July 2016

DNA-Based Epigenetic Changes in Recurrent and Tamoxifen-Resistant Breast Cancer

Kristin E. Williams
University of Massachusetts Amherst

DNA-BASED EPIGENETIC CHANGES IN RECURRENT AND TAMOXIFEN-RESISTANT BREAST CANCER

A Dissertation Presented

by

KRISTIN E. WILLIAMS

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

MAY 2016

Molecular and Cellular Biology

DNA-BASED EPIGENETIC CHANGES IN RECURRENT AND TAMOXIFEN-RESISTANT BREAST CANCER

A Dissertation Presented

by

KRISTIN E. WILLIAMS

Approved as to style and content by:

_____
Kathleen F. Arcaro, Chair

_____
Brian T. Pentecost, Member

_____
Sallie Smith Schneider, Member

_____
D. Joseph Jerry, Member

_____
Dominique Alfandari, Director
Molecular and Cellular Biology

DEDICATION

To Kara

May you always stay curious and creative
wherever your beautiful mind may take you.

ACKNOWLEDGMENTS

My advisor Kathleen Arcaro
Thank you for all your help and guidance over the years and for pushing me to design my own project and obtain my own funding. You really helped me grow as both a scientist and a person.

Brian Pentecost
I appreciate you answering all of my questions and helping me think like a scientist. Putting the last chapter together wouldn't have happened without your support.

My Committee Members
Your advice and suggestions helped shape my dissertation into what it is today. Thank you so much for everything.

My Labmates – Katerina Fagan-Solis, Eva Browne, Beth Punska, & Stephanie Zimmers
I feel very lucky to have become friends with all of you over the years. Thank you so much for everything you've helped me with - from experiments to life advice!

Dr. Christopher Otis and Dr. Rahul Jawale
I am sincerely grateful for all your help getting the tumor study off the ground and for your pathology expertise. I couldn't have completed this without your assistance.

Adam
I am so fortunate that you're my best friend. Thank you for being there through the thick and the thin. I love you.

Mom and Dad
Thank you for supporting my dreams and for always being there for me.

ABSTRACT

DNA-BASED EPIGENETIC CHANGES IN RECURRENT AND TAMOXIFEN-RESISTANT BREAST CANCER

MAY 2016

KRISTIN E. WILLIAMS, A.B., MOUNT HOLYOKE COLLEGE

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Kathleen F. Arcaro

Roughly two-thirds of all breast cancers are Estrogen Receptor α (ER)-positive and can be treated with an anti-estrogen such as Tamoxifen, however resistance occurs in 33% of women who take the drug for more than 5 years. In addition to this acquired antiestrogen resistance, *de novo*- or intrinsic-resistance occurs primarily in ER-negative tumors but also occasionally in ER-positive tumors. Aberrant DNA promoter methylation, a major epigenetic mechanism by which gene expression is altered in cancer, is thought to play a role in this resistance. To date, few studies have examined promoter methylation and Tamoxifen resistance in breast cancer. Of the studies conducted, one detected drug-specific promoter methylation and gene expression profiles in an ER-positive, Tamoxifen-selected MCF-7 derivative cell line. However, studies using both ER-positive and –negative, Tamoxifen-selected cell lines have not been described until now.

To develop an understanding of Tamoxifen-resistance and identify novel pathways and targets of aberrant methylation, I first analyzed two Tamoxifen-resistant clones of MCF-7, one that retained expression of ER (TMX2-11) and one that lost expression of the gene (TMX2-28) after 6-months of Tamoxifen treatment, by Illumina HumanMethylation450 BeadChip (HM450BC). I found that prolonged treatment with Tamoxifen induced hypermethylation and hypomethylation throughout the genome. Compared to MCF-7, the ER-positive line, TMX2-11 had 4,000 hypermethylated sites, while the ER-negative line, TMX2-28 had over 33,000. Analysis of CpG sites in both TMX2-11 and TMX2-28 revealed that the two Tamoxifen-selected lines share 3,000 hypermethylated CpG sites with 21% of those sites being located in the promoter region. Promoter methylation and expression of two genes, MAGED1 and ZNF350, in both Tamoxifen-resistant cell lines demonstrated cell line-specific responses to treatment with 5-aza-2'deoxycitidine (5-Aza). Sixteen additional genes involved in signal transduction, cell adhesion, transcriptional repression, inflammatory response, cell proliferation and hormone response were chosen for further analysis based on their shared hypermethylation or their reduced expression in TMX2-28 as detected in a previously completed expression array. Five genes, RORA, THBS1, CAV2, TGFβ2, and BMP2 had decreased expression in TMX2-28, but not TMX2-11 as compared to MCF-7, and 5-Aza

increased expression of the genes. This indicates that Tamoxifen is affecting a set of genes similarly in both the ER-positive and -negative breast cancer cell lines, however overall methylation changes are more pronounced in the ER-negative line. Our data as well as others suggest that DNA methylation may be contributing to Tamoxifen-resistance.

I hypothesized that both ER-positive and ER-negative second human breast tumors occurring after anti-estrogen treatment would be hypermethylated. I characterized the methylation profiles of 70 human breast tumor samples using the HM450BC. These data confirm previous findings that ER-positive breast tumors have more hypermethylated CpG sites than ER-negative tumors. Stratification of the tumors by ER-positive first and second tumor sets shows that methylation is greater in first tumors.. Additionally, I saw that first tumors from ipsilateral pairs had higher methylation than the second tumors; in contrast, second tumors from contralateral pairs had higher methylation than in the first tumor. These data, together with the fact that tumor progression is associated with an increase in methylation, are consistent with the prediction that ipsilateral, not contralateral, tumors are more likely to be a true recurrence.

Pathway analysis was conducted to provide insight into biomarkers associated with tumors that recur. Two pathways, 'homophilic cell adhesion via plasma membrane adhesion molecules' and 'cell fate commitment', were selected for further analysis. ER-positive first tumors that recurred as either ER-positive or ER-negative compared with non-recurrent tumors shared hypermethylated genes in the homophilic cell adhesion pathway. ER-positive first tumors that recurred as ER-negative compared with ER-positive first tumors that recurred as ER-positive were associated with a unique set of hypermethylated genes in the cell fate commitment pathway. To examine the association of methylation changes in my tumor data set with breast cancer patient survival data, Kaplan-Meier plots were created using TGCA breast cancer data available online. Expression of the genes only hypermethylated in each individual comparison group in the homophilic cell adhesion pathway was linked to overall survival. These data suggest that the genes hypermethylated only in ER-positive tumors recurring as ER-negative are a potential signature for poor survival.

The underlying mechanisms of anti-estrogen resistance are poorly understood. Variable responses to breast cancer therapy highlights the need for biomarkers that can effectively guide treatment. The findings presented here underscore the potential use of breast tumor stratification based on methylation biomarkers in guiding treatment.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS v

ABSTRACT vi

LIST OF FIGURES xiv

CHAPTER

1. INTRODUCTION 1

The Biology of Breast Cancer 1
Molecular Subtypes of Breast Cancer 2
The Estrogen Receptor 3
Breast Cancer Treatments 5
DNA Methylation 7
Tamoxifen Resistant Breast Cancer Cell Line Model 10
HumanMethylation450 BeadChip 10

2. HIGH DENSITY ARRAY ANALYSIS OF DNA METHYLATION IN TAMOXIFEN-RESISTANT BREAST CANCER CELL LINES 22

Introduction 22
Materials and Methods 24
Cell Culture, RNA and DNA Purification 24
Illumina HumanMethylation450 BeadChip (HM450BC) 24
5-Aza-2'deoxycitidine Treatment of Cells 25
Pyrosequencing 26
Quantitative Real Time Reverse Transcriptase-PCR (qRT-PCR) 27
Data Analysis 27
Results 28
Tamoxifen-selection results in extensive changes in DNA methylation 28
Differentially methylated CpG (dmCpG) sites in the Tamoxifen-selected lines are primarily hypermethylated 29
dmCpG sites are found primarily in the intergenic, body and promoter regions 31
Sensory perception is among the top pathways affected by Tamoxifen selection 32
Promoter methylation of ZNF350 and MAGED1 33
Treatment with 5-Aza reverses DNA methylation in TMX2-28 35

Decreasing methylation results in increased expression of ZNF350 and MAGED1 in TMX2-28 35
Discussion 36

3. FURTHER CHARACTERIZATION OF GENES ABERRANTLY METHYLATED AND DIFFERENTIALLY EXPRESSED IN TAMOXIFEN-RESISTANT CELL LINES 50

Introduction 50
Materials and Methods 54
Cell Culture, RNA and DNA Purification 54
5-Aza-2'deoxycitidine (5-Aza) and Tamoxifen Treatment of Cells 54
S-adenosylmethionine (SAM) Treatment of MDA-MB-231 Cells 55
Quantitative Real Time Reverse Transcriptase-PCR (qRT-PCR) 55
Pyrosequencing 56
Data Analysis 57
Results 57
AKAP12, REG1A, and HDAC9 have low expression in MCF-7 and Tamoxifen-resistant cell lines 58
Decreased expression of GREB1 and increased expression of LZTS1 58
Expression of TGFβ2 and CAV2 increases in TMX2-28 cells with 5-Aza treatment 59
RORA is hypermethylated in Tamoxifen-resistant lines 60
THBS1 expression and methylation are affected by 5-Aza treatment in TMX2-28 cells 60
Expression and methylation of BMP2 are increased in TMX2-28 cells 61
5-Aza treatment does not result in re-expression of ERα in TMX2-28 or MDA-MB-231 62
5-Aza treatment results in decreased cell proliferation of TMX2-28 62
Discussion 64

4. HIGH-DENSITY ARRAY ANALYSIS OF DNA METHYLATION IN HUMAN BREAST TUMOR SAMPLES 87

Introduction 87
Materials and Methods 90
Human Tissue 91
Immunohistochemistry (IHC) 92
DNA Purification 93
Illumina HumanMethylation450 BeadChip (HM450BC) 94
Data Analysis 95
Results 96

ER-positive versus ER-negative tumors: greater hypermethylation in ER-positive tumors 100
ER-status and differential hypermethylation of 35 target CpGs: confirmation of previously identified dmCpGs 101
ER-associated trends in methylation within groups of first, and groups of recurrent tumors: methylation differences between ER-positive and ER-negative tumors are reduced in second tumors 102
Differential methylation based on both occurrence and ER-status of primary and recurrent tumors: ER-negative second tumors from women who had ER-positive primary tumors have greater hypermethylation 103
Differential methylation based on occurrence and side: second tumors from ipsilateral pairs have the greater hypermethylation 105
Differential methylation in ER-positive and ER-negative second tumors based on side: ER-negative tumors from contralateral pairs have greater hypermethylation 105
Further analysis of paired tumors stratified by ER-status of primary and recurrent and by location: greater methylation among second tumors from ipsilateral pairs and first tumors from contralateral pairs 106
Statistical differences in methylation by subgroups 107
Functional genomic location of methylation among comparison groups 109
Thirty-five previously identified CpG sites remain hypermethylated with respect to ER-status when groups are stratified by first and second tumor occurrence 110
Pathway Analysis of differentially methylated genes among ER-positive first tumors and non-recurrent tumors 111
Discussion 118
Future Directions 125

APPENDICES

A. HYPER AND HYPOMETHYLATED PATHWAYS SHARED BY TMX2-11 AND TMX2-28 182
B. HYPERMETHYLATED PATHWAYS SHARED BY TMX2-11 AND TMX2-28 IN THE PROMOTER 193
C. HYPERMETHYLATED PATHWAYS SHARED BY TMX2-11 AND TMX2-28 IN THE GENE BODY 197
D. GENES WITH TWO OR MORE HYPERMETHYLATED CPG SITES IN THE PROMOTER IN BOTH TMX2-11 AND TMX2-28 IN ORDER OF DECREASING NUMBER OF HYPERMETHYLATED CPGS 200
E. COMPARISON OF HM450BC AND PYROSEQUENCING DATA SHOWS CORRELATION BETWEEN ANALYSIS METHODS 203

F. METHYLATION OF RORA AND THBS1 ACROSS CPG SITES INTERROGATED BY PYROSEQUENCING 204
G. HM450BC METHYLATION OF THE ER GENE 205
H. CYTOTOXCICITY ASSAY TO DETERMINE 5-AZA CONCENTRATION 206
I. CYTOTOXCICITY ASSAY TO DETERMINE SAM CONCENTRATION 207
J. INDIVIDUAL PATIENT CHARACTERISTICS 208
K. DIAGRAM OF TISSUE MICRODISSECTION FOR DNA PURIFICATION 212
L. 35 HYPERMETHYLATED AND HYPOMETHYLATED CPG SITES AND ASSOCIATED GENES 213
M. SHARED CPG SITES BETWEEN HYPERMETHYLATED AND HYPOMETHYLATED PATHWAYS 215
N. SHARED GENES BETWEEN HYPERMETHYLATED AND HYPOMETHYLATED PATHWAYS 224
O. PATHWAYS SHARED BY A1, C1, AND NR GROUPS 231

BIBLIOGRAPHY 232

LIST OF TABLES

Table Page

1.1. Commonly used chemotherapeutics in breast Cancer treatment 12

2.1 CpG methylation changes in Tamoxifen-resistant cell lines as compared to the parental line 40

2.2 Hyper- and Hypomethylated Pathways shared by TMX2-11 and TMX2-28 41

2.3 Hypermethylated Pathways in the promoter and body regions shared by TMX2-11 and TMX2-28 42

2.4 Comparison of ZNF350 methylation by HumanMethylation450 BeadChip and Pyrosequencing in Tamoxifen-resistant and Parental Cell Lines 43

3.1 Summary of Genes and their functions 71

3.2 qRT-PCR Primers 72

3.3 Pyrosequencing Primers 73

3.4 Overview of methylation and expression data from genes selected based on HM450BC hypermethylation in TMX2-11 and TMX2-28 74

3.5 Overview of methylation and mRNA expression of genes underexpressed in TMX2-28 from Affymetrix array 75

4.1 DNA concentrations and bisulfite conversion controls for all 86 tumors 127

4.2 Patient and tumor characteristics 131

4.3 Sample group re-labeling for readability 133

4.4 Patient and tumor demographics by groups A, B, C and non-recurrent ER-positive 134

4.5 All tumors separated by first recurrence 137

4.6 Paired tumors separated by occurrence and ER-status 138

4.7 Group sizes[1] for first and second tumors stratified by location 139

4.8 Group sizes for all ER-positive and ER-negative tumors stratified by occurrence and location 140

4.9 Group sizes for paired tumors stratified by ER-status, occurrence and location..... 141

4.10 Hyper-and hypomethylation of CpG sites using Group A1 as reference 140

4.11 Hyper-and hypomethylation of CpG sites using Group B1 as reference 141

4.12 Hyper-and hypomethylation of CpG sites using Group A1 as reference 142

4.13 Hypermethylated CpG sites within gene regions..... 145

4.14 Hypomethylated CpG sites within gene regions..... 146

4.15 Differentially methylated CpG sites and genes within groups 147

4.16 Top 20 hypermethylated pathways in each comparison group..... 148

4.17 Top 20 hypomethylated pathways in each group 151

4.18 Homophilic cell adhesion via plasma membrane pathway genes..... 154

4.19 Association of survival with genes identified in ER-positive first tumors from pairs with ER-negative second tumors..... 156

4.20 Cell fate commitment genes..... 157

4.21 Association of survival with genes identified in ER-positive first tumors with either ER-negative or ER-positive second tumors..... 159

4.22 Genes shared in both cell fate commitment and neuron fate commitment pathways 160

LIST OF FIGURES

Figure Page

1.1 Structure of the human breast.. 13

1.2 Mammary epithelial hierarchy and the molecular subtypes of human breast cancer..... 14

1.3 Chemical structure of estrogens. 17β-estradiol (E2), estriol (E3), and estrone (E1). 15

1.4 Structure of the Estrogen Receptor. 16

1.5 Activation of estrogen responsive genes controlled by estrogen molecule binding to the estrogen receptor.. 17

1.6 Mechanism of Tamoxifen (antiestrogen) in an ER-positive breast cell. 18

1.7 Promoter methylation plays a key role in regulating transcription. 19

1.8 Methylation and nucleosome condensation. 20

1.9 Differences in ER-α gene and protein expression occur across Tamoxifen-selected cell lines.. 21

2.1. Visual representation of DNA methylation among the breast cancer cell lines. 44

2.2. Scatter plots indicate genome-wide methylation changes in Tamoxifen-resistant lines compared with the parental. 45

2.3. Location of aberrantly methylated CpG sites shared between TMX2-11 and TMX2-28..... 46

2.4. ZNF350 and MAGED1 are differentially methylated in Tamoxifen resistant cells. 47

2.5. CpG site methylation of ZNF350 and MAGED1 in Tamoxifen-resistant and parental cell lines..... 48

2.6. Comparison of gene expression and promoter methylation in ZNF350 and MAGED1. 49

3.1 Low levels of AKAP12, REG1A, and HDAC9 mRNA are expressed in Tamoxifen-resistant cell lines. 76

3.2. 5-Aza treatment does not affect mRNA expression of TCF12, SSRP1, GFI1, GREB1 and LZTS1. 78

3.3 mRNA expression of TGFβ2 and CAV2 increases after 5-Aza treatment in TMX2-28 cells. 79

3.4 RORA promoter methylation, but not mRNA expression increases after 5-Aza treatment in TMX2-28. 81

3.5. mRNA expression and promoter methylation of THBS1 are affected by 5-Aza treatment. 82

3.6. mRNA expression and promoter methylation of BMP2 are increased in TMX2-28 cells. 83

3.7 ERα is not re-expressed in TMX2-28 or MDA-MB-231 cells after treatment with 5-Aza. 84

3.8 5-Aza treatment decreased cell proliferation in TMX2-28. 85

3.9 mRNA expression and DNA methylation of uPA in MDA-MB-231 cells is unaffected by S-adenosylmethionine (SAM). 86

4.1. Flow chart explaining tumor selection and criteria for analysis by HM450BC. 161

4.2. Methylation of ER-positive breast tumors compared with ER-negative breast tumors. 162

4.3. Analysis of 35 CpG sites previously shown to be differentially methylated between ER-positive and ER-negative breast tumors. 163

4.4. Methylation analysis of 23 CpG sites previously identified as hypermethylated in ER-positive tumors. 164

4.5. Methylation analysis of 12 CpG sites previously identified as hypermethylated in ER-negative tumors. 165

4.6. Methylation analysis of tumors stratified by both ER-status and occurrence. 166

4.7. Visual analysis of paired tumors stratified by ER-status of the primary and second. 167

4.8. Comparison of first and second tumors stratified by side of second tumor occurrence. 168

4.9. Analysis of second tumors stratified by ER-status and side of second tumor occurrence. 169

4.10. Analysis of paired tumors from Group A (ER-positive primary and ER-positive second) stratified by side of location. 170

4.11. Analysis of paired tumors from Group C (ER-positive primary and ER-negative second) stratified by side of location. 171

4.12 CpG site hypermethylation within gene regions. 172

4.13 CpG site hypomethylation within gene regions. 173

4.14. Group methylation analysis of twenty-three CpG sites previously identified as hypermethylated in ER-positive breast tumors. 174

4.15. Group methylation analysis of twelve CpG sites previously identified as hypermethylated in ER-negative breast tumors. 175

4.16. Ancestor chart for homophilic cell adhesion via plasma membrane pathway. 176

4.17. Ancestor chart for cell fate commitment pathway. 177

4.20. Ancestor chart for neuron fate commitment. 180

4.21. SATB2 is differentially methylated between A1, C1 and NR. 181

CHAPTER 1

INTRODUCTION

The Biology of Breast Cancer

The breast is composed of three basic structures, lobules, ducts, and stroma (Figure 1.1) [1, 2]. Lobules consist of small, hollow, grape-like structures called alveoli, which are responsible for the production of milk during lactation. A grouping of 15-20 lobules is called a lobe. Lobules are connected by ducts, which assist in carrying milk to the nipple where it is discharged during lactation. Lobules are composed of both luminal cells, the secretory cells that line the inside of the duct, and myoepithelial cells, the contractile cells that lie behind the luminal cells [3-5]. Ducts also contain both luminal and myoepithelial cells and recent research has suggested the presence of mammary stem cells (MaSCs), which give rise to both the luminal and myoepithelial cells of the breast through a series of progenitor cell intermediates (Figure 1.2) [3]. Both cell layers sit on top of a basement membrane, which is surrounded by adipocytes and fibroblasts that comprise the stroma of the breast. Stroma, the fatty and connective tissue, encloses the areas between the lobules and ducts.

Breast cancer, a disease characterized by abnormal and uncontrolled cell growth, can occur in either the cells lining the lobules (lobular cancers) or those lining the ducts (ductal cancers). The two most common types of breast cancer occur in the ducts. One in five new cases of non-invasive cancer is diagnosed as Ductal Carcinoma *In Situ* (DCIS). This type of breast cancer is highly treatable and many diagnoses are a result of advanced

screening methods, such as mammograms [1]. Infiltrating Ductal Carcinoma (IDC), the most common type of invasive breast cancer, makes up between 70-80% of all breast cancer cases, and if left untreated it is likely metastasis will occur [1].

It was estimated that in 2015 around 232,000 new cases of breast cancer occurred and of those affected approximately 40,000 were expected to die from the disease, making it the second leading cause of all cancer deaths in women [1].

Molecular Subtypes of Breast Cancer

Through the use of high-throughput gene expression analyses, breast cancer has been classified into five subtypes: Luminal A, Luminal B, HER2-enriched, Basal-like and Claudin-low (see Figure 1.2) [3, 6]. These classifications are based on specific markers either present or absent from the tumor cells. Luminal A tumors have the best prognosis and are least likely to see patient relapse. Tumors from this subgroup are estrogen receptor (ER) & progesterone receptor (PR)-positive, human epidermal growth factor receptor 2 (HER2) negative, and express low levels of Ki67, a proliferation marker associated with higher grade tumors at increasing levels [7]. Similarly to Luminal A, Luminal B tumors are also ER and PR-positive, but can either be HER2 negative and Ki67 high or are positive for HER2. Additionally, Luminal B tumors have a higher expression of genes involved in proliferation as compared with Luminal A tumors and are classified as having a high risk of recurrence. Luminal tumors stain positive for the cytokeratins (CK) 8 and 18. The HER2 subgroup consists of tumors that are HER2 positive, do not express basal genes, have low expression of luminal genes and are highly proliferative. Basal-like tumors are ER and PR-negative and HER2-negative and a

combined data set of 400 tumors showed that the basal-like subset had the highest percentage of triple negative tumors at 49% when compared with the other subsets [6]. They are also defined as having positive staining for CK5/6 and EGFR. The Claudin-low subgroup is the least prevalent subtype; 12-14% of all breast tumors are classified as Claudin-low. In addition to having low expression of Claudin, a tight junction component, it is most similar in terms of hormone expression to the Basal-like subgroup with most tumors having the ER & PR-negative/HER2-negative phenotype. However Claudin-low tumors differ from Basal-like tumors in that have low expression of a cluster of cell-cell adhesion proteins and high expression of a cluster enriched for immune system response genes [6]. Both Basal-like and Claudin-low tumors are classified as having a high risk of recurrence.

The Estrogen Receptor

Estrogen Receptor-α (ERα) is encoded by the *ESR1* gene and plays a critical role in cell proliferation during both a woman's normal menstrual cycle and preparing the breasts for lactation during pregnancy. Natural estrogens, steroids produced and secreted by the ovaries, travels through the blood stream and binds to ERα in the breast and uterus [8]. Aromatase, the enzyme responsible for conversion of androstenedione and testosterone, which are secreted by the adrenal zona fasculata and the ovarian stroma, is present in high concentrations in the ovaries of pre-menopausal, but not post-menopausal women [9]. In both pre- and post-menopausal women, aromatase is also present in subcutaneous fat, the liver, muscle and the normal breast [10, 11]. There are three major forms of estrogen produced by the ovaries: estrone (E1), which is converted from

androstenedione, 17β-estradiol (E2), converted from testosterone, and estriol (E3) (Figure 1.3A) [12]. ERα is located within the nucleus of the cell and when bound to estrogen, induces a conformational change in the protein allowing activation of ERα through dimerization (Figure 1.4) [13]. This activation results in transcription of genes controlled by the Estrogen Response Element (ERE), many of which are involved in cell proliferation (Figure 1.5) [8, 12, 14]. In the normal human non-lactating breast, only 7% of cells are ERα -positive, with the majority concentrated in lobules [15].

In contrast to ERα, the mechanism and role of Estrogen Receptor-β (ERβ) in breast cancer, has remained a mystery. ERβ is encoded in the *ESR2* gene and the protein is not expressed endogenously at detectable levels in any breast cancer cell lines [16]. Hypermethylation of the ERβ promoter has been shown to downregulate mRNA expression of the gene in both breast cancer cell lines and tumors, specifically in those with an unfavorable prognosis [16]. Treatment of the cell lines with demethylating agents resulted in increased mRNA expression of ERβ, suggesting a potential biomarker for poor prognosis breast tumors [16-18].

In breast cancer, ERα is expressed in the tumor and is a determining factor in establishing a treatment regime, however PR expression is used in conjunction as a prognostic marker. ER-positive tumors with PR expression have a positive prognostic outcome as opposed to those with no PR expression [19]. The PR, is a transcription factor and recent research has shown that the functional importance of the protein in disease outcome is the cross-talk between the PR and ER in gene expression [19]. This cross-talk is believed to regulate the expression of a group of genes that are associated with better disease prognosis [19].

Breast Cancer Treatments

After a breast cancer diagnosis there are multiple options for treatment of the tumor including radiation, chemotherapy, hormonal therapy and targeted therapy. The suggested treatment path depends on the molecular subtype, size, and tumor stage. Radiation therapy uses high-energy rays to target cancer cells left in the breast after surgery and is given to both lumpectomy and mastectomy patients [20].

Neoadjuvant therapy is chemotherapy given before tumor excision. It is given to patients with larger tumors with the goal of shrinking the tumor for easier excision. Adjuvant therapy is given post-surgery to prevent reoccurrence of the tumor, it is recommended for low risk tumors with lymph node metastases larger than 2 cm and aggressive tumors that are larger than 0.5 cm with any lymph node metastases [21]. Commonly used chemotherapeutics are highlighted in Table 1. These chemotherapies can be given in alone or combination with another drug [20, 22]. Administration of adjuvant therapy depends on whether the woman was pre-menopausal or post-menopausal at diagnosis. Pre-menopausal women are given the anti-estrogen, Tamoxifen for 5 years whereas post-menopausal women are given either an aromatase inhibitor or Tamoxifen for 5 years [23].

Roughly 70% of all breast cancer cases are ER-positive and a majority are easily treated with, Tamoxifen (4-hydroxytamoxifen). Tamoxifen, a selective estrogen receptor modulator (SERM), has been extensively used to treat breast cancer for over 40 years [24]. A competitive inhibitor of estrogen, Tamoxifen functions in the cell by preventing E2 from binding to the ERα and ultimately blocking cell growth in ERα-positive breast tumors (Figure 1.6) [8]. The binding of Tamoxifen to ERα induces a conformational

change, different from that caused by estrogen, which inhibits the dimerization of the subunits and ultimately leads to inhibition of the protein (Figure 1.4B) [13, 25].

While the majority of women respond well to Tamoxifen, approximately 33% of all patients treated with the drug for more than 5 years have disease recurrence [26]. Determining the mechanisms responsible for this acquired resistance has been the subject of investigation in recent years. Researchers have examined possible mechanisms of resistance including loss of expression or function of ERα, changes in expression of co-repressor and co-activator proteins, increase in growth factor activity, autophagy, stress responses, and more recently epigenetic modifications, such as histone modifications and DNA methylation [25, 27-31]. Additionally, the extent to which the acquired Tamoxifen resistance is due to biological changes in the cells or forced selection of cells resistant to the drug is unclear [29]. Though considerable research has been done in the field of Tamoxifen resistance, the mechanism of resistance is proving to be a complex one and more research is needed to determine additional factors involved.

Additional anti-estrogen drugs are available, but not widely used in treatment regimes. Toremifene (chloro-tamoxifen) and droloxofine (3-hydroxytamoxifen) are not generally given due to their direct similarities to Tamoxifen as derivatives of the drug [32]. Raloxifene, another anti-estrogen, was found to decrease the incidence of invasive breast cancer and is widely given as both an osteoporosis treatment as well as a chemopreventive drug for aggressive breast cancers [32]. Faslodex (Fulvestrant), is given to post-menopausal women due to its ability to act as a pure antagonist to the estrogen receptor, by binding to and degrading it [32].

The aromatase inhibitors (AIs), anastrazol, exemestane, and letrozole, function by blocking aromatase enzyme activity in the cell. As described previously, the ovaries produce aromatase in pre-menopausal women, whereas aromatase in post-menopausal women is created by fat and muscle cells. In both cases estradiol is produced from conversion of testosterone by aromatase [10]. AIs are generally given to post-menopausal women with hormone receptor-positive breast cancer as the primary treatment. Treatment with Tamoxifen has a 33% relapse rate regardless of whether chemotherapy was used and that rate increases if a metastasis is present [10]. AIs have been shown to decrease disease progression and increase survival rates among post-menopausal women as compared with other hormonal therapies [10].

Finally, targeted therapies are available for those tumors that are positive for HER2 and negative for ER and PR. Treatment with trastuzumab (Herceptin), a monoclonal antibody, targets the extracellular component of the HER2 protein and prevents dimerization [33]. This inhibits cell growth, reduces angiogenic factors and induces apoptosis in the cells [33]. In addition to being used for HER2 positive tumors, trastuzumab is also used with metastatic breast cancer as a secondary treatment.

DNA Methylation

Epigenetics is the study of reversible and heritable modifications that affect gene expression, but do not cause changes in the DNA sequence [34]. Non-tumorigenic cells use a host of epigenetic mechanisms such as acetylation, methylation, sumolyation, phosphorylation, and ubiquitination to regulate expression of genes [35]. In normal cells, these transient epigenetic modifications play important roles by communicating to the

cell the need for gene expression during critical times in development, in response to environmental toxins or chemicals, and in aging [36]. However in response to certain environmental factors, the epigenetics of the cell can be altered and ultimately lead to diseases such as cancer. An epigenetic mechanism identified in breast carcinogenesis is DNA methylation.

DNA methylation is the addition of a methyl group (-CH_3) by DNA methyltransferases (DMNTs) to the cytosine of a 5'-**C**ytosine-**p**hosphate-**G**uanine-3' (CpG) site in the DNA. During replication, DNA on the parent strand remains methylated while the newly replicated daughter strand is unmethylated, this is known as hemimethylation and it is important for maintenance of DNA methylation [37]. DNA methylation across the human genome is heterogeneous and segments of unmethylated DNA are mixed with methylated ones [38]. The CpG rich regions of the DNA, called CpG islands, are found in about half of all genes in the human genome and span anywhere from 500-5000 bp [38]. In the last 30 years researchers have discovered that these CpG islands, particularly in promoter regions proximal to the transcriptional start site (TSS), play a critical role in transcriptionally silencing genes (Figure 1.7) [39-41].

Methylation of CpG islands is a highly regulated biological process important for controlling expression of genes that require activation during a specific point in the cell cycle [41]. In normal cells, CpG islands directly upstream of the TSS are unmethylated. Stretches of DNA with unmethylated CpG islands upstream of the TSS are not wound tightly around nucleosomes, leaving the DNA in an open arrangement and accessible to transcription factors (Figure 1.8) [41, 42]. Alternatively, in the case of cancer, CpG sites in promoter regions can be aberrantly methylated, leading to silencing of tumor

suppressor, cell-adhesion, and growth regulatory genes [42]. The hypermethylation of CpG sites in promoters attracts nucleosomes to the DNA, which irreversibly silences genes with the addition of histone methylation markers on histone H3 at lysine residues 9 and 27 (H3K9 and H3K27) or histone H4 at lysine 20 (H4K20), and DNMT3A and DNMT3B to the nucleosome (Figure 1.8) [43]. Treatment with a demethylating agent such as 5-aza-2'-deoxycitidine (5-Aza) has been shown to inhibit the irreversible methylation that occurs in cancer cells [42]. 5-Aza replaces the cytosine that is normally incorporated into the DNA strand, thereby trapping DNMTs on the DNA and halting methylation [42]. Demethylation of CpG islands in cancer cells using agents such as 5-Aza is a promising therapeutic target, as non-cancerous DNA is largely unmethylated and thus unaffected [42].

Until recently, gene silencing by DNA methylation was focused solely on the promoter region, however recent research has shown that the body region of the gene also plays an important role. One recent study by Yang et al. found methylation in the body region correlates with increased gene expression and that treatment with 5-Aza decreased expression of genes overexpressed in colon cancer [44]. The function of body methylation is still unknown, however it has been suggested that it blocks initiation of intergenic promoters, affecting repetitive DNAs, or affects the rate of transcription by forming an ordered structure [44, 45].

In breast cancer, hypermethylation has been found to affect the expression of several genes, such as p16, RARB2, GSTP1, RASSF1, and 14-3-3γ [46]. Therefore, DNA promoter methylation is considered one of the most promising breast cancer

biomarkers, as the changes in promoter methylation are thought to occur early in disease progression and are potentially reversible with treatment.

Tamoxifen Resistant Breast Cancer Cell Line Model

Our laboratory has been studying the changes in gene expression and signaling pathways that occur between ERα-negative and ERα-positive breast cancers for the last 10 years [31, 47, 48]. In order to observe these changes, we utilized the ERα-negative, Tamoxifen-resistant breast cancer cell line, TMX2-28. These cells were cloned from the parent MCF-7 breast cancer cell line, which was maintained in Tamoxifen for six months [49]. In contrast to the parent cell line, the TMX2-28 cells are basal-like in their cytokeratin-gene expression, aggressive, and invasive [47, 50]. We also obtained TMX2-4 and TMX2-11, two additional cell lines that were cloned from MCF-7 alongside TMX2-28. However, unlike the TMX2-28 cell line, TMX2-4 and TMX2-11 express ERα mRNA and protein and are non-invasive or migratory (Figure 1.9) [48, 49](Fagan-Solis, unpublished data). Together these cell lines, ERα -positive MCF-7 and Tamoxifen-selected derivatives, ERα -positive TMX2-4 and TMX2-11 and ERα -negative TMX2-28, provided a unique opportunity to investigate epigenetic changes related to Tamoxifen exposure and ERα expression.

HumanMethylation450 BeadChip

For this research, we used the Illumina HumanMethylation450 BeadChip (HM450BC) as a high-throughput way to analyze DNA methylation in both breast cancer cell lines and breast tumors. The HM450BC analyzes the methylation of 482,421

individual CpG dinucleotides in the human genome. It is enriched for CpG sites in both the promoter region with over 140,000 CpG sites located in the transcriptional start site regions 200 and 1500 (TSS200 & TSS1500) and CpG islands with 96% coverage (150,000 CpG sites) [51]. While there are other methylation analysis approaches available, they have limitations. The HumanMethylation27 BeadChip (HM27BC), which analyzed 27,000 CpG sites in the human genome, was the predecessor to the HM450BC and was phased out at the time of our analysis. Methyl specific PCR (MSP) is a PCR based analysis that requires primer design around areas in the promoters of genes, which can contain one or more CpG sites and therefore it may not distinguish the methylation of individual CpG sites [52]. Methylated DNA immunoprecipitated-sequencing (MeDIP-seq) is most similar to HM450BC with the number of CpG sites interrogated, however the protocol is not automated and is therefore less favorable to use with a large number of samples as it is cost-prohibitive [53].

Drug (Abbreviation)	Brand Name
Cyclophosphamide (C)	Cytoxan
Docetaxel (T)	Taxotere
Doxorubicin (A)	Adriamycin
Epirubicin (E)	Ellence
Methotrexate (M)	Maxtrex
Paclitaxel (T)	Taxol

Table 1.1. Commonly used chemotherapeutics in breast Cancer treatment [22].

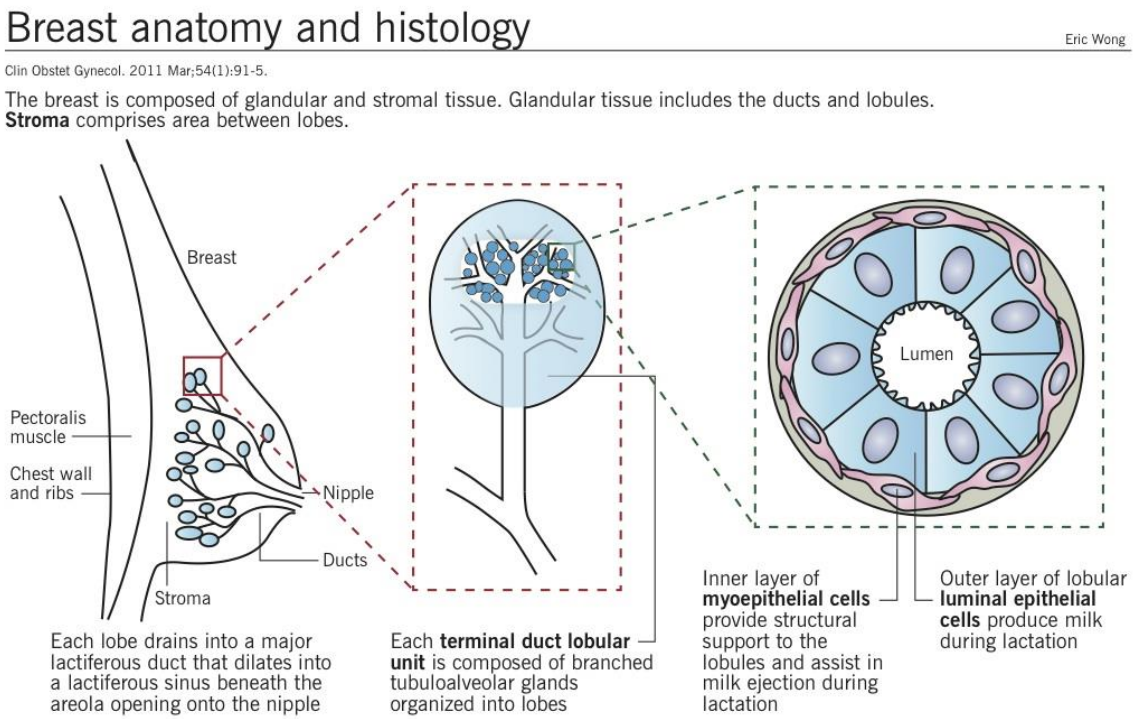


Figure 1.1 Structure of the human breast.

Lobules and ducts are surrounded by fatty connective tissue called the stroma. Luminal and epithelial cells line the lobules and ducts [2].

