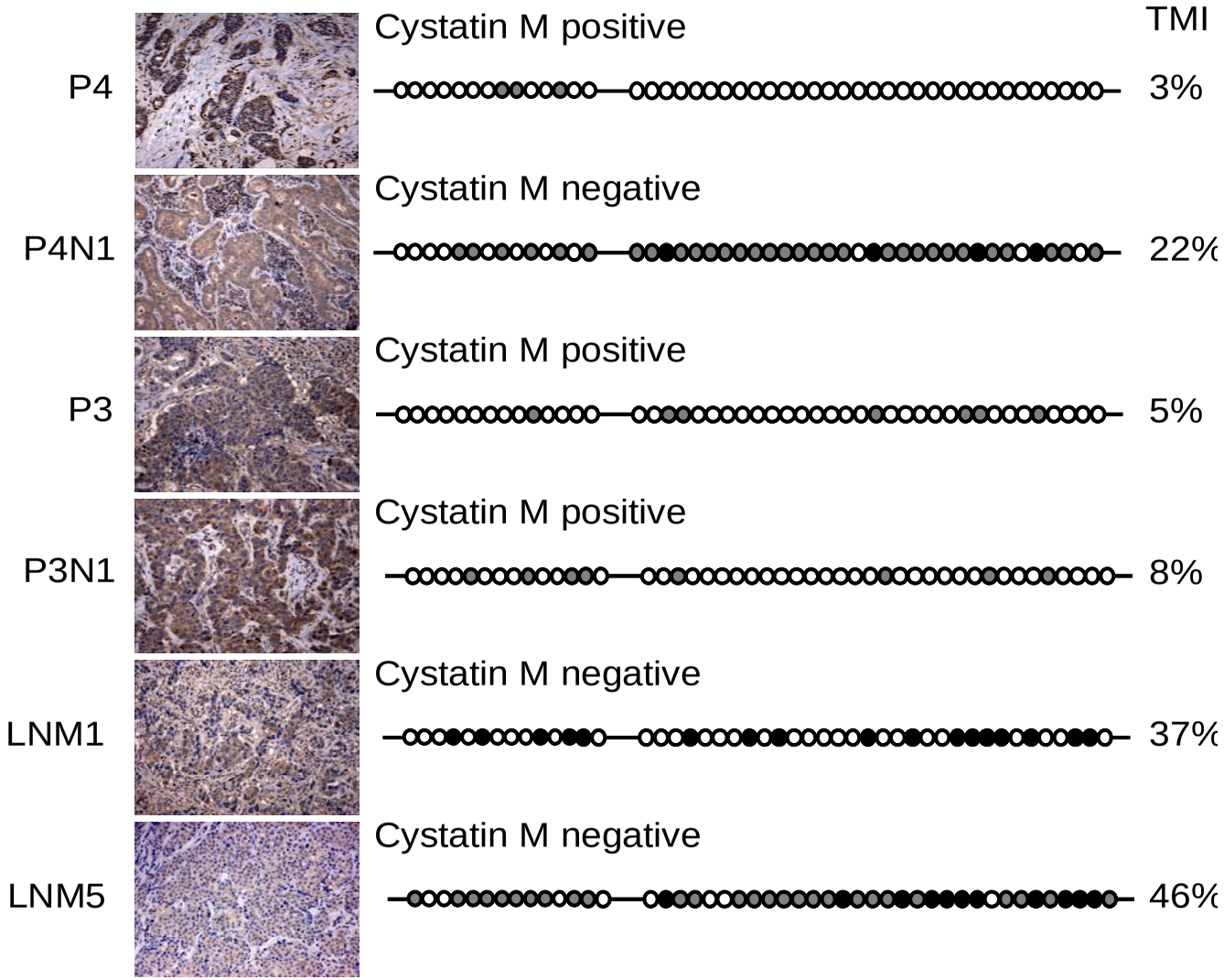
The majority (7/10, 70%) of metastatic lesions that are negative for cystatin M expression exhibit hypermethylation of the *CST6* promoter/exon 1 region, with TMI values ranging from 11% to 46% (average TMI 17 $\pm$ 4%). The TMIs for cystatin M-negative lymph node metastases were found to be statistically increased relative to that determined for normal breast ($P$<0.0001). Figure 20B shows representative methylation analyses. LNM1 and LNM5 were methylated at 17/46 (37%) and 38/46 (83%) CpGs, respectively, resulting in TMIs of 37% and 46% (Figure 20B, LNM1 and LNM5). P4N1 and LNM2 were 100% or

intermediately methylated at 35/46 (76%) and 38/46 (83%) CpGs respectively, with the remaining CpGs unmethylated, reflecting TMIs of 22% and 24%. Cystatin M negative nodes P2N1, LNM3, and LNM4 were intermediately methylated at 12/46 (26%), 34/46 (74%), and 18/46 (39%) CpGs, resulting in TMI values of 11%, 13%, and 13% respectively. There were 3 cystatin M-negative lymph nodes that displayed TMI values ranging from 1% to 3%. Two lymph node metastases were positive for cystatin M and exhibit low levels of methylation. In lymph node P1N1, 23/46 (50%) CpGs were methylated at an intermediate level and 23/46 (50%) CpGs were unmethylated. In lymph node P3N1, 8/46 (17%) CpGs were methylated at an intermediate level, and 38/46 (83%) CpGs were unmethylated. In total, 8/12 (67%) metastatic lesions from 5/7 (71%) patients displayed *CST6* promoter hypermethylation.

***CST6 Gene Methylation Correlates with Loss of Cystatin M Expression in a Subset of Primary Breast Tumors and Lymph Node Metastases***

Cystatin M expression is associated with methylation status in the majority (12/23, 52%) of tumor tissues (Figure 21). In most cases (4/7, 57%) cystatin M-positive tumors show a lack of *CST6* promoter methylation (Figure 21, P3, P3N1, and P4). In contrast, 3/7 (43%) cystatin M-positive tumors exhibit *CST6* hypermethylation. This result suggests the existence of other epigenetic or genetic mechanisms that can bypass promoter hypermethylation. Eight of 16 (50%) cystatin M-negative tumors exhibit *CST6* promoter hypermethylation with TMI values ranging from 11% to 46% (Figure 21, P4N1, LNM1, and LNM5). These include one primary breast tumor and 7 lymph node metastases that lack expression of cystatin M. The remaining 8/16 (50%) cystatin M-negative tumors show very low levels of *CST6* promoter methylation (TMI values ranging from 0% to 3%), including 5

**Figure 21. Correlation analysis of cystatin M expression and *CST6* methylation status in primary breast tumors and lymph node metastases.** Panels show cystatin M immunostaining (on left) and a summary of the methylation analysis of the *CST6* promoter/exon 1 (46 CpGs) is show on the right. Black circles correspond to fully (100%) methylated CpGs, gray circles correspond to CpGs with intermediate methylation, and open circles correspond to unmethylated CpGs. TMI values for the entire promoter/exon 1 region (46 CpGs) are given for each tissue sample. P4 and P3 primary breast tumors, and lymph node metastasis P3N1 express cystatin M. Metastatic lesions P4N1, LNM1, and LNM5 show reduced expression of cystatin M. (Original objective lens magnification 10x).


TMI
P4
Cystatin M positive
3%
P4N1
Cystatin M negative
22%
P3
Cystatin M positive
5%
P3N1
Cystatin M positive
8%
LNM1
Cystatin M negative
37%
LNM5
Cystatin M negative
46%

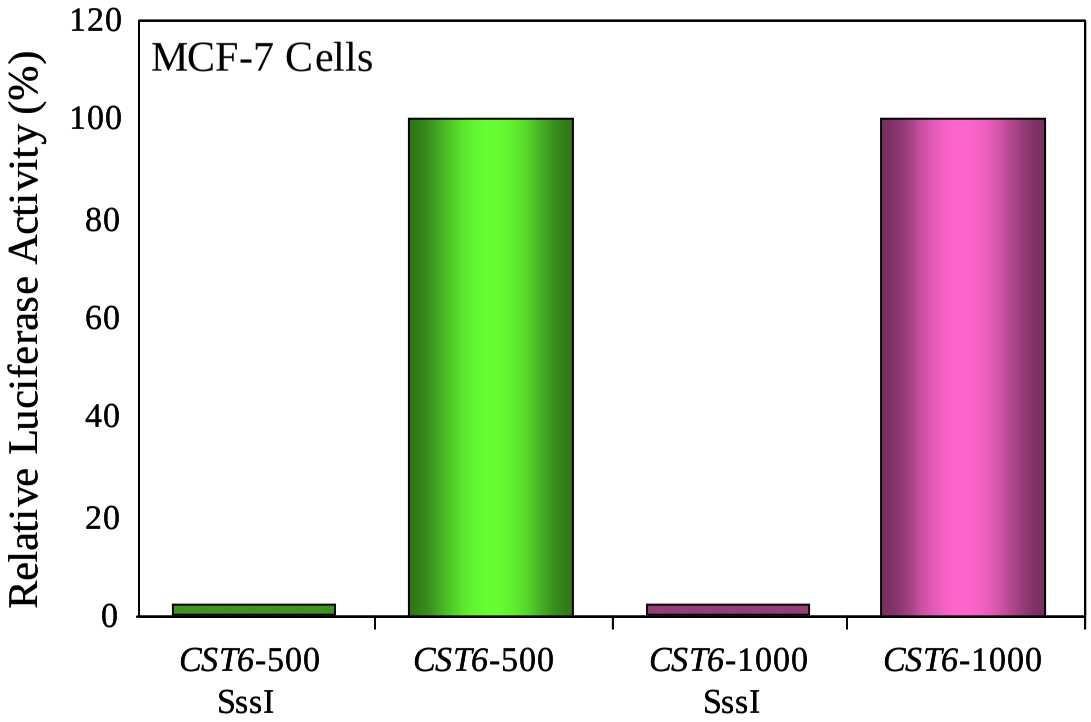
primary tumors and 3 metastatic lesions. Overall, the extent of *CST6* methylation corresponds with the expression of cystatin M expression in the majority (52%) of breast neoplasms, suggesting that methylation-dependent silencing of *CST6* may represent an important mechanism for loss of cystatin M in a subset of breast neoplasms.

## F. *CST6* CpG Methylation Requires an Upstream DNA Sequence Element that Directs Promoter CpG Island Methylation

### *Cloning the Promoter Region of CST6 and Analysis of Luciferase Activity*

To address the question of whether *cis*-acting elements direct DNA methylation of specific promoter target sequences, two truncations (designated *CST6*-500 and *CST6*-1000) of the *CST6* promoter region (encompassing the CpG island) were cloned into the pGL4.17[luc2/Neo] reporter vector. Luciferase reporter constructs were transiently transfected into MCF-7 cells to assess promoter activity and to ensure that construction of the reporter gene cassette did not compromise that ability of the truncated *CST6* promoter to drive luciferase activity. Both *CST6* reporter gene constructs produced good luciferase activity after transient transfection: *CST6*-500, 604 units, and *CST6*-1000, 2572 units. *CST6* reporter gene constructs were treated with SssI methylase in order to test if the *CST6*-500 and *CST6*-1000 promoter constructs are sensitive to methylation (Figure 22). Both luciferase reporter constructs were transiently transfected into MCF-7 cells to assess promoter activity after SssI methylase treatment. Methylated *CST6* reporter gene constructs produced minimal luciferase activity after transient transfection relative to the luciferase activity of the unmethylated constructs, suggesting that both promoter constructs are sensitive to silencing