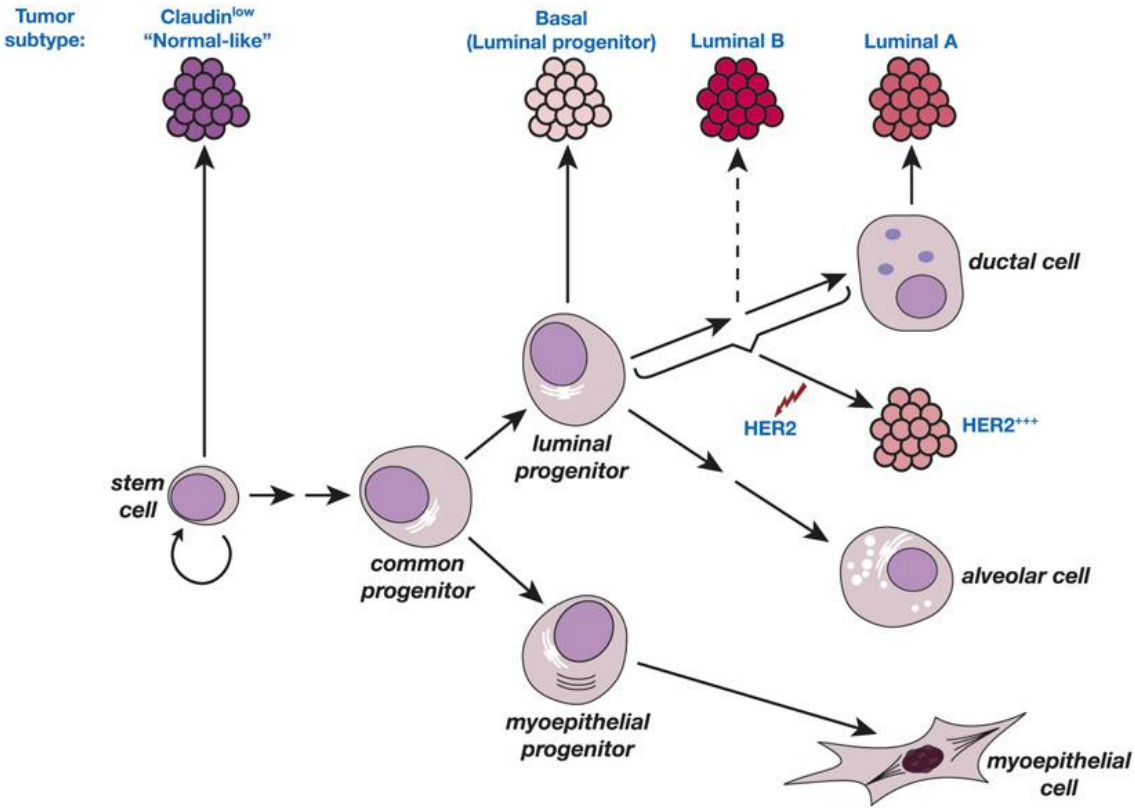
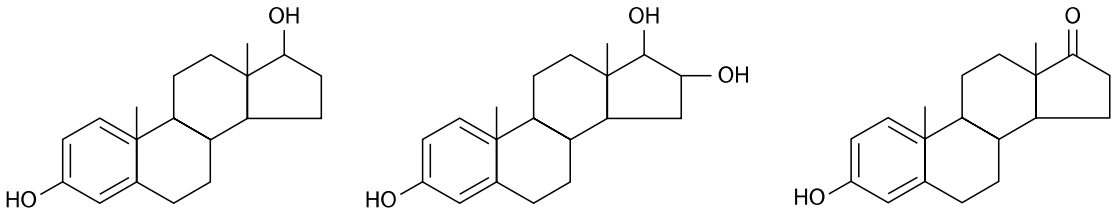


Figure 1.2 Mammary epithelial hierarchy and the molecular subtypes of human breast cancer.

Mammary stem cells give rise to progenitor cells from two pathways, luminal and myoepithelial. Luminal progenitor cells give rise to mature ductal and alveolar cells whereas myoepithelial progenitor cells give rise to mature myoepithelial cells. Tumor subtypes listed show the associated cell type based on expression analysis **[3]**



17β-Estradiol **Estriol** **Estrone**

Figure 1.3 Chemical structure of estrogens.

17β-estradiol (E2), estriol (E3), and estrone (E1).

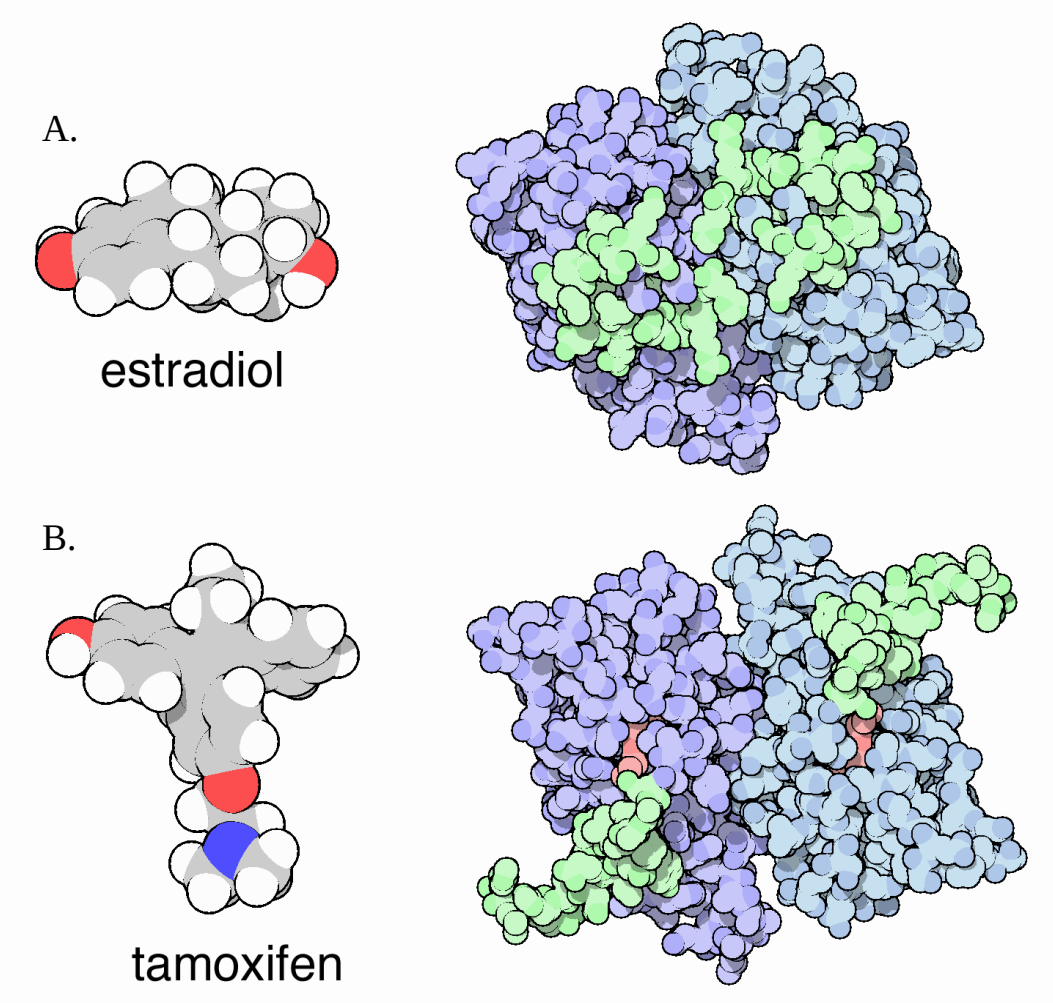


Figure 1.4 Structure of the Estrogen Receptor.

Structure of ER bound to A) 17-β-estradiol and B) Tamoxifen. When estradiol is bound, the signaling loop (colored green) of the estrogen receptor is activated. Binding of Tamoxifen causes the loop to take an inactive shape [13].

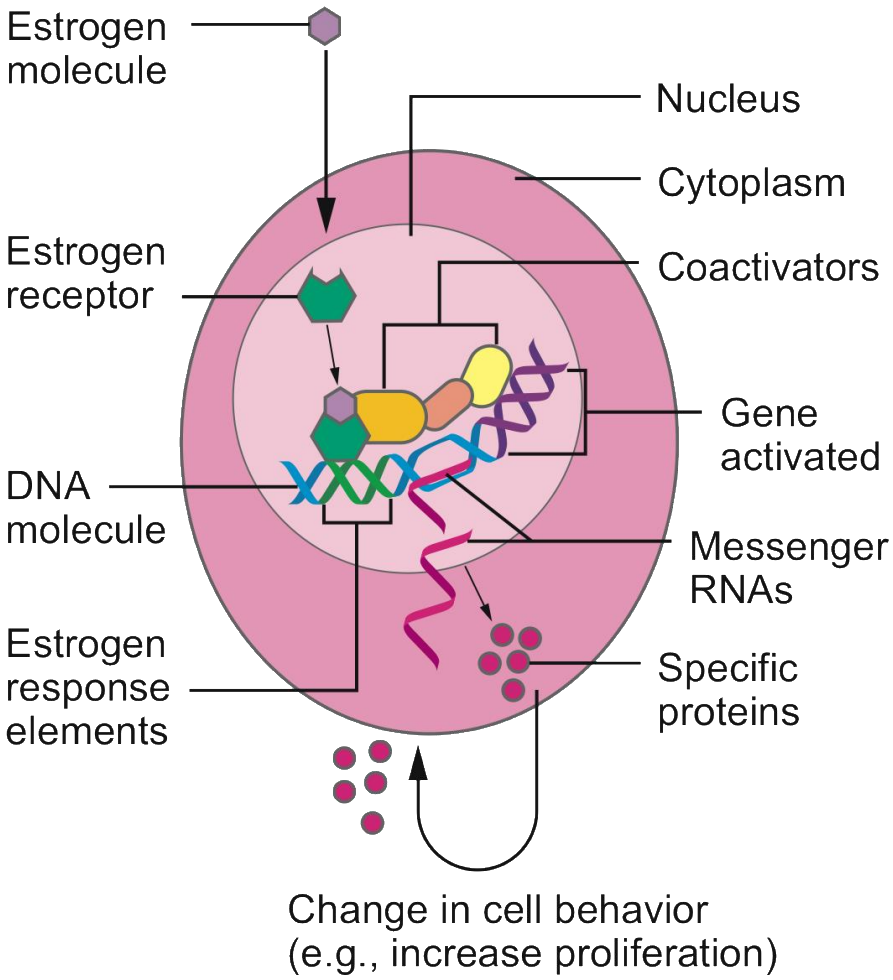


Figure 1.5 Activation of estrogen responsive genes controlled by estrogen molecule binding to the estrogen receptor.

Briefly, the estrogen molecule binds to the estrogen receptor and activates the estrogen response elements (ERE) and genes involved in cell proliferation [8].

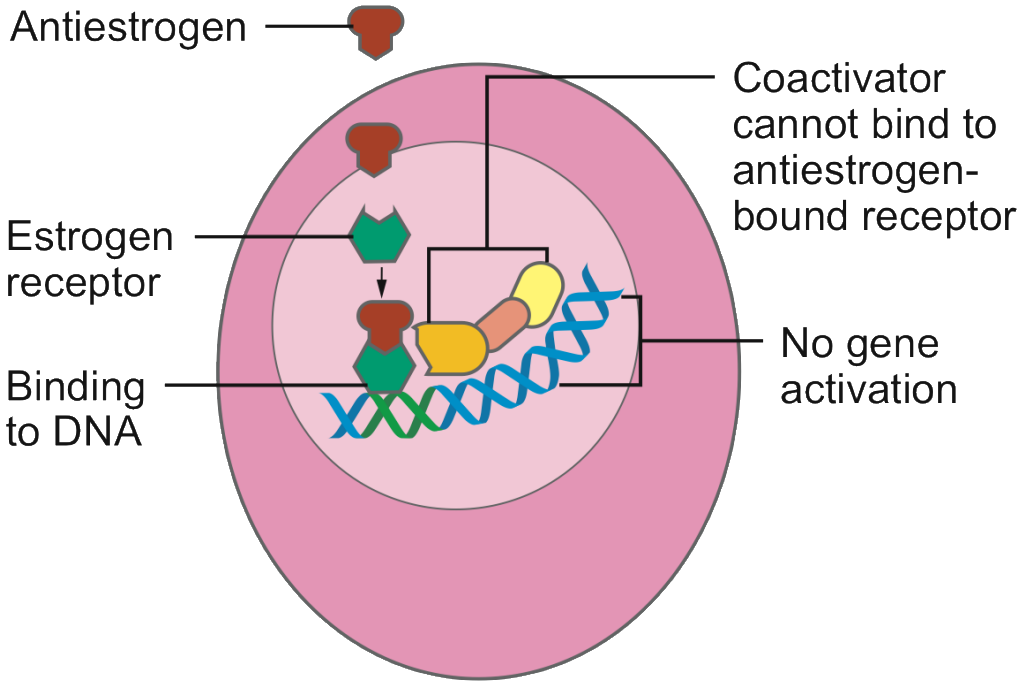


Figure 1.6 Mechanism of Tamoxifen (antiestrogen) in an ER-positive breast cell.

Binding of Tamoxifen to the ER blocks binding of the ER to the DNA and inhibits gene activation [8].

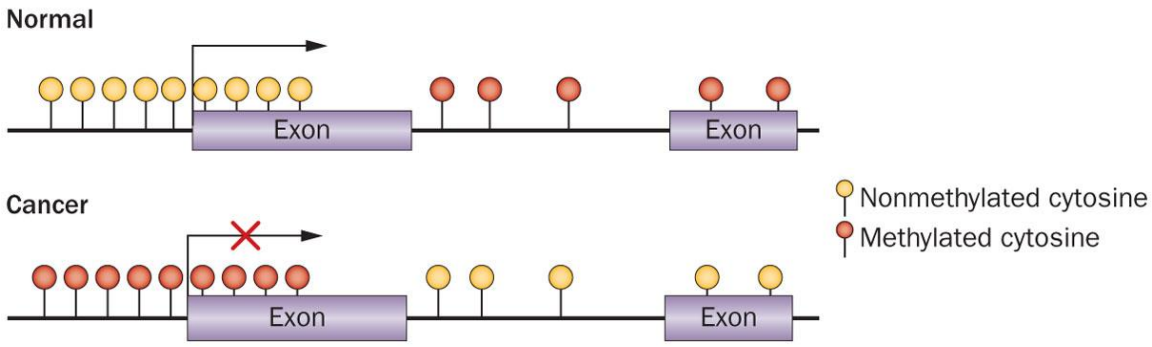


Figure 1.7 Promoter methylation plays a key role in regulating transcription.

A) CpG islands in the promoter region of a normal cell are unmethylated allowing for binding of transcription machinery and activation of transcription B) In a cancer cell, CpG islands in the promoter are methylated and transcription cannot occur **[54]**.

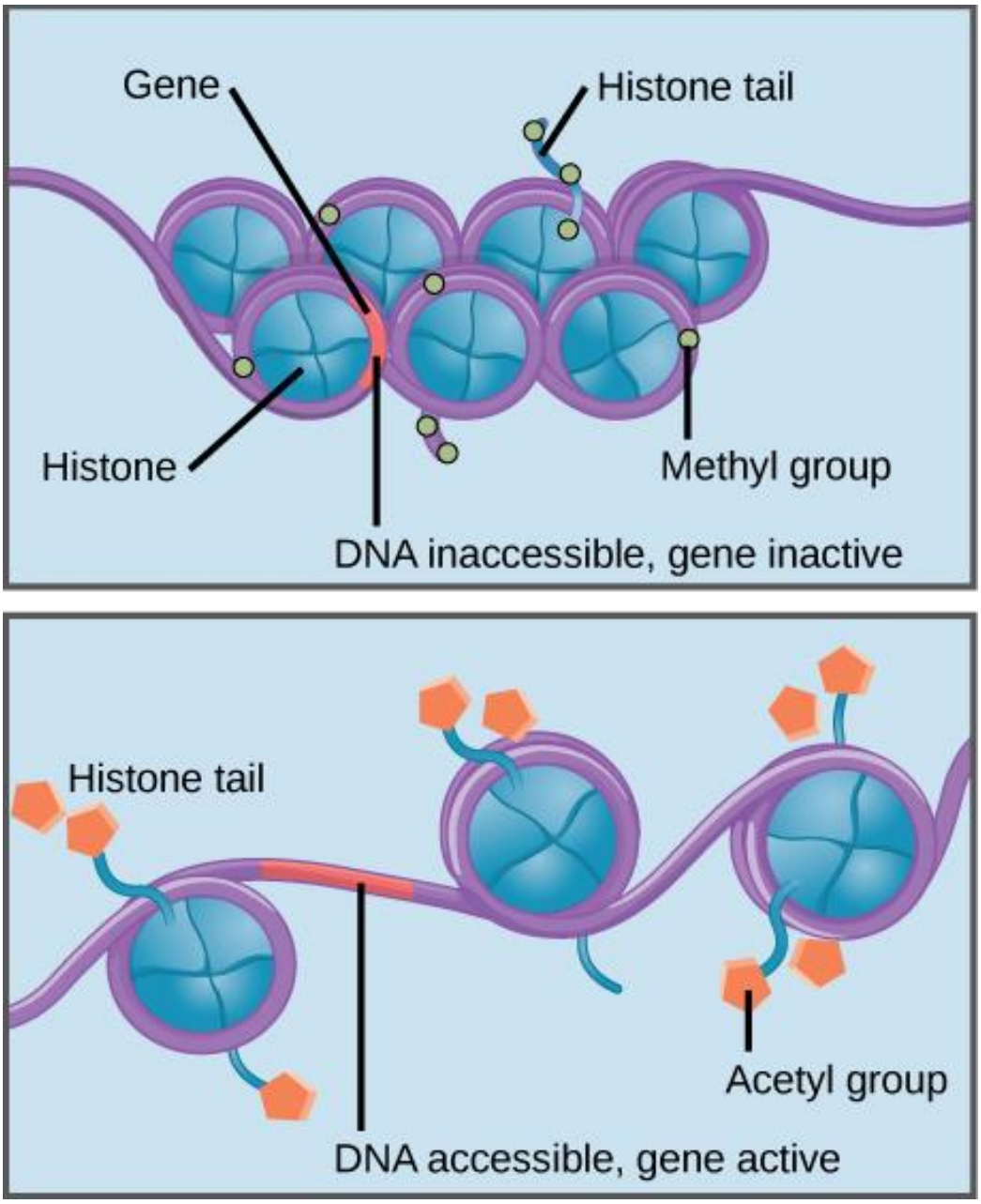


Figure 1.8 Methylation and nucleosome condensation.

The addition of methyl groups to the CpG sites of DNA results in tight winding of the DNA around histones and transcriptional repression. Conversely, demethylation of the DNA allows for transcription to occur **[55]**.

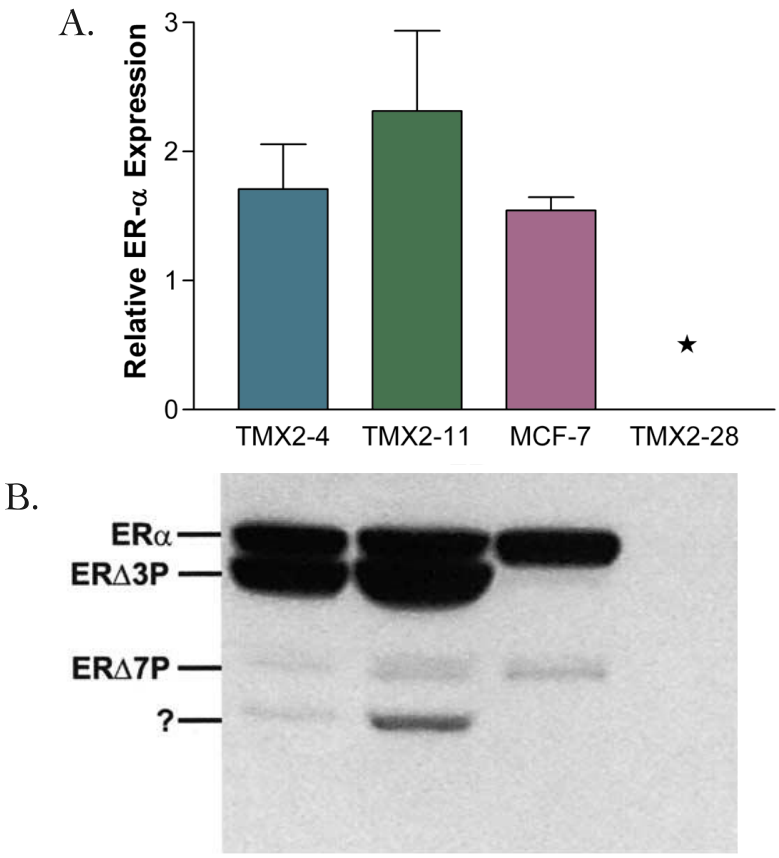


Figure 1.9 Differences in ER-α gene and protein expression occur across Tamoxifen-selected cell lines.

A) Gene expression analysis shows that ERα is expressed similarly to the parental strain, MCF-7, in two of the tamoxifen-selected cell lines, TMX2-4 and TMX2-11. ERα gene expression in tamoxifen-selected TMX2-28 is abolished. B) Protein expression by western blot confirms gene expression [49].

CHAPTER 2

HIGH DENSITY ARRAY ANALYSIS OF DNA METHYLATION IN TAMOXIFEN-RESISTANT BREAST CANCER CELL LINES

Introduction

Estrogen receptor-α (ERα) status remains one of the most important breast cancer diagnostic and prognostic biomarkers. Roughly 70% of all breast cancers are ERα-positive and can be treated with an antihormone such as Tamoxifen. However, a meta-analysis of 150,000 women from 200 randomized trials found that 33% of women receiving Tamoxifen for five years had recurrence (acquired resistance) within 15 years and 26% died [26]. In addition to this acquired antiestrogen resistance, *de novo*- or intrinsic-resistance occurs primarily in ERα-negative tumors but also occasionally in ERα-positive tumors [32]. Studies of endocrine resistance and global gene expression in Tamoxifen-resistant cell cultures and human tumors have detected alterations in numerous pathways including ER-signaling, growth factor receptor and cytoplasmic signaling, cell cycle, apoptosis and cell survival signaling [56]. A recent proteomics analysis of an ERα-negative, Tamoxifen-resistant MCF-7 derivative showed changes in expression of genes involved in metastasis, tumorigenesis, and ER-signaling pathways [57]. However, knowledge of the specific molecular mechanisms that cause these changes and determine the endocrine-resistance is far from complete.

DNA promoter methylation is a major epigenetic mechanism by which gene expression is altered in cancer. DNA methyltransferases (DMNTs) are responsible for the addition of methyl groups to the cytosine of a CpG site [39]. In normal adult tissue, CpG islands remain largely unmethylated; however, in the case of cancer, hypermethylation of normally unmethylated cytosines in promoter CpG islands frequently results in gene silencing, while hypomethylation of normally methylated cytosines in CpGs outside promoter regions leads to genetic instability [46]. Suppression of genes involved in cell cycle control, DNA repair, apoptosis and cell survival, and toxicant metabolism is thought to play a major role in the etiology and progression of cancer.

To date, few studies have examined promoter methylation and Tamoxifen resistance in breast cancer. Of the studies conducted, one detected drug-specific promoter methylation and gene expression profiles in an ERα-positive, Tamoxifen-resistant MCF-7 derivative cell line [30]. Another study demonstrated that promoter hypermethylation was not the cause of decreased progesterone receptor expression in a Tamoxifen-resistant but estrogen-dependent MCF-7 derived clone [58]. However, methylation analysis of both ERα-positive and ERα-negative Tamoxifen-resistant cell lines derived from a single parental line have not been reported until now.

In the present study, we examine DNA methylation in two Tamoxifen-resistant clones of MCF-7, TMX2-11 and TMX2-28. TMX2-11 retained expression of ERα, while TMX2-28 lost expression of the gene [49]. We found that prolonged treatment with Tamoxifen induced hypermethylation and hypomethylation throughout the genome. Analysis of methylation and expression of two genes with promoter methylation in both

Tamoxifen-resistant cell lines demonstrated cell line-specific responses to treatment with 5-aza-2'deoxycitidine (5-Aza).

Materials and Methods

Cell Culture, RNA and DNA Purification

TMX2-11 and TMX2-28 were kindly provided by John Gierthy (Wadsworth Center Albany, NY). MCF-7 cells were purchased from the American Type Culture Collection (ATCC). Cell lines were grown in Dulbecco's modified eagle medium (without phenol red) supplemented with 5% cosmic calf serum (Hyclone), 2.0 mM of L-glutamine, 0.1 mM of nonessential amino acids, 10 ng/mL of insulin, 100 units/mL of penicillin, and 100 µg/mL of streptomycin. Cells were maintained at 37°C with 5% CO_2 in a humidified incubator and media was changed every 2 days. MCF-7 cells were cultured with and without 10^{-10} M E2 (Sigma-Aldrich) added to the media for 14 days.

RNA was purified in triplicate for each cell line using TriReagent (Molecular Research Center, Inc) and DNA was purified using QIAamp DNA Mini kit (Qiagen) as per manufacturer suggestion and protocols previously described [47, 50]. Purified RNA and DNA samples were quantified using a NanoDrop 8000 (Thermo Scientific).

Illumina HumanMethylation450 BeadChip (HM450BC)

Samples purified from MCF-7, TMX2-11, TMX2-28, and short-term (14 day) 10^{-10} M E2-treated MCF-7 using the QIAmp DNA Mini kit (Qiagen) were sent to the core

facility at Roswell Park Cancer Institute for HM450BC (Illumina) analysis. Briefly, DNA sent to Roswell Park Cancer Institute was quantitated by PicoGreen (Molecular Probes) prior to bisulfite treatment with the EZ DNA methylation kit (Zymo). Bisulfite-treated DNA was amplified at 37°C for 20-24 hours after treatment with 0.1N NaOH. The DNA was then fragmented at 37°C for 1 hour using an enzymatic process and subsequently precipitated in 100% 2-propanol at 4°C for 30 minutes followed by centrifugation at 3000xg at 4°C for 20 minutes. Dried pellets were resuspended in hybridization buffer, 48°C for 1 hour followed by 95°C for 20 minutes, then loaded onto the HM450BC and incubated at 48°C for 16-24 hours. Following hybridization of DNA to the primers on the BeadChip, unhybridized and non-specific DNA was removed using wash buffers to prepare the chip for staining. After a single base extension of the hybridized primers using labeled nucleotides, the BeadChip was stained with Cy-3 and Cy-5 fluorescent dyes and read using the Illumina iScan Reader. The image data were then analyzed using Illumina GenomeStudio to assess efficiency of the reaction. Methylation of the interrogated CpG loci were calculated as the ratio of the fluorescent signals of methylated to unmethylated sites (beta values).

5-Aza-2'deoxycitidine Treatment of Cells

Cells were seeded into 6-well plates at varying concentrations (MCF-7 and TMX2-11: 150,000 cells/well; TMX2-28 and MDA-MB-231: 100,000 cells/well) and allowed to attach overnight at 37°C and 5% CO_2. Two experiments were completed 9 months apart. Triplicate replicate wells were treated with either 0.1% DMSO (vehicle control) or 2.5 µM 5-aza-2'deoxycitidine (Sigma-Aldrich) in 0.1% DMSO for 4 days,

refreshing the media every other day. On the fourth day, DNA and RNA were purified from the cells as described above and concentration and quality were evaluated using the NanoDrop 8000 (Thermo Scientific).

Pyrosequencing

DNA (1 µg) was bisulfite treated using the EpiTect Bisulfite kit (Qiagen) and PCR Primers were designed using the Pyromark Assay Design Software (Qiagen). One µL of bisulfite treated DNA was amplified using the Pyromark PCR kit (Qiagen) in a BioRad MyCycler and the following gene specific primers designed to target CpG sites in the TSS200 promoter region of the gene analyzed by the BeadChip: ZNF350 (NM_021632) GRCh37 HG19 Map position (MAPINFO) Ch19 coordinates: 52490101, 52490120, 52490127, and 52490173; Primers for pyrosequencing: FWD Biot-5'-TTGGTTTTTGGTTTAAAAATTTGTTAT-3', REV 5'-ACACTAACCTCTATTTTCTCCAAATACACAA-3', SEQ 5'-ACTCCTACTTCTAAAATCCT-3'; MAGED1 (NM_001005332) MAPINFO ChX coordinate: 51546021; Primers for pyrosequencing: FWD 5'-GAGGTTTGAGTTAAGGGATTAAGATGA-3', REV 5'-Biot-TACCCCCTCCTTCACTTCAA-3', SEQ 5'- AGATGAAGGGAGATATTT-3'. Additional CpG sites not analyzed by the BeadChip were assessed in the pyrosequencing assay due to their proximity to the CpG sites of interest. Single stranded products were prepared for pyrosequencing by PyroMark vacuum prep tool (Biotage). Pyrosequencing reactions were performed using a Pyromark Q24 system (Biotage) and manufacturers

protocol (Qiagen). Data were analyzed using Pyromark Q24 Software for percent methylation at the CpG sites interrogated.

Quantitative Real Time Reverse Transcriptase-PCR (qRT-PCR)

Primers for qRT-PCR were designed using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and the UCSC RefGene Accession number associated with the CpG site of interest (MAGED1 NM_001005332, ZNF350 NM_021632) or as previously described [59]: MAGED1 FWD 5'-CCTTCTTCGTCAAGCCCCCAG-3', REV 5'- AGGCAGCATTTGGACCCTTT-3'; ZNF350 FWD 5'-CCCAGTTGAATGCTGTCTTCC-3', REV 5'-CCACTCCTCCCAAGTGAAGTC-3'. qRT-PCR analysis was carried out as previously described on a Roche LightCycler using the Qiagen OneStep RT-PCR kit (Qiagen) and SYBR green I nucleic acid stain (Invitrogen) [47, 50]. Total RNA (75 ng) was combined with OneStep RT-PCR master mix, dNTPs, SYBR green (2X), and primers (25 μM each) described above in chilled capillaries (Roche). RNA was reverse transcribed for 30 minutes at 50°C and subsequent amplification was assayed for 45 cycles using fluorescence generated by intercalating SYBR green dye into the resulting DNA product. Relative mRNA expression levels were normalized to hypoxanthine ribosyltransferase (HPRT) as described previously [47].

Data Analysis

Using the Minfi package for R [60], a beta MDS was created from the beta values of the top 1000 CpG sites that deviate the greatest most among the samples in the

HM450BC data files. GenomeStudio Methylation Module (v.1.9) was used to analyze the beta values of the methylation data obtained from the HM450BC. CpG sites with detection p-values of <0.01 were selected to ensure statistically significant CpG site data were analyzed. Average beta value of >0.3 (range from 0-1) was used as a cutoff for hypermethylated CpG site divergence in the Tamoxifen resistant cell lines and an average beta value of <0.3 for hypomethylated CpG sites. Lastly, to discern differences in the CpG site methylation data in Tamoxifen-resistant cell lines as compared to the parental line, MCF-7, a positive fold change was calculated as average beta of Tamoxifen resistant cell line over parental cell line. A negative fold change was calculated as the parental cell line over the Tamoxifen resistant clone. GraphPad Prism (GraphPad Software Inc.) was used to analyze and graph the biological replicate statistical results from pyrosequencing and qRT-PCR and to calculate a Pearson correlation coefficient for HM450BC and pyrosequencing data. Unpaired students t-test with a p-value of <0.05 were considered statistically significant. The Database for Annotation, Visualization and Integrated Discovery (https://david.ncifcrf.gov) was used to conduct pathway analysis from a list of genes associated with CpG sites described above as hyper- or hypomethylated.

Results

Tamoxifen-selection results in extensive changes in DNA methylation

To compare DNA methylation among the Tamoxifen-selected cell lines, the 17β-estradiol (E2)-treated cells and the non-treated parental cell line, we used

Multidimensional Scaling (MDS) to analyze results from the HM450BC. Beta values of the top 1000 CpG sites that varied most among samples were plotted using the Minfi package for R. As illustrated in the MDS plot, one of the two Tamoxifen-selected cell lines and the E2-treated MCF-7 do not deviate from the parental MCF-7 on Dimension 1 (Figure 2.1). These three groups all have a value of -7.5 relative units (RUs) on the X-axis (Dimension 1), and all are positive for ERα. In contrast, the Tamoxifen-selected, ERα-negative cell line, TMX2-28, falls about 28 RUs from the other samples. The deviation in Dimension 1 was restricted to the ERα-negative cell line suggesting that the methylation in this Dimension may be secondary to the loss of ERα and not a direct consequence of Tamoxifen-selection. In contrast, the deviation in Dimension 2, while significantly less than that of Dimension 1, may reflect methylation changes directly related to the E2 and Tamoxifen treatments. Both of the Tamoxifen-selected cell lines show a modest deviation in the same direction on Dimension 2. TMX2-11 is roughly 4 RUs from MCF-7, while TMX2-28 is one RU from the parent cell line. We treated MCF-7 cells with E2 for 14 days to examine the overall effect that short-term treatment with a known ERα agonist had on methylation. Interestingly, the E2-treated MCF-7 cells deviate by 1.5 RUs from the untreated MCF-7 but in the opposite direction as the two Tamoxifen-selected cell lines.

Differentially methylated CpG (dmCpG) sites in the Tamoxifen-selected lines are primarily hypermethylated

To further assess the effects of prolonged Tamoxifen treatment on DNA methylation we prepared scatter plots comparing all CpG sites among the Tamoxifen-

selected cell lines and the parental line (Figure 2.2). The areas outlined in blue on each of the scatterplots in Figure 2.2 include data points for dmCpG sites that show a 2-fold change and have average beta values >0.3. The beta value cut-off point of 0.3 was chosen based on previous literature demonstrating significant changes in CpG site methylation between Tamoxifen-resistant and parental cell lines [61-63]. The scatter plots confirm and expand the results illustrated in the MDS plots; prolonged treatment with Tamoxifen results in methylation changes that are more pronounced in the cell line that lost expression of ERα (Figure 2.2A &B). Additionally, for both cell lines the majority of dmCpGs are hypermethylated. Roughly eight times more CpG sites are hypermethylated in TMX2-28 as compared to TMX2-11 (33,752 versus 4,309; Table 2.1). While hypomethylation was less common, there are twice as many hypomethylated CpG sites in TMX2-28 as compared to TMX2-11 (5,252 versus 2,593; Table 2.1). The methylation patterns of TMX2-11 and control MCF-7 (Figure 2.2A) are more similar than those of the two Tamoxifen-resistant lines TMX2-11 and TMX2-28 (Figure 2.2C). In contrast to prolonged Tamoxifen treatment, 14 days of treatment with E2 resulted in few dmCpGs, and these are primarily hypomethylation changes (Figure 2.2D and Table 2.1).

To assess the effects of Tamoxifen on DNA methylation while limiting the potential bias due to loss of ERα in TMX2-28, we restricted the next set of analyses to CpG sites with methylation changes in similar directions (both hyper- or hypomethylated) in both TMX2-11 and TMX2-28 as compared to the parent cell line, MCF-7. The Tamoxifen-selected cell lines share roughly 3,000 hypermethylated (>0.3

average beta value and >2-fold change from MCF-7) and 200 hypomethylated (<0.3 average beta value and <-2-fold change from MCF-7) CpG sites (Table 2.1).

dmCpG sites are found primarily in the intergenic, body and promoter regions

To obtain a better understanding of the effect of Tamoxifen on breast cancer cells, differential methylation of TMX2-11 and TMX2-28 as compared to MCF-7 was examined over the entire genome. Figure 2.3A shows the number of CpG sites included on the HM450BC in each of five regions: promoter (TSS200 and TSS1500 regions; 29%), 5'UTR/1st Exon (12%), body (31%), 3'UTR (3%) and intergenic (areas not included in the previous four regions; 25%) [51]. The functional genomic distribution of dmCpGs in the Tamoxifen-selected lines is shown in Figure 2.3B and C. In general the distribution of hyper- and hypomethylated CpG sites reflects their representation on the BeadChip. Thirty-two percent of CpG sites with hypermethylation are found in intergenic regions followed closely by the body (30%) and promoter regions (21%; Figure 2.3B). Results are similar for hypomethylated CpG sites with 32% located in the body, 30% in the promoter, and 22% in intergenic regions (Figure 2.3C). A single CpG site may be counted several times if there are multiple transcripts or gene-overlap, so that the total number of methylated CpG sites in Figure 2.3B and C do not add up to those in Table 2.1.

Figure 2.3D summarizes neighborhood location of all CpG sites on the HM450BC as described in the GenomeStudio Methylation Module user guide (Illumina, San Diego, CA); shores (23%) are located 0-2 kb and shelves (10%) are 2-4 kb from the canonical CpG islands, while the remainder of the sequence is defined as open sea (36%;

Figure 2.3D) [51, 64]. The relationship of shared hyper- and hypomethylated CpG sites in the Tamoxifen-selected lines to the canonical CpG islands is shown in Figure 2.3E and F. The pattern of the hypermethylated sites deviates from their representation on the HM450BC. Only 10% of hypermethylated CpG sites lie within the CpG islands, while 31% of the CpGs included on the BeadChip are within an island (Figure 2.3D and E). The open sea region has the greatest number of hypermethylated sites (68% of all hypermethylated CpGs) and deviates the greatest from the representation on the BeadChip (36% of all CpGs). In contrast, the pattern of the hypomethylated genes reflects their representation on the BeadChip.

Sensory perception is among the top pathways affected by Tamoxifen selection

Pathway analyses were conducted on genes with dmCpG sites in both TMX2-11 and TMX2-28 as compared to MCF-7. The first DAVID analysis separately examined genes with either hyper- or hypomethylated sites occurring anywhere in the gene. The top 20 pathways with hypermethylated genes, out of an extensive list of statistically significant pathways, and the top 5 statistically significant pathways with hypomethylated genes are shown in Table 2.2. The hypermethylated pathway with the highest statistical significance is sensory perception of smell, which includes 100 olfactory receptor genes (see Appendix A). This is followed closely by the cell surface linked signal transduction pathway, which includes many of the same olfactory receptor genes as described above, as well as genes involved in the WNT and TGFβ signaling pathways. Sixty-four genes in the cell adhesion pathway have increased methylation and the majority of these genes are involved in ECM-receptor interaction pathways (Table 2.2). The hypomethylated gene

list was less associated with any specific pathway, presumably due to the small number of hypomethylated dmCpGs.

Next we conducted DAVID analyses restricted to hypermethylated genes in either the promoter or the body regions. The top pathways with promoter hypermethylated genes are sensory perception of smell and sensory perception of chemical stimulus (70 and 72 genes respectively) with the majority being olfactory receptor genes (Appendix B). In comparison, the top pathways with body hypermethylated genes are ion and metal ion transport (51 and 34 genes respectively), followed by cell adhesion (42 genes; Appendix C). The promoter and body regions share only six out of the top 20 hypermethylated gene pathways (Table 2.3).

Promoter methylation of ZNF350 and MAGED1

Given that promoter methylation (TSS200 and TSS1500 regions; Figure 2.4A) can alter gene expression in cancer [39, 46], we wanted to further examine the role of promoter methylation in the Tamoxifen selected cell lines. We selected two genes with at least two dmCpG sites that had beta values above 0.3 and a >2-fold change in the promoter region in both TMX2-11 and TMX2-28 as compared to MCF-7 from the HM450BC. Expression of both genes has been shown to be downregulated in breast cancer, yet DNA promoter methylation has not been suggested as a potential mechanism of decreased expression [59, 65]. The first gene, zinc finger protein 350 (ZNF350), a DNA damage response protein, has increased methylation in 7 out of 10 promoter CpG sites represented on the HM450BC in TMX2-11 and in 8 out of 10 in TMX2-28 (Figure 2.4B). The second gene, melanoma antigen family D1 (MAGED1), a tumor antigen and

putative regulator of p53 transcription has five CpG promoter sites in transcript variant 3 represented on the HM450BC. Of these five sites, all are hypermethylated in TMX2-28 and four are hypermethylated in TMX2-11 (Figure 2.4C). There were 120 additional genes that also displayed hypermethylation in at least two CpG sites in the promoter, approximately 40% of which are thought to play a role in cancer (Appendix D).