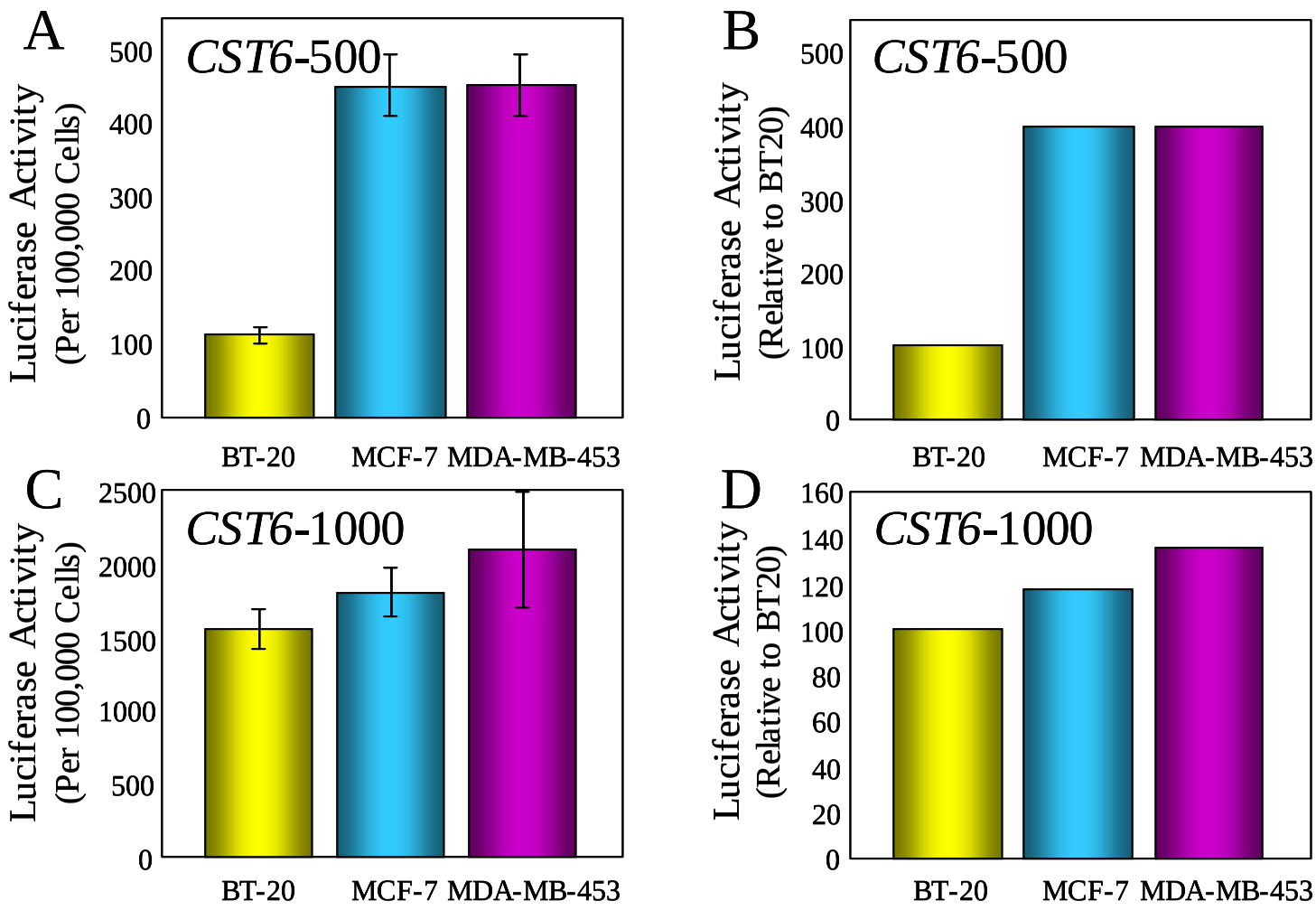
**Figure 22. Analysis of luciferase activity of *CST6* promoter reporter constructs treated with SssI methylase.** The bars depict levels of luciferase activity from the SssI methylase treated *CST6*-500 (green diagonal bar) and *CST6*-1000 (pink diagonal bar) promoter reporter constructs, and untreated *CST6*-500 (green bar) and *CST6*-1000 (pink bars) constructs following transient transfection into MCF-7 breast cancer cells. Levels of luciferase activity from the methylated reporter constructs relative to untreated reporter constructs are shown, where untreated reporter constructs are set as 100%.


MCF-7 Cells
Relative Luciferase Activity (%)
120
100
80
60
40
20
0
CST6-500
SssI
CST6-500
CST6-1000
SssI
CST6-1000

by methylation (Figure 22). Subsequently, these reporter constructs were stably transfected into BT-20 breast cancer cells (which express *CST6* and lack promoter methylation), and MCF-7 and MDA-MB-453 cells (that lack *CST6* expression and exhibit promoter hypermethylation). After stable transfection and cell propagation, firefly luciferase activity expressed from the *CST6*-500 promoter construct was assayed in cell lysates prepared 2, 4, and 5 weeks after transfection into BT-20 and MCF-7 cells, and 5 weeks after transfection into MDA-MB-453 cells. Likewise, firefly luciferase activity expressed from the *CST6*-1000 promoter construct was determined 3, 4, and 5 weeks after transfection into BT-20 and MCF-7 cells, and 5, 6, and 7 weeks after transfection into MDA-MB-453 cells for firefly luciferase activity. Temporal analysis of luciferase activities enables examination of progressive methylation-dependent changes in reporter construct expression. Both *CST6*-500 and *CST6*-1000 reporter constructs were analyzed for DNA methylation by bisulfite sequencing after the last luciferase determination (5-7 weeks following initial transfection).

In BT-20 cells, the *CST6*-500 construct demonstrated a low level of luciferase activity (Figure 23A). In contrast, MCF-7 and MDA-MB-453 cells, expressed significantly higher levels of luciferase activity from the *CST6*-500 construct compared to BT-20 cells ($P<0.01$) (Figure 23A). When normalized to BT-20 cells, MCF-7 and MDA-MB-453 cells express 4-fold more luciferase activity from the *CST6*-500 construct than BT-20 cells (Figure 23B). This difference may reflect greater general transcriptional activity in MCF-7 and MDA-MB-453 cells, compared to BT-20. These results show that the *CST6*-500 construct, which contains the majority of the *CST6* promoter CpG island, can drive transcription of the luciferase gene. This observation suggests that the *CST6*-500 construct represents a minimal/essential promoter for *CST6*. The level of luciferase activity resulting from *CST6*-

**Figure 23. Analysis of luciferase activity of *CST6* promoter reporter constructs.** The bars depict levels of luciferase activity from the *CST6*-500 and *CST6*-1000 promoter reporter constructs following stable transfection into BT-20 (yellow bars), MCF-7 (blue bars), and MDA-MB-453 (pink bars) breast cancer cell lines. Luciferase activities represent the mean of 3-9 independent determinants. Error bars reflect S.E.M. (A) Levels of luciferase activity per 100,000 cells (arbitrary units) from the *CST6*-500 reporter construct. (B) Levels of luciferase activity from the *CST6*-500 reporter construct normalized to BT-20, where BT-20 is set as 100. (C) Levels of luciferase activity per 100,000 cells (arbitrary units) from the *CST6*-1000 reporter construct. (D) Levels of luciferase activity from the *CST6*-1000 reporter construct normalized to BT-20, where BT-20 is set as 100.


A
CST6-500
Luciferase Activity
(Per 100,000 Cells)
500
400
300
200
100
0
BT-20
MCF-7
MDA-MB-453
B
CST6-500
Luciferase Activity
(Relative to BT20)
500
400
300
200
100
0
BT-20
MCF-7
MDA-MB-453
C
CST6-1000
Luciferase Activity
(Per 100,000 Cells)
2500
2000
1500
1000
500
0
BT-20
MCF-7
MDA-MB-453
D
CST6-1000
Luciferase Activity
(Relative to BT20)
160
140
120
100
80
60
40
20
0
BT-20
MCF-7
MDA-MB-453

500 remained consistent over time in each of the three cell lines. The continued expression of *CST*-500 in the hypermethylator cell lines (MCF-7 and MDA-MB-453) over a period of five weeks suggests that the minimal/essential promoter does not become methylated in these cells.

The *CST6*-1000 construct produced increased levels of luciferase activity in each of the three breast cancer cell lines (Figure 23C) compared to the *CST*-500 construct. This overall increase in promoter activity from the *CST6*-1000 construct (measured as increased luciferase activity) may reflect the presence of positive regulatory elements in the sequence region upstream of the CpG island. MDA-MB-453 cells produced the highest levels of luciferase activity from the *CST6*-1000 construct, but there was no significant difference in luciferase activities among the cell lines examined (Figure 23C and D). The *CST6*-1000 construct exhibited a consistent level of luciferase activity over time in each of the three cell lines. The continued expression of *CST6*-1000 in hypermethylator cell lines over a period of five to seven weeks, suggests that the *CST6* promoter region contained in this reporter construct does not become methylated in these cell lines.

### *CST6 Promoter Construct Methylation Analysis in Differentially Expressing Breast Cancer Cell Lines*

Cystatin M contains a large CpG island (424 bp) including 54 CpG dinucleotides that spans the proximal promoter and exon 1, encompassing the start site for transcription. We examined the methylation status of 34 CpGs that are contained within the *CST6*-500 and *CST6*-1000 constructs by bisulfite sequencing using primers designed to amplify a region of the firefly luciferase gene and the proximal region of the constructs. BT-20 cells transfected

with the *CST6*-500 construct demonstrated luciferase activity and lacked methylation five weeks after transfection, exhibiting a TMI value of 4% (Figure 24). Likewise, methylation analysis of the *CST6*-500 construct from MCF-7 and MDA-MB-453 cells five weeks after transfection exhibited a lack of methylation, producing TMI values of 5% and 0% respectively (Figure 24). These results suggest that truncation of the *CST6* promoter disassociates an upstream *cis* regulatory element that directs DNA methylation of the CpG island region of the *CST6*-500 construct. This finding also establishes that the *CST6* promoter CpG island is not sufficient to recruit the DNA methylation machinery and direct its own methylation.

BT-20 cells transfected with the *CST6*-1000 construct produced strong expression of luciferase activity and lacked methylation five weeks after transfection, exhibiting a TMI value of 7% (Figure 25). Likewise, methylation analysis of the *CST6*-1000 construct transfected into MCF-7 and MDA-MB-453 cells after five weeks exhibited a lack of methylation, demonstrating TMI values of 1% and 0% respectively (Figure 25). Thus, there is perfect correspondence between expression of luciferase activity and lack of promoter CpG island methylation using this construct and these cell lines. These observations suggest that the *CST6*-1000 promoter truncation dissociated a critical regulatory (directive) sequence (*cis* element) from the CpG island target sequence within the *CST6* promoter. This result suggests that the putative instructional *cis* element must be located >1200 bp upstream of the transcription start site, and >1000 bp upstream of the target CpG island.

**Figure 24. Methylation analysis of the *CST6*-500 promoter reporter construct.** The distribution of the CpG dinucleotides in the *CST6*-500 promoter reporter construct inserted proximal to the firefly luciferase gene (pink box) is depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). Methylation analysis was performed on a region containing 34 CpG dinucleotides from the *CST6*-500 promoter reporter construct following transfection into BT-20, MCF-7, and MDA-MB-453 breast cancer cell lines. All clones analyzed for methylation of the *CST6*-500 promoter reporter construct are shown. Black circles correspond to methylated CpGs and open circles correspond to unmethylated CpGs. TMI values for the entire region (34 CpGs) are given for each breast cancer cell lines.

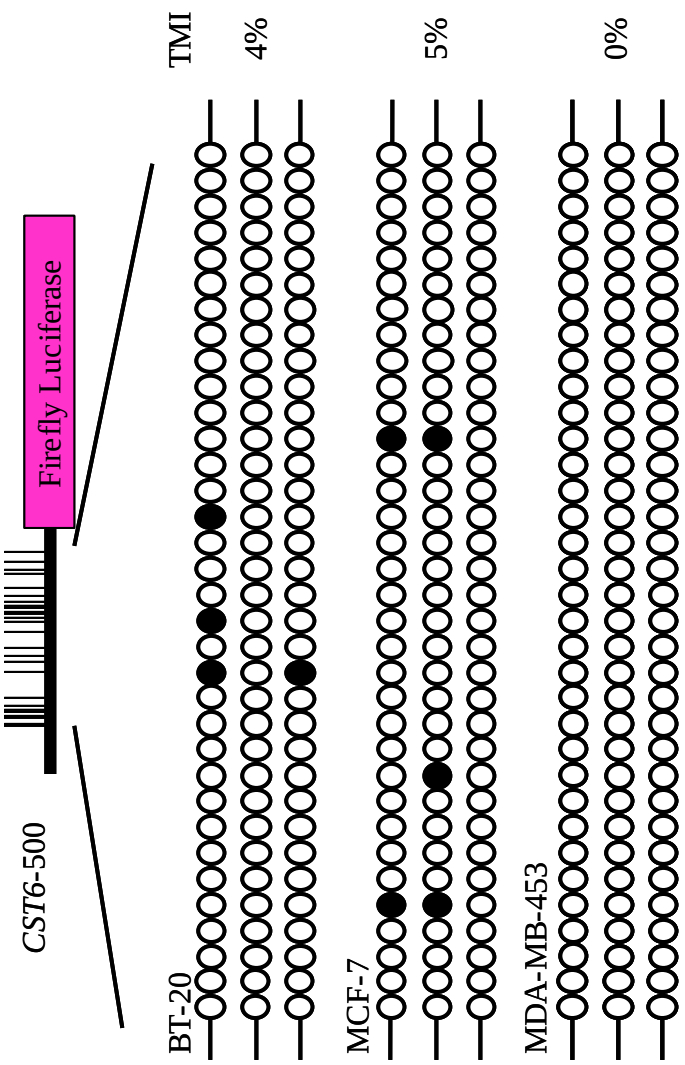
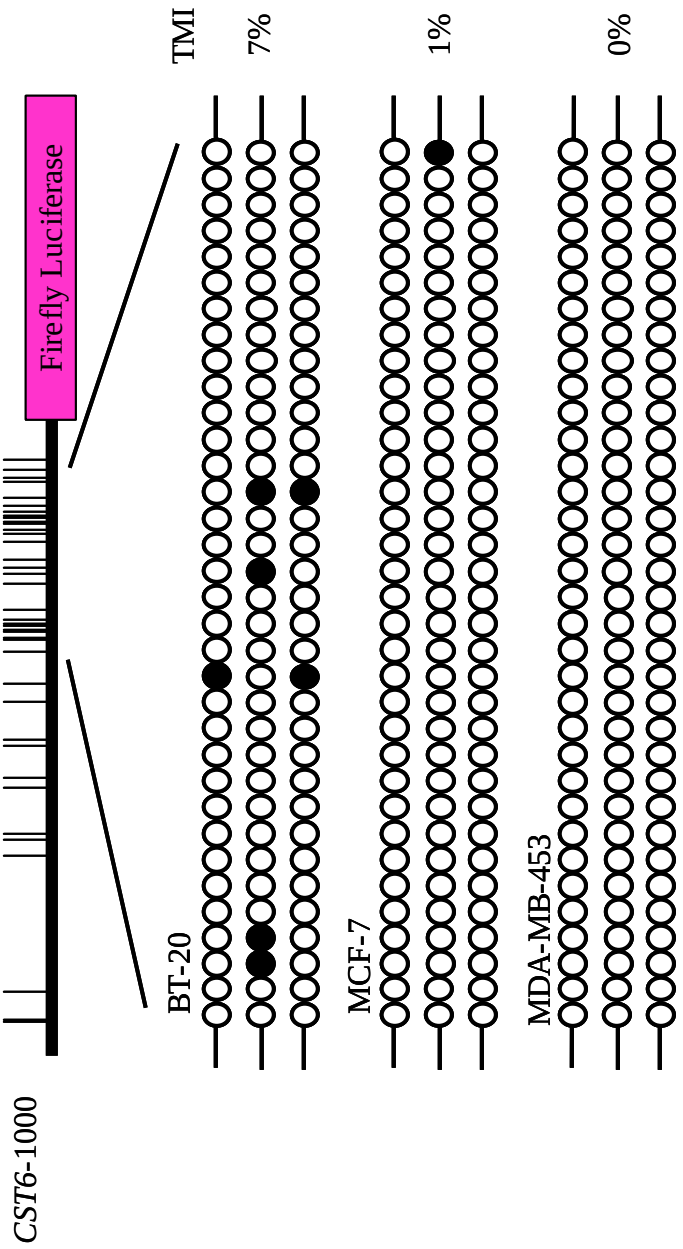

CST6-500
Firefly Luciferase
TMI
4%
5%
0%
BT-20
MCF-7
MDA-MB-453

**Figure 25. Methylation analysis of the *CST6*-1000 promoter reporter construct.** The distribution of the CpG dinucleotides in the *CST6*-1000 promoter reporter construct inserted proximal to the firefly luciferase gene (pink box) is depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). Methylation analysis was performed on a region containing 34 CpG dinucleotides from the *CST6*-1000 promoter reporter construct following transfection into BT-20, MCF-7, and MDA-MB-453 breast cancer cell lines. All clones analyzed for methylation of the *CST6*-1000 promoter reporter construct are shown. Black circles correspond to methylated CpGs and open circles correspond to unmethylated CpGs. TMI values for the entire region (34 CpGs) are given for each breast cancer cell lines.


CST6-1000
Firefly Luciferase
TMI
7%
1%
0%
BT-20
MCF-7
MDA-MB-453

# IV. DISCUSSION

## A. Identification of Putative Epigenetically-regulated Genes in MCF-7 Breast Cancer Cells

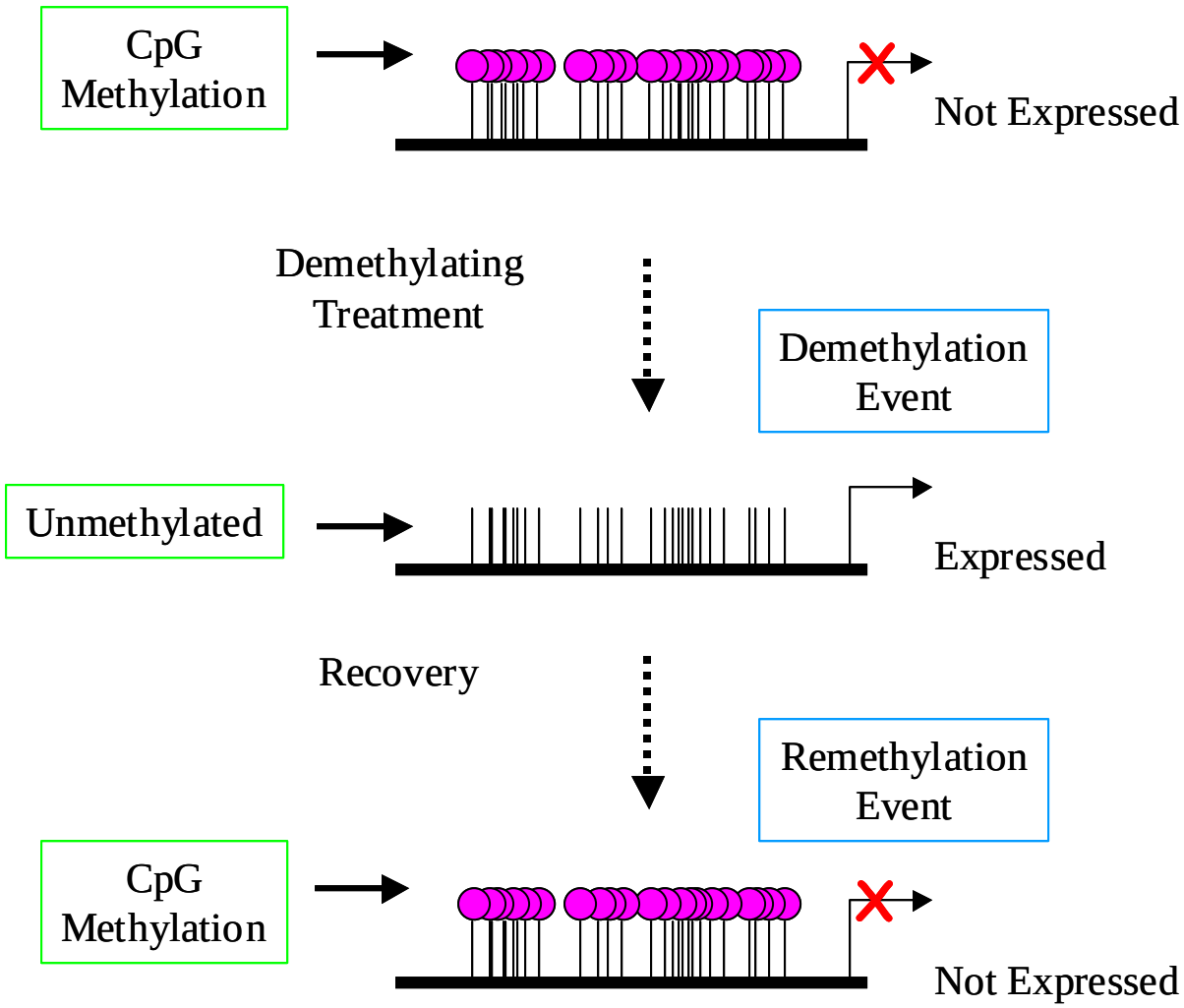
Recent evidence suggests that epigenetic mechanisms play a major role in breast carcinogenesis, contributing to genetic instability in breast cancer, as well as to the silencing of specific genes (24,125). A number of different genes have been shown to be inactivated in breast cancer through methylation-dependent gene silencing, including cell cycle control genes (*p16$^{INK4a}$*), steroid receptor genes (*ERα, PR, RARα2*), tumor suppressor genes (*BRCA1*), genes associated with cancer metastasis (*E-cadherin, TIMP-3*), and others (24,40-43). In an attempt to more comprehensively catalogue methylation-sensitive genes in breast cancer, several recent studies have employed treatment of cells in culture with demethylating drugs and microarray analysis of gene expression (72-74). We have utilized a similar strategy to identify epigenetically-regulated genes in human breast cancer cells. In our study, MCF-7 cells were treated with a low concentration of 5-aza (250 nM) or 5-aza + TSA (50 nM) for 3 weeks, followed by a 5 week recovery period after treatment withdrawal (Figure 7). The concentration of 5-aza and TSA, utilized in this study was 4-fold to 6-fold lower than traditional methods (126-129), eliminating the typically encountered cytotoxic effects (72,130) and allowing prolonged exposure of MCF-7 cells to the demethylating drugs. Treatment of MCF-7 cells resulted in increased and decreased expression of numerous genes

(Figure 7), many of which may not be directly regulated by DNA methylation. Therefore, to enrich for genes that are putatively epigenetically regulated, we identified subsets of genes that demonstrated a significant increase in expression level in response to treatment (5-aza or 5-aza + TSA), but then returned to steady-state levels of expression following a recovery period (Figure 26). Increased gene expression in response to treatment presumably reflects a demethylation event, resulting in an induction or enhancement of gene expression (Figure 26). Likewise, the return of gene expression to control levels following treatment withdrawal presumably reflects remethylation of the promoter sequence, resulting in gene silencing or down-regulation of expression (Figure 26). This analysis identified a group of 20 putative epigenetically-regulated genes for further study, some of which have been suggested to be epigenetically regulated by other investigators (72,123,124). These observations combine to suggest that our strategy for selection of putative epigenetically-regulated genes was sound. In addition to genes that exhibited increased expression, we identified genes that showed decreased expression upon demethylating treatment. These genes may be important as well, but we chose to focus on the genes that demonstrated an increase in gene expression level in response to treatment, possibly reflecting a demethylation event.

## B. Classification of Epigenetically-regulated Genes Based Upon Promoter CpG Features

It is well known that methylation affecting the promoter and downstream proximal sequences can result in gene silencing, but that methylation elsewhere in a gene will not hinder transcription (68). Therefore, we expected that many of the putative epigenetically-

**Figure 26. Alteration of gene expression and CpG methylation status during demethylation treatment.** Representation of a gene promoter CpG island located proximal to the transcription start site (indicated by the bent arrow) is depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides). The gene is silenced by promoter CpG methylation (represented by pink lollipops) resulting in a lack of expression. After demethylating treatment the gene promoter becomes unmethylated resulting in gene expression. Treatment withdrawal results in CpG remethylation and consequently loss of gene expression.


CpG Methylation
Not Expressed
Demethylating Treatment
Demethylation Event
Unmethylated
Expressed
Recovery
Remethylation Event
CpG Methylation
Not Expressed

regulated genes identified in MCF-7 cells would contain CpG islands within their promoter sequences. However, we found that only 45% of putative epigenetically-regulated genes contain typical CpG islands (65) in their promoter and/or exon 1, consistent with other studies reporting that genes lacking CpG islands are frequently induced in response to demethylating drugs (72-77,131). Some (or all) of these genes lacking CpG islands may respond to demethylating drugs as a result of indirect regulation by DNA methylation. That is, these genes may be regulated directly by the protein products (transcription factors, etc.) of genes containing CpG islands. Alternatively, novel CpG targets of DNA methylation may function to confer methylation-sensitivity to genes lacking CpG islands. In the present study, 40% of putative epigenetically-regulated genes contained weak CpG islands, whereas three genes were identified that lack features expected for epigenetically-regulated genes (typical or weak CpG islands). It is intriguing to suggest that novel CpG target sequences may confer methylation-sensitivity to these genes. Direct evidence for methylation-dependent regulation of genes lacking CpG islands has emerged from a few investigations (79,132,133). Based upon our observations, we propose that putative epigenetically-regulated genes can be classified based upon their promoter sequence characteristics related to CpG frequency and distribution: (i) genes with typical CpG features (CpG islands within the promoter or exon 1), (ii) genes with intermediate CpG features (weak CpG islands within the promoter or exon 1), and (iii) genes with atypical CpG features (no CpG islands).

### *Genes with Typical CpG Features*

Putative epigenetically-regulated genes with typical CpG features contain conventionally-defined CpG islands (65), and in some cases weak CpG islands as well. *CST6* represents the

prototype of a gene with typical CpG features, with a large CpG island that spans the promoter and exon 1, encompassing the start site for transcription. *CST6* is a member of a family of proteins that represent physiological inhibitors of lysosomal cysteine proteases that are expressed in normal and premalignant breast epithelium, but not in metastatic breast cancer cell lines (103). Ectopic expression of *CST6* suppresses the neoplastic phenotype of MDA-MD-435S breast cancer cells, reducing their cell proliferation, migration, and invasion *in vitro* (113), and delaying tumor growth and reducing metastatic tumor burden *in vivo* (116). *CST6* expression is significantly diminished in primary human breast cancers (116), which is unrelated to gene deletion (103) but may be due to transcriptional silencing through methylation of its CpG island (134). Our methylation analysis of the *CST6* promoter shows that this gene is subject to DNA methylation in MCF-7 cells, and that there is an inverse correlation between *CST6* expression and methylation of its promoter CpG island (Figure 9). These results strongly suggest that *CST6*, a putative breast cancer tumor suppressor gene (116), is sensitive to DNA methylation and that methylation-dependent epigenetic silencing may represent an important mechanism for loss of this gene during breast carcinogenesis and/or tumor progression.