To confirm the methylation observed with the BeadChip and evaluate the TSS200 region (flanking region upstream of the TSS) in greater depth, we designed pyrosequencing assays to interrogate CpG sites in both ZNF350 and MAGED1. The pyrosequencing assay for ZNF350 examines seven CpG sites, four of which were represented on the HM450BC (Figure 2.4B, orange box). The pyrosequencing assay for MAGED1 examines four CpG sites, one of which was included on the HM450BC (Figure 2.4C, orange box). Results obtained from pyrosequencing of bisulfite-modified DNA (percent methylated) confirm the increased promoter methylation discovered on the HM450BC for both ZNF350 and MAGED1 (Table 2.4). For ZNF350, the percent methylation in MCF-7 cells is remarkably similar to beta values for all four CpG sites examined with both methods. Likewise, the percent methylation in ZNF350 in the Tamoxifen-selected cell lines is highly comparable to the beta values. A similar trend is observed for MAGED1. A strong correlation is seen between HM450BC beta values and pyrosequencing values for all CpG sites assayed (Pearson $r = 0.931$, $p = <0.0001$; Appendix E).

Figure 2.5 shows the detailed pyrosequencing results for MCF-7 and the Tamoxifen-selected cell lines. For both ZNF350 and MAGED1 the CpG-site specific pattern is highly reproducible in DNA isolated nine months apart. Pyrosequencing across

all sites confirm greater mean methylation in TMX2-11 (30% increase) and TMX2-28 (17% increase) as compared to MCF-7 for ZNF350 (Figure 2.5A). Results for MAGED1 also confirm greater mean methylation, TMX2-11 (3% increase) and TMX2-28 (30% increase; Figure 2.5B).

Treatment with 5-Aza reverses DNA methylation in TMX2-28

To assess whether promoter methylation of ZNF350 and MAGED1 could be reversed to the levels of MCF-7, cell cultures were treated with 2.5 μM of 5-Aza or vehicle control for 4 days. Pyrosequencing of CpG sites in the TSS200 region of ZNF350 reveals a significant, 23% decrease in methylation (from 27 to 20) in TMX2-28 treated with 5-Aza ($p=0.006$; Figure 2.6A). Likewise, a 31% decrease in methylation is observed in the promoter of MAGED1 in TMX2-28 ($p = 0.0002$; Figure 2.6B). A small (10%) but significant decrease in methylation is also observed in TMX2-11 cells (Figure 2.6B).

Decreasing methylation results in increased expression of ZNF350 and MAGED1 in TMX2-28

After determining that treatment with 5-Aza decreased promoter methylation, we asked whether the 5-Aza treatment also increases mRNA expression levels. We compared mRNA levels of ZNF350 and MAGED1 in treated and control cell lines. Treatment with 5-Aza significantly increases the expression of ZNF350 in TMX2-28 (5.6 fold) as compared to the untreated cell cultures (Figure 2.6A). Interestingly, the expression of ZNF350 also increases in TMX2-11 (2 fold) even though there is no change in promoter methylation (Figure 2.6A). In TMX2-11, ZNF350 expression levels

are equivalent to those of MCF-7, while in the TMX2-28 cells ZNF350 is significantly overexpressed (TMX2-28 5-Aza vs. MCF-7 Control: p=0.04; Figure 2.6A). Treatment with 5-Aza increases expression of MAGED1 in TMX2-28 (442 fold) to a level significantly above that of MCF-7 (p= 0.028; Figure 2.6B). In contrast, the expression of MAGED1 is not increased in TMX2-11, despite the significant decrease in methylation (Figure 2.6B).

Analysis of the ERα-negative line, MDA-MB-231 is included for comparison with TMX2-28. Treatment with 5-Aza increases the expression of ZNF350 in MDA-MB-231 (p=0.008) to levels similar to TMX2-28, but has no effect on MAGED1. Promoter methylation of ZNF350 and MAGED1 is low in MDA-MB-231 and not altered by 5-Aza treatment.

Discussion

Acquired Tamoxifen resistance occurs in approximately 33% of all women who are given the drug for 5 years [26]. The mechanism of this acquired resistance by the cells is largely unknown, however DNA methylation has been shown to differ between Tamoxifen-resistant and Tamoxifen-sensitive cell lines [30, 58]. Past studies examined methylation changes in ERα-positive, Tamoxifen-resistant cell lines. Here we present methylation data on both ERα-positive and ERα-negative Tamoxifen-resistant cell lines derived concurrently from the parental cell line, MCF-7.

We found substantial overall changes in methylation, suggesting that DNA methylation is contributing to Tamoxifen resistance in both ERα-positive and -negative

cell lines. Interestingly, the loss of ERα expression in TMX2-28 does not appear to be controlled by changes in methylation. TMX2-28 ERα has an average of 3% methylation in the promoter region as analyzed by pyrosequencing and treatment with 5-Aza does not cause re-expression (see Chapter 3, Figure 3.7). Further studies examining histone modifications and other epigenetic changes will likely provide insight into the loss of ERα expression in TMX2-28.

Since the ERα-negative TMX2-28 cells show significantly greater methylation changes than the ERα-positive TMX2-11 cells, it is likely that a large percent of the observed DNA methylation is secondary to the loss of ERα expression. To eliminate the bias due to ERα loss and to focus on pathways most relevant to Tamoxifen-resistance, we examined CpG sites similarly methylated in both TMX2-11 and TMX2-28 as compared with the parental MCF-7 line. The number of hypermethylated sites in both cell lines is greater than the number of hypomethylated and the dmCpGs are distributed across the gene regions. Because of the importance of promoter methylation in controlling gene expression [46, 65], the HM450BC is enriched for CpG sites in the promoter region, with over 140,000 sites represented in the TSS200 and TSS1500 regions [51] Recent literature, however, suggests that body methylation may play an equally important role in controlling gene expression [66, 67]. Less than 1% of the promoter and body CpG sites represented on the HM450BC are hypermethylated in both TMX2-11 and TMX2-28 and of these dmCpGs, slightly more are in the body than in the promoter region (0.66% versus 0.52%).

ZNF350 is frequently underexpressed in primary breast cancer [59]. It functions as a transcriptional repressor by binding to its co-repressor, BRCA1, and silencing target

genes involved in DNA damage response [59]. Treatment with 5-Aza increased expression of ZNF350 in both Tamoxifen-resistant cell lines as well as MDA-MB-231, yet only in TMX2-28 was a significant decrease in promoter methylation observed. Expression of ZNF350 in TMX2-11 and MDA-MB-231 may be regulated by an upstream factor or by methylation outside of the CpGs examined. Published studies using 5-Aza to induce expression of genes downregulated in cancer indicate that multiple factors, such as location of CpG sites within the island regions, transcription factor promoter methylation and histone methylation play a role in controlling expression [68-72]. A further investigation into gene expression using array-based methods may help elucidate the genes affected specifically by promoter methylation.

MAGED1 is an adaptor protein involved in regulation of various cellular processes altered in cancer including apoptosis, proliferation and cell growth [65, 73]. MAGED1 is downregulated in cancer and it has been reported that transfection of the gene into breast cancer cells lacking MAGED1 inhibits proliferation and invasion of the cells [65]. Treatment with 5-Aza significantly decreases methylation of MAGED1 in both TMX2-11 and TMX2-28, but concomitant increased expression occurs only in TMX2-28. This suggests that methylation may be necessary, but not sufficient to re-express MAGED1 in TMX2-11 as the methylation decreases, but no change in expression is seen. No changes in either methylation or expression of MAGED1 were observed in MDA-MB-231. Cell line differences in response to 5-Aza highlight the difficulty of using agents, which target methylation to treat breast cancer. TMX2-28 are more sensitive to the effects of 5-Aza and their appearance is notably altered (flatter, rounder and larger in appearance) after four days of treatment (See Chapter 3, Figure 3.8). The

differences among the cell lines are analogous to the differences among breast cancers in patients. Not all breast tumors will respond similarly to treatment with demethylating agents and future emphasis must be placed on identifying markers that accurately predict response to treatment.

Table 2.1 CpG methylation changes in Tamoxifen-resistant cell lines as compared to the parental line

	TMX2-11 /MCF-7	TMX2-28 /MCF-7	MCF-7 E2 /MCF-7	TMX2-11 and TMX2-28 /MCF-7
Increased Methylation*	4,039	33,752	128	3,130
Decreased Methylation**	2,593	5,252	1,698	203
No Change in Methylation	472,153	436,113	479,003	431,909

*Increased methylation: >2-fold change, >0.3 beta-value in TMX2-11, TMX2-28, or E2 treated MCF-7; **Decreased methylation: >2-fold change, >0.3 beta-value in MCF-7; No change in methylation: <2-fold change in all lines. Detection p-value of < 0.01 was used to distinguish statistically significant methylation changes.

Table 2.2 Hyper- and hypomethylated pathways shared by TMX2-11 and TMX2-28

Hypermethylated*

Pathway	***p value***
sensory perception of smell	2.13E-34
cell surface receptor linked signal transduction	4.28E-33
neurological system process	1.34E-32
sensory perception of chemical stimulus	1.72E-32
G-protein coupled receptor protein signaling pathway	5.72E-30
cognition	2.95E-27
sensory perception	6.98E-26
ion transport	5.82E-08
cell-cell signaling	4.29E-07
transmission of nerve impulse	2.94E-06
synaptic transmission	1.13E-05
neuron differentiation	1.83E-05
metal ion transport	2.07E-05
behavior	5.32E-05
cell motion	1.41E-04
regulation of system process	2.78E-04
cell adhesion	3.77E-04
biological adhesion	3.91E-04
neuron projection development	4.75E-04
calcium ion transport	5.98E-04

Hypomethylated**

Pathway	***p value***
fear response	0.007
cell morphogenesis involved in differentiation	0.025
neuron development	0.028
multicellular organismal response to stress	0.029
neuron differentiation	0.030

* Top 20 hypermethylated pathways ** Top 5 hypomethylated pathways

Table 2.3 Hypermethylated pathways in the promoter and body regions shared by TMX2-11 and TMX2-28*

Promoter		***Body***	
Pathway	***p value***	***Pathway***	***p value***
sensory perception of smell	6.50E-32	ion transport	7.35E-09
sensory perception of chemical stimulus	8.96E-31	metal ion transport	4.77E-07
G-protein coupled receptor protein signaling pathway	5.19E-26	cell adhesion	2.95E-06
neurological system process	2.31E-22	biological adhesion	3.02E-06
sensory perception	6.40E-21	cell-cell signaling	1.78E-05
cognition	6.69E-21	cation transport	1.96E-05
cell surface receptor linked signal transduction	2.52E-20	multicellular organismal response to stress	7.55E-05
defense response to bacterium	0.00135	transmission of nerve impulse	8.78E-05
gamma-aminobutyric acid signaling pathway	0.00220	neurological system process	1.01E-04
ion transport	0.00229	appendage development	1.10E-04
transmission of nerve impulse	0.00257	limb development	1.10E-04
regulation of cell migration	0.00334	calcium ion transport	1.33E-04
behavior	0.00420	cell surface receptor linked signal transduction	1.35E-04
synaptic transmission	0.00533	response to pain	1.60E-04
chemotaxis	0.00619	neuron differentiation	1.66E-04
taxis	0.00619	muscle organ development	2.15E-04
cell-cell signaling	0.00738	cell motion	2.50E-04
regulation of locomotion	0.00908	regulation of system process	3.11E-04
response to drug	0.00917	di-, tri-valent inorganic cation transport	3.26E-04
regulation of cell motion	0.00944	synaptic transmission	5.24E-04

*Top 20 hypermethylated pathways

Table 2.4 Comparison of ZNF350 methylation by HumanMethylation450 BeadChip and Pyrosequencing in Tamoxifen-resistant and Parental Cell Lines

MAPINFO coordinate	***MCF-7***		***TMX2-11***		***TMX2-28***	
	*BeadChip**	*Pyroseq***	*BeadChip**	*Pyroseq***	*BeadChip**	*Pyroseq***
ZNF350						
52490101	9	8	41	34.6	46	27
52490120	12	12.3	51	48.6	37	33
52490127	7	10	49	47	30	27
52490173	7	9.7	45	31.3	34	21
MAGED1						
51546021	17	28.3	44	33	79	71

MAPINFO coordinate = genomic coordinate of C in CpG site; *Average beta value times 100; **Average percent methylation per site; n = 3 biological replicates

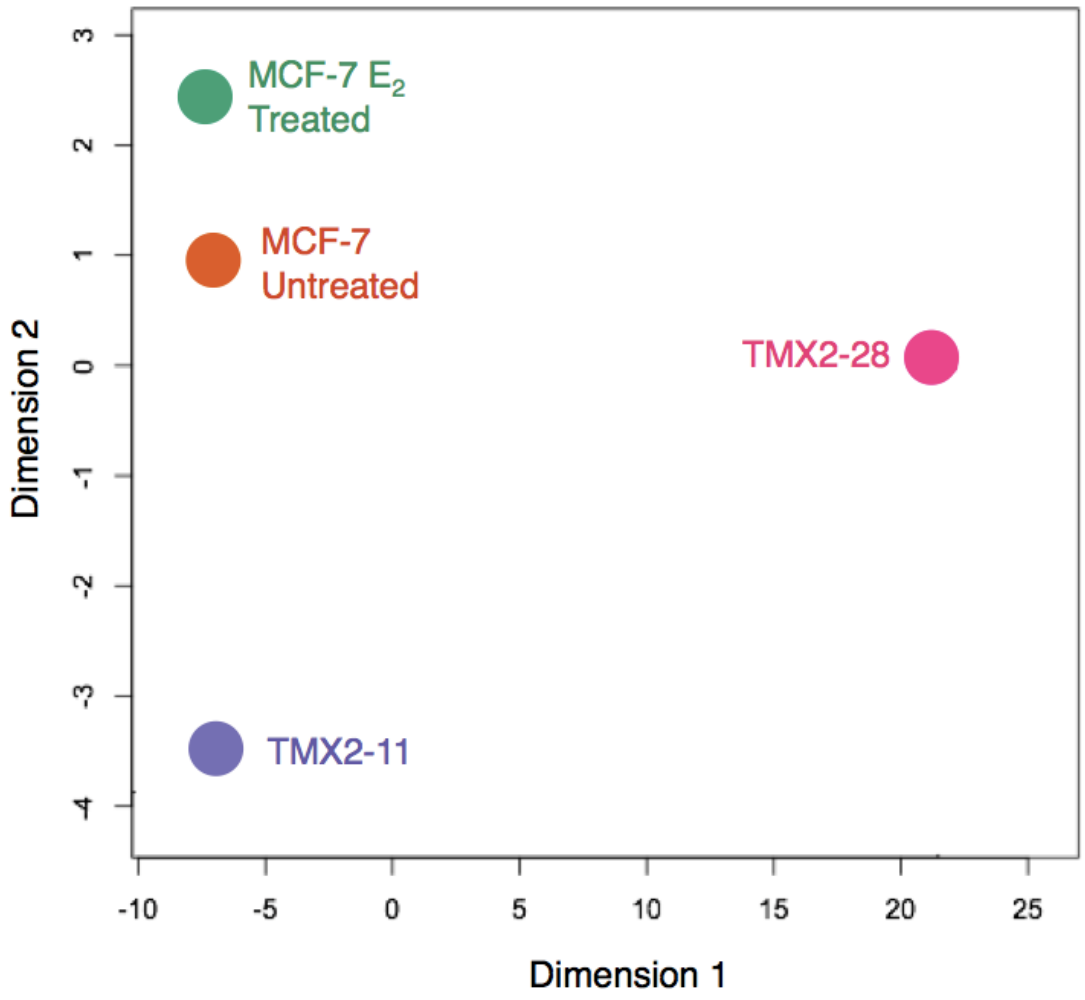


Figure 2.1. Visual representation of DNA methylation among the breast cancer cell lines.

Methylation beta values of the 1000 CpG sites that varied the most among the four groups were used with the Minfi package for R to prepare the multidimensional scaling (MDS) plot. Dimension 1 and 2 represent arbitrary distances among samples.

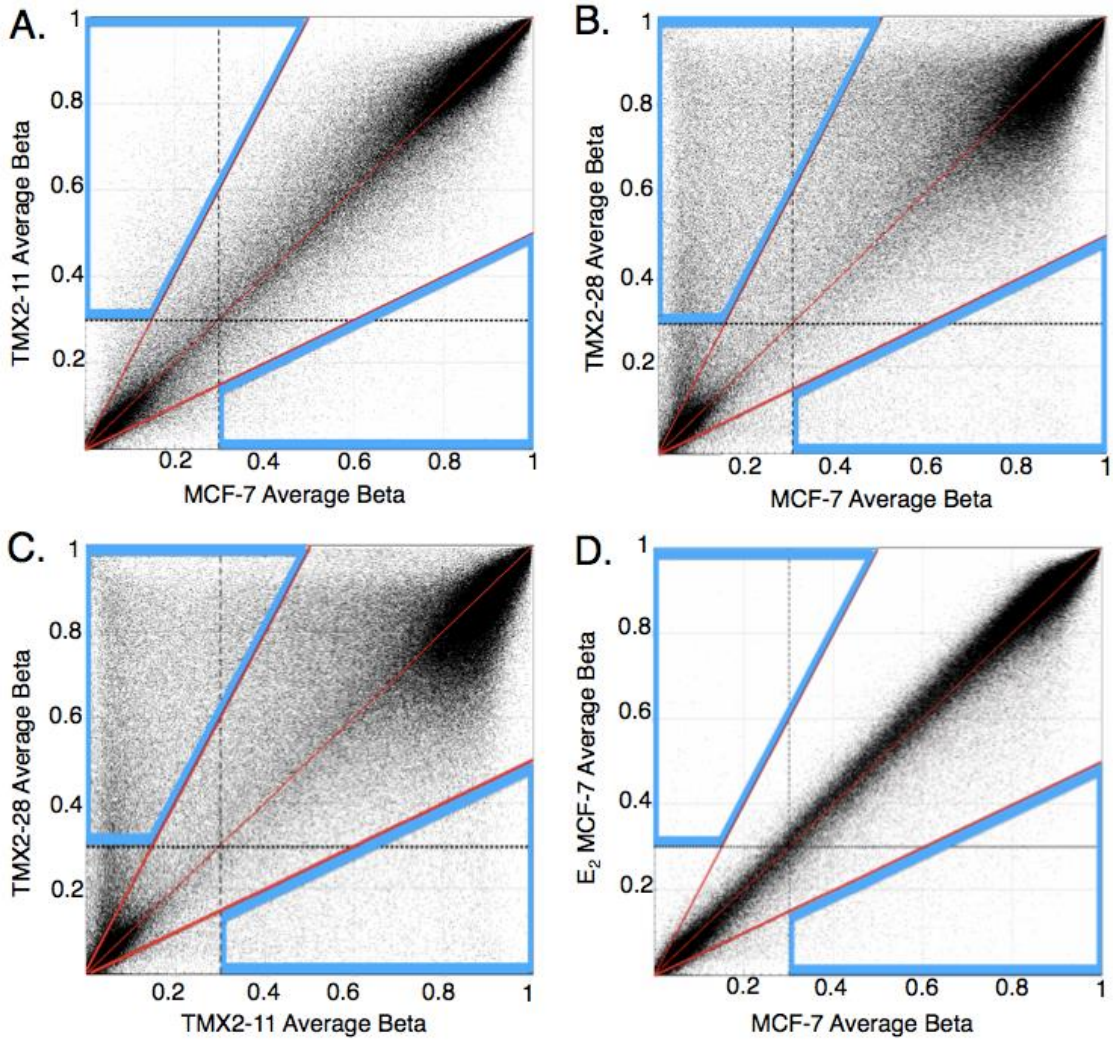


Figure 2.2. Scatter plots indicate genome-wide methylation changes in Tamoxifen-resistant lines compared with the parental.

TMX2-11 (A), TMX2-28 (B) and MCF-7 treated with 10^{-10} M E2 for 14 days (D) were compared with the parental line, MCF-7 and the Tamoxifen-resistant clones TMX2-11 and TMX2-28 (C) were compared against each other using GenomeStudio to determine the overall changes in methylation. Dashed lines mark the average beta cut-off value of 0.3 for each sample; center red line represents equal beta values in the two samples; outer red lines mark the two-fold change in average beta values for each sample; blue boxes enclose all CpG sites with average beta values >0.3 and a >2-fold change in methylation.

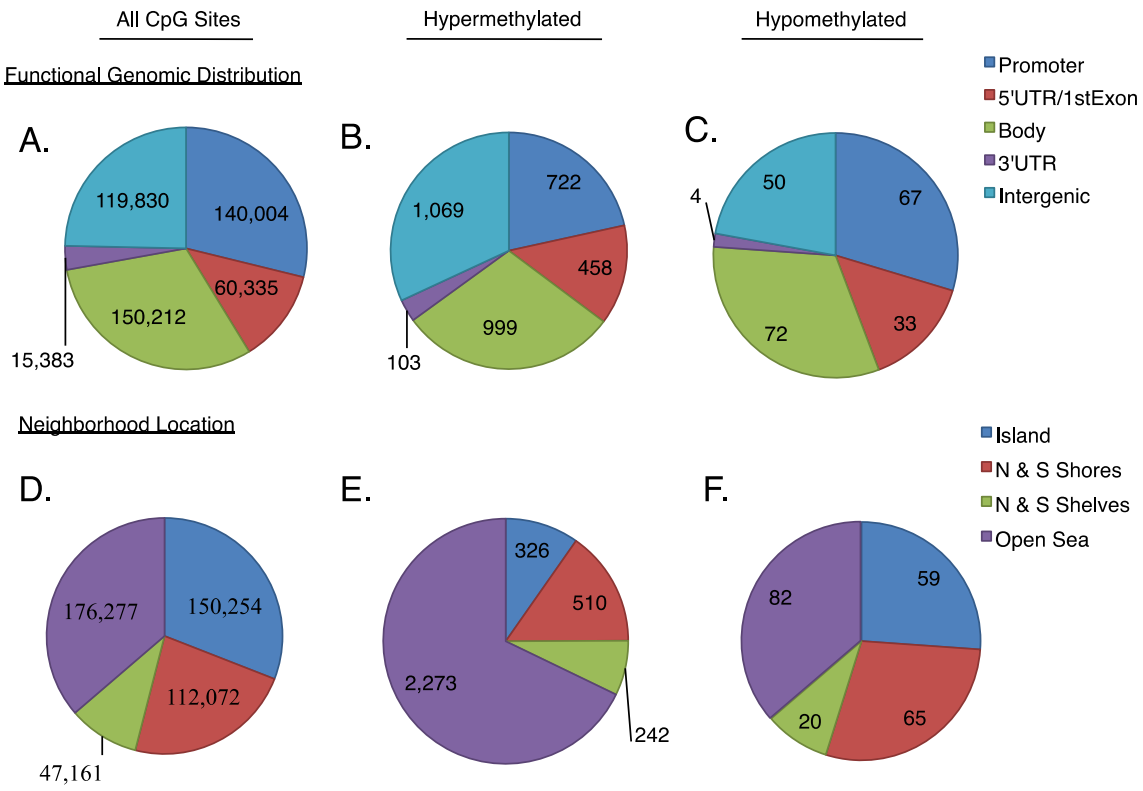


Figure 2.3. Location of aberrantly methylated CpG sites shared between TMX2-11 and TMX2-28.

Functional genomic location of all CpG sites on the BeadChip (A) hypermethylated (B) and hypomethylated (C) CpG sites. Neighborhood location of all CpG sites on the BeadChip (D) hypermethylated (E) and hypomethylated (F) CpG sites. Promoter is TSS200 and TSS1500 regions of the gene; Intergenic regions are undefined locations in GenomeStudio; shores, located 0-2 kb and shelves, 2-4 kb from the canonical CpG island; open sea is defined as the remainder of the sequence.

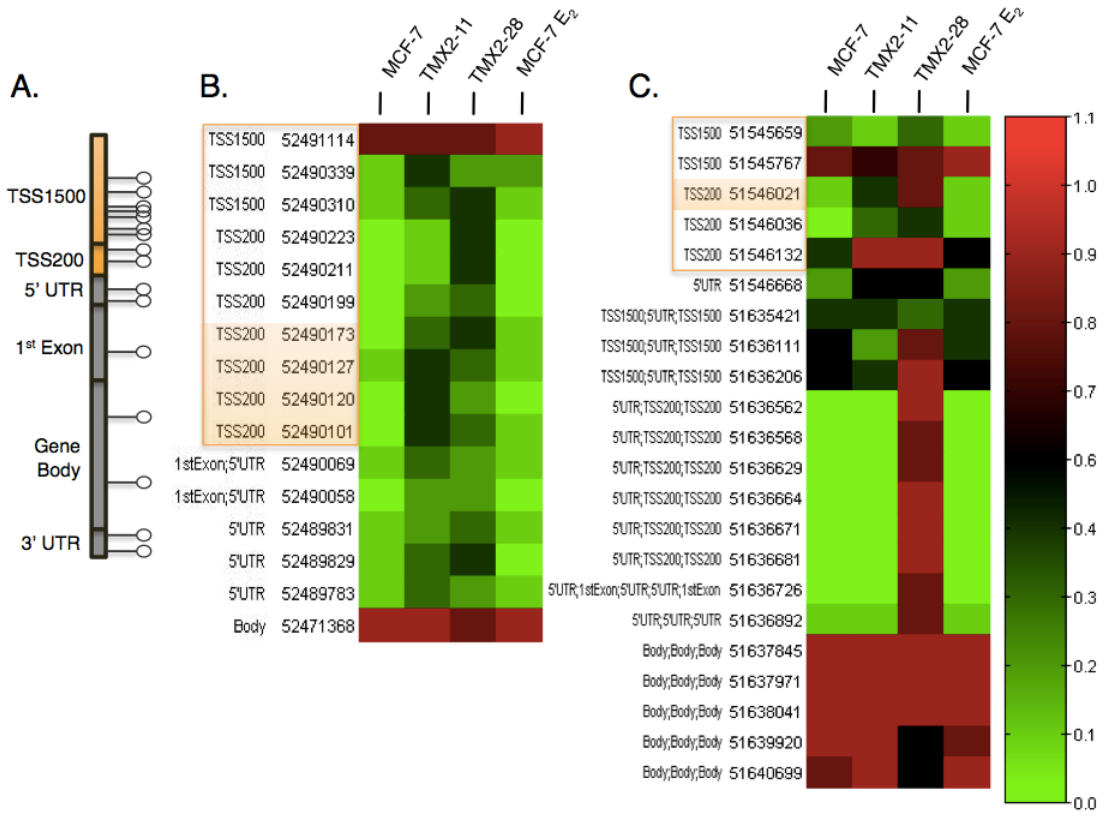


Figure 2.4. ZNF350 and MAGED1 are differentially methylated in Tamoxifen resistant cells.

A) A map of typical CpG site locations in a gene. Promoter region (TSS200 and TSS1500) is shown in orange. B) ZNF350 C) MAGED1 heat maps show average beta values of CpG sites interrogated across the gene. Functional genomic location and MAPINFO coordinate are shown for each CpG site. Orange boxes around MAPINFO and functional genomic location show promoter region CpG sites. Orange highlighted CpG sites indicate location of pyrosequencing primers. Average beta value is represented by the scale on right with the highest methylation value (1) in red and the lowest (0) in green.

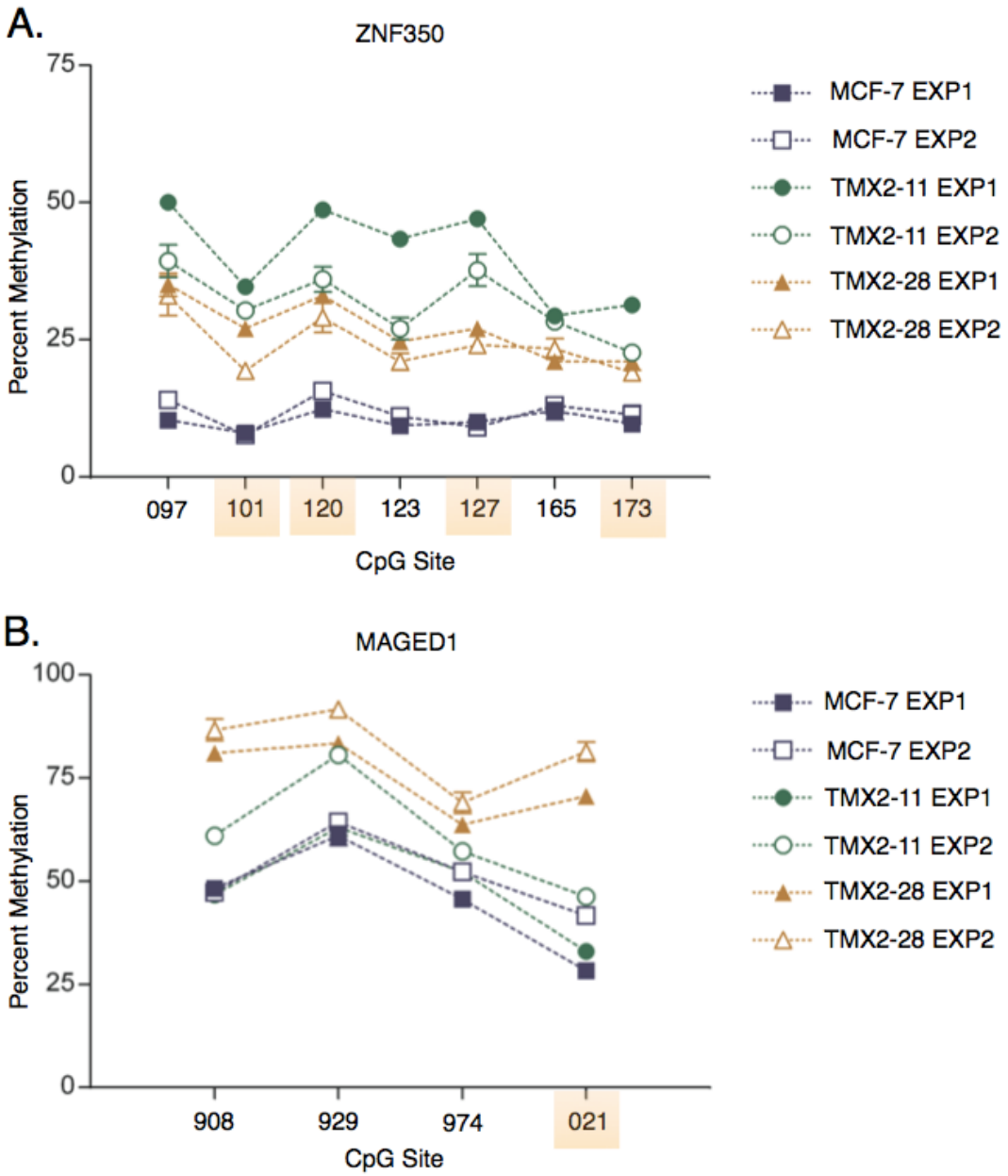


Figure 2.5. CpG site methylation of ZNF350 and MAGED1 in Tamoxifen-resistant and parental cell lines.

CpG sites in the TSS200 region of A) ZNF350 and B) MAGED1 were analyzed by pyrosequencing. CpG sites present on the BeadChip are highlighted in orange. Two experiments conducted 9 months apart demonstrate the permanence of methylation changed: Filled symbols indicate Experiment 1 (Exp1) and open symbols indicate Experiment 2 (Exp2). Each experiment consisted of three biological replicates for each cell line.

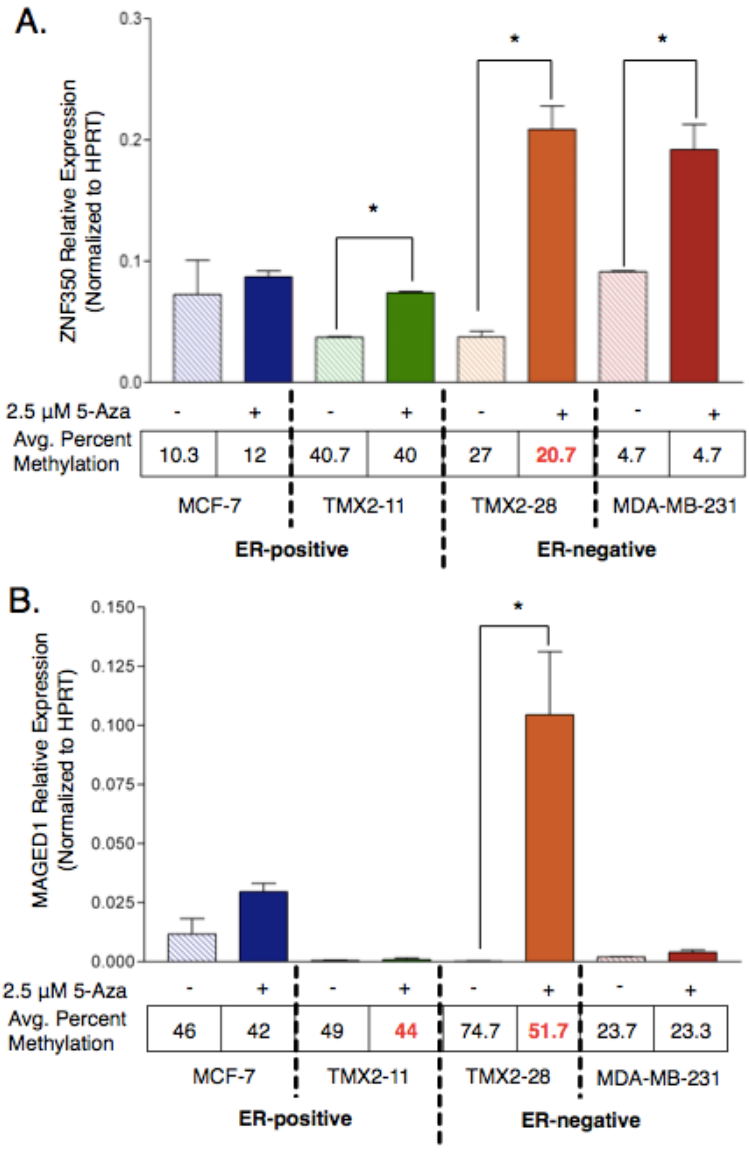


Figure 2.6. Comparison of gene expression and promoter methylation in ZNF350 and MAGED1.

Relative mRNA levels measured by qRT-PCR and average percent methylation of the TSS200 regions measured by pyrosequencing of A) ZNF350 and B) MAGED1 in control cultures and cultures treated with 5-Aza for four days. A) Treatment with 5-Aza resulted in significantly increased expression of ZNF350 in TMX2-11, TMX2-28 and MDA-MB-231; however, only TMX2-28 showed a corresponding significant decrease in promoter methylation (see text). B) Treatment with 5-Aza resulted in a significant increase in expression of MAGED1 in TMX2-28, however a significant decrease in methylation was observed in both TMX2-11 and TMX2-28 (see text). Comparisons were made on results from triplicate biological samples using an unpaired student's T-test; * = $p < 0.05$ and ** = $p < 0.01$. Two independent experiments were conducted nine months apart with similar results; results from experiment 1 are shown.

CHAPTER 3

FURTHER CHARACTERIZATION OF GENES ABERRANTLY METHYLATED AND DIFFERENTIALLY EXPRESSED IN TAMOXIFEN-RESISTANT CELL LINES

Introduction

The HumanMethylation450 BeadChip (HM450BC) interrogates 485,764 CpG sites in the human genome, 92,000 of which are located in both the proximal promoter (within 200-1,500 bp upstream of the transcription start site) and a CpG island [51]. In Chapter 2, it was shown that the Tamoxifen-selected cell lines, TMX2-11 and TMX2-28 share 3,000 hypermethylated and 200 hypomethylated CpG sites as compared with MCF-7. Two of the hypermethylated genes, MAGED1 and ZNF350, were discussed in more detail. Treatment with 5-Aza was found to increase expression of both MAGED1 and ZNF350 in the ERα-negative cell line TMX2-28, however only expression of ZNF350 was increased in the ERα-positive, TMX2-11 cells.

It is well known that treatment with the demethylase, 5-Aza restores expression of genes that are hypermethylated in cancer by inhibiting DNMT1 [30, 74-76]. Clinical trials targeting aberrantly methylated DNA using demethylases including 5-Aza are established for cancers of the blood, such as Acute Myeloid Leukemia (AML). Treating patients with 5-Aza has been shown to both successfully combat the disease and prolong patient life, especially in the case of AML, a disease that can result from the hypermethylation of genes in white blood cells [77-79]. Currently there are six clinical

trials using 5-Aza to target methylation in breast cancer (clinicaltrials.gov) [80, 81]. I postulated that long-term exposure to Tamoxifen causes alterations in the methylation profile of breast cancer cells therefore, treating cells with a combination of 5-Aza and Tamoxifen would reverse the cancer phenotype and restore sensitivity to Tamoxifen.

Using methyl donors such as S-adenosylmethionine (SAM) to reverse the cancer phenotype is not as conventional as using demethylating agents, however, Pakneshan et al. showed that treatment with SAM blocked proliferation of MDA-MB-231 cells [82]. This was accomplished by methylating urokinase (uPA), a gene implicated in the progression of breast cancer, using SAM [82]. Additionally, a study using human gastric and colon cancer cell lines showed that treatment with SAM heavily methylated the previously hypomethylated promoter regions of c-myc and H-ras [83]. This study also demonstrated that non-cancerous cells were not affected by SAM treatment [83].

I hypothesized that treatment of Tamoxifen-resistant breast cancer cells with a combination of 5-Aza and SAM would reverse the aberrantly methylated CpG sites, both hypermethylated and hypomethylated, and decrease proliferation of the rapidly growing TMX2-28 cells.

In this chapter, I present results from a series of experiments aimed at reversing the aberrant methylation associated with prolonged Tamoxifen treatment. Sixteen hypermethylated genes that were selected for further expression and methylation analysis are discussed. Twelve genes as described in Table 3.1 were hypermethylated in at least two CpG sites in the promoter region of both TMX2-11 and TMX2-28. Genes involved in signal transduction, cell adhesion, transcriptional repression, inflammatory response, cell proliferation and hormone response were chosen for analysis (see Table 3.1).

Two genes were chosen for their involvement in the signal transduction pathway; A kinase (PRKA) anchor protein 12 (AKAP12), a scaffold protein and regulator cell growth [84] and Bone Morphogenic Protein 2 (BMP2), a member of the cytokine family of proteins identified for their involvement in cartilage and bone formation and a positive regulator of the WNT pathway [85]. Collagen Type VI, Alpha 5 (Col29A1 or Col6A5) was selected for its association with cell adhesion and extracellular matrix formation [86]. Two growth regulators were chosen for their involvement in breast cancer; Regenerating islet-derived 1 (REG1A), a growth factor and indicator of prognosis in multiple cancer types including breast [87] and protachykinin 1 (TAC1), a neuropeptide with anti-proliferative and anti-apoptotic effects found to be differentially methylated in DCIS and IDC breast tumors as compared with normal tissue [88, 89].

Half of the genes selected for further analysis are transcriptional regulators in the cell. Three of the genes are involved in transcriptional repression. Growth factor independent 1 transcription factor (GFI1), is involved in hematopoiesis [90], histone deacetylase 9 (HDAC9), is a transcriptional regulator responsible for deacetylation of lysines on core histones and ultimately epigenetic repression in the cell [91] and transcription factor 12 (TCF12), is a transcriptional repressor that is upregulated in metastatic colon cancer [92]. Two genes act as tumor suppressors in breast cancer; Leucine zipper putative tumor suppressor 1 (LZTS1), a tumor suppressor associated with increased cell motility and invasion when downregulated in poor prognosis in breast cancer [93] and RAR-related orphan receptor A (RORA), a regulator of gene expression of various cellular functions and a putative tumor suppressor in breast cancer [94]. Finally, structure specific recognition protein 1 (SSRP1), the nuclear DNA binding

domain of histone chaperone p80, is a protein involved in nucleosome reorganization during transcription, elongation and DNA repair [95].

Four genes were selected for their reduced expression in TMX2-28 in a previously completed expression array [47, 50]. Of these genes caveolin 2 (CAV2), a membrane component of caveolae, vesicles involved in cellular transport is involved in signal transduction, lipid metabolism and cell growth [96]. The gene, growth regulation by estrogen in breast cancer 1 (GREB1), is an ER co-activator required for ER-positive breast cancer cell growth and proliferation and shown to be a good clinical outcome predictor [97]. Tumor growth factor beta 2 (TGFβ2) is a cytokine involved in breast cancer and overexpression of the protein has been connected with progression of the disease [98]. Lastly, Thrombospondin 1 (THBS1) is a glycoprotein involved in cellular functions related to tumor progression and metastasis and mediates cell-cell and cell-matrix interactions [99].

Gene expression and promoter methylation was examined in MCF-7, TMX2-11 and TMX2-28 both with and without treatment with 5-Aza. Treatment with the drug resulted in an increase in expression of five genes in TMX2-28: RORA, BMP2, THBS1, CAV2, and TGFβ2. Additionally, cell lines were treated with 5-Aza, Tamoxifen, or a combination of the drugs and assessed for proliferation. Proliferation was unaffected in MCF-7 and TMX2-11 after 5-Aza or Tamoxifen treatment, however, 5-Aza caused a significant decrease in proliferation in TMX2-28. Lastly, treatment of MDA-MB-231 cells with the methyl donor SAM, did not decrease expression or increase methylation of uPA, a gene shown previously to be methylated by SAM in vitro, indicating that

treatment with the molecule was ineffective [82]and therefore no further experiments with SAM were conducted.

Materials and Methods

Cell Culture, RNA and DNA Purification

TMX2-11 and TMX2-28 were kindly provided by John Gierthy (Wadsworth Center Albany, NY). MCF-7 cells were purchased from the American Type Culture Collection (ATCC). Cell lines were grown in Dulbecco's modified eagle medium (without phenol red) supplemented with 5% cosmic calf serum (Hyclone), 2.0 mM of L-glutamine, 0.1 mM of nonessential amino acids, 10 ng/mL of insulin, 100 units/mL of penicillin, and 100 µg/mL of streptomycin. Cells were maintained at 37°C with 5% CO_2 in a humidified incubator and media was changed every 2 days. MCF-7 cells were cultured with and without 10^{-10} M E2 (Sigma-Aldrich) added to the media for 14 days.

RNA was purified in triplicate for each cell line using TriReagent (Molecular Research Center, Inc) and DNA was purified using QIAamp DNA Mini kit (Qiagen) as per manufacturer suggestion and protocols previously described [47, 50]. Purified RNA and DNA samples were quantified using a NanoDrop 8000 (Thermo Scientific).

5-Aza-2'deoxycitidine (5-Aza) and Tamoxifen Treatment of Cells

Cells were seeded into 6-well plates at varying concentrations (MCF-7 and TMX2-11: 150,000 cells/well; TMX2-28 and MDA-MB-231: 100,000 cells/well) and allowed to attach overnight at 37°C and 5% CO_2. Two experiments were completed 9 months apart. Triplicate replicate wells were treated with either 0.1% DMSO (vehicle

control), 2.5 µM 5-Aza (Sigma-Aldrich) in 0.1% DMSO, 10^{-6} M Tamoxifen in 0.1% DMSO, or a combination of 5-Aza and Tamoxifen for 72-hours, refreshing the media every other day. After 72-hours, DNA and RNA were purified from the cells as described above and concentration and quality were evaluated using the NanoDrop 8000 (Thermo Scientific).

S-adenosylmethionine (SAM) Treatment of MDA-MB-231 Cells

MDA-MB-231 cells were seeded into 12-well plates at 50,000 cells/well and allowed to attach overnight at 37°C and 5% CO_2. Triplicate replicate wells were treated with either 0.1% EtOH (vehicle control) or 100 µM SAM (New England Biolabs) for 6 days, refreshing the media every other day. DNA and RNA were purified using the methods described previously and concentration and quality were assessed using the NanoDrop 8000 (Thermo Scientific).

Quantitative Real Time Reverse Transcriptase-PCR (qRT-PCR)

Primers for qRT-PCR were created using the UCSC RefGene Accession number associated with the CpG site of interest on the HM450BC and designed to span an exon-exon junction using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) for the following genes: AKAP12, BMP2, COL29A1, GFI1, HDAC9, IL7F, LZTS1, REG1A, RORA, SSRP1, TAC1, TCF12, CAV2, GREB1, and TGFβ (Table 3.2). Additionally, the following genes, THBS1, ERα, and uPA, were designed to span an exon-exon junction using Primer-BLAST but were not designed around a CpG site on the HM450BC (Table 3.2). All primers were purchased from Integrated DNA Technologies,

Inc. qRT-PCR analysis was carried out as previously described on a Roche LightCycler using the Qiagen OneStep RT-PCR kit (Qiagen) and SYBR green I nucleic acid stain (Invitrogen) [47, 50]. Total RNA (75 ng) was combined with OneStep RT-PCR master mix, dNTPs, SYBR green (2X), and primers (25 μM each) described above in chilled capillaries (Roche). RNA was reverse transcribed for 30 minutes at 50°C and subsequent amplification was assayed for 45 cycles (denaturation: 95ºC for 15 sec; annealing: 60ºC for 15 sec; extension 72ºC for 30 sec) using fluorescence generated by intercalating SYBR green dye into the resulting DNA product. Relative mRNA expression levels were normalized to hypoxanthine ribosyltransferase (HPRT) as described previously [47].

Pyrosequencing

DNA (1 μg) was bisulfite treated using the EpiTect Bisulfite kit (Qiagen) and PCR Primers were designed using the Pyromark Assay Design Software (Qiagen). One μL of bisulfite treated DNA was amplified using the Pyromark PCR kit (Qiagen) in a BioRad MyCycler and gene specific primers for ESR1, THBS1, RORA, and uPA were purchased from Integrated DNA Technologies, Inc. (Table 3.3). Pyrosequencing primers were designed to span a region of the promoter where the HM450BC showed hypermethylation in at least two CpG sites in both TMX2-11 and TMX2-28, except in the case of THBS1, which was hypermethylated only in TMX2-28. Additional CpG sites not analyzed by the BeadChip were assessed in the pyrosequencing assay due to their proximity to the CpG sites of interest. Single stranded products were prepared for pyrosequencing by PyroMark vacuum prep tool (Biotage). Pyrosequencing reactions were performed using a Pyromark Q24 system (Biotage) and manufacturers protocol

(Qiagen). Data were analyzed using Pyromark Q24 Software for percent methylation at the CpG sites interrogated.

Data Analysis

GraphPad Prism (GraphPad Software Inc.) was used to analyze and graph the biological replicate statistical results from pyrosequencing and qRT-PCR and to calculate a Pearson correlation coefficient for HM450BC and pyrosequencing data. Unpaired student's t-test with a p-value of <0.05 were considered statistically significant.

Results

To assess the extent to which methylation controls gene expression in Tamoxifen-resistant cell lines, a panel of 16 genes were analyzed by qRT-PCR. A summary of the differentially methylated CpG (dmCpG) sites hypermethylated on the HM450BC and the mean average beta value of the dmCpG sites for each gene can be seen in Table 3.4 and 3.5. Twelve of the 16 genes, AKAP12, BMP2, COL29A1, GFI1, HDAC9, IL17F, LZTS1, REG1A, RORA, SSRP1, TAC1, and TCF12, were selected out of a pool of 3,130 genes that met the following criteria: had an average β-value of >0.3 in TMX2-11 and TMX2-28, a >2-fold change in methylation in both TMX2-11 and TMX 2-28 as compared to the parental line MCF-7, a detection p-value of <0.01 on the HM450BC, and at least 2 hypermethylated CpG sites in either the TSS200 or TSS1500 region (Table 3.4).

An additional four genes, CAV2, GREB1, TGFB2, and THBS1 were chosen from a list of downregulated genes on a cDNA microarray analysis comparing TMX2-28 with MCF-7 [50]. Additionally, these genes matched the HM450BC criteria mentioned above

for TMX2-28, but not TMX2-11, as compared with MCF-7. All of the genes examined here have been shown previously to be involved in metastasis, cell proliferation, transcriptional repression, or apoptosis, and for seven of the fourteen genes, expression in cancer is known to be affected by methylation (Table 3.5).

AKAP12, REG1A, and HDAC9 have low expression in MCF-7 and Tamoxifen-resistant cell lines

Three genes, AKAP12, REG1A, and HDAC9, had extremely low expression levels in MCF-7, TMX2-11, and TMX2-28 (Figure 3.1). Of these genes, REG1A had significantly higher expression in TMX2-11 as compared with MCF-7 ($p=0.0131$), however expression remained extremely low (0.000026 and 0.000004 respectively) (Table 3.4). Two additional genes, Col29A1 and TAC1, had no expression in any of the cell lines (see Table 3.4).

Decreased expression of GREB1 and increased expression of LZTS1

No expression differences were seen between the Tamoxifen-resistant cell lines and MCF-7 in SSRP1, GFI1 or TCF12 (Figure 3.2A-C). Treatment of MCF-7, TMX2-11, TMX2-28, and the ER-negative cell line MDA-MB-231 with 5-Aza did not result in an expression change in these genes. An increase in expression of TCF12 was seen in the 5-Aza treated MCF-7 as compared with the control, but this increase was not significant ($p=0.1295$). Similarly, a non-significant decrease in expression was seen in GFI1 5-Aza treated MCF-7 as compared to control ($p=0.2556$). In contrast, significant changes in expression were observed in GREB1 and LZTS1. A significant decrease in expression of

GREB1 as compared to the control MCF-7 was seen in both TMX2-11 and TMX2-28 ($p=0.0001$ for both), however 5-Aza treatment did not result in an expression increase (Figure 3.2D). Additionally, a non-significant increase in expression of GREB1 is seen in 5-Aza treated MCF-7 as compared with the control ($p=0.1126$). LZTS1 had a significant increase in expression in the control TMX2-28 samples as compared with the MCF-7 control ($p=0.0037$), however 5-Aza treatment did not alter expression further (Figure 3.2E).

Expression of TGFβ2 and CAV2 increases in TMX2-28 cells with 5-Aza treatment

Compared with MCF-7, expression of TGFβ2 decreased significantly in TMX2-28 cells ($p=0.009$), however no change was seen in TMX2-11 cells (Figure 3.3A). Treatment with 5-Aza resulted in a significant increase in expression in MCF-7 ($p=0.0191$) and MDA-MB-231 ($p=0.048$) (Figure 3.3A), but not TMX2-11. Expression of TGFβ2 increased significantly in 5-Aza treated TMX2-28 as compared with the control, yet expression of the gene remained low ($p=0.0002$) (Figure 3.3A).

CAV2 had a significant decrease in expression in TMX2-28 cells as compared with MCF-7 ($p=0.0111$) (Figure 3.3B). A slight, but non-significant decrease was also seen in TMX2-11 as compared with MCF-7 (Figure 3.3B). Treatment with 5-Aza significantly increased expression of the gene in both TMX2-28 ($p=0.0013$) and MDA-MB-231 ($p=0.0004$), but did not affect the expression of the gene in TMX2-11 (Figure 3.3B).

RORA is hypermethylated in Tamoxifen-resistant lines

RORA was hypermethylated in 40 out of 111 CpG sites on the HM450BC with the majority of the sites being located in the body of the gene (Figure 3.4A). Seven CpG sites were located within the promoter region in four of the six transcript variants of the gene (Figure 3.4A). A total of seven promoter CpG sites in transcript variant 4 were chosen for further interrogation by pyrosequencing, two of which, 60885258 and 60885322 (Figure 3.4A orange highlighted CpG sites in orange box) were included on the HM450BC. A significant increase in the average percent methylation from MCF-7 (22%) was seen in both TMX2-11 (34.5%) and TMX2-28 (38.8%) ($p=0.0001$ for both cell lines). Treatment with 5-Aza decreased methylation 26% (from 51.8% to 38.8%) in TMX2-28 in the region analyzed by pyrosequencing, however no significant change in expression occurred ($t=0.8706$ and $p=0.4044$) (Figure 3.4B). Further investigation shows that the methylation decrease seen in 5-Aza treated TMX2-28, was not specific to a few CpG sites; all seven CpG sites interrogated by pyrosequencing had decreased methylation (Appendix F, panel A). No change in methylation was seen in TMX2-11 (Figure 3.4B). Low expression of RORA was seen in all cell lines and while 5-Aza treatment increased expression in TMX2-28, it was not significant (Figure 3.4B).

THBS1 expression and methylation are affected by 5-Aza treatment in TMX2-28 cells

THBS1 was downregulated in TMX2-28 in cDNA microarray analysis previously conducted [50]. DNA methylation was evaluated on the HM450BC and half of the promoter CpG sites for THBS1 were hypermethylated in TMX2-28 cells (Figure 3.5A).

Two promoter CpG sites, including one site interrogated by the HM450BC, were analyzed by pyrosequencing and a significant increase in methylation (93%, p=0.0001) was seen in TMX2-28 (54.3%) as compared with MCF-7 (3.3%, Figure 3.5 B; Figure 3.5A, orange highlighted CpG in orange box). Methylation of both TMX2-11 and MCF-7 was low (5.5% and 3.3% respectively) and 5-Aza treatment resulted in a slight increase in methylation of both of these cells lines (10.8% and 8.3% respectively). Interestingly, there was a significant 34% decrease in methylation of TMX2-28 after 5-Aza treatment (p=0.0001). This decrease was seen in both CpG sites probed by pyrosequencing in 5-Aza treated TMX2-28 cells (Appendix F, panel B). qRT-PCR analysis confirmed cDNA microarray results showing expression of THBS1 in TMX2-28 was significantly decreased as compared to MCF-7 (p=0.0001). Treatment of the cells with 5-Aza significantly increased expression, however expression remained exceedingly low (p=0.0001). Expression of THBS1 was significantly increased in ER-negative, MDA-MB-231, after treatment with 5-Aza despite no change in methylation (p=0.0001).

Expression and methylation of BMP2 are increased in TMX2-28 cells

The HM450BC revealed that BMP2 was hypermethylated in four out of five promoter CpG sites (40-60% methylation) in both TMX2-11 and TMX2-28 (Figure 3.6A, orange box). qRT-PCR analysis showed a significant increase in expression of BMP2 in TMX2-28 control as compared with the MCF-7 control (p=0.0001) (Figure 3.6B). Expression of BMP2 in TMX2-11 was low and comparable with expression levels in MCF-7 (Figure 3.6B). Treatment with 5-Aza did not significantly affect expression of BMP2 in either TMX2-11 or TMX2-28 (p=0.7229) (Figure 3.6B). Pyrosequencing of

BMP2 was attempted, however due to the overabundance of CpG sites within the promoter region no primer set could be designed using the PSQ Assay Design Software.

5-Aza treatment does not result in re-expression of ERα in TMX2-28 or MDA-MB-231

Previous results from Gozgit et al., Fagan-Solis et al., and Fasco et al. show that TMX2-28 cells do not express ERα mRNA or protein, while TMX2-11 has increased expression of the gene [47-49]. HM450BC methylation results showed hypermethylation in 26 out of 38 promoter CpG sites among 4 different gene transcripts (Appendix G, blue boxes). To determine whether the decrease in expression in TMX2-28 was due to methylation of the gene, MCF-7, TMX2-11, TMX2-28 and MDA-MB-231 cells were analyzed for ERα expression changes after treatment with 5-Aza. TMX2-28 had a significant increase in mRNA expression of ERα, but the expression was exceedingly low in both control (0.00048) and 5-Aza treated (0.00078) samples ($p=0.036$) (Figure 3.7). No significant increase in expression was seen in MCF-7, TMX2-11 or the ER-negative line, MDA-MB-231 (Figure 3.7). Pyrosequencing of the control samples revealed low methylation in the promoter region in TMX2-11 and TMX2-28 (1.3 and 2.7% respectively) therefore methylation was not assayed in 5-Aza treated samples (Figure 3.7).

5-Aza treatment results in decreased cell proliferation of TMX2-28

I observed 3,000 hypermethylated genes in our Tamoxifen-resistant cell lines on the HM450BC and wanted to determine if Tamoxifen-sensitivity could be restored with short-term 5-Aza treatment. MCF-7, TMX2-11, and TMX2-28 cells were treated with 5-

Aza, Tamoxifen, a combination of 5-Aza and Tamoxifen or the vehicle control and analyzed by MTS assay for cell proliferation (cytotoxicity assay data shown in Appendix H). TMX2-28 was significantly affected by 5-Aza, which reduced cell proliferation by 35% ($p=0.0001$) and caused the cells to take on a rounder, flatter appearance (Figure 3.8). In contrast, 5-Aza did not affect the proliferation of TMX2-11 (Figure 3.8). As expected, treatment with Tamoxifen alone had no effect on either TMX2-11 or TMX2-28 (Figure 3.8). Treatment with a combination of 5-Aza and Tamoxifen concurrently had no effect on TMX2-11. Furthermore, it did not result in an additional decrease in the proliferation of TMX2-28 (43% decrease) (Figure 3.8). An unusual, but not unprecedented finding was that treatment with Tamoxifen resulted in a more rapid proliferation of MCF-7 as compared to control (136% increase, $p=0.0001$) (Figure 3.8).

To determine if treatment with a methylating agent would also affect proliferation, MDA-MB-231 cells were treated with S-adenosylmethionine (SAM) using the protocol described in Pakneshan et al., a cytotoxicity assay was completed to obtain a suitable concentration (Appendix I) [82]. The results showed no significant decrease in expression ($p=0.45$) or increase in methylation ($p=0.44$) of the uPA gene after 6 days of treatment with the molecule (Figure 3.9A & B). Upon further investigation, New England Biolabs (NEB) does not recommend using SAM to treat cells in vitro as it is neutral pH labile and unlikely to be active when it enters the cell (personal communication). Additionally, if methyltransferase activity in the cell is altered, the addition of SAM alone may not be effective. Therefore, this experiment was not completed using MCF-7, TMX2-11 and TMX2-28.

Discussion

In Chapter 2 overall methylation changes that occur between the parental cell line and the Tamoxifen-resistant cell lines were discussed. It was found that two clones, TMX2-11 and TMX2-28 derived from MCF-7 cells treated long-term with Tamoxifen resulted in methylation changes and a greater number were seen in the ERα-negative line, TMX2-28 as compared with the ERα-positive line, TMX2-11. The two Tamoxifen-resistant lines shared over 3,000 hypermethylated and 200 hypomethylated CpG sites with the promoter region representing 22% of the total. The significant increase in methylation suggests that DNA promoter methylation is playing a role in acquired Tamoxifen-resistance.

To address whether hypermethylation of CpG sites in the promoter affects mRNA expression in Tamoxifen-resistant cell lines, a panel of 18 hypermethylated genes involved in pathways previously shown to be affected by long-term Tamoxifen treatment (metastasis, cell proliferation, transcriptional repression, and apoptosis) were selected for qRT-PCR analysis. Two genes, MAGED1 and ZNF350 were examined earlier in Chapter 2. Five of the remaining 16 genes analyzed, AKAP12, REG1A, HDAC9, Col29A1 and TAC1 had little or no expression in MCF-7, TMX2-11 and TMX2-28 control samples. 5-Aza treated RNA was not analyzed for these genes as the purpose of this experiment was to determine whether expression of the Tamoxifen-resistant cell lines differed from MCF-7. Promoter methylation did not appear to affect expression of these genes as MCF-7 had low overall methylation on the HM450BC and did not express the genes.

An additional five genes, LZTS1, SSRP1, GFI1, TCF12, and GREB1 were expressed in all of the cell lines analyzed, however expression of SSRP1, GFI1, and

TCF12 did not differ between the Tamoxifen-resistant lines and the parental line. The tumor suppressor LZTS1 is downregulated in many cancers by promoter methylation and Wang et al found a decrease in expression to be correlated with an increase in invasive characteristics and poor prognosis in breast cancer [93]. Our results for MDA-MB-231 expression mimic those of Wang et al. who found that MDA-MB-231 cells express low levels of LZTS1 and ectopically expressing the gene inhibits cell migration, however no comparison was made to a non-invasive, non-migratory line such as MCF-7. We saw that expression of LZTS1 in the highly invasive MDA-MB-231 line was equivalent to the levels seen in MCF-7. Treatment of MDA-MB-231 cells with 5-Aza did not increase expression indicating that methylation does not play a role in controlling expression of the gene. Contrary to the finding of Wang et al. [93], that promoter methylation was not related to invasiveness in the ER-positive MDA-MB-231 cells, we found that both promoter methylation and expression of LZTS1was increased in our invasive ER-negative TMX2-28 as compared with the non-invasive MCF-7. However, 5-Aza treatment did not further increase expression of LZTS1 in TMX2-28. Therefore, the gene was not analyzed by pyrosequencing for changes in promoter methylation after 5-Aza treatment.

Comparison of data from a cDNA microarray previously run in our lab with the HM450BC data allowed for selection of a panel of 4 genes that were both downregulated and hypermethylated in TMX2-28 as compared with MCF-7 [50]. TGFβ2, a tumor suppressor frequently inactivated in breast cancer, has been previously found to be downregulated in a Tamoxifen-resistant, MCF-7 derivative when compared with the parental line [30, 98]. Our data resemble those of Fan et al. who showed that TGFβ2 was

downregulated in an ER-positive, Tamoxifen-resistant cell line [30]. We see that TGFβ2 is expressed at low levels in the ER-negative TMX2-28, but not the ER-positive TMX2-11. We also see that in TMX2-28 treatment with 5-Aza significantly increases expression of the gene. Similarly, CAV2 expression was found to be downregulated in the ER-positive, Tamoxifen-resistant cell line in the Fan et al. paper as well as in TMX2-28 and we saw that treatment with 5-Aza increased expression of the gene in TMX2-28. CAV2 has been investigated in breast cancer and decreased expression is associated with poor prognosis, and basal-like tumors [96]. Interestingly, a stronger association was seen in tumors lacking ER and PR, those expressing basal cytokeratin markers (5/6, 14, and 17), and those with higher proliferation rates, characteristics TMX2-28, but not TMX2-11 possesses. It is also noteworthy that the expression of both TGFβ2 and CAV2 is higher in the aggressive line, MDA-MB-231, than in MCF-7, and that the MDA-MB-231 cells behave similarly to TMX2-28 in response to 5-Aza treatment. This expression increase after 5-Aza treatment suggests the need for a case-based treatment regimen in breast cancer.

RORA (or RORα) is a steroid hormone receptor used in a wide range of transcriptional regulation processes in the body, from controlling inflammation to participation in carcinogenesis through angiogenesis [100]. The primary isoforms of RORA expressed in normal breast tissue are isoforms 1 and 4 and mRNA expression of RORA is downregulated in a number of breast cancer cell lines, including MCF-7 [101]. Our data show that mRNA expression of RORA transcript variant 4 in TMX2-28 is low, however expression is greater than that of MCF-7. Treatment of TMX2-28 with 5-Aza resulted in a non-significant upregulation of RORA expression. Methylation in the

promoter region of RORA transcript variant 4 was significantly increased in both TMX2-11 and TMX2-28 in comparison to MCF-7 and treatment with 5-Aza decreased methylation in TMX2-28, but not TMX2-11. Interestingly, the expression of RORA in TMX2-28 was similar to that of MDA-MB-231, another ERα-negative cell line, yet methylation in the control MDA-MB-231 cells was only 12% compared with 51.8% methylation in the TMX2-28 control. Treatment with 5-Aza did not affect the methylation or expression of MCF-7 or TMX2-11, suggesting that expression of the gene may not be directly controlled by promoter methylation in those cell lines.

A third gene downregulated in TMX2-28 on the cDNA microarray, THBS1, has been identified as downregulated in breast cancer cell lines and breast tumors [102, 103]. Methylation analysis of nine-breast cancer cell lines, including MCF-7 and MDA-MB-231, using bisulfite genomic sequencing showed no methylation of the gene [102]. Using pyrosequencing, we see that methylation of THBS1 in the promoter region of both MCF-7 and MDA-MB-231 cells is low (3.3% and 3.8% average respectively) and a slight, but non-significant increase in methylation is seen in TMX2-11 (5.5%). THBS1 is hypermethylated in TMX2-28 cells at 54.3% and treatment with 5-Aza significantly decreases promoter methylation and increases expression of the gene, indicating that methylation plays a role in gene expression. The extent of this role however is unknown and will need to be examined further as the expression increase, although significant, is very slight. A longer treatment with 5-Aza may be the key to determining whether demethylating THBS1 in TMX2-28 cells leads to expression levels similar to MCF-7 and confirming that methylation is playing a role in suppressing expression.

The final gene analyzed for expression changes based on methylation data from the HM450BC was BMP2. Recent research has shown that expression of BMP2 is frequently decreased in breast cancer, however downregulation is not associated with tumor type, and re-expression of the gene inhibits proliferation and encourages apoptosis in breast cancer cell lines [104, 105]. In this study, we found that the TSS200 region of BMP2 is more highly methylated in the Tamoxifen-resistant cell lines than the parental line in the HM450BC data. Unfortunately, the overabundance of CpG sites in the promoter region was prohibitive and we were unable to analyze the region of interest by pyrosequencing as primers could not be created. While pyrosequencing is considered a gold standard for analyzing a region of interest in the promoter, in Chapter 2 we compared pyrosequencing methylation percentages with beta values of individual CpG sites on the HM450BC and confirmed that the HM450BC methylation results are sufficient to provide an assessment of the area (Chapter 2, Figure 2.5). Interestingly, the expression of BMP2 in TMX2-28 is significantly higher than MCF-7 and is similar to that of ER-negative, MDA-MB-231. Treatment with 5-Aza does not affect expression levels in TMX2-28, however expression is significantly higher in the 5-Aza treated MDA-MB-231 as compared with the control. We see that the body of BMP2 has low methylation in MCF-7 and TMX2-11, whereas in TMX2-28 it is highly methylated. Emerging research suggests that body methylation may play a role in expression of the gene.

To confirm that TMX2-11 retained ERα expression and TMX2-28 lost expression of the gene qRT-PCR analysis was run. We saw that expression of the ERα is increased in TMX2-11 and that TMX2-28 lost expression of the gene, confirming what has

previously been shown in the literature [49]. Average methylation of the gene was less than 4% in the parental and Tamoxifen-resistant lines, as compared with the ERα-negative MDA-MB-231 cells, which had an average methylation of 16.3%. We also examined the cell lines post-treatment with 5-Aza and saw that TMX2-28 had a slight, but significant increase in expression. The knowledge that methylation of the gene was low in both Tamoxifen-resistant lines and that 72-hour 5-Aza treatment did not restore expression levels of TMX2-28 to those in MCF-7, suggests that expression of ERα in TMX2-28 is not directly controlled by methylation of the gene. Additional analysis spanning the promoter would give a more thorough assessment of the role that methylation is playing in expression of the gene, as the primer set used for pyrosequencing only analyzed a small region of the promoter and it was not a region included in the HM450BC.

To determine whether the Tamoxifen-resistant cell lines retained resistance to the drug, cells were treated for 72-hours with 10^{-6} M Tamoxifen and proliferation was measured. As expected, Tamoxifen had no effect on TMX2-11 and TMX2-28. Conversely, the drug increased proliferation of MCF-7 and decreased proliferation of MDA-MB-231. Previous studies have suggested that long-term culturing of cells in low estrogen media can elicit increased sensitivity to the weak estrogenic effects of Tamoxifen and result in increased proliferation [106, 107]. As our MCF-7 cells have been cultured in low-estrogen media for >7 years, further investigation of this effect is needed to determine if treating the cells with estrogen both short term and long term prior to 5-Aza and Tamoxifen treatment reduces the observed Tamoxifen-induced cell proliferation. Treatment of the cells with the demethylating agent 5-Aza had no effect on

the proliferation of TMX2-11, MCF-7 or MDA-MB-231 cells. Conversely, TMX2-28 had a decrease in proliferation, however the addition of Tamoxifen to the media concurrently did not strengthen the reduction. Generally, breast cancer cell lines are treated simultaneously with 5-Aza and Tamoxifen for proliferation assessment, however pre-treatment of our Tamoxifen-resistant lines with 5-Aza to initiate demethylation before adding Tamoxifen may be required as we did not see an additive effect.

While treatment with 5-Aza had an effect on methylation, expression and cell growth, SAM did not. SAM is thought to inhibit demethylation of the DNA by acting as a methyl donor for DMNTs to stimulate hypermethylation [108]. In clinical applications, SAM is prescribed as a dietary supplement to combat the effects of liver disease, depression, and colon cancer which often result from DNA hypomethylation [83, 109, 110]. In laboratory studies, SAM is used along with the CpG methyltransferase SssI to methylate DNA in vitro through a PCR reaction (NEB). However, SAM is unstable at pH 7.5 and 37ºC for an extended period, which results in the molecule being unsuitable for cell culture use (NEB, personal communication). This may explain why only a few studies have reported using SAM to methylate cell cultures in vitro [82, 83, 108]. To further assess the use of SAM as a methylating agent in TMX2-11 and TMX2-28, it is important to analyze the activity of DMNTs in the cell lines as altered activity could result in an imbalance in the methylation pathway.

Table 3.1 Summary of Genes and their functions

Genes chosen based on hypermethylation of TMX2-11 and TMX2-28 in HM450BC

Full Gene Name	Gene Acronym	RefSeq #	Function
A-kinase Anchor Protein	AKAP12	NM_144497	Signal transduction
Bone Morphogenic Protein 2	BMP2	NM_001200	Positive regulator of WNT pathway, bone and cartilage formation
Collagen, Type VI, Alpha 5	COL29A1	NM_153264	Epidermal collagen involved in cell adhesion and extracellular matrix formation
Growth Factor Independent 1 Transcription Repressor	GFI1	NM_001127215	Transcriptional repressor
Histone Deacetylase 9	HDAC9	NM_058176	Histone deacetylase and transcriptional repressor
Interleukin 17F	IL17F	NM_052872	Cytokine involved in inflammatory response and TGFB signaling pathway
Leucine Zipper, Putative Tumor Suppressor 1	LZTS1	NM_021020	Transcriptional regulation
Regenerating Islet-Derived 1 Alpha	REG1A	NM_002909	Regulator of cell proliferation
RAR-Related Orphan Receptor A	RORA	NM_134262	Transcriptional regulation and angiogenesis
Structure Specific Recognition Protein 1	SSRP1	NM_003146	Transcriptional regulation
Tachykinin, Precursor 1	TAC1	NM_013998	Encodes hormones that invoke vasodilation
Transcription Factor 12	TCF12	NM_207037	Transcriptional regulation

Genes chosen based on mRNA expression array

Full Gene Name	Gene Acronym	RefSeq #	Function
Caveolin 2	CAV2	NM_001233	Protein binding involved in signal transduction, lipid metabolism, cell growth and apoptosis
Growth Regulation By Estrogen In Breast Cancer 1	GREB1	NM_149803	Estrogen response gene in cell proliferation
Transforming Growth Factor, Beta 2	TFGB2	NM_001135599	Receptor signaling
Thrombospondin 1	THBS1	NM_003246	Mediates cell-cell and cell-matrix interactions

Table 3.2 qRT-PCR Primers

Genes Chosen Based on HM450K Hypermethylation		
Gene	**Forward Sequence 5'→3'**	**Reverse Sequence 3'→5'**
AKAP12	ATCTGGGGAAATGCATCCGC	TCTCTGTCCAACTGTGATGGTG
BMP2	AGAATAACTTGCGCACCCCA	ACCATGGTCGACCTTTAGGAG
Col29A1	CCACCCTCTGGATCATCACT	GTTTTCTGTGCCACCGTTCT
GFI1	GAGCGTGGCCACCGGCTGCACGCA	GCTCCGTTCCTGCGAGTGCACGGC
HDAC9	GAGCCCCAAATGAGGTTGGA	TGCCGTCACTTTGTACCCTC
IL17F	CTGCTGTCGATATTGGGGCT	AGTGACAGTGTAATTCCAGGGG
LZTS1	TTTGGACTGCTTCTCTCAGTTCCTGC	TTTGACAATGTGTTGCCCAACCAAAG
REG1A	GACCTCAAGCACAGGATTCCA	CCAGGTTGAGTTGAGTTGGAGA
RORA	CTTCTTTCCCTACTGTTCGTTC	GCTCTTCTCTCAAGTATTGGC
SSRP1	GTGCACCACAGGCAAGAATG	GCTTGGGTTCATGCCCTCTT
TAC1	ACTGTCCGTCGCAAAATCCA	AGCATCCCGTTTGCCCATTA
TCF12	TAATCGGGGTGGTTGGATGC	TGGGGAAAACATCGCACTGA
Genes Chosen Based on Expression Array		
CAV2	AGACGGAGAAGGCGGACGTA	ATTAAAATCCAGATGTGCAGACAGC
GREB1	GCTTAGCCTCTTGGCTGGTT	CACTCGGCTACCACCTTCTA
TGFB2	TCTTCCCCTCCGAAAATGCC	GCAATAGGCCGCATCCAAAG
THBS1	TTGTCTTTGGAACCACACCA	CTGGACAGCTCATCACAGGA
Other Genes of Interest		
ERα	ATGATCAACTGGGCGAAGAG	GATCTCCACCATGCCCTCTA
uPA	GCTCACCACAACGACATTGC	CACCTGCCCTCCTTGGAA

Table 3.3 Pyrosequencing Primers

Gene	Forward 5'→3'	Reverse 3'→5'	Sequencing
THBS1	BIOT-TGAAATTTAATAAATATGTGTTTTAGGAATAT	CCCATCTTAACACTTAAACCTAACAAAA	3'-ACATTCATCAAACAACAAT-5'
RORA	AGGTGTAGATTAGGATTTTGGTTATTGGTATA	BIOT-TCTAACCACTTTCTACCCCCACT	5'-GAGTTTTTTTAGAAAGAT-3'
ER□	GAGGTGTATTTGGATAGTAGTAAGTT	BIOT-CTATTAAATAAAAAAAAACCCC	5'-GAGGGYGTYGTTTAYGAGTTTA-3'
uPA	TGTTGGTGATATTTGGGGATTGTTATT	BIOT-CCCAACTCTAAAACCTCCTAAAATCCTCT	5'-TGTGTTTTTGGGAGAGTA-3'

Table 3.4 Overview of methylation and expression data from genes selected based on HM450BC hypermethylation in TMX2-11 and TMX2-28

	CpG site (MAPINFO)	HM450BC CpG Site Hypermethylation (Avg. Beta)			HM450BC Mean CpG Site Hypermethylation (Avg. Beta)			qRT-PCR (relative expression)					
								Untreated			5-Aza Treated		
Gene Acronym		MCF-7	TMX2-11	TMX2-28	MCF-7	TMX2-11	TMX2-28	MCF-7	TMX2-11	TMX2-28	MCF-7	TMX2-11	TMX2-28
AKAP12	151646552	0.335	0.747	0.834	0.364	0.841	0.872	0.00075	0.00093	0.0003			
AKAP12	151646540	0.393	0.935	0.911									
BMP2	6748710	0.075	0.458	0.474	0.075	0.471	0.491	0.0046	0.0037	0.196	0.0051	0.01	0.241
BMP2	6748730	0.029	0.602	0.650									
BMP2	6748719	0.127	0.347	0.388									
BMP2	6748712	0.071	0.476	0.452									
COL29A1	130063469	0.094	0.529	0.406	0.146	0.693	0.562	0	0	0	0	0	0
COL29A1	130064261	0.121	0.776	0.742									
COL29A1	130063505	0.221	0.773	0.538									
GFI1	92950836	0.113	0.443	0.667	0.105	0.515	0.729	0.114	0.029	0.024	0.024	0.029	0.02
GFI1	92950711	0.107	0.483	0.583									
GFI1	92950698	0.095	0.620	0.937									
HDAC9	18535232	0.221	0.601	0.796	0.248	0.602	0.806	0.00046	0.00068	0.0017			
HDAC9	18534872	0.387	0.825	0.807									
HDAC9	18535263	0.134	0.380	0.816									
IL17F	52109385	0.201	0.471	0.445	0.198	0.442	0.475	0.092	0.13	0.27			
IL17F	52109347	0.194	0.413	0.505									

CpG sites are hypermethylated CpGs in TMX2-11 and TMX2-28 as compared with MCF-7 in TSS200 and TSS1500

Table 3.5 Overview of methylation and mRNA expression of genes underexpressed in TMX2-28 from Affymetrix array

	CpG site (MAPINFO)	HM450BC CpG Site Hypermethylation (Avg. Beta)			HM450BC Mean CpG Site Hypermethylation (Avg. Beta)			qRT-PCR (relative expression)					
								Untreated			5-Aza Treated		
Gene Acronym		MCF-7	TMX2-11	TMX2-28	MCF-7	TMX2-11	TMX2-28	MCF-7	TMX2-11	TMX2-28	MCF-7	TMX2-11	TMX2-28
CAV2	116139322	0.093	0.097	0.517	0.129	0.101	0.664	0.103	0.044	0.0049	0.183	0.084	0.277
CAV2	116139367	0.184	0.159	0.748									
CAV2	116139374	0.194	0.135	0.834									
CAV2	116139391	0.103	0.104	0.663									
CAV2	116139414	0.097	0.061	0.645									
CAV2	116139425	0.103	0.051	0.575									
GREB1	11679605	0.222	0.309	0.800	0.121	0.139	0.873	0.067	0.018	0	0.131	0.057	0
GREB1	11679845	0.045	0.023	0.807									
GREB1	11679872	0.043	0.056	0.917									
GREB1	11679879	0.087	0.122	0.890									
GREB1	11680020	0.083	0.069	0.844									
GREB1	11680057	0.243	0.256	0.980									
TFGB2	218518468	0.101	0.092	0.938	0.044	0.045	0.677	1.612	0.824	0.0032	3.19	1.246	0.017
TFGB2	218518579	0.033	0.032	0.907									
TFGB2	218518675	0.034	0.036	0.905									
TFGB2	218518963	0.039	0.044	0.443									
TFGB2	218519232	0.013	0.022	0.193									
THBS1	39873209	0.006	0.013	0.015	0.040	0.032	0.080	0.053	0.052	0.00012	0.053	0.051	0.00028
THBS1	39873258	0.075	0.052	0.145									