***Genes with Intermediate CpG Features***

Putative epigenetically-regulated genes with intermediate CpG features contain small regions of CpG density (weak CpG islands), but lack typical CpG islands. *C8orf4* is an example of a gene with intermediate CpG features. While the specific function of *C8orf4* is not known, its loss of expression in primary tumors, metastases, and cancer cell lines (135), along with its expression/involvement in the TGFβ-suppressive pathway, suggest that this

gene is a growth suppressive gene associated with negative regulation of cell proliferation. Thus, decreased *C8orf4* expression could contribute to the loss of TGFβ responsiveness in breast cancer (136). The relative paucity of CpG dinucleotides within the promoter/exon 1 of *C8orf4* argues against its direct regulation by DNA methylation. Nonetheless, *C8orf4* is responsive to demethylating drugs in RKO colorectal carcinoma cells (72), as well as MCF-7 breast cancer cells (this study), suggesting the possibility that this gene contains novel CpG targets for methylation. The two strongest possibilities for novel methylation targets include the weak CpG island contained in exon 1, and individual CpG dinucleotides contained in the gene promoter. The weak CpG island is substantially methylated when *C8orf4* is silent, and becomes demethylated in response to treatment, coordinate with reexpression of the gene. However, treatment withdrawal results in gene silencing with only partial remethylation of these CpG dinucleotides, suggesting that methylation events in the promoter rather than exon 1 may be responsible for regulation of *C8orf4* in MCF-7 breast cancer cells. In fact, the greatest concentration of methylated CpGs in the *C8orf4* promoter was localized to a 103 bp region containing 5 CpG dinucleotides. Treatment of MCF-7 cells resulted in demethylation of this region and coordinate expression of *C8orf4*, while treatment withdrawal resulted in silencing of *C8orf4* expression and remethylation of these CpGs (Figure 10). These results suggest that *C8orf4* is subject to methylation-dependent epigenetic silencing in MCF-7 breast cancer cells through discrete promoter methylation events, possibly resulting in loss of TGFβ responsiveness.

***Genes with Atypical CpG Features***

The third class of genes identified in this study lack all features expected for epigenetically-regulated genes (including CpG islands). *IFI27* is a prototype for genes with atypical CpG features and an example of an interferon α-inducible gene (137), which have been implicated in primary breast tumors (138) and breast cancer cell lines (139) suggesting their importance in breast carcinogenesis. Untreated MCF-7 cells lack expression of *IFI27* and most (93%) CpG dinucleotides within the promoter region are methylated. Treatment of MCF-7 cells results in demethylation of the majority of CpG dinucleotides and concurrent expression of the gene, while treatment withdrawal leads to remethylation and loss of gene expression (Figure 10). These results suggest that *IFI27* is subject to epigenetic regulation in MCF-7 breast cancer cells, related to methylation of individual CpG dinucleotides contained in its promoter.

## C. An Expanded Model for Methylation-dependent Epigenetic Regulation of Gene Expression

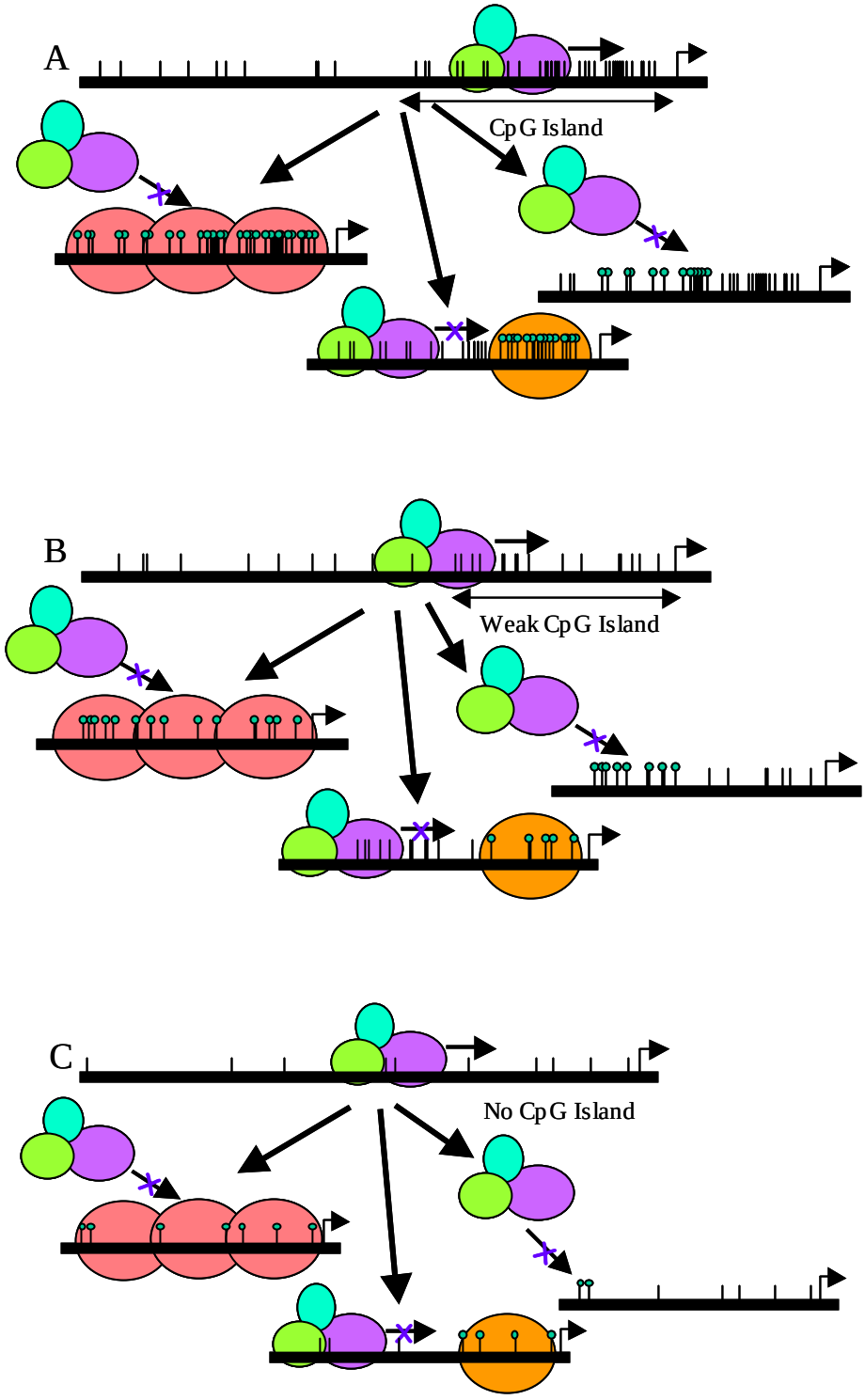
DNA methylation has three major effects on gene promoter sequences: (i) direct interference with the binding of transcription factors, (ii) attraction of methylated-DNA binding proteins, and (iii) alteration in chromatin packaging (26,140). Each of these effects results in diminished promoter activity as a consequence of impaired interactions between the transcription machinery and the promoter DNA sequence. To this point in time, it has been thought that a promoter CpG island was necessary to effectively catalyze methylation-dependent gene silencing through one of these mechanisms. However, recent evidence

suggests that genes lacking CpG islands can be directly regulated through methylation-dependent mechanisms (79,132,133). Therefore, based upon our results and studies from the literature, we propose expansion of the current model for DNA methylation-dependent epigenetic regulation of gene expression to include genes lacking typical CpG islands (Figure 27). The expanded model we propose recognizes (a) that all promoter CpG dinucleotides represent legitimate targets for methylation, (b) that sites for methylation may represent regional targets (CpG islands), local CpG density (weak CpG islands), or isolated CpGs, (c) that discrete methylation events occurring within CpG target sequences can contribute to gene silencing, and (d) that CpG methylation can contribute to gene silencing (or diminished expression) through several different mechanisms. This expanded model highlights the importance of the CpG characteristics of individual gene promoters and the targets for methylation that they contain, the nature of specific methylation events, and how these factors combine to regulate gene expression and/or silencing.

***Mechanisms of DNA Methylation-dependent Silencing***

Promoter CpG islands represent a recognized target for methylation leading to gene silencing. In most cases, methylation of a CpG island is considered to be a regional methylation event, where methylation of specific CpG dinucleotides is less important than the overall methylation of the CpG-dense region. CpG island methylation can result in gene silencing through several different mechanisms, including recruitment of methylated DNA binding proteins and/or direct interference with transcription factor complex binding to the promoter region. A number of methylated DNA binding proteins have been identified (141-143), several of which have particular affinity for CpG-rich heterochromatin (144). The

**Figure 27. Expanded model for methylation-dependent epigenetic regulation of gene expression.** The distribution of CpG dinucleotides proximal to the transcription start site (indicated by the bent arrow) is depicted schematically (vertical lines indicate the relative position of individual CpG dinucleotides) for theoretical genes with variable promoter CpG density. CpG sequence features (CpG islands, weak CpG islands) are indicated. Green lollipops correspond to methylated CpG dinucleotides. Genes with typical CpG features (A) are subject to both regional and discrete methylation events. Regional methylation can affect transcription through recruitment of methylated DNA binding proteins (pink ball). Likewise, focal methylation within a larger CpG island may attract methylated DNA binding proteins (orange ball) that might inhibit transcription by blocking the procession of the transcription machinery to the transcriptional start site. Alternatively, focal or regional methylation of a CpG island at a transcription factor binding sequence may directly interfere with binding by the transcription machinery (represented as associated green, blue, and purple balls). Similar or identical consequences could result from methylation events involving promoters of genes with intermediate CpG features (B) or atypical CpG features (C). This proposed model highlights the importance of the CpG characteristics of individual gene promoters and the targets for methylation that they contain, the nature of specific methylation events, and how these factors combine to regulate gene expression and/or silencing.


A
CpG Island
B
Weak CpG Island
C
No CpG Island

binding of these proteins to methylated target sequences results in inhibition of transcription (145), possibly through the specific function of transcription repression domains (146). In addition, several transcription factors have been shown to be sensitive to methylation of their recognition sequence (147). However, some investigators have suggested that transcriptional silencing requires association of proteins to methylated sequences (148,149). Nonetheless, there is strong evidence that methylation can directly interfere with the binding of some transcription factors to their recognition site (80,150). While several mechanisms for inhibition of transcription involving methylation of CpG islands have been established or proposed, DNA methylation-dependent mechanisms of regulation of genes lacking well-defined CpG islands are more elusive. One possibility is that genes lacking CpG islands are not truly epigenetically regulated, but that their expression is governed by indirect methylation-dependent mechanisms, secondary to the epigenetic regulation of CpG island-containing regulatory genes (encoding transcription factors or other regulatory proteins). While in some cases this may be true, in other cases there is direct evidence for methylation-dependent regulation of genes lacking CpG islands (62). We propose that mechanisms for methylation-dependent gene silencing similar to those suggested for genes with typical CpG features may also apply to genes with intermediate or atypical CpG features. Hence, methylation of weak CpG islands or discrete methylation events affecting specific CpG dinucleotides may (i) recruit methylated DNA binding proteins resulting in a blockade of transcription factor access to crucial recognition sequences, or (ii) directly inhibit transcription factor binding to the promoter region. Methylated DNA binding proteins that require very few methylated CpG sites or only a single methylated CpG dinucleotide for binding have been described (151). Likewise, methylation of specific CpG dinucleotides

within or proximal to transcription factor binding sequences can lead to loss of proper transcription factor interaction with its target sequence, negatively impacting on gene transcription (80,150). There are several examples in the literature of methylation-dependent silencing of genes with intermediate or atypical CpG features. Well-characterized examples of methylation-sensitive genes with intermediate CpG features include *E-cadherin* (25,78), *RAR-β2* (79), and *APC* (80). Our survey of the literature identified only one example of a well characterized gene with atypical CpG features, *LAMB3*, which is silenced through promoter methylation in cancers of the prostate, breast, lung, and bladder (81-84). Whereas there are only a few examples of methylation-sensitive genes lacking CpG islands in the current literature, numerous genes with intermediate or atypical CpG features are likely to have been identified in microarray studies aimed at cataloguing cancer-related epigenetically-regulated genes (72-74). The results from the current study suggest that genes lacking CpG islands from these previous studies should be rigorously evaluated to characterize their methylation status in breast cancer, and to examine the possible mechanisms through which methylation of weak CpG islands or discrete methylation targets (individual CpG dinucleotides) produce gene silencing.

## D. *CST6* is Silenced by DNA Methylation in Breast Cancer Cells

### *CST6: A Candidate Breast Tumor Suppressor Gene*

*CST6* (which encodes cystatin M) was originally identified as a gene whose expression is lost in metastatic breast cancer, suggesting a possible role for this gene in suppression of the invasive/metastatic phenotype (103). Consistent with this suggestion, normal human breast

epithelial cells express high levels of cystatin M, while invasive ductal carcinoma cells do not express or express very low levels of this protein (116). Cystatin M is a member of a family of proteins that function as inhibitors of lysosomal cysteine proteases, which include the cathepsin proteases (134,152). Increased levels of cathepsin protease activity have been noted in breast cancer (153), and predict poor prognosis among these patients (154). Thus, breast cancer invasiveness may be a direct consequence of inappropriate cathepsin protease activity in the absence of their inhibitory molecules, which include cystatin M (155). While *CST6* expression is known to be lost in many primary breast tumors and cancer cell lines, the mechanism that accounts for loss of *CST6* expression has only recently been investigated. Given a role for *CST6* in tumor suppression and/or metastasis suppression, possible mechanisms for loss of expression include genetic alterations (mutation or deletion) and epigenetic silencing. Southern blotting in breast cancer cell lines failed to identify gross structural rearrangements of the *CST6* gene or deletion of the gene locus (103), leading to the more recent suggestion that *CST6* expression may be lost due to gene silencing, either in response to a transcriptional repressor protein or as a consequence of promoter hypermethylation (134).

### *CST6 is Silenced by Methylation in Multiple Cancers*

The *CST6* gene contains a large CpG island that encompasses the proximal promoter and exon 1, consisting of 54 CpG dinucleotides distributed over a 424 bp region (CpGPlot, http://www.ebi.ac.uk/emboss/cpgplot/). We identified *CST6* as a target for methylation-dependent gene silencing in MCF-7 breast cancer cells by a microarray-based gene expression study, and showed that loss of *CST6* expression in these cells is related to

hypermethylation of its CpG island (156). More recently, Ai *et al.* showed that 12/20 (60%) primary breast tumors exhibit *CST6* promoter hypermethylation, and microdissection of individual cells from select tumors revealed that methylation occurs in both DCIS and IDC cells (8). Additionally, Kim *et al.* (112) reported *CST6* to be frequently methylated in glioma cell lines and primary brain tumors. We have established a direct association between *CST6* promoter hypermethylation and gene silencing in a panel of human breast cancer cell lines that differentially express *CST6*. The majority of *CST6*-negative cell lines examined were originally isolated from invasive/metastatic breast neoplasms (157-159), whereas the *CST6*-positive BT-20 cells (which express very high levels of *CST6*) were derived from a primary breast carcinoma (160). These results suggest strongly that (i) CpG island hypermethylation contributes to *CST6* silencing in breast cancer cell lines, and (ii) the loss of *CST6* expression is associated with the invasive/metastatic phenotype of the breast cancer cell line.

### *Epigenetic Mechanisms of CST6 Gene Silencing by DNA Methylation*

Epigenetic silencing of gene expression is a consequence of DNA hypermethylation and/or chromatin remodeling related to direct interference with the binding of transcription factors to their recognition sequences (80,150), indirect mechanisms associated with recruitment and binding of methylated DNA binding proteins (141-144), and/or modification of histone proteins and alteration of chromatin structure (140). In the current study, we generated evidence for CpG island hypermethylation in the epigenetic silencing of *CST6* in human breast cancer cell lines. However, the precise mechanism and the temporal order of events related to *CST6* gene silencing have not been determined. Nonetheless, we were able to gain significant insight into the process through comparison of the natural methylation

pattern of *CST6*-positive and -negative cell lines. *CST6*-negative cell lines are characterized by extensive CpG island methylation, suggesting the possibility that regional methylation across the *CST6* proximal promoter and exon 1 may be required for gene silencing. However, a subset of *CST6*-positive cell lines contain a significant level of methylation in exon 1, indicating that methylation in this portion of the CpG island does not negatively impact on *CST6* transcription. In contrast, methylation of CpG dinucleotides in the proximal promoter is strongly associated with loss of *CST6* expression, suggesting that regional methylation or specific methylation events affecting this portion of the promoter contribute to gene silencing. There is evidence that discrete methylation events within a larger methylation target (CpG island) can negatively affect gene expression. The AP-2α tumor suppressor gene is an example of a gene that is silenced in response to CpG methylation of a discrete region that is contained within a larger CpG island (71). Discrete methylation events may negatively impact on promoter function by direct interference with transcription factor binding or through indirect interference related to binding by methylated DNA binding proteins. Among the 55 CpGs that comprise the *CST6* CpG island, seven CpGs in the proximal promoter were found to be unmethylated in all cell lines that express the gene, and these CpGs were frequently methylated in *CST6*-negative cell lines. Specific methylation events affecting these CpG dinucleotides may be required for *CST6* gene silencing, but it is not clear if these specific methylation events occur in isolation or if they always take place in conjunction with more extensive regional methylation. Methylation of these CpGs may directly impact on the binding of specific transcription factors to their recognition sequence. Analysis of the *CST6* promoter using ProSpector (http://prospector.nci.nih.gov) identified 16 CpGs within the proximal promoter that directly impinge upon transcription factor binding

sites and are associated with loss of *CST6* expression when methylated, including sequences for AP-2, AP-4, Egr-1, MEIS1, NF-κB, Sp1, and YY1. Methylated CpGs in the binding sites for AP-2 and Sp1 transcription factors have been shown to directly down-regulate gene expression (161,162). Of note, CpG dinucleotide 13, which impinges on three transcription factor binding sites, is never methylated in *CST6*-positive cell lines. The transcription factor requirements for expression of the *CST6* promoter have not been determined. Nonetheless, CpG methylation of the proximal promoter of *CST6* is likely to inhibit or impair gene transcription by either direct or indirect interference with the transcription machinery.

***Methylation Events Leading to CST6 Silencing in Breast Cancer***

The differential CpG island methylation profile among *CST6*-postive and -negative breast cancer cell lines indicates that certain methylation events and/or specific promoter regions are strongly associated with gene silencing. Nevertheless, *CST6*-negative cell lines tend to be heavily methylated across the entire promoter/exon 1 CpG island. We posit that there is a succession of methylation events that progressively leads to *CST6* gene silencing in metastatic breast cancer cell lines: (i) individual CpG dinucleotides within the promoter region are preferentially methylated resulting in transient silencing of *CST6*, (ii) methylation spreads throughout the promoter/exon 1 CpG island, (iii) chromatin remodeling occurs resulting in stable silencing of *CST6*. Alternatively, chromatin remodeling might occur prior to the completion of regional methylation affecting the entire CpG island. Additional studies will be required to establish the temporal order of events and the nature of chromatin alterations that accompany *CST6* silencing in breast cancer cell lines. Likewise, additional

studies will be required to determine if methylation-dependent gene silencing accounts for loss of *CST6* expression in primary breast tumors and their metastatic lesions.

## E. Methylation-dependent Silencing of *CST6* in Human Breast Cancer

### *Role of Tumor Suppressor CST6 in Human Breast Cancer*

Cystatin M was originally described as exhibiting diminished expression in metastatic breast cancer, suggesting a role in suppression of the invasive/metastatic phenotype (103). It has been shown that *CST6* is epigenetically regulated by DNA methylation-dependent silencing in breast cancer cell lines (8,113-115) and primary invasive ductal carcinomas (8,115). In a recent study, Schagdarsurengin *et al.* showed that 24/40 (60%) breast carcinomas exhibited *CST6* promoter hypermethylation, and that estrogen-receptor positive tumors were more frequently methylated than estrogen-receptor negative tumors (115). While *CST6* is suggested to be epigenetically regulated through DNA methylation-dependent mechanisms in breast cancer cell lines and primary breast tumors that lack cystatin M protein expression, tumor metastases have not been examined for cystatin M expression or methylation status. Given a putative role for *CST6* in suppression of tumor invasion and metastasis, loss of cystatin M expression may be one mechanism that enables tumor cells to spread from the primary site and invade adjacent tissues (or distant sites) during breast cancer progression. Furthermore, evidence from breast cancer cell lines suggests that *CST6* promoter hypermethylation leading to gene silencing may represent one major mechanism for loss of cystatin M in breast cancer. *CST6* is located in the chromosomal region 11q13, which is subject to amplification or loss of heterozygosity in several cancers (94,118,163).

Previously, we reported that the majority of *CST6*-negative breast cancer cell lines were originally established from metastatic lesions (pleural effusions) rather than primary breast tumors and that *CST6*-positive breast cancer cell line (BT-20) was derived from a primary breast carcinoma (114). These observations argue that the loss of *CST6* expression is strongly associated with the invasive/metastatic phenotype of the breast cancer cell line and that *CST6* promoter hypermethylation may be frequently involved in gene silencing/loss. In the current study, we present evidence that metastatic breast cancers exhibit lower levels of cystatin M protein expression and increased *CST6* promoter hypermethylation compared to primary breast tumors.

***Proposed Mechanism for Epigenetic Silencing of CST6 in Human Breast Cancer***

The differential CpG island methylation of *CST6* between primary breast tumors and lymph node metastases indicates that certain individual methylation events occur during or following stromal invasion and tumor spread. We envision that there is a succession of methylation events that lead to *CST6* gene silencing in metastatic breast cancer: (i) individual CpG dinucleotides within the promoter region are preferentially methylated resulting in decreased expression of cystatin M in DCIS and/or primary breast carcinomas, (ii) methylation spreads throughout the CpG island during surrounding stromal invasion of tumor cells and metastasis to the regional lymph nodes resulting in a complete loss of cystatin M protein expression, and (iii) chromatin remodeling occurs resulting in the stable silencing of *CST6*. However, breast tumors that exhibit silencing of cystatin M but lack DNA methylation could achieve this silencing through histone deacetylation or through a putative transcription repressor binding to the promoter regulatory regions of *CST6*.

Likewise, we have observed areas of tumor sections that show strong staining for cystatin M and areas that show weak staining. This finding may indicate that cystatin M silencing can be heterogeneous within a single breast tumor and can reflect different levels of DNA methylation.

## F. *CST6* CpG Methylation Requires an Upstream DNA Sequence Element that Directs Promoter Methylation Events

### *Regulation of DNA Methylation by Directive DNA Sequences*

Genomic DNA exhibits a specific pattern of CpG methylation in normal cells, and a different pattern in cancer cells, both of which reflect nonrandom hypermethylation of specific regions of DNA resulting in silencing of specific genes. The mechanisms that control this nonrandom distribution of CpG methylation are poorly understood. We suggest that the instructional signals that govern (direct) DNA methylation at specific CpG targets will be contained within the DNA sequence (*cis* elements) of regulatory regions of methylation-sensitive genes. Possible mechanisms that direct CpG methylation include, (i) CG-rich regions such as CpG islands within the promoter of a gene that can direct methylation to the CG-dense region, and (ii) that CpG methylation can be signaled by a *cis*-acting DNA sequence element. The promoter region of many genes contain CpG islands, and sequences have been identified that can protect CpG islands from DNA methylation (87,164). Likewise, several lines of evidence support the notion that *cis*-acting sequence elements exist that regulate *de novo* methylation, including directive (methylation-promoting) instructions (85,86). It is conceivable that these directive and protective elements

coexist in the promoter regions of epigenetically-regulated genes, and that a balance between these forces dictates the methylation status of the promoter in specific cell types or under specific cellular conditions.

### *Methylation Directing Cis-acting Elements in DNA Sequences*

Recognizing that methylated genes are nonrandom in normal and cancer cell types, a number of studies have addressed the question of whether *cis* elements direct DNA methylation of specific target genes. A *cis* element responsible for aberrant methylation of the APRT promoter was localized to a 838 bp region approximately 1.3 kbp upstream of the transcription start site (90,91). Deletion analysis of the chloramphenicol acetyltransferase (*CAT*) gene localized two *cis*-acting elements to 775 bp and 1.3 kbp upstream of the transcription start site (165). The effects of these elements appeared to be exerted in *cis*, and dependent on proximity, but not on orientation (165). The ideal size of the *cis* elements was between 500-700 bp and small retrotransposon sequences within the larger *cis*-acting element sequences show greater efficiency in attracting methylating enzymes (165). In addition, a number of sequence elements have been identified that predict methylation of promoter sequences with high discrimination potential (92). However, it has not been determined if these sequences function to direct or promote methylation. In a recent study, a computational epigenetics approach was utilized to discriminate between CpG islands that are methylation-prone from those that remain unmethylated (166). Bock *et al.* showed that the methylation state of CpG islands (methylated and unmethylated) were determined by a complex combination of the presence or absence of sequence motifs found within the DNA sequence, and proposed that the methylation pattern of an individual CpG island can be assigned a degree based on DNA sequence (166). These findings support the idea that promoter regions

in genes containing a CpG methylation target contain specific DNA sequences involved in the regulation of DNA methylation.

### *Disassociation of Methylation Directing Cis-acting Elements From the Upstream Promoter Sequence of CST6*

We have utilized the breast tumor suppressor *CST6* as an index gene for the identification of *cis* elements that direct promoter CpG island methylation. The *CST6* gene contains a large CpG island that encompasses sequences within the proximal promoter and exon 1. This CpG island represents the target for methylation in our model system. To examine the existence of instructional *cis* regulatory elements upstream of this promoter CpG island target sequence, two regions of the *CST6* gene promoter (-1187 to +33 and -438 to +33), including the proximal promoter CpG island, were cloned into luciferase reporter constructs and transfected into cell lines known to methylate and silence the *CST6* gene (114). Using this model system, truncation of the *CST6* promoter region disassociated a putative instructional *cis* element from the target CpG island, resulting in a lack of methylation of the downstream target sequence. Thus, breast cancer cell lines that hypermethylate and silence the endogenous *CST6* gene fail to hypermethylate the exogenous *CST6*-luciferase reporter construct. Consistent with several previous reports on other methylation-sensitive genes (85,86,90,91,165), this observation suggests that the *CST6* promoter CpG island does not direct its own methylation. This result supports the suggestion that regulatory sequences are required to direct CpG island methylation by the DNA methylation machinery. A few investigators have identified and characterized *cis* regulatory sequences that appear to direct CpG island methylation. Hasse *et al.* localized a *cis*-acting regulatory element responsible

for the methylation of the *CAT* gene approximately 775 bp upstream of the transcriptional start site (165). However, our studies of the *CST6* promoter indicate that the putative *cis* regulatory element is >1200 bp upstream of the transcription start site, similar to the findings of other published studies (90,91). While the results of the current study are intriguing, a more extensive *CST6* promoter truncation analysis needs to be performed to precisely identify the location of the *cis* regulatory element that governs CpG island methylation of the *CST6* promoter and to characterize its properties.