CpG sites are hypermethylated CpGs in TMX2-28 as compared with MCF-7 in TSS200 and TSS1500

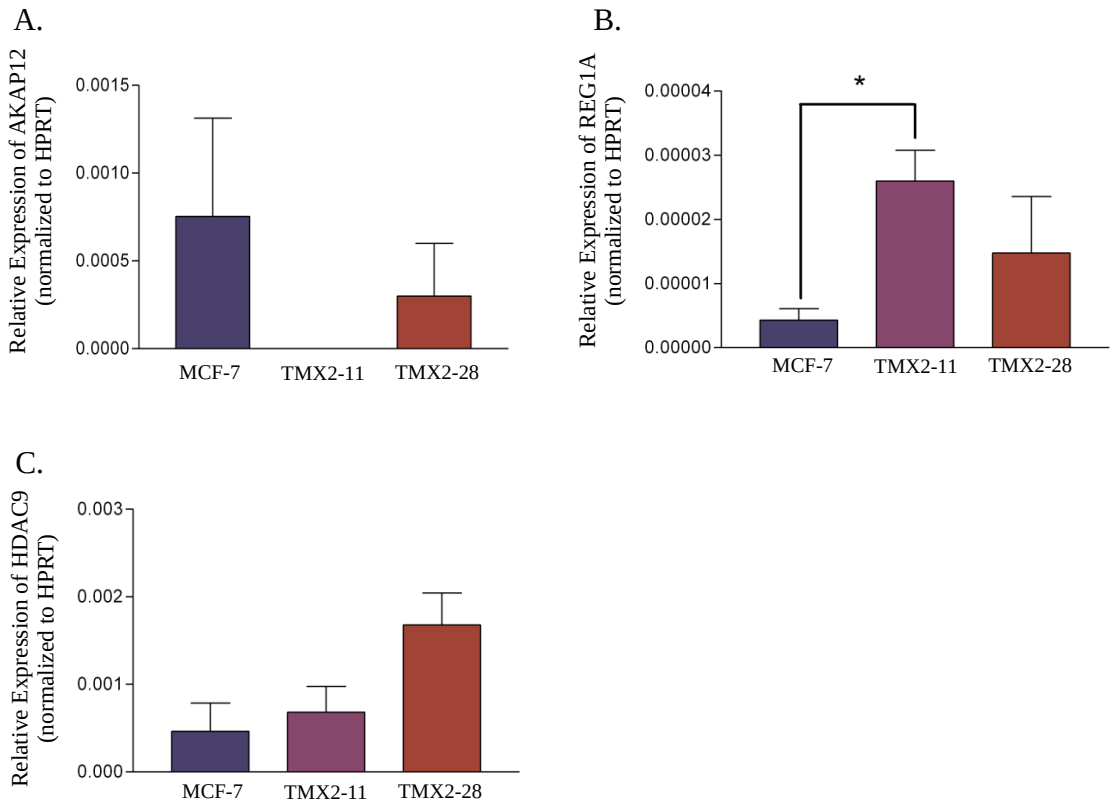
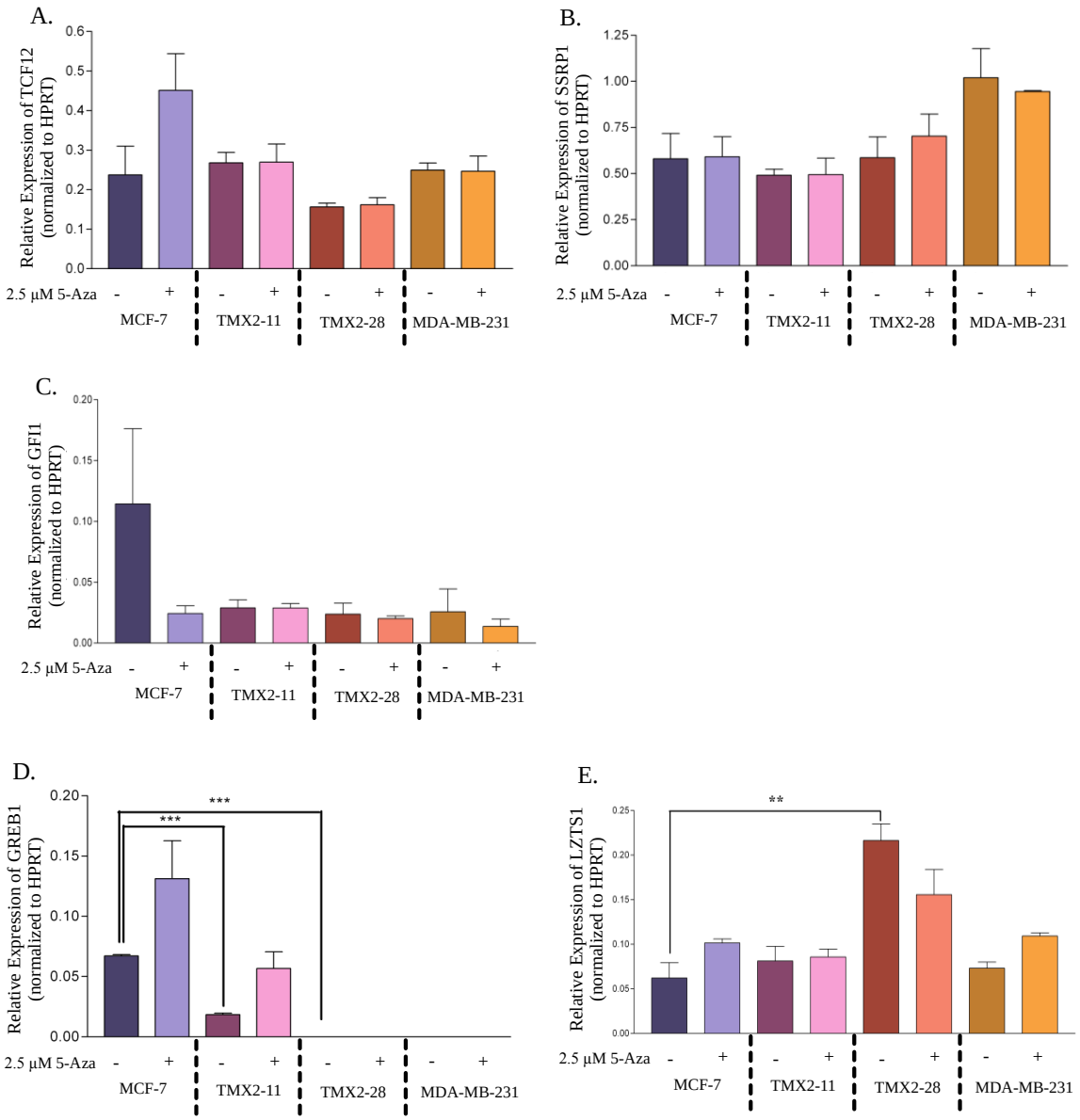


Figure 3.1 Low levels of AKAP12, REG1A, and HDAC9 mRNA are expressed in Tamoxifen-resistant cell lines.

Three genes, (A) AKAP12 (B) REG1A and (C) HDAC9 are expressed at extremely low levels in MCF-7, TMX2-11 and TMX2-28. REG1A (B) is significantly more expressed in TMX2-11 as compared with MCF-7 ($p=0.0131$). Comparisons were made from triplicate biological replicates using unpaired students t-test.


A.
Relative Expression of TCF12 (normalized to HPRT)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
2.5 µM 5-Aza - + - + - + - +
MCF-7
TMX2-11
TMX2-28
MDA-MB-231
B.
Relative Expression of SSRP1 (normalized to HPRT)
1.25
1.00
0.75
0.50
0.25
0.00
2.5 µM 5-Aza - + - + - + - +
MCF-7
TMX2-11
TMX2-28
MDA-MB-231
C.
Relative Expression of GFI1 (normalized to HPRT)
0.20
0.15
0.10
0.05
0.00
2.5 µM 5-Aza - + - + - + - +
MCF-7
TMX2-11
TMX2-28
MDA-MB-231
D.
Relative Expression of GREB1 (normalized to HPRT)
***
***
0.20
0.15
0.10
0.05
0.00
2.5 µM 5-Aza - + - + - + - +
MCF-7
TMX2-11
TMX2-28
MDA-MB-231
E.
Relative Expression of LZTS1 (normalized to HPRT)
**
0.25
0.20
0.15
0.10
0.05
0.00
2.5 µM 5-Aza - + - + - + - +
MCF-7
TMX2-11
TMX2-28
MDA-MB-231

(Figure 3.2)

Figure 3.2. 5-Aza treatment does not affect mRNA expression of TCF12, SSRP1, GFI1, GREB1 and LZTS1.

Expression of (A) TCF12 (B) SSRP1 (C) GFI1 and (D) GREB1 (E) LZTS1 is shown pre and post treatment with 2.5 mM 5-Aza. (A-C) No significant effect is seen in either control or 5-Aza treated samples for SSRP1 or GFI1.
(D) A significant decrease in expression of GREB1 is seen in TMX2-11 and TMX2-28 control samples ($p = <0.0001$ for both). (E) Expression of LZTS1 in TMX2-28 differs significantly from MCF-7 ($p = 0.0037$). Comparisons were made from triplicate biological replicates using unpaired students t-test.

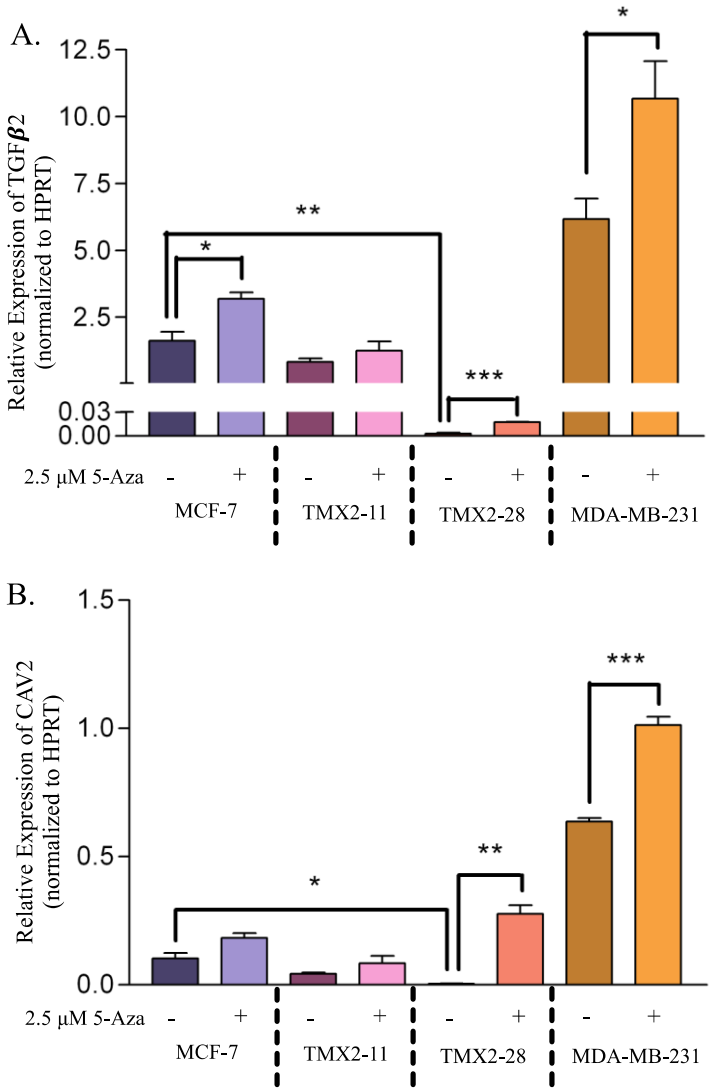
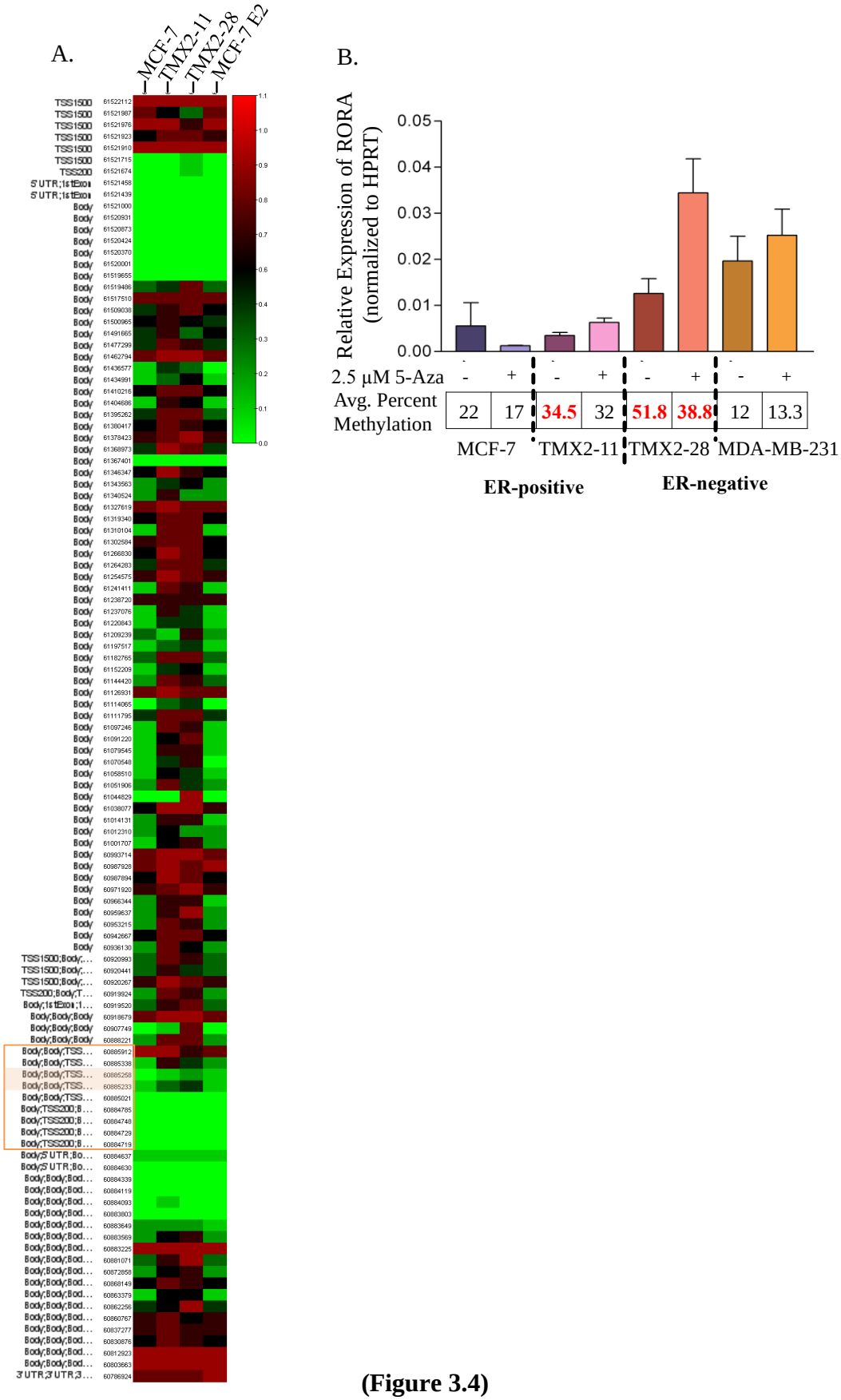


Figure 3.3 mRNA expression of TGFβ2 and CAV2 increases after 5-Aza treatment in TMX2-28 cells.

(A) Expression of TGFB2 is significantly decreased in TMX2-28 as compared with MCF-7 ($p = 0.009$). Treatment with 2.5 μM 5-Aza results in a significant increase in expression in MCF-7 and MDA-MB-231, and TMX2-28 ($p = 0.0191$, 0.048, and 0.0002, respectively). (B) Expression of CAV2 is significantly decreased in TMX2-28 as compared with MCF-7 and treatment with 5-Aza significantly increases expression of the gene ($p = 0.0111$ and 0.0013, respectively). A significant increase is also seen in MDA-MB-231 cells after 5-Aza treatment ($p = 0.0004$). Comparisons were made from triplicate biological replicates using unpaired students t-test.


A.
MCF-7
TMX2-11
TMX2-28
MCF-7 E2
B.
Relative Expression of RORA
(normalized to HPRT)
0.05
0.04
0.03
0.02
0.01
0.00
2.5 µM 5-Aza
Avg. Percent
Methylation
22
17
34.5
32
51.8
38.8
12
13.3
MCF-7
TMX2-11
TMX2-28
MDA-MB-231
ER-positive
ER-negative

(Figure 3.4)

Figure 3.4 RORA promoter methylation, but not mRNA expression increases after 5-Aza treatment in TMX2-28.

(A) A heat map shows methylation of CpG sites in RORA interrogated by the HM450BC. More than half of the CpG sites have increased methylation (β-value) in both TMX2-11 and TMX2-28 as compared with MCF-7. The orange box includes the promoter region of the transcript interrogated. Sites highlighted in orange (MAPINFO 60885258 & 60885233) are those further interrogated by pyrosequencing. (B) No expression change is seen between TMX2-11, TMX2-28 and MCF-7. A significant increase in methylation is seen in both TMX2-11 and TMX2-28 as compared with MCF-7 ($p = 0.0001$ for both). Treatment with 2.5 μM 5-Aza results in an additional methylation increase in TMX2-28 only ($p = 0.0001$). Comparisons were made from triplicate biological replicates using unpaired students t-test.

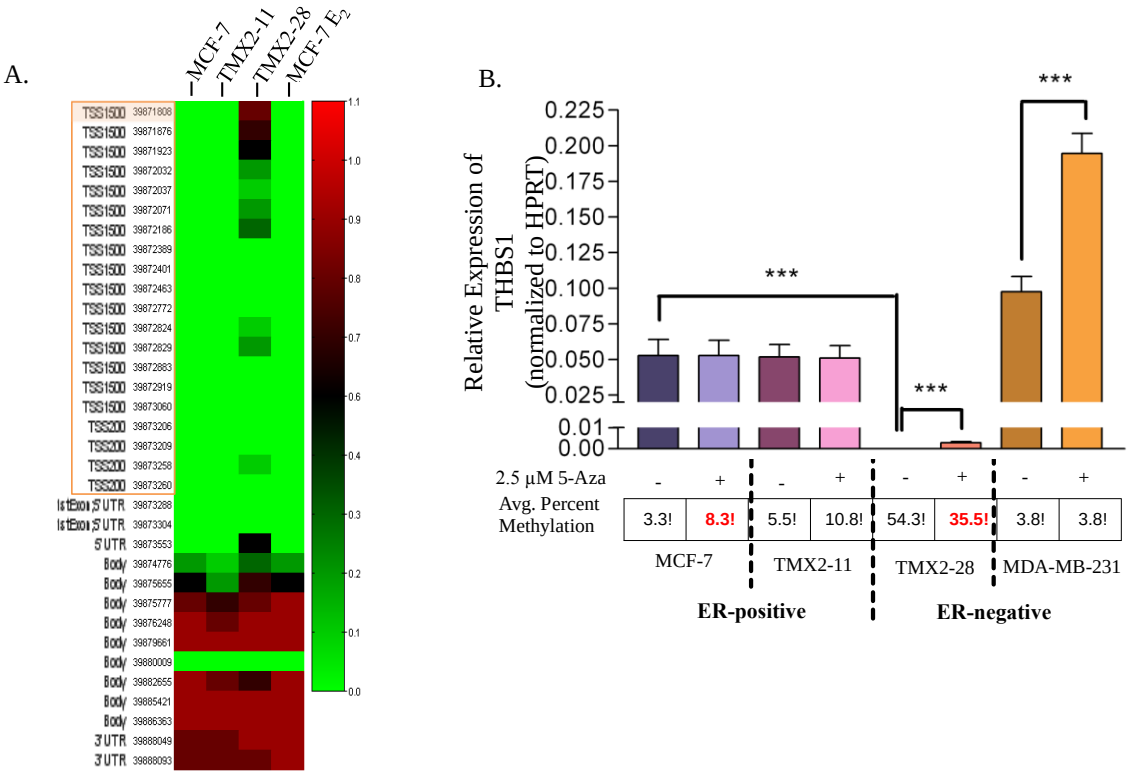


Figure 3.5. mRNA expression and promoter methylation of THBS1 are affected by 5-Aza treatment.

(A) The heat map of THBS1 shows an increase in methylation (β-value) in 10 CpG sites in the promoter region (orange box). MAPINFO site 39871808 (highlighted in orange) was analyzed further by pyrosequencing analysis. (B) A significant decrease in expression is seen only in TMX2-28 as compared to MCF-7. Treatment with 5-Aza resulted in a significant increase in expression and decrease in methylation in TMX2-28 ($p = 0.0001$ for both). A significant increase in promoter methylation occurs in 5-Aza treated MCF-7 ($p = 0.0253$). A significant change in expression, but not methylation is seen in MDA-MB-231 ($p = 0.0001$). Comparisons were made from triplicate biological replicates using unpaired students t-test.

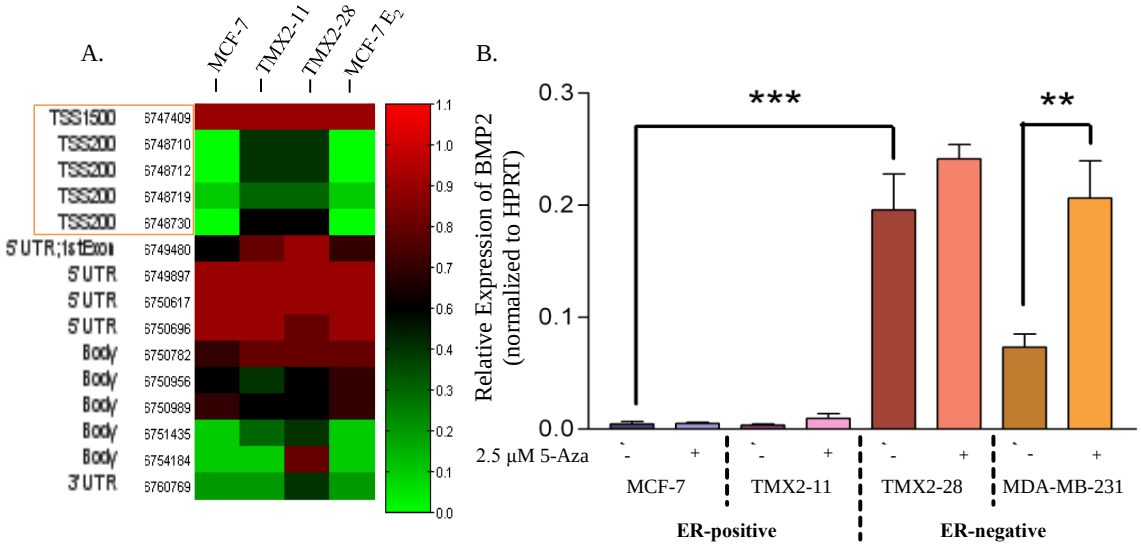


Figure 3.6. mRNA expression and promoter methylation of BMP2 are increased in TMX2-28 cells.

(A) The heat map of BMP2 shows an increase in methylation in four out of five CpG sites in the promoter region (orange box). (B) A significant increase in expression is seen in TMX2-28 as compared with MCF-7 ($p = 0.0001$). No change in expression is seen after treatment with 5-Aza. A significant increase in expression is seen in MDA-MB-231 after treatment with 5-Aza. Comparisons were made from triplicate biological replicates using unpaired students t-test.

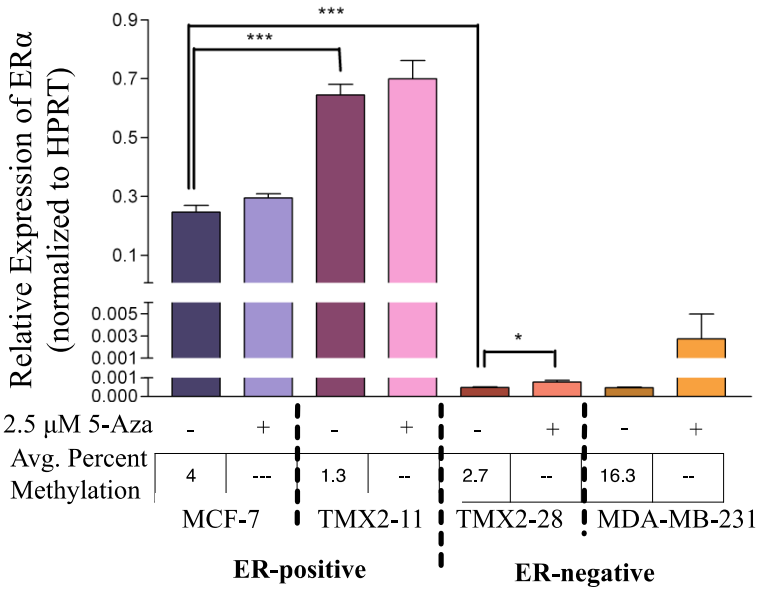


Figure 3.7 ERα is not re-expressed in TMX2-28 or MDA-MB-231 cells after treatment with 5-Aza.

TMX2-11 has a significant increase in expression of ERα as compared with MCF-7 (p=0.0008). TMX2-28 (p=0.0004) and MDA-MB-231 (p=0.0004) have a significant decrease in expression as compared with MCF-7. Treatment with 2.5 μM 5-Aza for 48 hours resulted in a small, but significant increase in expression in TMX2-28 (p = 0.036). Average percent methylation is low (<4%) in MCF-7, TMX2-11, and TMX2-28 control cells and therefore 5-Aza treated samples were not analyzed for methylation changes. Comparisons were made from triplicate biological replicates using unpaired students t-test.

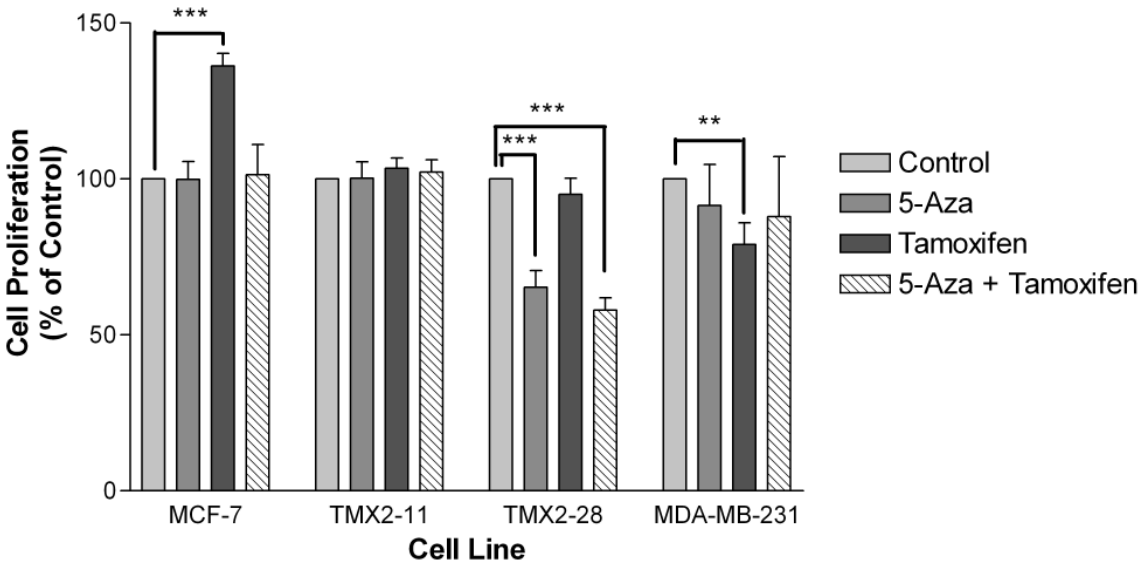


Figure 3.8 5-Aza treatment decreased cell proliferation in TMX2-28.

Cells were seeded at varying concentrations in 96-well plates and treated with 2.5 μm 5-Aza in DC5 for 4 days and cell proliferation was assessed by MTS assay. 5-Aza results are represented as percent of control. TMX2-28 cell proliferation is decreased by 35% ($p = 0.0001$) in 5-Aza treated samples and 43% ($p = 0.0001$) in 5-Aza + Tamoxifen treated samples compared with the control. The change in proliferation is non-significant between the two treatments ($p = 0.36$). MCF-7 Tamoxifen treated samples have a 136% ($p = 0.0001$) increase in cell proliferation compared with the control. Replicates from two separate experiments were combined ($n = 8$) and an unpaired student's t-test was completed to determine significance.

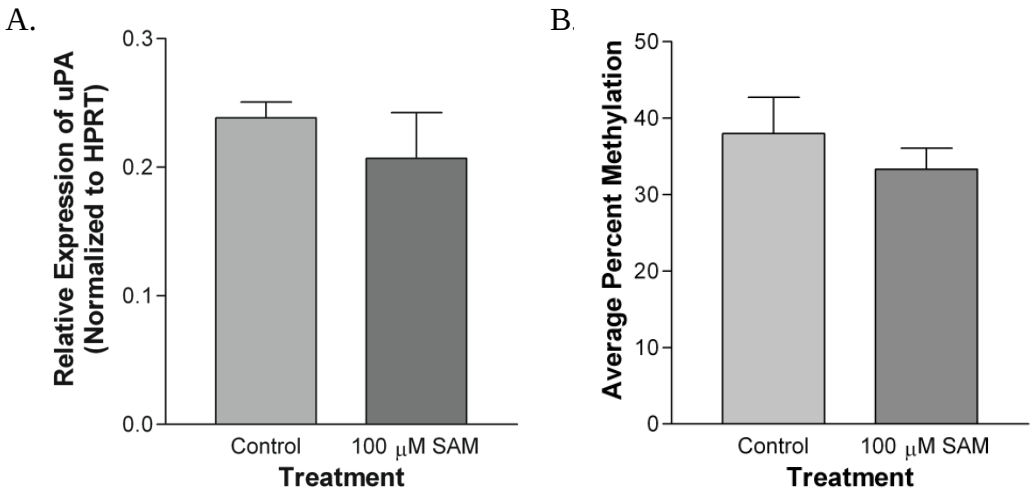


Figure 3.9 mRNA expression and DNA methylation of uPA in MDA-MB-231 cells is unaffected by S-adenosylmethionine (SAM).

Cells were treated with 100 μM SAM or a vehicle control for 6 days and (A) expression and (B) methylation were measured. No significant change is seen between the control and the SAM treated cells in (A) expression ($p = 0.44$) or (B) methylation ($p = 0.45$) assays. Comparisons were made from triplicate biological replicates using unpaired students t-test.

CHAPTER 4

HIGH-DENSITY ARRAY ANALYSIS OF DNA METHYLATION IN HUMAN BREAST TUMOR SAMPLES

Introduction

In Chapter 2, I reported that both the ER-positive Tamoxifen-selected cell line, TMX2-11, and the ER-negative Tamoxifen-selected cell line, TMX2-28, had significant hypermethylation as compared with the ER-positive parental line, MCF-7. I predicted that both ER-positive and ER-negative second breast tumors occurring in women after anti-estrogen treatment would be hypermethylated as compared with primary ER-positive tumors. Furthermore, based on the greater methylation of the ER-negative, TMX2-28 cell line, I predicted that second ER-negative tumors would be more methylated than second ER-positive tumors. Finally, I suggested that differences in methylation profiles among individual ER-positive primary tumors could be used to predict the hormonal status of the second tumor, and thereby allow a more precise and targeted treatment for individual women.

At this point it is necessary to digress and discuss the terminology used for second breast tumors. In some breast cancer literature, all tumors occurring after or during treatment of the primary tumor may be referred to as a recurrence [111]; i.e. the woman has a recurrence of breast cancer. However, for tailored effective treatment, it is considered important to distinguish between a true new primary or *de novo* tumor, and a recurrence originating from cells of the first primary. Classifying second tumors as new

primaries or true recurrences is primarily based on histopathologic and clinical characteristics. Typically, 75% of second tumors occurring in the same breast, *ipsilateral tumors*, are classified as recurrences while only 15% of second tumors occurring in the opposite breast, *contralateral tumors*, are classified as recurrences [111]. True recurrent contralateral tumors represent metastasis and therefore, in general are associated with a less favorable diagnosis than a new primary contralateral tumor [111]. In practice, however, it remains difficult to distinguish true recurrences from new primaries and the literature uses "recurrence" broadly to indicate a second tumor that may or may not be a clonal descendant of the primary tumor. Accordingly, I use the terms second and recurrent tumor interchangeably and not to distinguish clonal origin. It is important to note that almost all second tumors in women who had ER-positive primary tumors develop under anti-estrogen selection regardless of whether they are true recurrences or new primaries.

In general ER-positive tumors have been shown to have greater DNA methylation than do ER-negative tumors [112-115]. For example, a study by Fackler and colleagues [114] of 103 primary invasive breast tumors and 21 normal breast tissue samples using HumanMethylation27 BeadChip (HM27BC) data found that ER-positive tumors were more highly methylated than ER-negative ones. Examination of methylation with regards to gene location showed that ER-positive tumors had a greater number of hypermethylated CpG sites than ER-negative tumors in the promoter region. The authors went on to identify a set of 40 CpG sites with the greatest amount of hypermethylation in ER-positive (27 probes) and ER-negative (13 probes) breast tumors and found that these sites were capable of classifying additional tumor samples into hormone receptor status

subgroups [114]. While this study confirms that hormone receptor status plays a role in breast tumor methylation, it did not examine methylation in recurrent tumors. To date, little is known regarding the DNA methylation profiles of second tumors and the extent to which methylation patterns are associated with ER status.

The present human study was designed to include both the primary and second tumors from women who had ER-positive primary tumors and received anti-estrogen treatment before the second, either ER-positive or ER-negative, tumor occurred. The major goals of this design were to i) determine the extent to which methylation data from ER-positive primary tumors could be used to further categorize these tumors and predict the ER-status of the second tumor, and ii) determine whether the group of ER-negative second tumors that occurred in women who had ER-positive primary tumors were highly methylated. In addition, the study design included primary and recurrent tumors from a group of women who had only ER-negative tumors as well as primary tumors from women who did not have a recurrence. Inclusion of these groups allowed me to i) expand and refine our knowledge of the DNA methylation differences between ER-positive and ER-negative tumors, and ii) determine the extent to which DNA methylation of primary tumors can be used to predict recurrence.

My data confirm previous findings that ER-positive breast tumors have more hypermethylated CpG sites than ER-negative tumors. When the tumors were stratified by occurrence, it was found that first tumors, not second tumors, are predominantly responsible for the hypermethylation of the ER-positive group as compared with the ER-negative group. Further stratification by side of tumor recurrence revealed that first tumors from pairs in which the second tumor occurred in the contralateral breast were

more hypermethylated than those from ipsilateral pairs. In contrast, second tumors recurring in the ipsilateral breast were more hypermethylated than those recurring in the contralateral breast. These data suggest that methylation is playing a role in progression of the tumor particularly if the second tumor is more likely to be a recurrence rather than a new tumor. Additionally, paired tumors stratified by ER-status of the second tumor indicated that ER-negative second tumors with ER-positive first tumors had the greatest overall methylation differences, particularly in the second tumor. Pathway analysis was completed to provide insight into biomarkers associated with tumors that recur. Two pathways, 'homophilic cell adhesion via plasma membrane adhesion molecules' and 'cell fate commitment', were selected for further analysis. ER-positive first tumors that recurred as either ER-positive or ER-negative compared with non-recurrent tumors shared hypermethylated genes in the homophilic cell adhesion pathway. ER-positive first tumors that recurred as ER-negative compared with ER-positive first tumors that recurred as ER-positive were associated with a unique set of hypermethylated genes in the cell fate commitment pathway. Kaplan-Meier plots from TCGA data showed that expression of the genes only hypermethylated in each individual comparison group in the homophilic cell adhesion pathway was linked to overall survival. Similarly those genes shared between ER-positive first tumors with either ER-negative or ER-positive second tumors in the cell fate commitment pathway were also linked to survival.

Materials and Methods

Human Tissue

Institutional Review Board (IRB) approval was obtained from Baystate Medical Center for the study. Identifiers were removed and samples were labeled sequentially to ensure patient anonymity. Ninety-one samples were identified as available in the database maintained by the Department of Surgical Pathology at Baystate Medical Center. The following blocks for 86 of the 91 tumors were located and pulled: tumors from women who had an ER-positive first tumor (19) and an ER-positive second tumor (20); from women who had an ER-negative first tumor (7) and an ER-negative second tumor (8); from women who had an ER-positive first tumor (7) and an ER-negative second tumor (8). In 5 cases, blocks for either first or second tumors were not available from the same woman, however those cases were included as we had demographic data on the missing blocks (see Table 4.1 and Figure 4.1). Additionally, 4 third tumors were collected from women with an ER-positive primary and an ER-positive second tumor, 2 third tumors from women with an ER-negative primary and an ER-negative second tumor, and 1 third tumor from a woman with an ER-positive primary and an ER-negative second tumor. Eight non-recurrent ER-positive and two non-recurrent ER-negative breast tumors were also collected for a total of 86 tumors (Table 4.1). Clinical and pathological data for the 86 tumors are in Appendix J. Of these 86 tumors, HM450BC data was collected for 70 tumors and clinical and pathological data are summarized in Table 4.2.

Primary tumors were matched to a second tumor from the same woman in either the ipsilateral or contralateral breast, where possible. The 70 tumors for which HM450BC data were obtained included 50 paired tumors as follows: 14 paired tumors from women with ER-positive primary and ER-positive second tumors, 5 tumor pairs from women

with ER-negative first tumors and ER-negative second tumors, and 6 tumor pairs from women with ER-positive first tumors and ER-negative second tumor. Tumors were given group labels for reading ease: ER-positive first tumors (A1) with ER-positive second tumors (A2); ER-negative first tumors (B1) with ER-negative second tumors (B2); ER-positive first tumors (C1) with ER-negative second tumors (C2) (Table 4.3). Tissue was prepared for DNA purification in 10μm thick sections from the corresponding formalin-fixed paraffin embedded (FFPE) blocks. A summary of characteristics for the 50 paired tumors from groups A, B and C as well as the 8 non-recurrent tumors can be found in Table 4.4.

Immunohistochemistry (IHC)

FFPE tissue blocks from 70 breast tumors were sectioned at a thickness of 4μm and placed on slides (7 slides for each tumor). The slides were stained for ER, PR, HER2, CK5/6, CK14, and p40 using the UltraView Universal DAB Detection Kit on the Ventana BenchMark Ultra platform. CK5 and Ki67 were performed on the Dako autostainer platform using the Dako Advanced HRP Polymer System. Additionally, hemotoxylin and eosin (H&E) slides were prepared for tumor verification. Antibodies used were optimized for each protein: ER (Ventana CONFIRM anti-estrogen receptor SP1 rabbit monoclonal primary antibody), PR (Ventana CONFIRM anti-progesterone receptor 1E2 rabbit monoclonal primary antibody), HER2 (Ventana PATHWAY anti-HER2/neu antibody 4B5 rabbit monoclonal antibody), CK5/6 (Ventana anti-Cytokeratin 5/6 (D5/16B4) mouse monoclonal primary antibody), CK14 (Ventana anti-Cytokeratin 14 mouse monoclonal primary antibody), p40 (Biocare anti-p40 mouse monoclonal

antibody), and Ki67 (Dako mouse monoclonal MIB-1 antibody). Slides were dehydrated in ethanol and xylene post-staining and coverslips were added. One anatomical pathologist (Rahul Jawale) conducted scoring of the slides. Allred scores were recorded for ER & PR ranging from 0-8 with a score of 3 and above considered as positive for receptor status and tumors were considered HER2 positive if 30% of the cells contained 3+ membrane staining. CKs were scored on the intensity of immunoreactivity ranging from a score of 0-4 (negative to strong staining) and percent of positive cells was recorded. Ki67 was scored as percent of positive cells within the area of invasive cells with 0% being the lowest score and 100% being the highest.