## G. Conclusions and Impact

The studies contained in this dissertation are relevant to breast cancer research and DNA methylation-dependent gene regulation in many significant ways. These studies (i) characterize a group of putative methylation-sensitive genes identified in MCF-7 breast carcinoma cells, (ii) validates that the methylation-sensitive gene, *CST6* is subject to methylation-dependent regulation in multiple breast cancer cell lines, primary breast tumors, and lymph node metastases, and (iii) identifies critical promoter methylation targets. Thus, the completion of these studies has established a role for methylation-dependent epigenetic mechanisms in the silencing of important genes in the molecular pathogenesis of breast cancer. Furthermore, this dissertation has begun to establish several distinct classes of epigenetically-regulated genes, and that these classes can be distinguished based upon the CpG content and CpG organization of their promoters. Thus, a new definition for an epigenetically-regulated gene that recognizes the importance of all CpG targets has been

proposed. In addition, this dissertation identifies the existence of *cis* regulatory sequences located in the 5' upstream promoter region of *CST6* that function to direct CpG methylation. Consequently, these results advance our understanding of mechanisms governing DNA methylation in breast carcinogenesis.

## V. REFERENCES


1. Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., and Thun, M. J. (2007) Cancer statistics, 2007. *CA Cancer J Clin* **57**, 43-66

2. Reis, L. A. G., Harkins, D., Krapcho, M., Mariotto, A., Miller, B. A., Feuer, E. J., Clegg, L., Eisner, M. P., Horner, M. J., Howlader, N., Hayat, M., Hankey, B. F., and Edwards, B. K. (2007) *SEER Cancer Statistics Review, 1975-2003*, National Cancer Institute, Bethesda, MD

3. Coleman, W. B., and Tsongalis, G. J. (2002) Cancer Epidemiology: Incidence and Etiology of Human Neoplasms in *The Molecular Basis of Human Cancer* (Coleman, W. B., and Tsongalis, G. J., eds) pp. 3-22, Humana Press, Totowa

4. Charpentier, A., and Aldaz, C. M. (2002) The Molecular Basis of Breast Carcinogenesis in *The Molecular Basis of Human Cancer* (Coleman, W. B., and Tsongalis, G. J., eds), pp. 347-363, Humana Press, Totowa, N.J.

5. Ponten, J., Holmberg, L., Trichopoulos, D., Kallioniemi, O., Kvale, G., Wallgren, A., and Taylor-Papadimitriou, J. (1990) Biology and the natural history of breast cancer. *Int. J. Cancer* **5**, 5-21

6. Symmans, W. F. (2005) Histopathology of Breast Cancer: Correlation with Molecular Markers in *Molecular Oncology of Breast Cancer* (Ross, J. S., and Hortobagyi, G. N., eds) pp. 106-116, Jones and Bartlett, Sudbury, M.A.

7. Simpson, P. T., Reis-Filho, J. S., Gale, T., and Lakhani, S. R. (2005) Molecular evolution of breast cancer. *J Pathol* **205**, 248-254

8. Ai, L., Kim, W. J., Kim, T. Y., Fields, C. R., Massoll, N. A., Robertson, K. D., and Brown, K. D. (2006) Epigenetic silencing of the tumor suppressor cystatin M occurs during breast cancer progression. *Cancer Res* **66**, 7899-7909

9. Muller, V., Kasimir-Bauer, S., and Pantel, K. (2005) Sentinel Lymph Node Dissection and Micrometastasis Detection in Bone Marrow and Lymph Nodes in *Molecular Oncology of Breast Cancer* (Ross, J. S., and Hortobagyi, G. N., eds) pp. 117-127, Jones and Bartlett, Boston

10. Hortobagyi, G. N. (2005) Introduction and Background in *Molecular Oncology of Breast Cancer* (Ross, J. S., and Hortobagyi, G. N., eds) pp. 1-11, Jones and Bartlett Publishers, Sudbury, M.A.

11. Reed, W., Hannisdal, E., Boehler, P. J., Gundersen, S., Host, H., and Marthin, J. (2000) The prognostic value of p53 and c-erb B-2 immunostaining is overrated for patients with lymph node negative breast carcinoma: A multivariate analysis of

prognostic factors in 613 patients with a follow-up of 14-30 years. *Cancer* **88**, 804-813

12. Mirza, A. N., Mirza, N. Q., Vlastos, G., and Singletary, S. E. (2002) Prognostic factors in node-negative breast cancer: A review of studies with sample size more than 200 and follow-up more than 5 years. *Ann Surg* **235**, 10-26

13. Lau, R., Grimson, R., Sansome, C., Tornos, C., and Moll, U. M. (2001) Low levels of cell cycle inhibitor p27kip1 combined with high levels of Ki-67 predict shortened disease-free survival in T1 and T2 invasive breast carcinomas. *Int J Oncol* **18**, 17-23

14. Sutcliffe, S., Pharoah, P. D., Easton, D. F., and Ponder, B. A. (2000) Ovarian and breast cancer risks to women in families with two or more cases of ovarian cancer. *Int. J. Cancer* **87**, 110-117

15. Gayther, S. A., Pharoah, P. D., and Ponder, B. A. (1998) The genetics of inherited breast cancer. *J. Mammary Gland Biol.* **3**, 365-376

16. Sotiriou, C., Desmedt, C., Durbecq, V., Dal Lago, L., Lacroix, M., Cardoso, F., and Piccart, M. (2005) Genomic and Molecular Classification of Breast Cancer in *Molecular Oncology of Breast Cancer* (Ross, J. S., and Hortobagyi, G. N., eds) pp. 81-95, Jones and Bartlett, Sudbury, M.A.

17. Lakhani, S. R., Gusterson, B. A., Jacquemier, J., Sloane, J. P., Anderson, T. J., van de Vijver, M. J., Venter, D., Freeman, A., Antoniou, A., McGuffog, L., Smyth, E., Steel, C. M., Haites, N., Scott, R. J., Goldagar, D., Neuhausen, S., Daly, P. A., Ormiston, W., McManus, R., Scherneck, S., Ponder, B. A., Futreal, P. A., Peto, J., Stoppa-Lyonnet, D., Bignon, Y. J., and Stratton, M. R. (2000) The pathology of familial breast cancer: Histological features of cancers in families not attributable to mutations in BRCA1 and BRCA2. *Clin. Cancer Res.* **6**, 782-789

18. Srivastava, S., Zou, Z. Q., Pirollo, K., Blattner, W., and Chang, E. H. (1990) Germ-line transmission of a mutated p53 gene in a cancer-prone family with Li-Fraumeni syndrome. *Nature* **348**, 747-749

19. Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F., Jr., Nelson, C. E., Kim, D. H., Kassel, J., Gryka, M. A., Bischoff, F. Z., Tainsky, M. A., and Friend, S.H. (1990) Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250**, 1233-1238

20. Kelsey, J. L., and Berkowitz, G. S. (1988) Breast cancer epidemiology. *Cancer Res.* **48**, 5617-5623

21. Hsieh, C. C., Trichopoulos, D., Katsouyanni, K., and Yuasa, S. (1990) Age at menarche, age at menopause, height and obesity as risk factors for breast cancer:

Associations and interactions in an international case-control study. *Int. J. Cancer* **46**, 796-800

22. Kelsey, J. L., Gammon, M. D., and John, E. M. (1993) Reproductive factors and breast cancer. *Epidemiol Rev* **15**, 36-47

23. Lipworth, L. (1995) Epidemiology of breast cancer. *Eur J Cancer Prev* **4**, 7-30

24. Yang, X., Yan, L., and Davidson, N. E. (2001) DNA methylation in breast cancer. *Endocr Relat Cancer* **8**, 115-127

25. Widschwendter, M., and Jones, P. A. (2002) DNA methylation and breast carcinogenesis. *Oncogene* **21**, 5462-5482

26. Herman, J. G., and Baylin, S. B. (2003) Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med* **349**, 2042-2054

27. Baylin, S. (2001) DNA methylation and epigenetic mechanisms of carcinogenesis. *Dev Biol (Basel)* **106**, 85-87

28. Baylin, S. B., Herman, J. G., Graff, J. R., Vertino, P. M., and Issa, J. P. (1998) Alterations in DNA methylation: a fundamental aspect of neoplasia. *Adv Cancer Res* **72**, 141-196

29. Feinberg, A. P., and Vogelstein, B. (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**, 89-92

30. Feinberg, A. P., and Vogelstein, B. (1987) Alterations in DNA methylation in human colon neoplasia. *Semin Surg Oncol* **3**, 149-151

31. Goelz, S. E., Vogelstein, B., Hamilton, S. R., and Feinberg, A. P. (1985) Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science* **228**, 187-190

32. Feinberg, A. P., and Vogelstein, B. (1983) Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun* **111**, 47-54

33. Narayan, A., Ji, W., Zhang, X. Y., Marrogi, A., Graff, J. R., Baylin, S. B., and Ehrlich, M. (1998) Hypomethylation of pericentromeric DNA in breast adenocarcinomas. *Int J Cancer* **77**, 833-838

34. Eden, A., Gaudet, F., Waghmare, A., and Jaenisch, R. (2003) Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* **300**, 455

35. Gaudet, F., Hodgson, J. G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J. W., Leonhardt, H., and Jaenisch, R. (2003) Induction of tumors in mice by genomic hypomethylation. *Science* **300**, 489-492

36. Momparler, R. L. (2003) Cancer epigenetics. *Oncogene* **22**, 6479-6483

37. Jones, P. A., and Laird, P. W. (1999) Cancer epigenetics comes of age. *Nat Genet* **21**, 163-167

38. Tsou, J. A., Hagen, J. A., Carpenter, C. L., and Laird-Offringa, I. A. (2002) DNA methylation analysis: a powerful new tool for lung cancer diagnosis. *Oncogene* **21**, 5450-5461

39. Hanahan, D., and Weinberg, R. A. (2000) The hallmarks of cancer. *Cell* **100**, 57-70

40. Yan, P. S., Shi, H., Rahmatpanah, F., Hsiau, T. H., Hsiau, A. H., Leu, Y. W., Liu, J. C., and Huang, T. H. (2003) Differential distribution of DNA methylation within the RASSF1A CpG island in breast cancer. *Cancer Res* **63**, 6178-6186

41. Dammann, R., Yang, G., and Pfeifer, G. P. (2001) Hypermethylation of the CpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21.3 locus, occurs in a large percentage of human breast cancers. *Cancer Res* **61**, 3105-3109

42. Widschwendter, M., Berger, J., Muller, H. M., Zeimet, A. G., and Marth, C. (2001) Epigenetic downregulation of the retinoic acid receptor-beta2 gene in breast cancer. *J Mammary Gland Biol Neoplasia* **6**, 193-201

43. Widschwendter, M., Berger, J., Hermann, M., Muller, H. M., Amberger, A., Zeschnigk, M., Widschwendter, A., Abendstein, B., Zeimet, A. G., Daxenbichler, G., and Marth, C. (2000) Methylation and silencing of the retinoic acid receptor-beta2 gene in breast cancer. *J Natl Cancer Inst* **92**, 826-832

44. Herman, J. G., Merlo, A., Mao, L., Lapidus, R. G., Issa, J. P., Davidson, N. E., Sidransky, D., and Baylin, S. B. (1995) Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res* **55**, 4525-4530

45. Woodcock, D. M., Linsenmeyer, M. E., Doherty, J. P., and Warren, W. D. (1999) DNA methylation in the promoter region of the p16 (CDKN2/MTS-1/INK4A) gene in human breast tumours. *Br J Cancer* **79**, 251-256

46. Lapidus, R. G., Nass, S. J., and Davidson, N. E. (1998) The loss of estrogen and progesterone receptor gene expression in human breast cancer. *J Mammary Gland Biol Neoplasia* **3**, 85-94

47. Ottaviano, Y. L., Issa, J. P., Parl, F. F., Smith, H. S., Baylin, S. B., and Davidson, N. E. (1994) Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells. *Cancer Res* **54**, 2552-2555

48. Lapidus, R. G., Ferguson, A. T., Ottaviano, Y. L., Parl, F. F., Smith, H. S., Weitzman, S. A., Baylin, S. B., Issa, J. P., and Davidson, N. E. (1996) Methylation of estrogen and progesterone receptor gene 5' CpG islands correlates with lack of estrogen and progesterone receptor gene expression in breast tumors. *Clin Cancer Res* **2**, 805-810

49. Merajver, S. D., Pham, T. M., Caduff, R. F., Chen, M., Poy, E. L., Cooney, K. A., Weber, B. L., Collins, F. S., Johnston, C., and Frank, T. S. (1995) Somatic mutations in the BRCA1 gene in sporadic ovarian tumours. *Nat Genet* **9**, 439-443

50. Dobrovic, A., and Simpfendorfer, D. (1997) Methylation of the BRCA1 gene in sporadic breast cancer. *Cancer Res* **57**, 3347-3350

51. Magdinier, F., Ribieras, S., Lenoir, G. M., Frappart, L., and Dante, R. (1998) Down-regulation of BRCA1 in human sporadic breast cancer; analysis of DNA methylation patterns of the putative promoter region. *Oncogene* **17**, 3169-3176

52. Rice, J. C., Massey-Brown, K. S., and Futscher, B. W. (1998) Aberrant methylation of the BRCA1 CpG island promoter is associated with decreased BRCA1 mRNA in sporadic breast cancer cells. *Oncogene* **17**, 1807-1812

53. Bianco, T., Chenevix-Trench, G., Walsh, D. C., Cooper, J. E., and Dobrovic, A. (2000) Tumour-specific distribution of BRCA1 promoter region methylation supports a pathogenetic role in breast and ovarian cancer. *Carcinogenesis* **21**, 147-151

54. Catteau, A., Harris, W. H., Xu, C. F., and Solomon, E. (1999) Methylation of the BRCA1 promoter region in sporadic breast and ovarian cancer: correlation with disease characteristics. *Oncogene* **18**, 1957-1965

55. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O. P., Wilfond, B., Borg, A., and Trent, J. (2001) Gene-expression profiles in hereditary breast cancer. *N Engl J Med* **344**, 539-548

56. Magdinier, F., Billard, L. M., Wittmann, G., Frappart, L., Benchaib, M., Lenoir, G. M., Guerin, J. F., and Dante, R. (2000) Regional methylation of the 5' end CpG island of BRCA1 is associated with reduced gene expression in human somatic cells. *FASEB J* **14**, 1585-1594

57. Miyamoto, K., Fukutomi, T., Asada, K., Wakazono, K., Tsuda, H., Asahara, T., Sugimura, T., and Ushijima, T. (2002) Promoter hypermethylation and post-transcriptional mechanisms for reduced BRCA1 immunoreactivity in sporadic human breast cancers. *Jpn J Clin Oncol* **32**, 79-84

58. Bringuier, P. P., Umbas, R., Schaafsma, H. E., Karthaus, H. F., Debruyne, F. M., and Schalken, J. A. (1993) Decreased E-cadherin immunoreactivity correlates with poor survival in patients with bladder tumors. *Cancer Res* **53**, 3241-3245

59. Nass, S. J., Herman, J. G., Gabrielson, E., Iversen, P. W., Parl, F. F., Davidson, N. E., and Graff, J. R. (2000) Aberrant methylation of the estrogen receptor and E-cadherin 5' CpG islands increases with malignant progression in human breast cancer. *Cancer Res* **60**, 4346-4348

60. Uria, J. A., Ferrando, A. A., Velasco, G., Freije, J. M., and Lopez-Otin, C. (1994) Structure and expression in breast tumors of human TIMP-3, a new member of the metalloproteinase inhibitor family. *Cancer Res* **54**, 2091-2094

61. Bachman, K. E., Herman, J. G., Corn, P. G., Merlo, A., Costello, J. F., Cavenee, W. K., Baylin, S. B., and Graff, J. R. (1999) Methylation-associated silencing of the tissue inhibitor of metalloproteinase-3 gene suggest a suppressor role in kidney, brain, and other human cancers. *Cancer Res* **59**, 798-802

62. Jones, P. A., and Takai, D. (2001) The role of DNA methylation in mammalian epigenetics. *Science* **293**, 1068-1070

63. Takai, D., and Jones, P. A. (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A* **99**, 3740-3745

64. Bird, A. P. (1986) CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209-213

65. Gardiner-Garden, M., and Frommer, M. (1987) CpG islands in vertebrate genomes. *J Mol Biol* **196**, 261-282

66. Razin, A., and Riggs, A. D. (1980) DNA methylation and gene function. *Science* **210**, 604-610

67. Razin, A., and Szyf, M. (1984) DNA methylation patterns. Formation and function. *Biochim Biophys Acta* **782**, 331-342

68. Jones, P. A. (1999) The DNA methylation paradox. *Trends Genet* **15**, 34-37

69. Csankovszki, G., Nagy, A., and Jaenisch, R. (2001) Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *J Cell Biol* **153**, 773-784

70. Ferguson, A. T., Evron, E., Umbricht, C. B., Pandita, T. K., Chan, T. A., Hermeking, H., Marks, J. R., Lambers, A. R., Futreal, P. A., Stampfer, M. R., and Sukumar, S.

(2000) High frequency of hypermethylation at the 14-3-3 sigma locus leads to gene silencing in breast cancer. *Proc Natl Acad Sci U S A* **97**, 6049-6054

71. Douglas, D. B., Akiyama, Y., Carraway, H., Belinsky, S. A., Esteller, M., Gabrielson, E., Weitzman, S., Williams, T., Herman, J. G., and Baylin, S. B. (2004) Hypermethylation of a small CpGuanine-rich region correlates with loss of activator protein-2alpha expression during progression of breast cancer. *Cancer Res* **64**, 1611-1620

72. Suzuki, H., Gabrielson, E., Chen, W., Anbazhagan, R., van Engeland, M., Weijenberg, M. P., Herman, J. G., and Baylin, S. B. (2002) A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nat Genet* **31**, 141-149

73. Sato, N., Fukushima, N., Maitra, A., Matsubayashi, H., Yeo, C. J., Cameron, J. L., Hruban, R. H., and Goggins, M. (2003) Discovery of novel targets for aberrant methylation in pancreatic carcinoma using high-throughput microarrays. *Cancer Res* **63**, 3735-3742

74. Liang, G., Gonzales, F. A., Jones, P. A., Orntoft, T. F., and Thykjaer, T. (2002) Analysis of gene induction in human fibroblasts and bladder cancer cells exposed to the methylation inhibitor 5-aza-2'-deoxycytidine. *Cancer Res* **62**, 961-966

75. Soengas, M. S., Capodieci, P., Polsky, D., Mora, J., Esteller, M., Opitz-Araya, X., McCombie, R., Herman, J. G., Gerald, W. L., Lazebnik, Y. A., Cordon-Cardo, C., and Lowe, S. W. (2001) Inactivation of the apoptosis effector Apaf-1 in malignant melanoma. *Nature* **409**, 207-211

76. Eggert, A., Grotzer, M. A., Zuzak, T. J., Wiewrodt, B. R., Ho, R., Ikegaki, N., and Brodeur, G. M. (2001) Resistance to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis in neuroblastoma cells correlates with a loss of caspase-8 expression. *Cancer Res* **61**, 1314-1319

77. Ueki, T., Takeuchi, T., Nishimatsu, H., Kajiwara, T., Moriyama, N., Narita, Y., Kawabe, K., Ueki, K., and Kitamura, T. (2001) Silencing of the caspase-1 gene occurs in murine and human renal cancer cells and causes solid tumor growth in vivo. *Int J Cancer* **91**, 673-679

78. Graff, J. R., Gabrielson, E., Fujii, H., Baylin, S. B., and Herman, J. G. (2000) Methylation patterns of the E-cadherin 5' CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. *J Biol Chem* **275**, 2727-2732

79. Arapshian, A., Kuppumbatti, Y. S., and Mira-y-Lopez, R. (2000) Methylation of conserved CpG sites neighboring the beta retinoic acid response element may mediate

retinoic acid receptor beta gene silencing in MCF-7 breast cancer cells. *Oncogene* **19**, 4066-4070

80. Deng, G., Song, G. A., Pong, E., Sleisenger, M., and Kim, Y. S. (2004) Promoter methylation inhibits APC gene expression by causing changes in chromatin conformation and interfering with the binding of transcription factor CCAAT-binding factor. *Cancer Res* **64**, 2692-2698

81. Sathyanarayana, U. G., Maruyama, R., Padar, A., Suzuki, M., Bondaruk, J., Sagalowsky, A., Minna, J. D., Frenkel, E. P., Grossman, H. B., Czerniak, B., and Gazdar, A. F. (2004) Molecular detection of noninvasive and invasive bladder tumor tissues and exfoliated cells by aberrant promoter methylation of laminin-5 encoding genes. *Cancer Res* **64**, 1425-1430

82. Sathyanarayana, U. G., Padar, A., Huang, C. X., Suzuki, M., Shigematsu, H., Bekele, B. N., and Gazdar, A. F. (2003) Aberrant promoter methylation and silencing of laminin-5-encoding genes in breast carcinoma. *Clin Cancer Res* **9**, 6389-6394

83. Sathyanarayana, U. G., Padar, A., Suzuki, M., Maruyama, R., Shigematsu, H., Hsieh, J. T., Frenkel, E. P., and Gazdar, A. F. (2003) Aberrant promoter methylation of laminin-5-encoding genes in prostate cancers and its relationship to clinicopathological features. *Clin Cancer Res* **9**, 6395-6400

84. Sathyanarayana, U. G., Toyooka, S., Padar, A., Takahashi, T., Brambilla, E., Minna, J. D., and Gazdar, A. F. (2003) Epigenetic inactivation of laminin-5-encoding genes in lung cancers. *Clin Cancer Res* **9**, 2665-2672

85. Turker, M. S. (1999) The establishment and maintenance of DNA methylation patterns in mouse somatic cells. *Semin Cancer Biol* **9**, 329-337

86. Turker, M. S., and Bestor, T. H. (1997) Formation of methylation patterns in the mammalian genome. *Mutat Res* **386**, 119-130

87. Butcher, D. T., Mancini-DiNardo, D. N., Archer, T. K., and Rodenhiser, D. I. (2004) DNA binding sites for putative methylation boundaries in the unmethylated region of the BRCA1 promoter. *Int J Cancer* **111**, 669-678

88. Turker, M. S., Mummaneni, P., and Bishop, P. L. (1991) Region- and cell type-specific de novo DNA methylation in cultured mammalian cells. *Somat Cell Mol Genet* **17**, 151-157

89. Turker, M. S., Mummaneni, P., and Cooper, G. E. (1994) The mouse APRT gene as a model for studying epigenetic gene inactivation. *Adv Exp Med Biol* **370**, 647-652

90. Mummaneni, P., Bishop, P. L., and Turker, M. S. (1993) A cis-acting element accounts for a conserved methylation pattern upstream of the mouse adenine phosphoribosyltransferase gene. *J Biol Chem* **268**, 552-558

91. Mummaneni, P., Walker, K. A., Bishop, P. L., and Turker, M. S. (1995) Epigenetic gene inactivation induced by a cis-acting methylation center. *J Biol Chem* **270**, 788-792

92. Feltus, F. A., Lee, E. K., Costello, J. F., Plass, C., and Vertino, P. M. (2003) Predicting aberrant CpG island methylation. *Proc Natl Acad Sci U S A* **100**, 12253-12258

93. Hayashi, H., Tokuda, A., and Udaka, K. (1960) Biochemical study of cellular antigen-antibody reaction in tissue culture. I. Activation and release of a protease. *J. Exp. Med.* **112**, 237-247

94. Keppler, D. (2006) Towards novel anti-cancer strategies based on cystatin function. *Cancer Lett* **235**, 159-176

95. Dubin, G. (2005) Proteinaceous cysteine protease inhibitors. *Cell. Mol. Life Sci.* **62**, 653-669

96. Katunuma, N., Towatari, T., Kominami, E., Hashida, S., Takio, K., and Titani, K. (1981) Rat liver thiol porteinases: cathepsin B, cathepsin H, and catepsin L. *Acta Biol. Med. Ger.* **40**, 1419-1425

97. Takio, K., Towatari, T., Katunuma, N., Teller, D., and Titani, K. (1983) Homology of amino acid sequences of rat liver cathepsins B and H with that of papain. *Proc. Natl Acad. Sci. USA* **80**, 3666-3670

98. Levicar, N., Strojnik, T., Kos, J., Dewey, R., Pilkington, G., and Lah, T. (2002) Lysosomal enzymes, cathepsins in brain tumour invasion. *J. of Neuro-Oncol.* **58**, 21-32

99. Kos, J., and Lah, T. (1998) Cysteine proteinases and their endogenous inhibitors: Target proteins for prognosis, diagnosis, and therapy in cancer (Review). *Oncol Rep* **5**, 1349-1361

100. Kopitar-Jerala, N. (2006) The role of cystatins in cells of the immune system. *FEBS Letters* **580**, 6295-6301

101. Rawlings, N., Tolle, D., and Barrett, A. (2004) Evolutionary families of peptidase inhibitors. *Biochem J.* **378**, 705-716

102. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J., and Turk, V. (1988) The 2.0 A X-ray crystal structure of chicken egg white

cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **7**, 2593-2599

103. Sotiropoulou, G., Anisowicz, A., and Sager, R. (1997) Identification, cloning, and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer. *J Biol Chem* **272**, 903-910

104. Ni, J., Abrahamson, M., Zhang, M., Fernandez, M. A., Grubb, A., Su, J., Yu, G. L., Li, Y., Parmelee, D., Xing, L., Coleman, T. A., Gentz, S., Thotakura, R., Nguyen, N., Hesselberg, M., and Gentz, R. (1997) Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. *J Biol Chem* **272**, 10853-10858