DNA Purification

DNA was purified from FFPE breast tissue sections using the BiOstic FFPE tissue DNA isolation kit (Mo Bio, Carlsbad, CA). Briefly, the breast tumor size was measured using the H&E stained slide to determine the number of sections needed to purify a minimum of 500ng of DNA. For samples exceeding 4mm x 4mm, a single tissue section was purified. For samples smaller than 4mm x 4mm, two to four 5-10μm tissue sections were combined into a single tube prior to purification. As illustrated in Appendix K, sections were prepared for DNA extraction by arranging the unstained slide on top of the tumor marked H&E stained slide. Tissue was carefully removed from the slide by scraping using a sterile needle. Purification and concentration was completed as described by manufacturer.

Illumina HumanMethylation450 BeadChip (HM450BC)

DNA from breast tumor samples purified using the BiOstic FFPE tissue DNA isolation kit (Mo Bio, Carlsbad, CA) were sent to the core facility at the University of Southern California (USC) for HM450BC (Illumina) analysis. DNA sent to USC was bisulfite treated with the EZ DNA methylation kit (Zymo) and a MethylLight PCR reaction was completed to analyze the quality of the DNA. Two different assessments of the DNA were completed. The first assessment was an ALU-C4 bisulfite control reaction to test for sample quality and integrity. An ALU-C4 (HB-313) Ct value below 19 was considered passing (see Table 4.1). The second test, Conv 100% (HB-365), was to determine bisulfite conversion completeness and tested, unconverted (0%) and fully converted (100%) DNA. Of the 86 tumors collected, 70 passed QC (see Table 4.1). The age of the FFPE tumor block appeared to have an effect on the quality of DNA that was extracted from the sample. Tumors that failed QC were among the oldest tumor blocks and had poor quality DNA after extraction. As expected, the size of the tumor was also important, tumors less than 4mm x 4mm provided little DNA. Bisulfite treated DNA that passed QC analysis was amplified at 37°C for 20-24 hours after treatment with 0.1N NaOH. The DNA was then fragmented at 37°C for 1 hour using an enzymatic process and subsequently precipitated in 100% 2-propanol at 4°C for 30 minutes followed by centrifugation at 3000xg at 4°C for 20 minutes. Dried pellets were resuspended in hybridization buffer, 48°C for 1 hour followed by 95°C for 20 minutes, then loaded onto the HM450BC and incubated at 48°C for 16-24 hours. Following hybridization of DNA to the primers on the BeadChip, unhybridized and non-specific DNA was removed using wash buffers to prepare the chip for staining. After a single base extension of the

hybridized primers using labeled nucleotides, the BeadChip was stained with Cy-3 and Cy-5 fluorescent dyes and read using the Illumina iScan Reader. The image data were then analyzed using Illumina GenomeStudio to assess efficiency of the reaction. Methylation of the interrogated CpG loci were calculated as the ratio of the fluorescent signals of methylated to unmethylated sites (beta values). A beta value of 1 corresponds with complete methylation of DNA at the probe site and a beta value of 0 is no methylation DNA at the probe site.

Data Analysis

GenomeStudio methylation module software was used to analyze methylation data from the HM450BC. To ensure that only CpGs with complete bead hybridization were used in the analysis, only those sites with a detection p-value of ≤ 0.05 were included in the analysis (Illumina). Detection p-value is the likelihood that the measured signal is above background and therefore significant. CpG sites containing SNPs within 10 base pairs of the site were omitted, due to literature reports indicating SNPs in this region can affect probe binding [116]. For most analyses, a 2-fold cutoff in either direction was used to determine hypermethylated or hypomethylated CpG sites. This cutoff was selected based on previously reported literature values [51, 53, 114]. Additionally, for noted analyses a false discovery rate (FDR) of 0.05 was computed in GenomeStudio for each CpG site and only those sites with a DiffScore of ≥ 22 and a greater than 1.5-fold change for hypermethylated CpG sites or $\leq$-22 and a less than 0.67-fold change for hypomethylated CpG sites, corresponding with a p-value of ≤ 0.01, were considered statistically significant. The DiffScore is calculated using the Illumina custom

model, which assumes that the beta value, under a set of biological conditions such as tumor versus normal, is normally distributed among the replicates. STATA (http://www.stata.com) was used to compute unpaired student's t-test and one-way ANOVA. Those analyses with a p-value of <0.05 were considered statistically significant. Gene lists were obtained from the list of differentially methylated CpG sites. Panther Classification System (www.pantherdb.org) was used to conduct pathway analysis from a list of genes associated with hyper- and hypomethylated dmCpG sites described above. Kaplan-Meier plots (http://kmplot.com/) were created using genes associated with differential methylation in multiple comparison groups.

Results

Breast tumors were classified by ER-status and occurrence. Three major groups of tumors were used in most analyses. The first tumor group, A, was collected from women who had an ER-positive primary and their second tumor was ER-positive, the second group, B, was from women with an ER-negative primary tumor and an ER-negative second tumor, and the third group, C, was from women who had an ER-positive primary tumor and their second tumor was ER-negative. In addition to the 58 first and second tumors mentioned above, there were 4 third tumors from groups A & B and a group of 8 primary tumors (6 ER-positive and 2 ER-negative), from women who did not have a second tumor, for a total of 70 tumors that passed QC and were analyzed by HM450BC (Table 4.5 and Figure 4.1). The majority of analyses in this study were restricted to 50 paired tumors, including 14 paired tumors from women with ER-positive first tumors (A1) and ER-positive second tumors (A2), 5 paired tumors from women with ER-

negative first tumors (B1) and a second ER-negative tumor (B2), and 6 paired tumors from women with ER-positive first tumors (C1) and a second ER-negative tumor (C2) (Table 4.6). Third tumors and the 8 primary tumors with no recurrence were omitted from paired analyses.

Patient and tumor demographics based on all 70 tumors are shown in Table 4.2. The average patient age at first tumor occurrence was 56.6 years old, while the average age at recurrence was 64.1 years old; 20% more women were menopausal at recurrence (13 women versus 21 women). Tumors on average recurred after 67.8 months. PR status did not vary significantly between primary and recurrent tumors, however the number of HER2 positive tumors was slightly higher in the recurrent group (14.7% in primary versus 25% in recurrent). Interestingly, the number of Ki67 high tumors was 20% greater in recurrent tumors than primary ones (23.5% in primary versus 44.4% in recurrent). Tumors were also stained for basal markers CK 5/6 and 14. Half of the ER-negative tumors that recurred as ER-negative stained positive for both markers (Appendix J). Additionally, two ER-positive tumors that recurred as ER-negative stained positive for both markers and one tumor stained positive for CK14 only (Appendix J). Tumor grade ranged from 0-3 with the majority of tumors falling into categories 2 and 3 in primary tumors (33.3 and 36.3% respectively) and category 3 in recurrent tumors (63.6%). Invasive ductal carcinoma (IDC) was the most common type of tumor found in both primary and recurrent (67.6 and 72.2% respectively) followed by ductal carcinoma in situ (DCIS) at 14.7 and 13.9%, invasive lobular carcinoma (ILC) at 11.8 and 11.1%, and finally invasive carcinoma with a mix of ductal and lobular characteristics (IDLC) at 5.9 and 2.8%. All tumors that stained positive for basal markers were IDC (Appendix J).

Recurrent tumors were more likely to occur in the ipsilateral breast as the primary tumor (60%) than in the contralateral breast (40%). Tumor size did not vary between groups, with approximately half of the tumors being smaller than 20mm and half being larger than 20mm. Of the 34 primary tumors we collected, anti-hormonal treatment data were available for only 23 women. Of the 26 ER-positive primary tumors collected, 4 women were treated with only Tamoxifen, 7 were treated with an aromatase inhibitor (AI), and 5 were treated with a mixture of Tamoxifen and an AI. Chemotherapy data was available for 26 of the primary tumors and of those the majority were not treated with chemotherapeutics (46.1%). The remaining tumors were treated with Adriamycin (doxorubicin) and cyclophosphamide in combination with either paclitaxel (23.1%) or docetaxel (11.6%) or another chemotherapeutic that was not noted in the patients record (Table 4.2)

Table 4.4 shows the demographic data restricted to paired tumors and separated into groups based on the ER-status of the first and second tumors. Group A had 14, Group B had 5, and Group C had 6 tumor pairs (see Table 4.4 A, B, and C respectively). Additionally, demographic data on 6 unpaired ER-positive, non-recurrent tumors are included for comparison (see Table 4.4D). One-way ANOVA and student's t-tests were run for all comparisons, however only the statistically significant result is presented in this section. The average age of patients at primary and recurrent tumor occurrence in Group A were 56.8 and 64.1 years old and the TTR was 87.6 months (Table 4.4A). The average age of patients at primary tumor occurrence in Group B was the highest among all groups at 61.4 years old. However, average age of patients at tumor recurrence in Group B was similar to those in Group A (64.2 years old) (Table 4.4B). Group C patients

at primary tumor occurrence had an average age of 53 years old and at recurrent tumor occurrence were the youngest of all groups at 60.3 years old (Table 4.4C). Patients with non-recurrent tumors had an average age similar to those of Group A (56.5 years old) (Table 4.4D). In Group A the average ER staining score was 7.6 in both primary and recurrent tumors, which is unsurprising given that both groups of tumors are ER-positive (Table 4.4A). The Group C ER-positive primary tumors had similar ER staining as Group A (7.8) (Table 4.4A &C) and no samples in Group B or recurrent samples in Group C had ER staining (Table 4.4B & C). PR staining was highest in Group A primary tumors with an average score of 5.9 out of a possible 8 (Table 4.4A). HER2 expression was 2-fold higher in both primary and recurrent tumors in Group C than in both Groups A & B. Expression of Ki67 was significantly higher in ER-negative tumors in both Group B and Group C as compared with ER-positive tumors from Group A and recurrent ER-positive tumors in Group C ($p=0.0138$). ER-negative tumors in Group B were all Grade 3 IDC whereas tumors in Groups A and C ranged from grade 0-3 and IDC was the most commonly found type. Second tumors in Groups A & C were both ipsilateral and contralateral to the primary tumor, however Group B second tumors were only found in the ipsilateral breast. Primary tumors in Group C had a tendency to be larger than all other tumor groups with 83% of the tumors being ≥20mm. Finally, chemotherapy records were not available for many of the tumors. ER-positive primary tumors from Group A had the most information available. No treatment was given to 7 women in Group A, AC + paclitaxel was given to 3 women, and 2 women had other forms of treatment not previously described. Three women with ER-negative primary tumors in Group B received AC + docetaxel, and AC + paclitaxel or no treatment was given to only one

woman each. Only one tumor treatment record was available for Group C and that woman received a form of chemotherapy that is not listed in the table, but is one of several different combinations of chemotherapeutics. In general I found that PR, HER2 and Ki67 were associated with ER-status, however the overall the sample size was too small to further stratify based on other tumor characteristics such as tumor type and chemotherapy received.

Next, I compared the ER-positive primary tumors that had recurrences with the ER-positive non-recurrent tumors. The average age of women with ER-positive, non-recurrent tumors (Table 4.4D) was 53.2 years, which was approximately the same age of women with primary tumors from Group C. ER levels in the non-recurrent tumors was slightly higher than primary tumors from Groups A and C with a mean score of 8. In the non-recurrent tumors, PR and Ki67 levels were the lowest of the three groups with a mean score of 3.7 and 2.7 respectively, and HER2 levels fell between Groups A and C with a mean score of 0.7. Finally, the distribution of tumor types among non-recurrent tumors was similar to ER-positive primary tumors from Groups A and C. IDC was the most common type of tumor, followed by ILC and DCIS.

ER-positive versus ER-negative tumors: greater hypermethylation in ER-positive tumors

Several research groups have shown that ER-positive tumors have a greater number of hypermethylated CpG sites as compared to ER-negative tumors [114, 115]. To determine whether this finding held true for the current tumor set, I initially compared 41 ER-positive tumors with 22 ER-negative tumors. First, second and third tumors labeled

as either ER-positive or ER-negative were included in this analysis. A visual examination of the data showed that overall, ER-positive tumors had slightly higher methylation than ER-negative tumors. ER-positive tumors had 3,046 CpGs with mean beta values 2-fold greater than ER-negative tumors, 1,098 CpGs with mean beta values 2-fold greater than ER-positive tumors (Figure 4.2A, outer red lines). Distribution analysis of the ER-positive and ER-negative tumor groups revealed that methylation is slightly higher in ER-positive tumors, a median of 0.44 as compared with a median of 0.41 in ER-negative tumors (Figure 4.2B).

ER-status and differential hypermethylation of 35 target CpGs: confirmation of previously identified dmCpGs

Fackler et al. found a subset of 40 CpG probes on the HumanMethylation27K BeadChip to be differentially methylated with regards to ER-status (27 hypermethylated in ER-positive and 13 hypermethylated in ER-negative). I compared the 41 ER-positive tumors with the 22 ER-negative tumors filtering for the 40 CpG sites. Of the 40 hypermethylated CpG sites, only 35, 23 hypermethylated in ER-positive and 12 hypermethylated in ER-negative tumors, were included on the HM450BC (Appendix L). Overall, methylation of the 35 CpG sites was low, with no sites having a mean average beta value >0.4 (Figure 4.3). Of the 23 sites identified by Fackler and colleagues to be hypermethylated in ER-positive as compared with ER-negative tumors, 12 CpGs (52%) had a 2-fold higher methylation in our data set and the remaining 11 were hypermethylated in ER-positive tumors to a lesser degree (see Figure 4.3, blue dots). Nine out of 12 CpGs (75%) previously shown to be hypermethylated in ER-negative

tumors exhibited a 2-fold greater methylation among ER-negative tumors in our data set (see Figure 4.3, green dots). Of the remaining sites, one was slightly hypermethylated (fold change < 2) and two did not deviate from the x=y line (Figure 4.3, center red line). Distribution analysis reveals a large methylation difference between ER-positive and ER-negative tumors for the 35 sites. The median average beta for the 23 CpG sites that are hypermethylated in ER-positive tumors is 0.19, which is considerably higher than the methylation of ER-negative tumors at 0.10 (Figure 4.4A). The heat map analysis shows that all but one of the 23 CpG sites (cg02755525) has higher methylation in the ER-positive tumors (Figure 4.4B). Distribution analysis of the CpG sites hypermethylated in ER-negative tumors shows that ER-negative tumors have a median average beta of 0.25, which is much higher than ER-positive tumors at 0.07 (Figure 4.5A). Two of the 12 CpG sites previously shown to be hypermethylated in ER-negative tumors (cg00720137 and cg08090772) do not differ in methylation from ER-positive tumors (Figure 4.5B). The overall methylation signature of the 35 CpG sites in our data is similar to the one presented in the data from Fackler et al. [114], however trends for some sites were weaker than previously reported. This was especially true of CpG sites shown as hypermethylated in ER-positive tumors.

ER-associated trends in methylation within groups of first, and groups of recurrent tumors: methylation differences between ER-positive and ER-negative tumors are reduced in second tumors

I next compared ER-positive and ER-negative tumors classified by occurrence (either first tumor or second tumor) to determine the extent to which ER-associated

methylation was altered by tumor occurrence. As shown in the scatterplot analyses, the trend for greater methylation among ER-positive tumors is predominantly driven by first tumors (Figure 4.6A). Among first tumors, 10,314 CpGs were more methylated with a 2-fold change in ER-positive tumors, and only 1,235 CpGs were more methylated in ER-negative tumors; ER-positive tumors had 8.35 times more methylated CpGs. In contrast, the tendency for greater methylation in ER-positive tumors was reduced when the analysis was limited to recurrent tumors. Among recurrent tumors, 3,110 CpGs were more methylated in ER-positive, and 2,198 CpGs were more methylated in the ER-negative tumors; ER-positive tumors had only 1.41 times more methylated CpGs. The distribution analyses (Figures 4.6 C & D), support the conclusion based on the scatterplot analyses. Median methylation of ER-positive first tumors was 0.44 as compared to a median methylation of 0.39 in ER-negative first tumors. In contrast, the methylation distributions of ER-positive and ER-negative second tumors were similar, with median methylation of 0.43 and 0.44, respectively (Figure 4.6D). It is noteworthy that among recurrent tumors, the ER-negative tumors had slightly higher median beta values than the ER-positive tumors.

Differential methylation based on both occurrence and ER-status of primary and recurrent tumors: ER-negative second tumors from women who had ER-positive primary tumors have greater hypermethylation

Next, I examined the three groups of paired first and second tumors from the same woman as described previously. Tumors were assigned group labels and occurrence labels for clarification. Group A is the ER-positive primary and ER-positive second,

Group B is the ER-negative primary and ER-negative second, and Group C is the ER-positive primary and ER-negative second. First (1) tumors are the primary tumor, and second (2) tumors are the first recurrence of the tumor in the woman (Table 4.3). Third tumors were not used for these analyses.

The first comparison, ER-positive second tumors with ER-positive first tumors (A2 versus A1), showed the smallest difference in methylation (Figure 4.7A). Only 69 CpG sites in the second tumor were greater than 2-fold higher as compared to the first tumor (Figure 4.7A). The second comparison, ER-negative first tumors with ER-negative second tumors (B2 versus B1) showed slightly higher methylation in the second tumor among CpGs with methylation values below an average beta of 0.4. A total 204 CpG sites had a 2-fold difference (Figure 4.7B). The third comparison, women with ER-positive first tumors and ER-negative second tumors (C2 versus C1) showed the greatest methylation differences. ER-negative second tumors (C2) were more methylated than ER-positive first tumors with 3,652 CpG sites having a 2-fold higher methylation (Figure 4.7C). A distribution analysis showed that the ER-positive first tumors do not vary much from their second tumors (median average beta 0.44 versus 0.44) (Figure 4.7D; red), nor do ER-negative first tumors vary from the second tumors (median average beta 0.38 versus 0.39) (Figure 4.7D; blue). However, ER-negative second tumors with ER-positive first tumors have higher methylation than their ER-positive first tumors (median average beta 0.46 versus 0.42) (Figure 4.7D; purple). Additionally, ER-negative second tumors with ER-positive first tumors have higher methylation than ER-negative tumors that recur as ER-negative. This indicates that ER-negative second tumors that occur after ER-

positive first tumors may have a different methylation profile than both ER-positive tumors and ER-negative tumors.

Differential methylation based on occurrence and side: second tumors from ipsilateral pairs have the greater hypermethylation

I considered that a second tumor occurring in the same or ipsilateral breast as the primary tumor was more likely to be a recurrence of the original tumor than a tumor occurring in the opposite or contralateral breast. Accordingly, I reorganized the 25 pairs of tumors, comprising groups A, B and C, into two groups: those for which both first and second tumors occurred in the same breast (n=15 ipsilateral pairs) and those for which the first and second tumors occurred in different breasts (n=10 contralateral pairs) (Table 4.7). I found that first tumors from the contralateral pairs had more hypermethylated CpG sites with a greater than 2-fold change (1147 CpG sites) than first tumors from ipsilateral pairs (133 CpG sites) (Figure 4.8A). In contrast, among second tumors, those in ipsilateral pairs had more hypermethylated CpG sites with a greater than 2-fold change (2007 CpG sites) than those in contralateral pairs (278) (Figure 4.8B) and the number of dmCpGs was more than that observed in the combined comparisons between first tumors (Figure 4.8A versus 4B).

Differential methylation in ER-positive and ER-negative second tumors based on side: ER-negative tumors from contralateral pairs have greater hypermethylation

I further stratified all second tumors (shown in Figure 4.8B) by ER-status to determine if the greater methylation observed in second tumors from ipsilateral pairs was

modified by the presence of ER. Table 4.8 provides the sample size and ER-status of primary and second tumors; only second tumors were used in the present analysis. Methylation patterns of second tumors from ipsilateral and contralateral pairs appear to be modified by ER-status. The pattern in ER-positive second tumors was similar to that observed in the analysis of all second tumors; i.e. there was a greater number of hypermethylated CpG sites with a greater than 2-fold change in tumors from ipsilateral pairs than in contralateral pairs (4203 vs. 467) (Figure 4.9A). This pattern, however, was reversed in ER-negative second tumors; the number of hypermethylated CpG sites with a greater than 2-fold change in tumors from contralateral pairs was three times more than that observed in tumors from ipsilateral pairs (10280 vs. 3042) (Figure 4.9B). It should be noted that all of the ER-negative second tumors from contralateral pairs occurred in women who had ER-positive primary tumors (see Group C2 in Table 4.9). Finally, side-based differences in methylation were 2.8 times greater in ER-negative second tumors than in ER-positive second tumors (13,322 and 4670 dmCpGs, respectively.

Further analysis of paired tumors stratified by ER-status of primary and recurrent and by location: greater methylation among second tumors from ipsilateral pairs and first tumors from contralateral pairs

Having observed some side-based differences in methylation related to ER-status, occurrence, and location, I next asked whether differences in methylation between the paired first and second tumors from Groups A, B, and C, were associated with location. As shown in Table 4.9, ER-positive and ER-negative, primary and second tumors were further stratified by location. A visual analysis of the paired tumors from Group A

showed that ER-positive second tumors from ipsilateral pairs have more hypermethylated sites with a greater than 2-fold change (560 CpGs) than the primary tumor in the pair (95 CpGs) (Figure 4.10A). In contrast, tumors from Group A from contralateral pairs had more hypermethylated CpGs with a greater than 2-fold change in the first tumors than in the second tumors (2862 vs. 204 CpG sites) (Figure 4.10B). Analysis of Group B was not warranted, as all tumors in this group were from ipsilateral pairs, and therefore the scatterplot and analysis is identical to Figure 4.7B. Comparisons of tumors from Group C show that ipsilateral pairs have the greatest number of dmCpGs of all groups (22853 CpGs hypermethylated with a greater than 2-fold change in the ER-negative second tumors (C2), and 9543 CpG sites hypermethylated in the ER-positive first tumors (C1). However limited information can be inferred due to the small sample size (n=2 pairs) (Table 4.9 and Figure 4.11A). Group C tumors from contralateral pairs had more CpGs hypermethylated in the ER-positive first tumors (6261) than in the ER-negative second tumors (1885) (Figure 4.11B). This analysis indicates that stratification of the samples into side of second tumor occurrence may be important to discriminate methylation changes that occur in tumors that are more likely to be true recurrences.

Statistical differences in methylation by subgroups

While the visual analyses were useful in obtaining a global overview of differences in methylation patterns between groups, statistical analyses are useful in identifying individual CpGs that are significantly differentially methylated. For the following sections, an FDR of 0.05 was computed in GenomeStudio for each CpG site and only those sites with a DiffScore of ≥ 22 and fold change of >1.5 for

hypermethylated CpGs or ≤-22 and a fold change of <0.67 for hypomethylated CpG sites, corresponding with a p-value of ≤ 0.01, were considered statistically significant. These criteria were used for all subsequent analyses.

In the first set of analyses, Group A2, B1, C1 and C2 were each compared to Group A1 (reference). As shown in Table 4.10, first and second ER-positive tumors from the same women (A2 versus A1) were compared and it was discovered that there are 1.7 times more hypomethylated CpG sites than hypermethylated ones (914 hypermethylated versus 1588 hypomethylated) (Table 4.10). ER-negative first tumors (B1) were compared to ER-positive first tumors (A1) to elucidate the methylation patterns in ER-negative tumors as compared with ER-positive ones. It was found that B1 had 4491 hypermethylated CpG sites and 38044 hypomethylated sites than A1 (Table 4.10). Lastly, analysis was completed to determine whether methylation of a distinct set of CpG sites or genes results in ER-positive first tumors that recur as ER-negative (C1) or ER-positive first tumors that recur as ER-positive (A1). These data show that there are 5056 hypermethylated CpG sites and 3131 hypomethylated CpG sites in C1 as compared to A1 suggesting that methylation plays a role in a tumors response to treatment with an antiestrogen such as Tamoxifen (Table 4.10).

Group B1 was used as a reference in the second set of analyses. The first analysis compared ER-negative second tumors with the ER-negative first tumors from the same women (B2 versus B1). This analysis revealed that there were almost twice as many hypomethylated CpG sites than hypermethylated CpG sites (5815 hypomethylated sites versus 3036 hypermethylated sites) (Table 4.11). Comparing ER-positive first tumors that recurred as ER-negative (C1) with ER-negative first tumors that recurred as ER-negative

(B1) revealed a substantial number of hypermethylated CpG sites (34117); almost 5 times as many as the number of hypomethylated sites (6989) (Table 4.11). ER-negative second tumors (C2) had a similar number of hypermethylated CpG sites as their ER-positive first tumors (C1) when compared to the ER-negative B1 group (35809). Comparing the two ER-negative second tumor groups, B2 and C2, to each other we see that there are almost 12 times as many hypermethylated CpG sites in group C2. This indicates that while both groups B2 and C2 are ER-negative, they have distinct methylation profiles.

The last set of analyses was completed using C1 as a reference group. ER-positive first tumors were compared to their ER-negative second tumors (C2 versus C1). As expected, there were more hypermethylated CpG sites in this comparison than in the previous two first versus second tumor analyses (A2 versus A1 and B2 versus B1). There were 2.4 times as many hypermethylated CpG sites than hypomethylated sites in C2 versus C1 (18770 hypermethylated sites versus 8731 hypomethylated sites) (Table 4.12).

Functional genomic location of methylation among comparison groups

To gain a better understanding of where hypermethylation and hypomethylation occurs within gene regions in our breast tumor samples, groups were compared against each other in two graphs. The first graph, hypermethylated CpG sites within gene regions, shows that ER-negative C2 versus ER-positive C1 had the greatest number of dmCpGs (Figure 4.12). Comparing C2 with C1, revealed that the body contained the greatest number of hypermethylated CpGs, followed by the intergenic region and the promoter (Table 4.13). The group with the highest number of hypermethylated CpGs in the promoter region was ER-positive C1 compared with ER-positive A1 (Figure 4.12 and

Table 4.13). This was particularly interesting, because these are two ER-positive breast tumor groups with two different outcomes in their second tumors, one group stayed ER-positive while the other became ER-negative. Comparison of hypomethylation among the groups shows that the ER-negative groups B1 & B2 have the highest number of altered CpG sites and that both body and promoter methylation are similar in these groups (Figure 4.13 and Table 4.14).

Thirty-five previously identified CpG sites remain hypermethylated with respect to ER-status when groups are stratified by first and second tumor occurrence

Distribution analysis was completed on groups A, B and C, to determine whether the 35 CpG sites previously identified by Fackler and colleagues [114] represented overall methylation of the tumors. We found that the 23 CpG sites found to be hypermethylated in ER-positive tumors were also hypermethylated in each ER-positive subgroup in our study: A1, A2, and C1 (Figure 4.14). Methylation of the 23 CpG sites in the ER-positive groups ranged from a median average beta of 0.15 in C1, 0.18 in A1 and 0.2 in A2, whereas the ER-negative tumor groups ranged from 0.07 in B1 and B2 to 0.09 in C2 (Figure 4.14). Twelve CpG sites identified as hypomethylated in ER-negative tumors were also hypomethylated in our study (Figure 4.15). ER-negative groups B1, B2, and C2 had median average beta values that ranged from 0.22 in B1, 0.27 in B2 and 0.24 in C2, where the ER-positive groups A1, A2 and C1 had values ranging from 0.07 in A1 and C1 to 0.08 in A2 (Figure 4.15).

Pathway Analysis of differentially methylated genes among ER-positive first tumors and non-recurrent tumors

I was interested in comparing ER-positive first tumors from women who went on to experience a second tumor (recurrence) and women who did not have a tumor recurrence (NR). To understand how methylation differs between recurrent tumors and those that do not recur, I compared ER-positive first tumors that recurred as either ER-positive (A1) or ER-negative (C1) to 6 ER-positive NR tumors and the number of dmCpG sites and associated genes were determined. Results from both hypermethylated and hypomethylated analyses are presented in Table 4.15. The ER-positive recurring as ER-positive A1 group was the least different from the NR group with only 2275 hypermethylated sites associated with 1367 hypermethylated genes, whereas the ER-positive recurring as ER-negative C1 group had 4321 hypermethylated sites and 2056 associated hypermethylated genes. The number of hypomethylated genes was similar between the two groups with the A1 group having 6266 hypomethylated CpG sites and 2592 associated genes and the C1 group having 6903 hypomethylated CpG sites and 3149 associated genes. Pooled analysis indicated that there were no duplicate CpG sites in both the hyper- and hypomethylated lists from A1 vs. NR or C1 vs. NR. However, pooled analysis of the genes showed that there were 302 duplicate genes between the hyper- and hypomethylated lists in A1 vs. NR and 930 duplicate genes in C1 vs. NR (Table 4.15).

In addition to identifying the differentially methylated CpGs between primary and non-recurrent tumors, I was interested in which CpGs were differentially methylated between ER-positive primary tumors from pairs that differed in the ER-status of the

recurrent tumor. Identifying these CpGs would allow me to determine the extent to which methylation patterns in primary ER-positive tumors could predict the ER-status of the recurrent tumor and potentially contribute to personalized treatment. Therefore, I compared ER-positive first tumors from pairs that recurred as ER-negative (C1) with ER-positive first tumors from pairs that recurred as ER-positive (A1) to determine which gene pathways were altered. The ER-positive primary tumors from pairs that recurred as ER-negative (C1) had 5054 hypermethylated CpGs as compared with A1 and those CpGs corresponded with 2008 hypermethylated genes (Table 4.15). The number of hypomethylated CpG sites was 3121 in C1 with 1714 genes associated with those sites (Table 4.15). Pooled analysis of the hyper- and hypomethylated CpG sites showed no duplicated CpG sites among the groups, however there were 243 genes present in both lists.

To determine potential genes and pathways of interest, hypermethylated, hypomethylated and pooled lists were compared between the three groups to find CpG sites and genes unique to each comparison group. C1 versus A1 had the greatest number of unique CpG sites and genes in the hypermethylated group set with 3289 CpG sites and 825 genes (64% and 48% of the total respectively. Interestingly, the group with the greatest number unique hypomethylated CpG sites was A1 versus NR with 3484 (56%), however C1 versus NR had the greatest number of unique genes (870; 38%). In the pooled group, C1 versus NR had the highest number of both hyper- and hypomethylated CpG sites (4677) and genes (1147) combined with 41% and 24% of total respectively (Table 4.15).

Pathway analysis was conducted on genes with dmCpG sites in each of the three groups, A1 versus NR, C1 versus NR and C1 versus A1. The analysis separately examined CpG sites that were either hyper- or hypomethylated anywhere in the gene. The top 20 hyper- and hypomethylated pathways out of an extensive list of biological process GO terms that had a >2-fold enrichment and a p-value of <0.01 for each comparison group are presented in Tables 4.16 and 4.17. The hypermethylated pathways with the greatest fold enrichment were neuron fate commitment for A1 versus NR (4.68; p = 0.00045), neuron fate specification for C1 versus NR (>5; p = 0.00024), and dorsal spinal cord development (>5; p = 0.00095) (Table 4.16). The pathways with highest fold enrichment in the hypomethylated groups were telencephalon regionalization in A1 versus NR (>5; p=0.049), regulation of cardiac muscle cell differentiation in C1 versus NR (>5; p=0.027), and positive regulation of cardiac muscle cell differentiation (>5; p=0.016) (Table 4.17).

An intergroup hypermethylated pathway comparison was done to determine which GO terms were seen across groups. The ER-positive first tumor A1, which recurred as ER-positive and the ER-positive C1, which recurred as ER-negative groups versus NR tumors shared five hypermethylated pathways: cell fate commitment, central nervous system neuron differentiation, cell-cell adhesion via plasma membrane adhesion molecules, homophilic cell adhesion via plasma membrane adhesion molecules, and neuron fate commitment (Table 4.16, blue highlighted). ER-positive C1 tumors versus NR tumors shared five hypermethylated pathways with C1 tumors versus A1. Those pathways are cell differentiation in spinal cord, cell fate determination, dorsal spinal cord development, forebrain neuron differentiation, and neural retina development (Table

4.16, red highlighted). Among the top 20 differentially methylated pathways, A1 versus NR shared no pathways with C1 versus A1 and there were no pathways shared between all three hypermethylated groups.

Similarly, an intergroup hypomethylated pathway comparison was made to determine hypomethylated shared GO terms between groups. Nine pathways were shared between A1 versus NR and C1 versus NR, autonomic nervous system development, cell differentiation in spinal cord, embryonic eye morphogenesis, forebrain regionalization, neuron fate commitment, neuron fate specification, regulation of cardiac muscle cell proliferation, regulation of heart growth, and ventral spinal cord development (Table 4.17, orange highlighted). C1 versus NR shared four pathways with C1 versus A1: neuron fate commitment, positive regulation of cardiac muscle tissue development, regulation of cardiac muscle cell differentiation and regulation of cardiac muscle tissue development (Table 4.17, green highlighted). Three pathways were shared by A1 versus NR and C1 versus A1, ear morphogenesis, inner ear morphogenesis, and neuron fate commitment (Table 4.17, purple highlighted). All hypomethylated groups, A1 versus NR, and C1 versus NR shared one pathway, neuron fate commitment (Tables 4.16 and 4.17, red text). It is important to note that out of 19000 dmCpG sites, approximately 10% (1802) are found in both lists (Appendix M). Out of 6462 genes present in all three groups, 1833 are present in both the hyper- and hypomethylated lists (Appendix N). As multiple CpG sites can represent a single gene on the HM450BC, further analysis is needed to determine whether hyper- and hypomethylation of these genes is occurring in the same or different gene regions. A comparison of the genes shared between all six groups can be seen in Appendix O.

In addition to pathways present among groups, I was also interested in pathways present in each individual tumor group analyzed. These pathways may be a valuable tool in defining a methylation profile that will determine whether a primary tumor is prone to recur and whether the second tumor is likely to be ER-positive or ER-negative. Hypermethylated pathways likely to contain genes that offer clues to tumor recurrence were mesenchyme development and cell-cell adhesion, which was unique to A1 versus NR and endocrine system development, unique to C1 versus NR (Table 4.16 and 4.17 yellow highlighted). A hypermethylated pathway unique to C1 versus A1 where the primary tumors are ER-positive, but the second tumors are either ER-negative (C2) or ER-positive (A2) is the non-canonical WNT signaling pathway, which plays a role in tumor progression (Table 4.16, yellow highlighted) [117]. The non-canonical WNT signaling pathway is also present in the A1 versus NR in the hypomethylated pathway set (Table 4.17, yellow highlighted). A hypomethylated pathway only found in C1 versus NR tumors is the mesenchymal to epithelial transition, genes from this pathway are found to be upregulated in metastatic tumors allowing new tumors to form in distal sites [118]. Lastly, four hypomethylated pathways in the C1 versus A1 group are likely to provide information on why tumors recur as either ER-negative or ER-positive. Regulation of morphogenesis of a branching structure, which includes genes involved in mammary gland development, regulation of BMP signaling pathway and positive regulation of ossification, which contain many of the same growth factors involved in tumor proliferation and migration [119], and homophilic cell adhesion via plasma membrane adhesion molecules, genes from which assist in linking cells through identical plasma membrane adhesion molecules (Table 4.17).

Two pathways, homophilic cell adhesion via plasma membrane adhesion molecules and cell fate commitment were hypermethylated in A1 versus NR and C1 versus NR, but hypomethylated in C1 versus A1. Homophilic cell adhesion via plasma membrane pathway is a child term of the cell-cell adhesion via plasma membrane pathway, the ancestor chart for which can be seen in Figure 4.16. Of the 67 genes altered among the three groups in the homophilic cell adhesion pathway, 43 genes (64%) are in the protocadherin family, a subgroup of the cadherin family, from which there were also 7 genes (Table 4.18). Table 4.18 shows the number of genes shared between each group. A1 versus NR and C1 versus NR shared the most genes between groups with 16 (blue highlighted), A1 versus NR shared 4 genes with C1 versus A1 (orange highlighted) and C1 versus NR shared 1 gene with C1 versus A1 (Table 4.18, green highlighted). Twenty genes were shared by all three groups, of which almost all were protocadherins (90%) (Table 4.18, purple highlighted). The other two genes shared by all three groups were Calsyntenin 2 (CLSTN2), the function of which is unknown although it is believed to be involved in calcium signaling (humanproteinatlas.org), and FRAS1 Related Extracellular Matrix Protein 2 (FREM2), amplification of which is found in gliasarcomas undergoing mesenchymal transition [120]. The gene CDH13 was found to be hypermethylated only in A1 versus NR whereas C1 versus NR had 13 genes and C1 versus A1 had 12 genes (Table 4.18). Interestingly, high expression of the genes hypermethylated (and presumably down-regulated) in C1 versus NR was associated with better survival as calculated with simple Kaplan-Meier plots using TCGA data (Table 4.19) (http://kmplot.com/). Likewise, high expression of the genes hypermethylated in C1 as compared to A1 was also associated with a better in survival (Table 4.19).

The other pathway, cell fate commitment, a child term of cell differentiation, had 116 genes altered in all three pathways combined (Figure 4.17 and Table 4.20). Twenty-one genes were in both A1 versus NR and C1 versus NR hypermethylated lists (blue highlighted), 14 genes were in both C1 versus NR hypermethylated and C1 versus A1 hypomethylated (green highlighted), and 8 genes were hypermethylated in A1 versus NR, but hypomethylated in C1 versus A1 (Table 4.20, orange highlighted). Kaplan-Meier plots of the TCGA data revealed that high expression of the genes hypermethylated in both A1 versus NR and C1 versus NR was linked to good survival (Table 4.21). Further analysis of these genes revealed that 21 out of 65 CpG sites identified were hypermethylated in both groups (Table 4.21, MAPINFO). One gene, PAX3, has three CpG sites shared between the two groups and two of the CpG sites are located within 400 bases of each other. Three other genes, NKX2-5, PAX6, and VSX2, have two shared CpG sites, however the CpG sites are located at least 2000 bp apart.

DNA transcription factors and signaling proteins were among the genes with altered methylation. One gene, RORA, which is shared by A1 versus NR hypermethylated and C1 versus A1 hypomethylated, was shown to be hypermethylated in the promoter region of both the ER-positive and ER-negative breast cell lines TMX2-11 and TMX2-28 as compared with the ER-positive parental MCF-7 line (Chapter 3). However, among the tumors, dmCpGs in RORA were located in the body region of the gene (Figure 4.18). Seven genes are shared by all three groups: GFI1, HES5, PAX2, PRDM1, SATB2, SOX8, and TLX3 (Table 4.20, purple highlighted). In Chapter 3 we also reported that GFI1 was hypermethylated in both TMX2-11 and TMX2-28 as compared with MCF-7. In agreement with the cell line observations, the heatmap in

Figure 4.19 shows that GFI1 is hypermethylated in the TSS1500 region in all three groups of tumors.

Lastly, all groups except the C1 versus A1 hypermethylated shared one pathway, neuron fate commitment. This pathway is a child term of cell fate commitment and accordingly, shared 46 genes with that pathway (Figure 4.20 and Table 4.22). AT-rich sequence-binding protein-2 (SATB2), a DNA transcription factor involved in chromatin remodeling, was the only gene shared by all 5 pathways. Clustering analysis based on the 36 SATB2 CpGs interrogated on the HM450BC shows that the group A1 is more similar to the NR group than the C1 group (Figure 4.21). C1 is hypomethylated as compared with NR in 6 CpG sites and hypermethylated in 5 CpG sites. A1 is hypermethylated as compared with NR in 8 CpG sites and hypomethylated in 7 CpG sites (Figure 4.21). Compared with A1, C1 is hypermethylated 4 CpG sites all located in the body of the gene, and hypomethylated in 11 CpG sites across the gene (Figure 4.21).

Discussion

While patients with specific breast cancer subtypes are treated similarly with regards to therapy, breast cancer is a heterogeneous disease. Recent literature is only beginning to provide an understanding as to why patients with similar subtypes do not have similar outcomes after chemotherapeutic treatment [114]. Currently, tumors are classified based on expression of ER, PR, and HER2 as well as expression of the luminal markers CK8/18 and the basal markers CK 5/6, CK14 and Ki67. The use of the anti-estrogen, Tamoxifen along with chemotherapy, extends long-term, disease free survival

for women with ER-positive tumors and reduces the risk of a contralateral tumor recurrence [121]. However not all tumors respond to treatment, approximately 30% of women have disease recurrence [26]. Aside from chemotherapy, there are limited treatments available for ER-negative tumors and consequently, women with ER-negative primary tumors are more likely to have recurrence within 5 years of treatment [121]. It has also been shown that up to 80% of ER-negative patients with tumors 1cm or less respond well to local therapy and remain disease free [121]. ER-negative tumors are also more unlikely to recur as contralateral tumors than ER-positive tumors [121]. DNA methylation has been suggested as a mechanism able to provide clues into tumor disparity and recurrence likelihood.

DNA methylation provides relatively stable biomarkers for exploration. High-throughput analysis from the HM450BC provides information on 485,000 CpG sites in the human genome, with enrichment in the promoter region. In Chapters 2 and 3, I discussed HM450BC analysis of human breast cancer cell lines treated long-term with Tamoxifen. I found that the ER-positive Tamoxifen-selected cell line, TMX2-11, and the ER-negative Tamoxifen-selected cell line, TMX2-28, were hypermethylated as compared to the parent, ER-positive cell line, MCF-7. Using this knowledge, I hypothesized that both ER-positive and ER-negative second human breast tumors occurring after anti-estrogen treatment would be hypermethylated. Additionally, I predicted that ER-negative second tumors occurring after ER-positive first tumors would be more hypermethylated than ER-positive second tumors occurring after ER-positive first tumors. I collected genome-wide methylation data on 70 tumors analyzed by HM450BC. Tumor characteristic data were collected for all 70 tumors, but given the small sample size for

each tumor group, it was not feasible to further stratify the tumors based on characteristics.

Recently, Fackler et al. used the HM27BC to analyze 103 sporadic invasive primary breast cancers and found that ER-positive tumors are more highly methylated than ER-negative tumors [114]. They then selected a panel of candidate 40 CpGs that were among those with the highest fold-change differences and hypermethylated in either only the ER-positive or only the ER-negative tumor sets as potential biomarkers. In the present study, I compared methylation of ER-positive tumors with ER-negative tumors and found a similar methylation pattern as previously reported; ER-positive tumors were more highly methylated than ER-negative ones. I then examined methylation of the 35 CpG sites available on the HM450BC out of the 40 CpG sites previously described as being hypermethylated in either ER-positive or ER-negative tumors. The overall methylation signature of the 35 CpG sites shows a similar trend with previously reported results; the median average beta of ER-positive tumors was higher than ER-negative tumors (0.19 vs. 0.1). However, unlike the previous study, methylation of the CpG sites in the present study is low and not all sites had a significant change in methylation.

Building on the knowledge that methylation is affected by ER-status; I considered that methylation was also influenced by tumor occurrence. Tumors were separated into two groups, ER-positive versus ER-negative first tumors and ER-positive versus ER-negative second tumors. I determined that greater methylation of ER-positive tumors is driven by a decrease in methylation among ER-negative first tumors, which is consistent with the literature as no other group has compared methylation of recurrent tumors [63, 114]. Interestingly, ER-negative second tumors had a similar median methylation as ER-

positive first and second tumors. This was an unexpected finding as overall methylation of the ER-negative TMX2-28 cell line was greater than that of the ER-positive TMX2-11, although both were derived from the same parental line after prolonged Tamoxifen treatment. I expected to see that ER-negative second tumors had greater methylation than the ER-positive second tumors. However, ER-negative second tumors did have higher methylation than ER-negative first tumors, which was presumed to be associated with either tumor progression or the inclusion of ER-negative second tumors recurring from ER-positive first tumors. Therefore, the next set of analyses focused on groups of 50 paired first and second tumors from the same woman stratified by ER-status and occurrence.

In fact, I did find that separation of the tumors by their first tumor and recurrent tumor mattered greatly. ER-negative second tumors as compared to their matched ER-positive primary tumors have the greatest number of hypermethylated sites among the three groups analyzed. Additionally, I found that Group B1, ER-negative first tumors, when compared separately to Group A1, ER-positive first tumors, had 38,000 hypomethylated CpG sites, which likely accounts for the overall hypomethylation of the ER-negative tumors. Hypermethylation of the ER-negative second tumors can be attributed to Group C2, which had 35,000 hypermethylated CpG sites when compared to Group B1. These data closely resemble those of TMX2-28 as compared with MCF-7, suggesting that methylation plays a role in tumor progression. Similarly to what is occurring in TMX2-28, it is important to note that loss of ER could be playing a role in the overall methylation increase in ER-negative second tumors recurring from ER-

positive first tumors. Future analysis would control for this by comparing the Groups C2 with A2 and determining the CpG sites and genes similarly methylated in both groups.

As described previously, second tumors are generally referred to as a recurrence as they occur after or during treatment of a primary tumor [111]. However, this terminology may not be appropriate in some cases as the second tumor may be a new tumor rather than a *de novo* tumor. Therefore, I sought to determine whether the methylation profiles of paired tumors occurring in the ipsilateral and contralateral breasts differ. A comparison of ER-positive second tumors showed that tumors recurring in the ipsilateral breast had greater methylation than those in the contralateral breast, whereas ER-negative second tumors saw greater methylation in the contralateral breast as compared to the ipsilateral breast. All tumors from the ER-negative second contralateral tumor group were from group C2, while ER-negative ipsilateral tumors are from group B2 and C2. Higher methylation in recurrent ER-negative tumors from women with ER-positive primary tumors is consistent with the TMX2-28 data. Tumors were then separated into first and second paired comparison groups as described above, controlling also for side of tumor recurrence. In general, I saw that second tumors from ipsilateral pairs had higher methylation than the first tumors; in contrast, first tumors from contralateral pairs had higher methylation than in the second tumor. These data taken with the fact that tumor progression is associated with an increase in methylation are consistent with the prediction that ipsilateral, not contralateral tumors are more likely to be a true recurrence.

Until recently, DNA methylation studies have primarily focused on the promoter region as it is well established that promoter methylation affects gene expression [122].

Emerging research in the field has shown that in addition to promoter methylation, body methylation also affects gene expression and treatment with a demethylating agent reduces aberrant overexpression of the gene [44]. To understand where methylation is occurring in the present study, comparison groups were assessed for dmCpGs in each gene region. I found that although the BeadChip is enriched for promoter region, ER-negative paired second tumors, the body has greater hypermethylation than the promoter regardless of the ER-status of the paired first tumor (ER-positive or ER-negative). Additionally, we see that the ER-negative second tumors in group B2 compared with its paired ER-negative first tumors, B1, and the other ER-negative second recurrent tumor group, C2, have more hypomethylated CpG sites in the body. To establish potential biomarkers for the primary and recurrent tumor groups discussed in this chapter, future studies would focus on methylation changes in relation to the functional genomic location. This will provide further insight into how methylation shapes each tumor group.

Pathway analysis focused on the ER-positive first tumor groups, A1 and C1, and ER-positive, non-recurrent tumors. This was done to gain understanding of the pathways and associated genes affected by methylation in i) tumors that recur as opposed to tumors that do not recur and ii) ER-positive tumors that recur as ER-negative compared to those that recur as ER-positive.

It was also important to understand how methylation of genes compare between the three analyses as it would provide insight into biomarkers associated with ER-positive tumors that recur as either ER-positive or ER-negative. Genes shared between hypermethylated in ER-positive first tumor groups A1 and C1 compared to NR and hypomethylated in C1 compared to A1 were determined. This was done to find genes that

while, hypermethylated in the ER-positive first tumors recurring as ER-negative as compared with the non-recurrent, are less methylated than those genes hypermethylated in ER-positive first tumors recurring as ER-positive. Two pathways homophilic cell adhesion via plasma membrane adhesion molecules and cell fate commitment fit these terms.

To associate methylation changes in my tumor data set with breast cancer patient survival data, Kaplan-Meier plots were created using TGCA breast cancer data available online for those genes in the homophilic cell adhesion pathway which were only hypermethylated in each of the 3 individual groups (Table 4.18, white highlighted genes). I found that low expression is linked to decreased survival in 10 of the 13 genes only hypermethylated in ER-positive first tumors that recur as ER-negative (C1) as compared with non-recurrent tumors. Likewise, 6 of 12 genes hypermethylated only in ER-positive first tumors recurring as ER-negative (C1) as compared with ER-positive first tumors recurring as ER-positive (A1) are associated with survival. A single gene, CDH13, was hypermethylated in A1 versus NR and expression was not associated with survival. Similar analysis was completed for genes shared between A1 versus NR and C1 versus NR in the cell fate commitment pathway. High expression was linked to good survival in 13 out of 21 shared genes. Preliminary analysis revealed that of the genes idenified as hypermethylated in both groups, one-third of the CpG sites were shared between the groups. Of those, only one gene had 2 CpG sites located within 500 bp of each other, indicating that individual sites, not clusters are affected. These data suggest that the genes hypermethylated only in ER-positive tumors recurring as ER-negative are a potential a

signature for poor survival, however additional research is needed to determine the extent to which these genes are playing a role.

Finally, two genes shared in the cell fate commitment pathway, RORA and GFI1 were also found to be hypermethylated in both TMX2-11 and TMX2-28 and were analyzed for expression in Chapter 3. While RORA was found to be hypermethylated in the promoter region of both TMX2-11 and TMX2-28, expression was low and did not increase with 5-Aza treatment. In the current study, I saw that differential RORA methylation was restricted to the body region of the gene, suggesting a different mechanism of control. In both the tumor study and the cell line data, methylation of GFI1 occurred in the promoter region. However, GFI1 expression analysis in TMX2-11 and TMX2-28 showed that although the gene was hypermethylated in the promoter, expression the gene was not significantly different from MCF-7.

Future Directions

I have presented data that suggest that methylation is a useful biomarker for determining whether a tumor is likely to recur and gives clues as to whether the recurrent tumor will be hormone receptor positive or negative. Small samples size is one limitation of the current study. Additional tumor samples from each of the groups would allow for further stratification based on characteristics such as tumor type and treatment. Further investigation into side of tumor recurrence is also important to determine whether a tumor is a true recurrence or a new tumor. A larger tumor database would also allow for

investigation of how methylation relates to patient demographics. While it has been established that DNA methylation can be used as a method of determining whether first and second tumors are clonally related, using array comparative genomic hybridization (aCGH) would confirm the relationship [111].

To further understand which genes are unique biomarkers in each tumor group, additional analyses are needed. One analysis would compare ER-negative second tumors that occur from either ER-positive or ER-negative primary tumors. While these tumors are both hormone receptor negative and the gross pathology would look similar, they arose from different primary tumor types. Uncovering these various molecular subtypes would lead to personalized care.

Finally, as recent literature has described the importance of body methylation in gene expression, it is prudent to examine the functional genomic location in relation to primary tumors and second tumor outcomes. Using HM450BC data to focus on different gene regions would provide an understanding as to how functional genomic location plays a role in methylation of each tumor subgroup.

Table 4.1 DNA concentrations and bisulfite conversion controls for all 86 tumors

ER+ to ER+											
First Tumor				Second Tumor				Third Tumor			
Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value	Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value	Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value
A1A	26.51	17.1	31.1	A1B	31.74	14.4	28.0				
A2A	21.15	18.6	33.3	A2B	79.95	14.2	29.1				
A3A	11.8	22.2	36.2	A3B	36.96	16.3	29.9				
A4A	25.36	21.9	35.5	A4B	No Block Available						
A5A	20.19	22.7	36.8	A5B	12.06	18.5	33.4				
A6A	No Block Available			A6B	31.91	22.1	35.9	A6C	42.66	23.8	26.5
A7A	22.8	18.2	34.2	A7B	15.47	16.3	29.5				
A8A	12.6	17.7	30.4	A8B	24.14	18.3	33.4	A8C	56	25.36	21.9
A9A	27.1	23.1	38.1	A9B	30.14	18.0	32.0	A9C	47.36	16.0	29.9
A11A	38.64	18.2	31.7	A11B	47.23	17.0	31.1	A11C	16.7	19.2	33.0
A10A	No Block Available			A10B	20.55	20.5	34.3				
A12A	23.16	25.3	41.3	A12B	29.15	16.0	29.9				
A13A	53.22	18.1	31.8	A13B	25.34	17.5	31.1				
A14A	22.88	14.7	28.8	A14B	30.08	18.0	31.9				
A15A	25.27	14.8	28.7	A15B	25.4	18.2	32.2				
A16A	47.01	14.9	28.1	A16B	26.45	17.4	31.0				
A17A	17.2	18.1	31.4	A17B	46.87	18.4	34.2				
A18A	16.46	19.9	32.7	A18B	24.68	17.3	30.2				
A19A	24.25	18.1	30.6	A19B	39.44	14.6	27.3				
A20A	37.82	17.2	30.9	A20B	32.39	13.5	27.7				
A21A	36.13	17.1	30.6	A21B	32.74	16.3	29.8				

Red highlighted samples did not pass QC, no block could be found for blue highlighted samples; DNA concentration in μg/μl

Table 4.1 Cont. DNA concentrations and conversion controls for all 86 tumors

ER- to ER-											
First Tumor				Second Tumor				Third Tumor			
Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value	Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value	Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value
B1A	17.91	20.8	33.9	B1B	18	17.3	31.7				
B2A	55.14	16.8	29.9	B2B	71.99	14.5	27.1				
B3A	57.13	17.7	31.2	B3B	44.37	15.1	28.7				
B4A	87.51	16.8	30.0	B4B	67.7	15.4	28.4	B4C	65.91	17.1	30.3
B5A	28.96	16.8	30.8	B5B	24.06	18.9	32.2	B5C	65.27	16.1	29.8
B7A	76.85	18.9	32.9	B7B	126.1	18.2	34.5				
B8A	45.55	16.4	30.2	B8B	68.49	19.7	34.4				
B9A	No Block Available			B9B	46.51	16.5	30.2				

Red highlighted samples did not pass QC, no block could be found for blue highlighted samples; DNA concentration in μg/μl

Table 4.1 Cont. DNA concentrations and conversion controls for all 86 tumors

ER+ to ER-											
First Tumor				Second Tumor				Third Tumor			
Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value	Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value	Patient & Tumor ID	DNA Concentration	ALU-C4 (HB-313) Ct value	CONV-100% (HB-365) Ct value
C1A	No Block Available			C1B	23.93	16.2	29.5				
C2A	49.72	15.7	28.9	C2B	22.79	15.5	28.4				
C3A	20.72	19.4	32.7	C3B	107.5	16.0	30.3				
C4A	23.62	17.6	31.0	C4B	43.8	14.7	28.2				
C5A	19.06	14.5	28.4	C5B	32.04	15.6	29.5				
C6A	21.89	18.7	34.2	C6B	120.7	16.9	30.9				
C7A	33.26	17.7	31.4	C7B	24.77	18.4	34.5	C7C	18.44	22.6	36.8
C8A	26.49	19.6	33.4	C8B	16.28	20.1	34.1				

Red highlighted samples did not pass QC, no block could be found for blue highlighted samples; DNA concentration in µg/µl

Table 4.1 Cont. DNA concentrations and conversion controls for all 86 tumors

No Recurrence			
First Tumor			
Patient & Tumor ID	DNA Concentration	**ALU-C4 (HB-313) Ct value**	**CONV-100% (HB-365) Ct value**
N1	28.57	18.2	33.6
N2	21.75	19.2	34.4
N3	26.24	15.9	31.2
N4	23.83	16.2	33.3
N5	21.61	16.8	32.1
N6	28.46	12.2	29.1
N7	103.5	17.3	34.7
N8	32.4	16.4	32.7
N9	36.85	19.8	None
N10	26.46	17.2	33.2

Red highlighted samples did not pass QC; DNA concentration in μg/μl

Table 4.2 Patient and tumor characteristics		
Age (in years) mean (SD) range		
At Primary	56.6 (12.3), 37 - 84	
At Recurrence	64.1 (12.6), 39 - 90	
Menopausal n (%)		
At Primary	13 (38%)	
At Recurrence	21 (58%)	
TTR[1] (in months) mean (SD) range	67.8 (59.5) 10-252	
	Tumors	
	Primary[2] **(34)**	**Recurrent (36)**
ER status n (%)		
+	26 (76.5)	19 (52.8)
-	8 (23.5)	17 (47.2)
PR status n (%)		
+	20 (58.8)	16 (44.4)
-	14 (41.2)	20 (55.6)
HER2 status n (%)		
+	5 (14.7)	9 (25)
-	29 (85.3)	27 (75)
Ki67 IHC n (%)		
low (≤15)	26 (76.5)	20 (55.6)
high (>15)	8 (23.5)	16 (44.4)
Tumor Grade n (%)[3]		
0	5 (15.1)	5 (15.1)
1	5 (15.1)	2 (6.1)
2	11 (33.3)	5 (15.1)
3	12 (36.3)	21 (63.6)
Tumor Type n (%)		
DCIS	5 (14.7)	5 (13.9)
IDC	23 (67.6)	26 (72.2)
ILC	4 (11.8)	4 (11.1)
IDLC	2 (5.9)	1 (2.8)
Location of Recurrence(s) n (%)[4, 5]		
Ipsilateral to primary	NA	21 (60)
Contralateral to primary	NA	14 (40)
Tumor Size n (%)[3]		
≥20 mm	13 (48.2)	13 (52)
<20 mm	14 (51.8)	12 (48)

Anti-hormonal Therapy n (%)[3]		
No	7	NA
Yes, Tam	4	NA
Yes, AI	7	NA
Yes, Tam & AI	5	NA
Chemotherapy Type n (%)[3, 6]		
AC + paclitaxel	6 (23.1)	NA
AC + docetaxel	3 (11.6)	NA
Other	5 (19.2)	NA
None	12 (46.1)	NA

[1]TTR = time to recurrence
[2]Primary tumors include the 8 non-recurrent tumors
[3]Indicates that data are missing for some samples; percentages are calculated on the available data
[4]Includes second recurrences; missing laterality for one tumor
[5]Follow-up identified two ER-positive samples as lung metastases; analyses are done under the assumption that the tumor occurred in the opposite breast
[6]AC = Adriamycin (doxorubicin) and C = cyclophosphamide
NA = not applicable
See text for scoring of ER, PR, HER2 and Ki67

Table 4.3 Sample group re-labeling for readability

	Group		
Tumor	**A**	**B**	**C**
First (1)	ER-positive	ER-negative	ER-positive
Second (2)	ER-positive	ER-negative	ER-negative

Table 4.4 Patient and tumor demographics by groups A, B, C and non-recurrent ER-positive

A. Paired ER-positive primary tumors to ER-positive recurrent tumors

		Primary Tumor (n=14)	**Recurrent Tumor (n=14)**
Age	mean (SD) range	56.8 (15.5) 37-84	64.1 (15.1) 40-90
TTR[1]			87.6 (60.1) 12-252
ER		7.6 (0.8) 5-8	7.6 (1.1) 4-8
PR		5.9 (2.5) 0-8	4.5 (2.8) 0-8
HER2		0.4 (0.8) 0-3	0.6 (1.1) 0-3
Ki67		6.8 (10.0) 0-30	8.9 (10.8) 0-30
Tumor Grade[2] (n)			
0		1	2
1		4	2
2		5	3
3		3	6
Tumor Type		DCIS (1), IDC (10), ILC (2), IDLC (1)	DCIS (2), IDC (8), ILC (3), IDLC (1)
Location[3]			Ipsi: 8 Contra: 6
Tumor Size		≥20 mm (6), <20 mm (8)	≥20 mm (8), <20 mm (6)
Chemotherapy Type[2, 4]			
AC + paclitaxel		3	
AC + docetaxel		0	
Other		2	
None		7	

B. Paired ER-negative primary tumors to ER-negative recurrent tumors

		Primary Tumor (n=5)	**Recurrent Tumor (n=5)**
Age	mean (SD) range	61.4 (12.9) 46-79	64.2 (13.6) 48-80
TTR[1]			34 (33.14) 10-90
ER		0 (0) 0-0	0 (0) 0-0
PR		0 (0) 0-0	0 (0) 0-0
HER2		0.6 (1.3) 0-3	0.6 (1.3) 0-3
Ki67		13.8 (11.0) 2-25	26 (6.5) 20-35
Tumor Grade (n)[2]			
0		0	0

	1	0	0
	2	0	0
	3	5	5
Tumor Type		IDC (5)	IDC (5)
Location			Ipsi: 5
Tumor Size		≥20 mm (3), <20 mm (2)	≥20 mm (3), <20 mm (2)
Chemotherapy Type[2, 4]			
AC		0	
AC + paclitaxel		1	
AC + docetaxel		3	
Other		0	
None		1	

C. Paired ER-positive primary tumors to ER-negative recurrent tumors

		Primary Tumor (n=6)	**Recurrent tumor (n=6)**
Age	mean (SD) range	53 (8.2) 42-65	60.3 (6.1), 53-68
TTR[1]			82.8 (77.7) 17-216
ER		7.8 (0.4) 7-8	0 (0) 0-0
PR		3.8 (4.2) 0-8	0 (0) 0-0
HER2		1.2 (1.5) 0-3	1.5 (1.6) 0-3
Ki67		5.7 (5.8) 0-15	16.2 (14.2) 0-35
Tumor Grade[2] (n)			
	0	2	2
	1	0	0
	2	3	0
	3	1	4
Tumor Type		DCIS (2), IDC (3), IDLC (1)	DCIS (2), IDC (4)
Location			Ipsi: 2 Contra: 4
Tumor Size		≥20 mm (5), <20 mm (1)	≥20 mm (4), <20 mm (2)
Chemotherapy Type[2, 4]			
AC + paclitaxel		0	
AC + docetaxel		0	
Other		1	
None		0	

D. ER-positive non-recurrent tumors (n=6)

Age	mean (SD) range	53.2 (8.9) 44-69
ER		8 (8) 8-8
PR		3.7 (3.6) 0-8

HER2	0.7 (1.2) 0-3
Ki67 IHC	2.7 (3.8) 0-10
Tumor Grade[2] (n)	
0	1
1	1
2	3
3	1
Tumor Type	DCIS(1), IDC(3), ILC(2)
Tumor Size	≥20 mm (3), <20 mm (3)
Chemotherapy Type[2]	
AC + paclitaxel	2
AC + docetaxel	0
Other	1
None	3

[1]TTR = time to recurrence
[2]Indicates that data are missing for some samples; percentages are calculated on the available data
[3] Follow-up identified two samples as lung metastases; analyses are done under the assumption that the tumor occurred in the opposite breast
[4]AC = Adriamycin (doxorubicin) and C = cyclophosphamide
See text for scoring of ER, PR, HER2 and Ki67

Table 4.5 All tumors separated by first recurrence

	ER-positive to ER-positive	ER-negative to ER-negative	ER-positive to ER-negative	Non-recurrent
First Tumor	14 (0)	6 (1)	6 (0)	8
Second Tumor	18 (4)	7 (2)	7 (1)	
Third Tumor	2 (1)	2	0	
Totals	34 (5)	15 (3)	13 (1)	8

Numbers in parentheses are unpaired samples, where no paired tumor was included in total.

Table 4.6 Paired tumors separated by occurrence and ER-status

		ER-positive to ER-positive	ER-negative to ER-negative	ER-positive to ER-negative
		A	**B**	**C**
First Tumor	**1**	14	5	6
Second Tumor	**2**	14	5	6
Totals		28	10	12

Red highlighted boxes represent ER-positive samples, blue highlighted boxes represent ER-negative samples.

Table 4.7 Group sizes[1] for first and second tumors stratified by location

	Ipsilateral	Contralateral
First Tumor	15	10
Second Tumor	15	10
Total	30	20

[1]Includes paired tumors only

Table 4.8 Group sizes[1] for all ER-positive and ER-negative tumors stratified by occurrence and location

		Ipsilateral	Contralateral
First Tumor	**ER-positive**	10	10
	ER-negative	5	0
Second Tumor	**ER-positive**	8	6
	ER-negative	7	4
Total		30	20

[1]Includes paired tumors only

[2]Only second tumors were included in the analysis

Table 4.9 Group sizes for paired tumors stratified by ER-status, occurrence and location

Group Separated by First and Second Tumor	Ipsilateral	Contralateral
A1 (ER-positive)	8	6
A2 (ER-positive)	8	6
B1 (ER-negative)	5	0
B2 (ER-negative)	5	0
C1 (ER-positive)	2	4
C2 (ER-negative)	2	4
Total	30	20

Table 4.10

Hyper- and hypomethylation of CpG sites using Group A1 as the reference

	ER-positive to ER-positive	**ER-negative to ER-negative**	**ER-positive to ER-negative**
	Group A	Group B	Group C
1st Tumor (1)	REFERENCE	4491 / 38044	5056 / 3131
2nd Tumor (2)	914 / 1588	---	14329 / 7936

Number in top half is hypermethylated CpG sites, bottom half is hypomethylated CpG sites hypermethylation: >1.5-fold change and DiffScore >22; hypomethylation: <0.67-fold change and DiffScore <-22. Detection p-value of < 0.01 was used to distinguish statistically significant methylation changes.*

Table 4.11

Hyper and hypomethylation of CpG sites using Group B1 as the reference

	ER-positive to ER-positive (n =14 pairs)	**ER-negative to ER-negative (n = 5 pairs)**	**ER-positive to ER-negative (n = 6 pairs)**
	Group A	Group B	Group C
1st Tumor (1)	---	REFERENCE	34117 / 6989
2nd Tumor (2)	==	3036 / 5815	35809 / 1504

Number in top half is hypermethylated CpG sites, bottom half is hypomethylated CpG sites hypermethylation: >1.5-fold change and DiffScore >22; hypomethylation: <0.67-fold change and DiffScore <-22. Detection p-value of < 0.01 was used to distinguish statistically significant methylation changes.*

Table 4.12

Number of hyper- and hypomethylated CpG sites using Group C1 as the reference

	ER-positive to ER-positive	**ER-negative to ER-negative**	**ER-positive to ER-negative**
	Group A	Group B	Group C
1st Tumor (1)	**---**	6813 / 34722	REFERENCE
2nd Tumor (2)	**---**	**---**	18770 / 8731

Number in top half is hypermethylated CpG sites, bottom half is hypomethylated CpG sites hypermethylation: >1.5-fold change and DiffScore >22; hypomethylation: <0.67-fold change and DiffScore <-22. Detection p-value of < 0.01 was used to distinguish statistically significant methylation changes.

Table 4.13 Hypermethylated CpG sites within gene regions

	A2/A1	B2/B1	C2/C1	C1/A1	B1/C1	B2/C2
Promoter	376	967	4541	2200	42	1279
5'UTR/first Exon	256	549	3146	1397	265	877
Body	245	843	6574	1166	927	775
3'UTR	3	47	658	59	56	63
Intergenic	164	894	5552	833	1006	689

Table 4.14 Hypomethylated CpG sites within gene regions

	A2/A1	B2/B1	C2/C1	C1/A1	B1/C1	B2/C2
Promoter	790	1116	3598	1103	11654	16818
5'UTR/first Exon	395	1225	2172	567	6052	8896
Body	330	1786	2014	916	9885	17456
3'UTR	31	109	124	64	738	1369
Intergenic	277	1234	1688	3131	9326	14708

Table 4.15 Differentially methylated CpG sites and genes within groups

	Hypermethylated			Hypomethylated			Pooled		
	A1-NR	**C1-NR**	**C1-A1**	**A1-NR**	**C1-NR**	**C1-A1**	**A1-NR**	**C1-NR**	**C1-A1**
Number of sites with altered methylation	2275	4321	5054	6266	6903	3121	8541	11224	8175
Number of gene names with altered methylation	1367	2056	2008	2592	3149	1714	3657	4700	3479
Number of altered sites unique[1] to the pairing (%[2])	1454 (64%)	1781 (41%)	3289 (65%)	3484 (56%)	2896 (42%)	1773 (57%)	3201 (37%)	4677 (41%)	3278 (40%)
Number of altered gene unique to the pairing (%[2])	661 (48%)	532 (26%)	825 (41%)	846 (32%)	870 (28%)	653 (38%)	634 (17%)	1141 (24%)	663 (19%)

[1]Refers to unique CpGs for parings within either the hypermethylated or hypomethylated set
[2]Percentage is the number of unique CpG sites or genes compared to the total number of CpG sites or genes with duplicate terms removed.

Table 4.16 Top 20 hypermethylated pathways in each comparison group

GO biological process complete	Genes in pathway (1189)	Fold Enrichment	P-value
A1 vs. NR hypermethylated			
neuron fate commitment (GO:0048663)	19	4.68	4.53E-04
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)	41	4.63	1.86E-11
cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742)	44	3.72	2.99E-09
cell fate commitment (GO:0045165)	44	3.2	3.86E-07
regulation of heart contraction (GO:0008016)	27	2.97	7.78E-03
mesenchyme development (GO:0060485)	29	2.93	4.06E-03
ear development (GO:0043583)	33	2.76	2.58E-03
central nervous system neuron differentiation (GO:0021953)	27	2.72	3.98E-02
inner ear development (GO:0048839)	28	2.64	4.67E-02
embryonic organ morphogenesis (GO:0048562)	41	2.46	2.19E-03
regulation of neuron differentiation (GO:0045664)	67	2.42	7.03E-07
sensory organ morphogenesis (GO:0090596)	35	2.39	2.81E-02
forebrain development (GO:0030900)	48	2.37	6.52E-04
pattern specification process (GO:0007389)	57	2.36	4.56E-05
regulation of neuron projection development (GO:0010975)	46	2.35	1.39E-03
regionalization (GO:0003002)	43	2.29	6.70E-03
cell-cell adhesion (GO:0098609)	80	2.28	1.84E-07
tube morphogenesis (GO:0035239)	45	2.21	9.75E-03
regulation of neurogenesis (GO:0050767)	75	2.19	4.76E-06
positive regulation of neurogenesis (GO:0050769)	44	2.18	1.95E-02
regulation of system process (GO:0044057)	53	2.18	1.84E-03

Table 4.16 Cont.			
C1 vs. NR hypermethylated			
neuron fate specification (GO:0048665)	16	> 5	2.38E-04
dorsal spinal cord development (GO:0021516)	11	> 5	4.42E-02
neuron fate commitment (GO:0048663)	28	4.51	9.16E-07
cell differentiation in spinal cord (GO:0021515)	22	4.41	1.24E-04
cell fate specification (GO:0001708)	26	4.25	1.41E-05
forebrain generation of neurons (GO:0021872)	24	4.22	6.50E-05
cell fate determination (GO:0001709)	16	4.07	2.99E-02
neural retina development (GO:0003407)	17	4.05	1.59E-02
forebrain neuron differentiation (GO:0021879)	19	3.88	7.66E-03
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)	50	3.69	1.26E-10
central nervous system neuron differentiation (GO:0021953)	56	3.68	3.75E-12
central nervous system neuron development (GO:0021954)	23	3.6	2.10E-03
dorsal/ventral pattern formation (GO:0009953)	29	3.42	1.91E-04
spinal cord development (GO:0021510)	31	3.41	6.83E-05
cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742)	61	3.37	9.37E-12
cell fate commitment (GO:0045165)	71	3.37	4.40E-14
cerebellum development (GO:0021549)	25	3.25	4.25E-03
endocrine system development (GO:0035270)	37	3.21	1.33E-05
camera-type eye morphogenesis (GO:0048593)	30	3.15	6.46E-04
pancreas development (GO:0031016)	23	3.13	2.16E-02
neuron migration (GO:0001764)	31	3.11	5.15E-04

Table 4.16 Cont.			
C1 vs. A1 hypermethylated			
dorsal spinal cord development (GO:0021516)	13	> 5	9.50E-04
cerebellar cortex formation (GO:0021697)	11	> 5	4.20E-02
cerebellar granular layer development (GO:0021681)	9	> 5	4.17E-02
cell differentiation in hindbrain (GO:0021533)	12	> 5	2.40E-02
dopamine receptor signaling pathway (GO:0007212)	13	> 5	1.36E-02
cerebral cortex neuron differentiation (GO:0021895)	12	> 5	1.04E-02
spinal cord association neuron differentiation (GO:0021527)	10	> 5	9.22E-03
calcium ion-dependent exocytosis of neurotransmitter (GO:0048791)	14	> 5	9.62E-04
cerebellar cortex morphogenesis (GO:0021696)	13	4.82	4.02E-02
reproductive behavior (GO:0019098)	13	4.82	4.02E-02
non-canonical Wnt signaling pathway (GO:0035567)	14	4.74	2.18E-02
embryonic camera-type eye development (GO:0031076)	15	4.66	1.18E-02
hindbrain morphogenesis (GO:0021575)	17	4.65	2.48E-03
cerebellum morphogenesis (GO:0021587)	15	4.54	1.62E-02
embryonic eye morphogenesis (GO:0048048)	14	4.47	4.13E-02
hindlimb morphogenesis (GO:0035137)	15	4.42	2.21E-02
cell fate determination (GO:0001709)	17	4.34	6.29E-03
cell differentiation in spinal cord (GO:0021515)	21	4.24	5.11E-04
regulation of catecholamine secretion (GO:0050433)	15	4.21	3.98E-02
forebrain neuron differentiation (GO:0021879)	20	4.11	1.68E-03
neural retina development (GO:0003407)	17	4.07	1.48E-02

Genes shaded blue are those shared by A1 vs. NR hypermethylated and C1 vs. NR hypermethylated, shaded red are shared by C1 vs. NR hypermethylated and C1 vs. A1 hypermethylated, shaded yellow are unique to each group, neuron fate commitment in red text is shared by 5 groups

Table 4.17 Top 20 hypomethylated pathways in each group

GO biological process complete	Genes in pathway (1189)	Fold Enrichment	P-value
A1 vs. NR hypomethylated			
telencephalon regionalization (GO:0021978)	10	> 5	4.85E-02
dorsal spinal cord development (GO:0021516)	13	> 5	1.65E-02
spinal cord association neuron differentiation (GO:0021527)	11	> 5	1.32E-02
forebrain regionalization (GO:0021871)	16	> 5	6.80E-04
regulation of mesonephros development (GO:0061217)	14	4.6	3.02E-02
regulation of cardiac muscle cell proliferation (GO:0060043)	16	4.58	6.74E-03
neuron fate specification (GO:0048665)	16	4.58	6.74E-03
regulation of organ formation (GO:0003156)	16	4.44	1.01E-02
neuron fate commitment (GO:0048663)	35	4.38	1.13E-08
autonomic nervous system development (GO:0048483)	21	4.24	5.20E-04
cell differentiation in spinal cord (GO:0021515)	27	4.2	8.82E-06
embryonic eye morphogenesis (GO:0048048)	17	4.19	1.02E-02
non-canonical Wnt signaling pathway (GO:0035567)	16	4.18	2.16E-02
cell fate specification (GO:0001708)	33	4.18	1.52E-07
olfactory lobe development (GO:0021988)	16	3.94	4.38E-02
regulation of heart growth (GO:0060420)	19	3.92	6.64E-03
inner ear morphogenesis (GO:0042472)	45	3.88	5.06E-10
ear morphogenesis (GO:0042471)	52	3.81	1.13E-11
ventral spinal cord development (GO:0021517)	21	3.8	2.94E-03
forelimb morphogenesis (GO:0035136)	17	3.77	4.05E-02
spinal cord development (GO:0021510)	44	3.76	2.75E-09

Table 4.17 Cont.			
C1 vs. NR hypomethylated			
regulation of cardiac muscle cell differentiation (GO:2000725)	13	> 5	2.64E-02
mesenchymal to epithelial transition (GO:0060231)	12	> 5	2.49E-02
regulation of transcription from RNA polymerase II promoter involved in heart development (GO:1901213)	11	> 5	2.06E-02
ventral spinal cord interneuron differentiation (GO:0021514)	13	> 5	7.99E-03
cardiac epithelial to mesenchymal transition (GO:0060317)	15	> 5	5.04E-03
neuron fate specification (GO:0048665)	21	4.97	3.62E-05
ventral spinal cord development (GO:0021517)	31	4.64	4.94E-08
positive regulation of cardiac muscle tissue development (GO:0055025)	17	4.62	2.78E-03
neuron fate commitment (GO:0048663)	43	4.44	2.05E-11
positive regulation of heart growth (GO:0060421)	15	4.4	2.35E-02
cell differentiation in spinal cord (GO:0021515)	34	4.38	2.35E-08
spinal cord motor neuron differentiation (GO:0021522)	21	4.28	4.42E-04
regulation of cardiac muscle cell proliferation (GO:0060043)	18	4.26	4.04E-03
embryonic hindlimb morphogenesis (GO:0035116)	18	4.26	4.04E-03
forebrain regionalization (GO:0021871)	15	4.23	3.74E-02
autonomic nervous system development (GO:0048483)	25	4.17	4.24E-05
embryonic eye morphogenesis (GO:0048048)	20	4.08	1.93E-03
regulation of cardiac muscle tissue development (GO:0055024)	27	4.04	2.01E-05
hindlimb morphogenesis (GO:0035137)	21	3.95	1.62E-03
regulation of heart growth (GO:0060420)	23	3.93	4.92E-04
proximal/distal pattern formation (GO:0009954)	18	3.89	1.46E-02