105. Vigneswaran, N., Wu, J., Nagaraj, N., James, R., Zeeuwen, P., and Zacharias, W. (2006) Silencing of cystatin M in metastatic oral cancer cell line MDA-686Ln by siRNA increase cysteine proteinases and legumain activities, cell proliferation and in vitro invasion. *Life Sciences* **78**, 898-907

106. Lah, T., and Kos, J. (1998) Cysteine proteinases in cancer progression and their clinical relevance for prognosis. *Biol Chem* **379**, 125-130

107. Kos, J., Werle, B., Lah, T., and Brunner, N. (2000) Cysteine proteinases and their inhibitors in extracellular fluids: markers for diagnosis and prognosis in cancer. *Int J Biol Markers* **15**, 84-89

108. Frosch, B., Berquin, I., Emmert-Buck, M., Moin, K., and Slone, B. (1999) Molecular regulation, membrane association and secretion of tumor cathepsin B. *APMIS* **107**, 28-37

109. Zeeuwen, P. L. (2004) Epidermal differentiation: the role of proteases and their inhibitors. *Eur J Cell Biol* **83**, 761-773

110. Zeeuwen, P. L., van Vlijmen-Willems, I. M., Egami, H., and Schalkwijk, J. (2002) Cystatin M/E expression in inflammatory and neoplastic skin disorders. *Br J Dermatol* **147**, 87-94

111. Zhong, S., Fields, C. R., Su, N., Pan, Y. X., and Robertson, K. D. (2006) Pharmacologic inhibition of epigenetic modifications, coupled with gene expression profiling, reveals novel targets of aberrant DNA methylation and histone deacetylation in lung cancer. *Oncogene*, Epub

112. Kim, T. Y., Zhong, S., Fields, C. R., Kim, J. H., and Robertson, K. D. (2006) Epigenomic profiling reveals novel and frequent targets of aberrant DNA methylation-mediated silencing in malignant glioma. *Cancer Res* **66**, 7490-7501

113. Shridhar, R., Zhang, J., Song, J., Booth, B. A., Kevil, C. G., Sotiropoulou, G., Sloane, B. F., and Keppler, D. (2004) Cystatin M suppresses the malignant phenotype of human MDA-MB-435S cells. *Oncogene* **23**, 2206-2215

114. Rivenbark, A. G., Jones, W. D., and Coleman, W. B. (2006) DNA methylation-dependent silencing of *CST6* in human breast cancer cell lines. *Lab Invest* **86**, 1233-1242

115. Schagdarsurengin, U., Pfeifer, G. P., and Dammann, R. (2006) Frequent epigenetic inactivation of cystatin M in breast carcinoma. *Oncogene*, Epub

116. Zhang, J., Shridhar, R., Dai, Q., Song, J., Barlow, S. C., Yin, L., Sloane, B. F., Miller, F. R., Meschonat, C., Li, B. D., Abreo, F., and Keppler, D. (2004) Cystatin M: A novel candidate tumor suppressor gene for breast cancer. *Cancer Res* **64**, 6957-6964

117. Song, J., Jie, C., Polk, P., Shridhar, R., Clair, T., Zhang, J., Yin, L., and Keppler, D. (2006) The candidate tumor suppressor *CST6* alters the gene expression profile of human breast carcinoma cells: Down-regulation of the potent mitogenic, motogenic, and angiogenic factor autotaxin. *Biochem Biophy Res Comm* **340**, 175-182

118. Srivatsan, E. S., Chakrabarti, R., Zainabadi, K., Pack, S. D., Benyamini, P., Mendonca, M. S., Yang, P. K., Kang, K., Motamedi, D., Sawicki, M. P., Zhuang, Z., Jesudasan, R. A., Bengtsson, U., Sun, C., Roe, B. A., Stanbridge, E. J., Wilczynski, S. P., and Redpath, J. L. (2002) Localization of deletion to a 300 Kb interval of chromosome 11q13 in cervical cancer. *Oncogene* **21**, 5631-5642

119. Chomczynski, P., and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**, 156-159

120. Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193

121. Zhang, L., Miles, M. F., and Aldape, K. D. (2003) A model of molecular interactions on short oligonucleotide microarrays. *Nat Biotechnol* **21**, 818-821

122. Grunau, C., Clark, S. J., and Rosenthal, A. (2001) Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* **29**, E65-65

123. Mank-Seymour, A. R., Murray, T. R., Berkey, K. A., Xiao, L., Kern, S., and Casero, R. A., Jr. (1998) Two active copies of the X-linked gene spermidine/spermine N1-acetyltransferase (SSAT) in a female lung cancer cell line are associated with an increase in sensitivity to an antitumor polyamine analogue. *Clin Cancer Res* **4**, 2003-2008

124. Widschwendter, M., Siegmund, K. D., Muller, H. M., Fiegl, H., Marth, C., Muller-Holzner, E., Jones, P. A., and Laird, P. W. (2004) Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. *Cancer Res* **64**, 3807-3813

125. Salisbury, J. L. (2001) The contribution of epigenetic changes to abnormal centrosomes and genomic instability in breast cancer. *J Mammary Gland Biol Neoplasia* **6**, 203-212

126. Paz, M. F., Fraga, M. F., Avila, S., Guo, M., Pollan, M., Herman, J. G., and Esteller, M. (2003) A systematic profile of DNA methylation in human cancer cell lines. *Cancer Res* **63**, 1114-1121

127. Robert, M. F., Morin, S., Beaulieu, N., Gauthier, F., Chute, I. C., Barsalou, A., and MacLeod, A. R. (2003) DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells. *Nat Genet* **33**, 61-65

128. Weisenberger, D. J., Velicescu, M., Cheng, J. C., Gonzales, F. A., Liang, G., and Jones, P. A. (2004) Role of the DNA methyltransferase variant DNMT3b3 in DNA methylation. *Mol Cancer Res* **2**, 62-72

129. Cameron, E. E., Bachman, K. E., Myohanen, S., Herman, J. G., and Baylin, S. B. (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat Genet* **21**, 103-107

130. Bender, C. M., Pao, M. M., and Jones, P. A. (1998) Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines. *Cancer Res* **58**, 95-101

131. Pompeia, C., Hodge, D. R., Plass, C., Wu, Y. Z., Marquez, V. E., Kelley, J. A., and Farrar, W. L. (2004) Microarray analysis of epigenetic silencing of gene expression in the KAS-6/1 multiple myeloma cell line. *Cancer Res* **64**, 3465-3473

132. Zhang, J., Yu, J., Gu, J., Gao, B. M., Zhao, Y. J., Wang, P., Zhang, H. Y., and De Zhu, J. (2004) A novel protein-DNA interaction involved with the CpG dinucleotide at -30 upstream is linked to the DNA methylation mediated transcription silencing of the MAGE-A1 gene. *Cell Res* **14**, 283-294

133. Shimamoto, T., Ohyashiki, J. H., and Ohyashiki, K. (2005) Methylation of p15 (INK4b) and E-cadherin genes is independently correlated with poor prognosis in acute myeloid leukemia. *Leuk Res* **29**, 653-659

134. Keppler, D. (2005) Towards novel anti-cancer strategies based on cystatin function. *Cancer Lett* **235**, 159-176

135. Friedman, J. B., Brunschwig, E. B., Platzer, P., Wilson, K., and Markowitz, S. D. (2004) C8orf4 is a transforming growth factor b induced transcript downregulated in metastatic colon cancer. *Int J Cancer* **111**, 72-75

136. Benson, J. R. (2004) Role of transforming growth factor beta in breast carcinogenesis. *Lancet Oncol* **5**, 229-239

137. Huang, C., Sloan, E. A., and Boerkoel, C. F. (2003) Chromatin remodeling and human disease. *Curr Opin Genet Dev* **13**, 246-252

138. Chung, C. H., Bernard, P. S., and Perou, C. M. (2002) Molecular portraits and the family tree of cancer. *Nat Genet* **32**, 533-540

139. Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J. C., Lashkari, D., Shalon, D., Myers, T. G., Weinstein, J. N., Botstein, D., and Brown, P. O. (2000) Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* **24**, 227-235

140. Fazzari, M. J., and Greally, J. M. (2004) Epigenomics: Beyond CpG islands. *Nat Rev Genet* **5**, 446-455

141. Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L., and Bird, A. P. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58**, 499-507

142. Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905-914

143. Hendrich, B., and Bird, A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* **18**, 6538-6547

144. Nan, X., Tate, P., Li, E., and Bird, A. (1996) DNA methylation specifies chromosomal localization of MeCP2. *Mol Cell Biol* **16**, 414-421

145. Boyes, J., and Bird, A. (1991) DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64**, 1123-1134

146. Nan, X., Campoy, F. J., and Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471-481

147. Lux, W., Klobeck, H. G., Daniel, P. B., Costa, M., Medcalf, R. L., and Schleuning, W. D. (2005) In vivo and in vitro analysis of the human tissue-type plasminogen activator gene promoter in neuroblastomal cell lines: evidence for a functional upstream kappaB element. *J Thromb Haemost* **3**, 1009-1017

148. Bird, A. P., and Wolffe, A. P. (1999) Methylation-induced repression--belts, braces, and chromatin. *Cell* **99**, 451-454

149. Kass, S. U., Pruss, D., and Wolffe, A. P. (1997) How does DNA methylation repress transcription? *Trends Genet* **13**, 444-449

150. Prendergast, G. C., and Ziff, E. B. (1991) Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region. *Science* **251**, 186-189

151. Nan, X., Meehan, R. R., and Bird, A. (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res* **21**, 4886-4892

152. Abrahamson, M. (1994) Cystatins. *Methods Enzymol* **244**, 685-700

153. Yano, M., Hirai, K., Naito, Z., Yokoyama, M., Ishiwata, T., Shiraki, Y., Inokuchi, M., and Asano, G. (2001) Expression of cathepsin B and cystatin C in human breast cancer. *Surg Today* **31**, 385-389

154. Foekens, J. A., Kos, J., Peters, H. A., Krasovec, M., Look, M. P., Cimerman, N., Meijer-van Gelder, M. E., Henzen-Logmans, S. C., van Putten, W. L., and Klijn, J. G. (1998) Prognostic significance of cathepsins B and L in primary human breast cancer. *J Clin Oncol* **16**, 1013-1021

155. Bervar, A., Zajc, I., Sever, N., Katunuma, N., Sloane, B. F., and Lah, T. T. (2003) Invasiveness of transformed human breast epithelial cell lines is related to cathepsin B and inhibited by cysteine proteinase inhibitors. *Biol Chem* **384**, 447-455

156. Rivenbark, A. G., Jones, W. D., Risher, J. D., and Coleman, W. B. (2006) DNA methylation-dependent epigenetic regulation of gene expression in MCF-7 breast cancer cells. *Epigenetics* **1**, 32-44

157. Cailleau, R., Olive, M., and Cruciger, Q. V. (1978) Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. *In Vitro* **14**, 911-915

158. Hackett, A. J., Smith, H. S., Springer, E. L., Owens, R. B., Nelson-Rees, W. A., Riggs, J. L., and Gardner, M. B. (1977) Two syngeneic cell lines from human breast tissue: The aneuploid mammary epithelial (Hs578T) and the diploid myoepithelial (Hs578Bst) cell lines. *J Natl Cancer Inst* **58**, 1795-1806

159. Soule, H. D., Vazguez, J., Long, A., Albert, S., and Brennan, M. (1973) A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst* **51**, 1409-1416

160. Lasfargues, E. Y., and Ozzello, L. (1958) Cultivation of human breast carcinomas. *J Natl Cancer Inst* **21**, 1131-1147

161. Harder, A., Rosche, M., Reuss, D. E., Holtkamp, N., Uhlmann, K., Friedrich, R., Mautner, V. F., and von Deimling, A. (2004) Methylation analysis of the neurofibromatosis type 1 (NF1) promoter in peripheral nerve sheath tumours. *Eur J Cancer* **40**, 2820-2828

162. Zhang, Q., Rubenstein, J. N., Jang, T. L., Pins, M., Javonovic, B., Yang, X., Kim, S. J., Park, I., and Lee, C. (2005) Insensitivity to transforming growth factor-beta results from promoter methylation of cognate receptors in human prostate cancer cells (LNCaP). *Mol Endocrinol* **19**, 2390-2399

163. Cromer, A., Carles, A., Millon, R., Ganguli, G., Chalmel, F., Lemaire, F., Young, J., Dembele, D., Thibault, C., Muller, D., Poch, O., Abecassis, J., and Wasylyk, B. (2004) Identification of genes associated with tumorigenesis and metastatic potential of hypopharyngeal cancer by microarray analysis. *Oncogene* **23**, 2484-2498

164. Szyf, M., Tanigawa, G., and McCarthy, P. L., Jr. (1990) A DNA signal from the Thy-1 gene defines de novo methylation patterns in embryonic stem cells. *Mol Cell Biol* **10**, 4396-4400

165. Hasse, A., and Schulz, W. A. (1994) Enhancement of reporter gene de novo methylation by DNA fragments from the alpha-fetoprotein control region. *J Biol Chem* **269**, 1821-1826

166. Bock, C., Paulsen, M., Tierling, S., Mikeska, T., Lengauer, T., and Walter, J. (2006) CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS Genet* **2**, e26