Table 4.17 Cont.			
C1 vs. A1 hypomethylated			
positive regulation of cardiac muscle cell differentiation (GO:2000727)	8	> 5	1.62E-02
regulation of cardiac muscle cell differentiation (GO:2000725)	10	> 5	1.58E-02
positive regulation of cardiac muscle tissue development (GO:0055025)	12	> 5	8.90E-03
cell fate determination (GO:0001709)	15	4.56	1.50E-02
regulation of cardiac muscle tissue development (GO:0055024)	16	4.47	9.02E-03
positive regulation of muscle organ development (GO:0048636)	16	4.3	1.50E-02
positive regulation of striated muscle tissue development (GO:0045844)	16	4.3	1.50E-02
positive regulation of muscle tissue development (GO:1901863)	16	4.21	1.91E-02
neuron fate commitment (GO:0048663)	21	4.05	1.05E-03
regulation of morphogenesis of a branching structure (GO:0060688)	16	3.98	3.83E-02
regulation of BMP signaling pathway (GO:0030510)	21	3.59	6.97E-03
positive regulation of ossification (GO:0045778)	21	3.55	8.46E-03
nerve development (GO:0021675)	19	3.52	3.09E-02
endoderm development (GO:0007492)	19	3.47	3.72E-02
inner ear morphogenesis (GO:0042472)	25	3.32	2.79E-03
cell fate commitment (GO:0045165)	58	3.3	1.08E-10
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)	37	3.27	7.69E-06
embryonic skeletal system morphogenesis (GO:0048704)	22	3.2	2.42E-02
ear morphogenesis (GO:0042471)	28	3.17	1.47E-03
regulation of striated muscle tissue development (GO:0016202)	24	3.16	1.12E-02
kidney epithelium development (GO:0072073)	30	3.16	5.81E-04

Genes shaded orange are shared by A1 vs. NR hypomethylated and C1 vs. NR hypomethylated, green are shared by C1 vs. NR hypomethylated and C1 vs. A1 hypomethylated, purple are shared by A1 vs. NR hypomethylated and C1 vs. A1 hypomethylated, yellow are those unique to each group, neuron fate commitment in red text is shared by 5 groups

Table 4.18 Homophilic cell adhesion via plasma membrane pathway genes		
A1 vs. NR Hypermethylated	**C1 vs. NR Hypermethylated**	**C1 vs. A1 Hypomethylated**
CDH12	CADM3	AMIGO2
CDH13	CDH11	CDH10
CDH16	CDH12	CDH18
CDH18	CDH16	CELSR3
CDH22	CDH22	CLSTN2
CELSR3	CDH23	FAT1
CLSTN2	CDH7	FREM2
FREM2	CLSTN2	ME3
PCDH9	DCHS2	NPTN
PCDHA1	DSCAML1	PCDH10
PCDHA10	FREM2	PCDH17
PCDHA11	ME3	PCDHA1
PCDHA12	PCDH9	PCDHA10
PCDHA13	PCDHA1	PCDHA11
PCDHA2	PCDHA10	PCDHA12
PCDHA3	PCDHA11	PCDHA13
PCDHA4	PCDHA12	PCDHA2
PCDHA5	PCDHA13	PCDHA3
PCDHA6	PCDHA2	PCDHA4
PCDHA7	PCDHA3	PCDHA5
PCDHA8	PCDHA4	PCDHA6
PCDHA9	PCDHA5	PCDHA7
PCDHAC1	PCDHA6	PCDHA8
PCDHAC2	PCDHA7	PCDHA9
PCDHB9	PCDHA8	PCDHAC1
PCDHGA1	PCDHA9	PCDHAC2
PCDHGA10	PCDHAC1	PCDHB14
PCDHGA2	PCDHB18	PCDHB17
PCDHGA3	PCDHGA1	PCDHB4
PCDHGA4	PCDHGA10	PCDHB5
PCDHGA5	PCDHGA11	PCDHB7
PCDHGA6	PCDHGA2	PCDHB9
PCDHGA7	PCDHGA3	PCDHGA1
PCDHGA8	PCDHGA4	PCDHGA2
PCDHGA9	PCDHGA5	PCDHGA3
PCDHGB1	PCDHGA6	PCDHGB1
PCDHGB2	PCDHGA7	PVRL4
PCDHGB3	PCDHGA8	
PCDHGB4	PCDHGA9	

PCDHGB5	PCDHGB1	
PCDHGB6	PCDHGB2	
	PCDHGB3	
	PCDHGB4	
	PCDHGB5	
	PCDHGB6	
	PCDHGB7	
	PTPRM	
	PTPRT	
	PVR	
	RET	

Genes shaded purple are those shared by all three paired analyses, blue are shared by A1 vs. NR hypermethylated and C1 vs. NR hypermethylated, green are shared by C1 vs. NR hypermethylated and C1 vs. NR hypomethylated, orange are shared by A1 vs. NR hypermethylated and C1 vs. A1 hypomethylated

Table 4.19 Association of survival with genes identified in ER-positive first tumors from pairs with ER-negative second tumors

Genes hypermethylated only in C1 versus NR			
Gene	**Affymetrix probe**	**p-value**	**Expression pattern linked to good survival**
CADM3	221921_s_at	0.0516	high
CDH11	236179_at	0.4105	---
CDH23	232845_at	0.0002	high
CDH7	220679_s_at	2.10E-07	high
DCHS2	220373_at	0.0012	high
DSCAML1	232059_at	0.0005	high
PCDHB18	pseudogene		---
PCDHGA11	211877_s_at	0.0248	high
PCDHGB7	1552661_at	0.0979	---
PTPRM	1555579_s_at	1.10E-07	high
PTPRT	205948_at	4.40E-16	high
PVR	1556582_at	1.70E-07	high
RET	211421_s_at	3.91E-02	high
Genes hypermethylated only in C1 versus A1			
Gene	**Affymetrix probe**	**p-value**	**Expression pattern linked to good survival**
AMIGO2	222108_at	0.696	---
CDH10	220115_s_at	1.90E-06	high
FAT1	201579_at	0.7623	---
NPTN	202228_s_at	0.0431	high
PCDH10	1552925_at	0.0134	high
PCDH17	228863_at	0.0036	low
PCDHB14	231726_at	0.0141	high
PCDHB17	216313_at	0.3612	---
PCDHB4	240317_at	0.0002	high
PCDHB5	223629_at	0.3683	---
PCDHB7	231738_at	0.5007	---
PVRL4	223540_at	0.3589	---
Genes hypermethylated only in A1 versus NR			
Gene	Affymetrix probe	p-value	**Expression pattern linked to good survival**
CDH13	204726_at	0.8509	---

Table 4.20 Cell fate commitment genes		
A1 vs. NR Hypermethylated	**C1 vs. NR Hypermethylated**	**C1 vs. A1 Hypomethylated**
BCL11B	ASCL1	ATOH1
BMP4	BARHL2	BARHL2
C17orf96	BCL11B	BMP4
DLL1	CDON	CDC73
FEV	CHD5	CHD5
FEZF2	CYP26B1	DLL1
FGF3	DLX2	DLX1
FOXA2	DSCAML1	DMRTA2
FOXN4	FEV	DOCK7
GAP43	FEZF2	EYA1
GATA4	FGF10	FGFR1
GFI1	FGF3	GAP43
GSC	FGF8	GATA2
HES5	FGFR1	GFI1
HEY2	FGFR2	HES5
HNF1B	FGFR3	HEY2
HOXA13	FOXA2	HIPK2
IFRD1	GATA4	HOXA13
NKX2-5	GDF7	HOXC10
NODAL	GDNF	HOXC11
NOTCH3	GFI1	ISL1
NOTCH4	GLI2	ISL2
NTRK3	GSC	LHX3
OLIG3	GSX2	MEF2C
PAX2	HES5	NKX2-1
PAX3	HNF1B	NKX2-2
PAX6	IHH	NKX6-2
PRDM1	ISL1	NR2E1
ROR2	ISL2	NRG1
RORA	MNX1	NTF3
SATB2	MYOD1	ONECUT1
SIX1	NEUROG1	ONECUT2
SIX3	NKX2-1	OTX2
SMO	NKX2-5	PAX2
SOX8	NKX6-1	POU4F1
SOX9	NOTCH4	POU6F2
TBR1	NR2E1	PRDM1
TFAP2C	NTF3	RAG2
TLX3	NTF4	RORA

TRIM15	NTRK3	SATB2
VSX2	OLIG2	SFRP1
WNT1	OLIG3	SMAD2
WNT10A	ONECUT1	SOX17
WNT7A	PAX2	SOX18
	PAX3	SOX2
	PAX6	SOX5
	PAX7	SOX8
	POU4F1	TBR1
	POU6F2	TBX3
	PRDM1	TGFB1I1
	PRDM14	TGFB2
	PTCH1	TLX3
	PTF1A	TRIM15
	SATB2	WNT2
	SFRP1	WNT3A
	SIX1	WNT5A
	SIX3	WNT7B
	SOX1	WT1
	SOX12	
	SOX8	
	SOX9	
	TBX1	
	TLX3	
	VSX2	
	WNT1	
	WNT10A	
	WNT3	
	WNT3A	
	WNT6	
	WNT7A	
	WT1	

Genes shaded purple are those shared by all three pathways, blue are shared by A1 vs. NR hypermethylated and C1 vs. NR hypermethylated, green are shared by C1 vs. NR hypermethylated and C1 vs. NR hypomethylated, orange are shared by A1 vs. NR hypermethylated and C1 vs. A1 hypomethylated

Table 4.21 Association of survival with genes identified in ER-positive first tumors with either ER-negative or ER-positive second tumors

Gene	Affymetrix probe	p-value	Expression pattern linked to good survival	Shared CpG sites (MAPINFO)
BCL11B	219528_s_at	2.330E-02	high	99708008
FEV	207260_at	1.100E-05	high	219847233
FEZF2	221086_s_at	6.400E-05	high	62359390
FGF3	214571_at	3.260E-02	high	69632883
FOXA2	40284_at	1.600E-04	high	22566143
GATA4	205517_at	9.878E-01	---	---
GSC	1552338_at	6.133E-01	---	---
HNF1B	205313_at	1.000E-10	high	36105064
NKX2-5	206578_at	7.246E-01	---	172660996 172662130
NOTCH4	205247_at	3.800E-08	high	32164503
NTRK3	206462_s_at	7.900E-12	high	88798448
OLIG3	1556371_at	7.993E-01	---	---
PAX3	231666_at	4.800E-03	high	223065390 223158076 223158408
PAX6	235795_at	6.698E-01	---	31824973 31828715
SIX1	228347_at	7.300E-04	high	61116382
SIX3	242054_s_at	5.500E-01	---	45171583
SOX9	202936_s_at	3.300E-01	---	70119120
VSX2	missing under 3 possible names			74704714 74725362
WNT1	208570_at	1.000E-04	high	49371987
WNT10A	223709_s_at	5.400E-08	high	---
WNT7A	210248_at	4.500E-04	high	---

Table 4.22 Genes shared in both cell fate commitment and neuron fate commitment pathways

Genes	Genes
ASCL1	NKX2-1
ATOH1	NKX2-2
BCL11B	NKX6-1
BMP4	NKX6-2
C17orf96	NOTCH3
DLX1	NRG1
DLX2	NTRK3
DMRTA2	OLIG2
EYA1	OLIG3
FEV	OTX2
FEZF2	PAX3
FOXA2	PAX6
FOXN4	PAX7
GATA2	POU4F1
GLI2	PTF1A
GSX2	**SATB2**
HES5	SIX1
HOXC10	SOX1
IHH	TBR1
ISL1	TFAP2C
ISL2	TGFB2
LHX3	TLX3
MNX1	WNT1

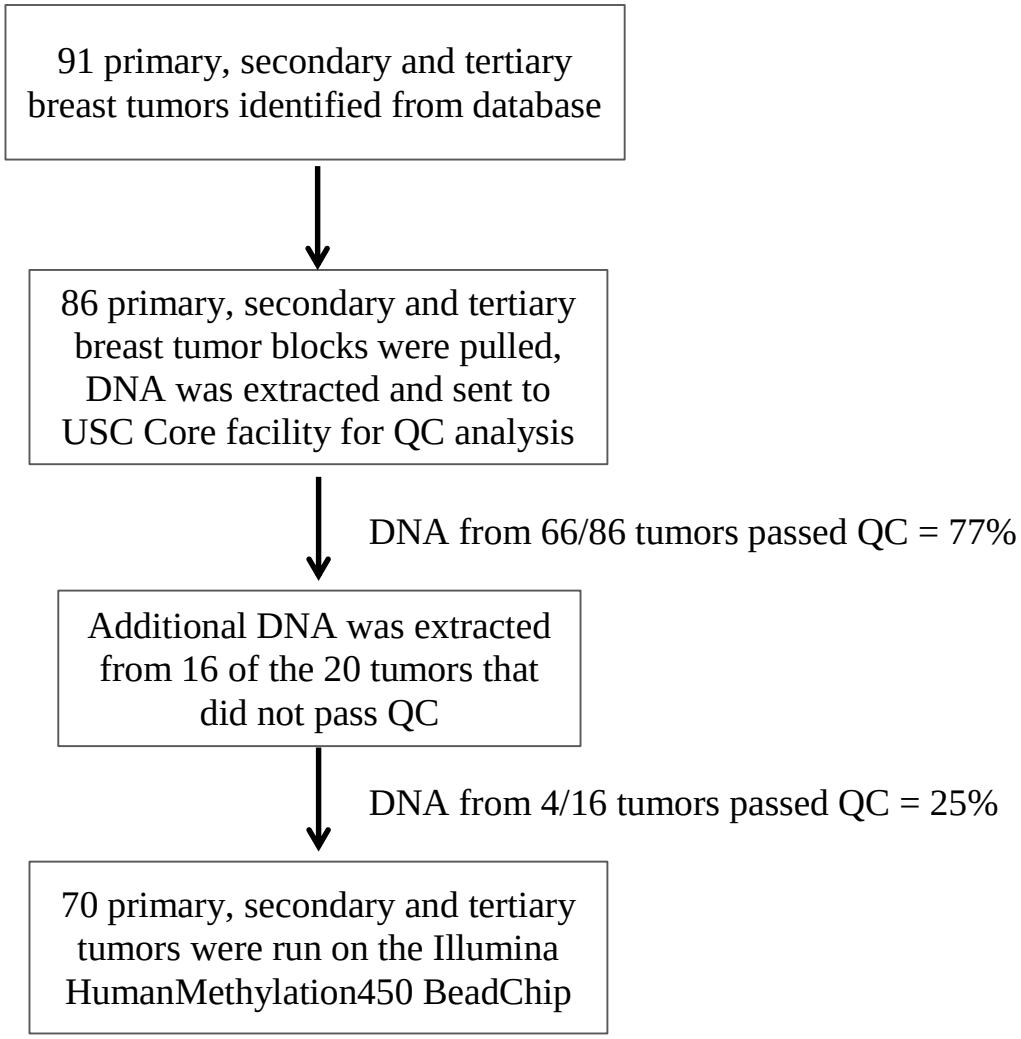


Figure 4.1. Flow chart explaining tumor selection and criteria for analysis by HM450BC.

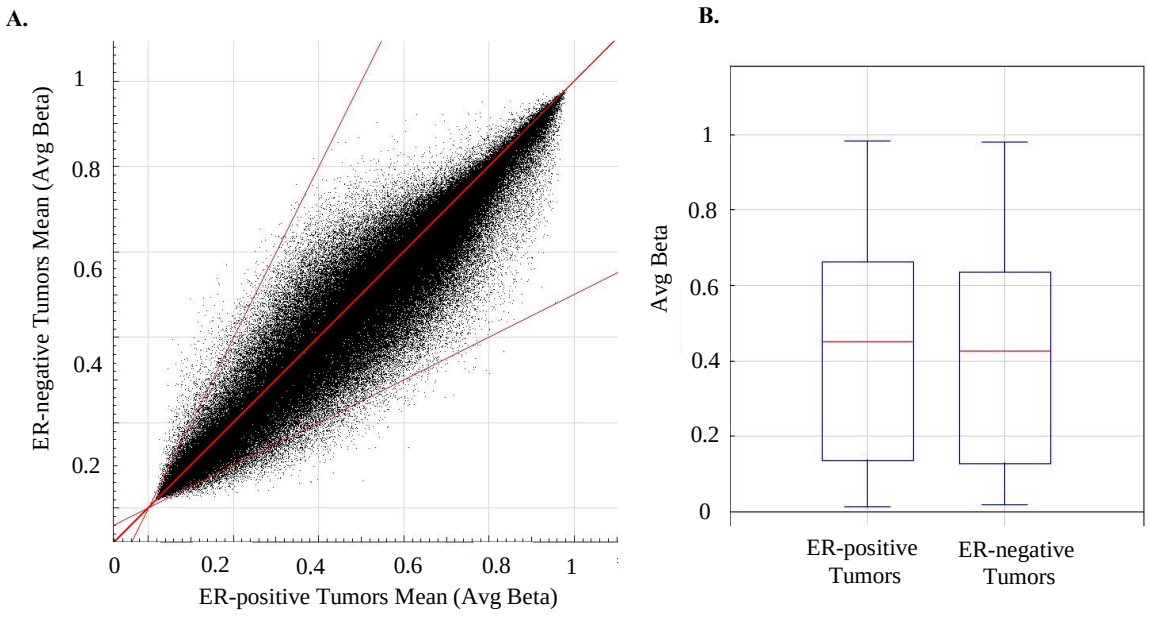


Figure 4.2. Methylation of ER-positive breast tumors compared with ER-negative breast tumors.

To visualize data, average beta values of 41 ER-positive tumors were compared with average beta values of 21 ER-negative breast tumors using GenomeStudio. A) Slightly greater mean methylation in ER-positive tumors was observed: 3,046 CpGs were hypermethylated in ER-positive tumors and 1,098 CpGs were hypermethylated sites in ER-negative tumors. Outer red lines mark the 2-fold change in methylation; center red line represents equal average beta values in the two groups. B) Average beta distribution analysis of 425,489 CpG sites by box plot. Box is 25th and 75th percentiles; red line is median average beta value.

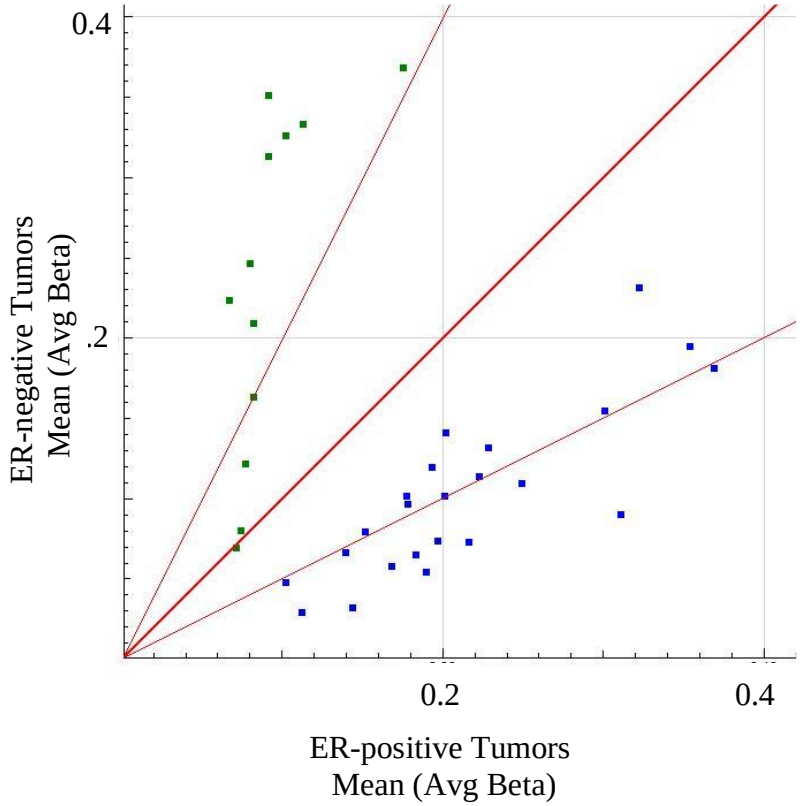


Figure 4.3. Analysis of 35 CpG sites previously shown to be differentially methylated between ER-positive and ER-negative breast tumors.

Twenty-three CpG sites previously identified as hypermethylated in ER-positive tumors (blue squares), and 12 CpG sites previously identified as hypermethylated in ER-negative tumors (green squares) were assessed in 41 ER-positive tumors and 21 ER-negative tumors. Outer red lines represent 2-fold difference in methylation; center red line represents equal beta values between groups.

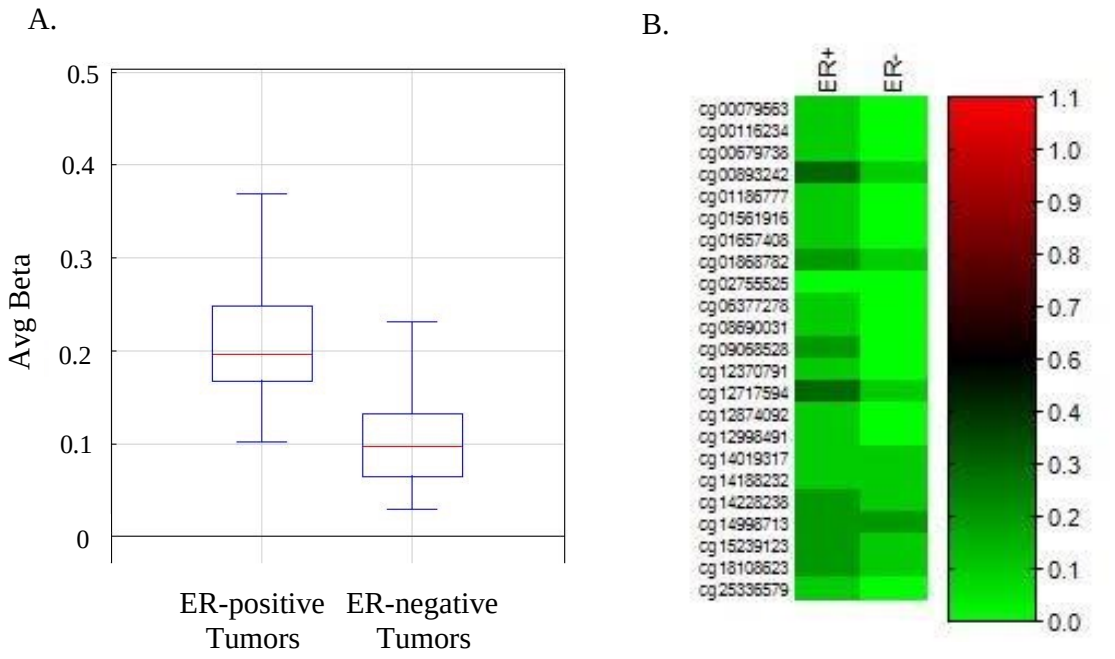


Figure 4.4. Methylation analysis of 23 CpG sites previously identified as hypermethylated in ER-positive tumors.

Average beta distribution analysis of all 23 CpG sites by box plot. Box is 25th and 75th percentiles; red line is median average beta value. B) Heat map shows average beta values of the 23 individual CpG sites. Only one CpG site, cg02755525 does not differ in methylation between ER-positive and ER-negative tumors. Green represents 0% methylated; red represents 100% methylated.

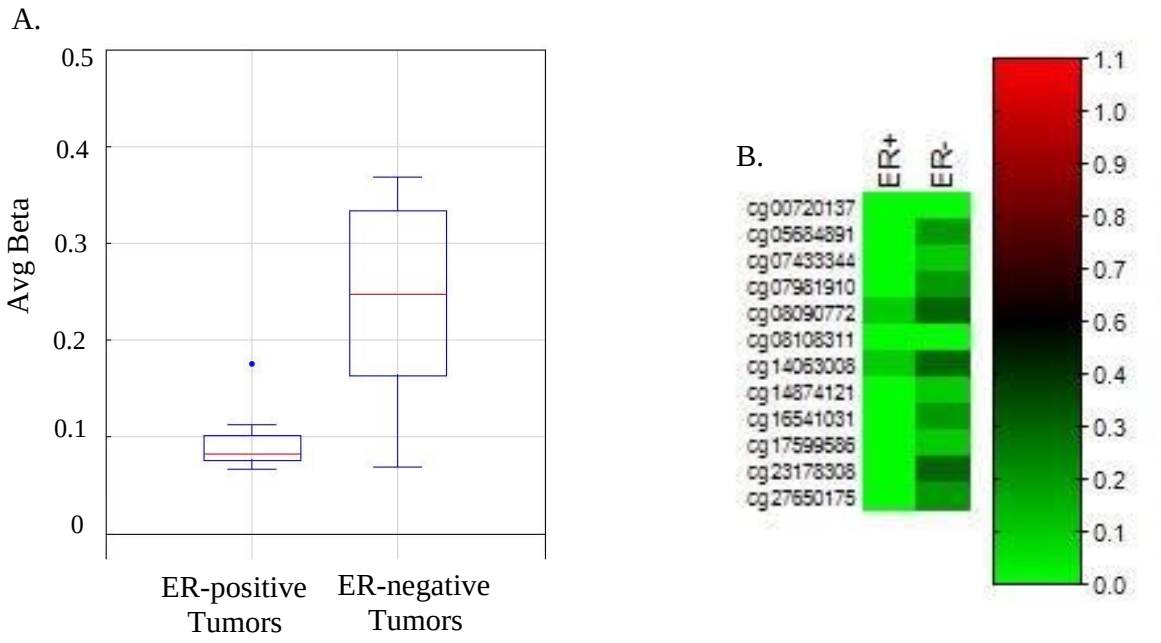


Figure 4.5. Methylation analysis of 12 CpG sites previously identified as hypermethylated in ER-negative tumors.

A) Average beta distribution analysis of all 12 CpG sites by box plot. Box is 25th and 75th percentiles; red line is median average beta value. B) Heat map shows average beta values of the 12 individual CpG sites. Two CpG sites, cg00720137 and cg0610631, had no change in methylation between ER-positive and ER-negative tumors. Green represents 0% methylated; red represents 100% methylated

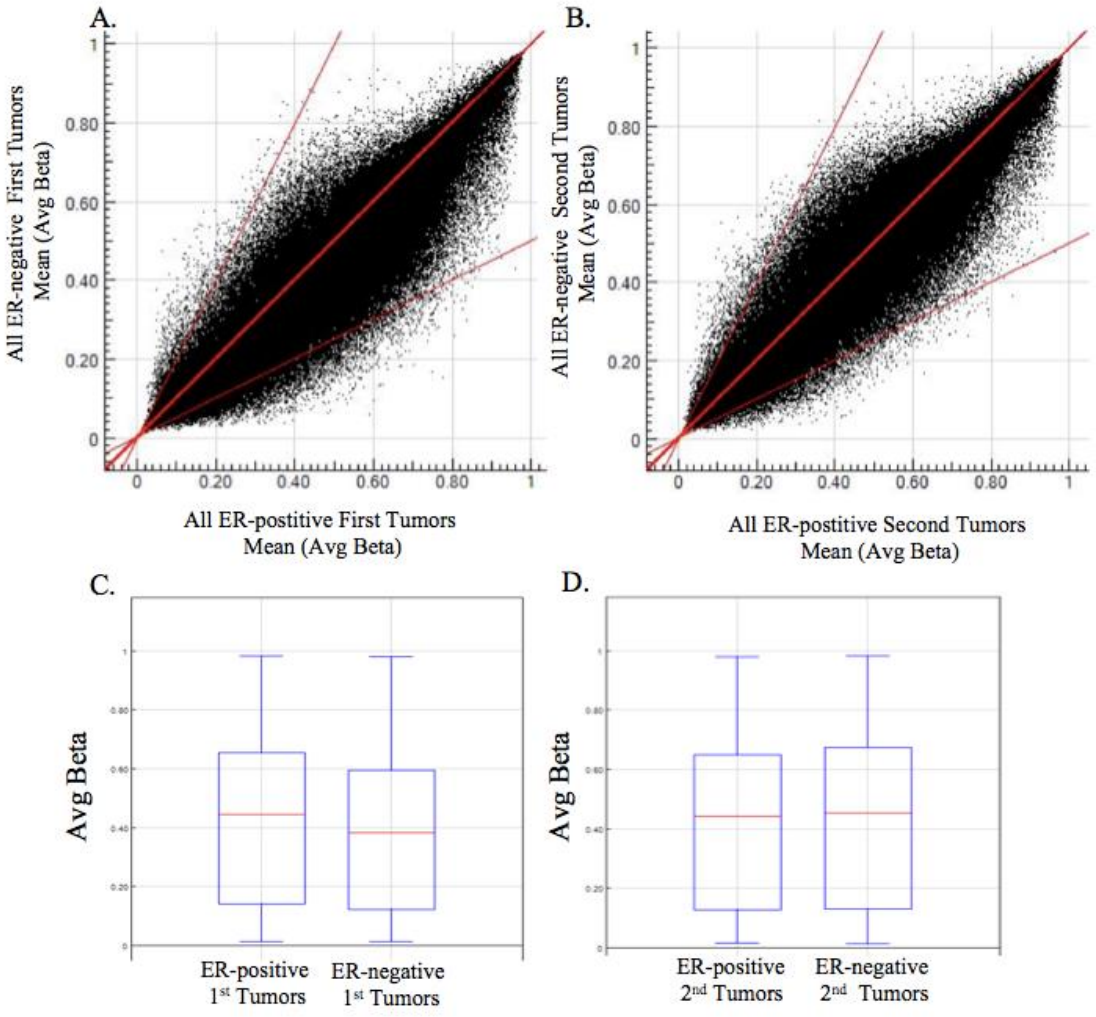


Figure 4.6. Methylation analysis of tumors stratified by both ER-status and occurrence.

A) 20 ER-positive first tumors had 10,314 hypermethylated CpG sites when compared to 6 ER-negative first tumors, which had 1,235 hypermethylated CpG sites. B) 14 ER-positive second tumors had 3,110 hypermethylated CpG sites when compared to 14 ER-negative second tumors, which had 2,198 hypermethylated CpG sites. Outer red lines represent a 2-fold change in methylation between the two groups and center red line represents equal methylation in both cases. Distribution analysis of C) ER-positive first tumors compared with ER-negative first tumors and D) ER-positive second tumors compared with ER-negative second tumors. Box is 25th and 75th percentiles; red line is median average beta value.

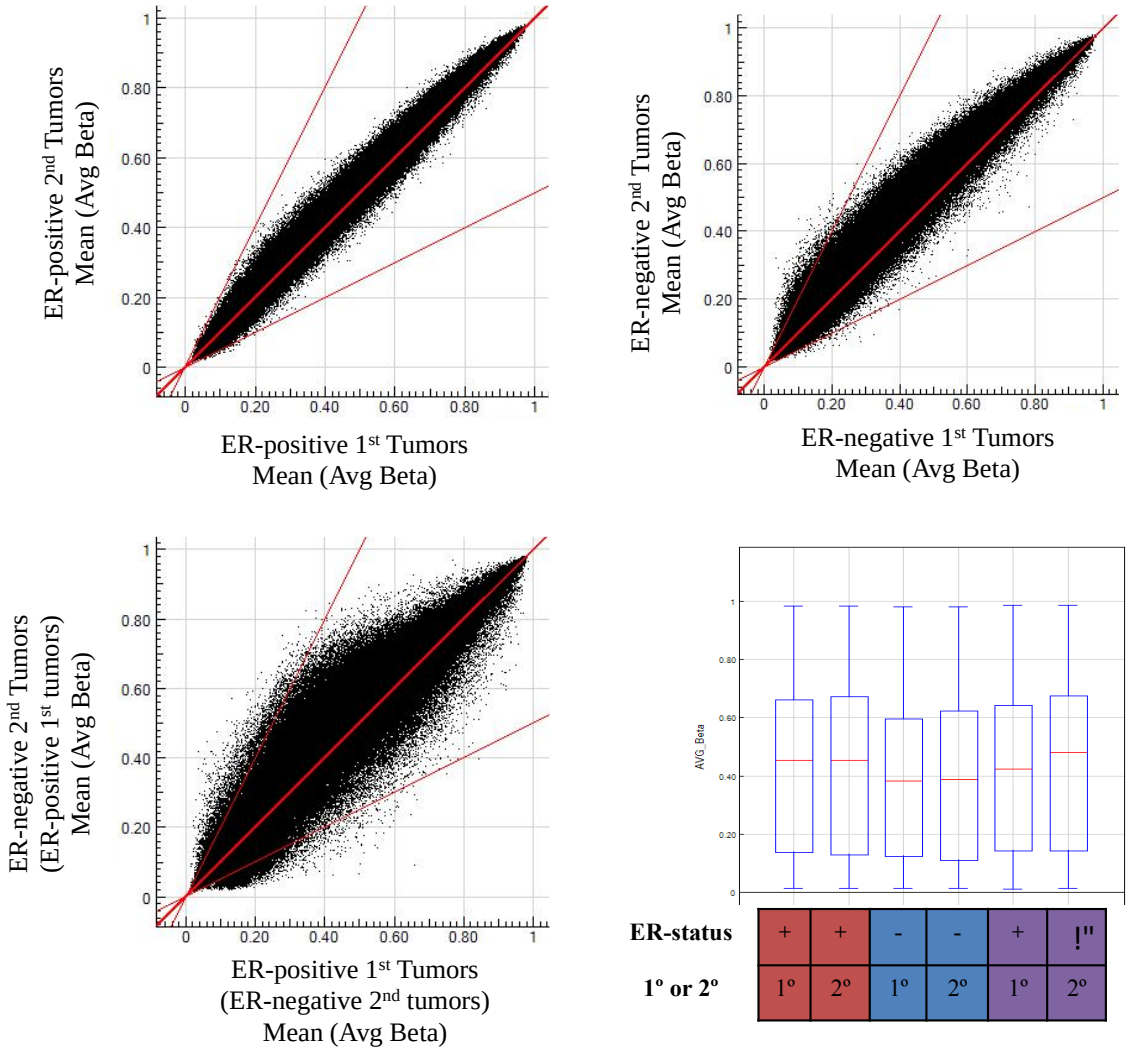


Figure 4.7. Visual analysis of paired tumors stratified by ER-status of the primary and second.

Methylation of A) 20 ER-positive first tumors were compared with 20 ER-positive second tumors from the same women, B) 5 ER-negative first tumors were compared with 5 ER-negative second tumors from the same women, and C) 6 ER-positive first tumors were compared with 6 ER-negative second tumors from the same women. Outer red lines represent 2-fold change in methylation between first and second tumors; center red line is equal methylation between samples. D) Distribution analysis of the three tumor groups in A, B and C above. Groups are indicated in colors: red is ER-positive first tumors with ER-positive second tumors; blue is ER-negative first tumors with ER-negative second tumors; and purple ER-positive first tumors with ER-negative second tumors. Box is 25th and 75th percentiles; red line is median average beta value.

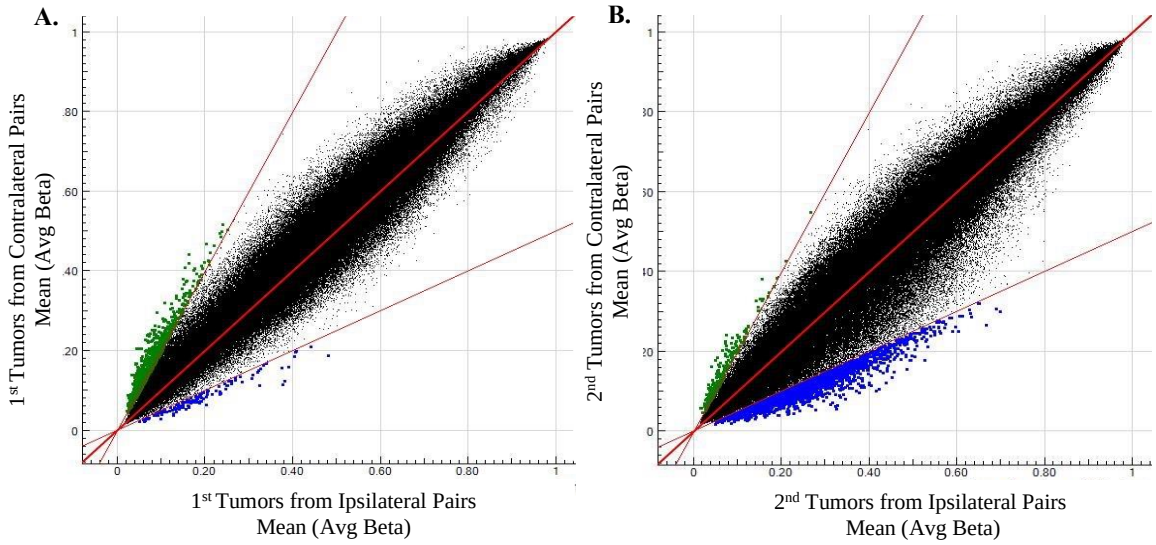


Figure 4.8. Comparison of first and second tumors stratified by side of second tumor occurrence.

A) First tumors from ipsilateral (n=15) and contralateral (n=10) pairs are compared. B) Second tumors from ipsilateral (n=15) and contralateral (n=10) pairs are compared. Blue dots are CpG sites hypermethylated in tumors (both first and second) from ipsilateral pairs; green dots are CpG sites hypermethylated in tumors (both first and second) from contralateral pairs. Outer red lines represent 2-fold change in methylation; center red line is no change in methylation between tumors.

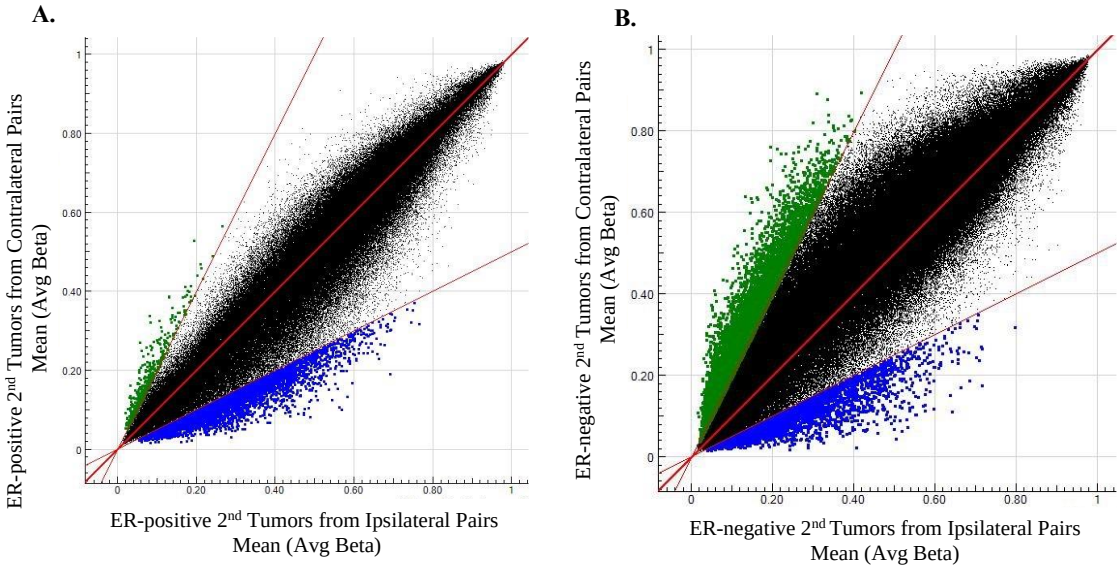


Figure 4.9. Analysis of second tumors stratified by ER-status and side of second tumor occurrence.

A) ER-positive second tumors from ipsilateral pairs (n=8) were compared to ER-positive second tumors from contralateral pairs (n=6). B) ER-negative second tumors from ipsilateral pairs (n=7) were compared to ER-negative second tumors from contralateral pairs (n=4). Blue dots represent CpG sites with a 2-fold change in methylation in second tumors (either ER-positive or ER-negative) from the ipsilateral pairs. Green dots represent CpG sites with a 2-fold change in methylation in second tumors (either ER-positive or ER-negative) from the contralateral pairs. Center red line is equal methylation in both group.

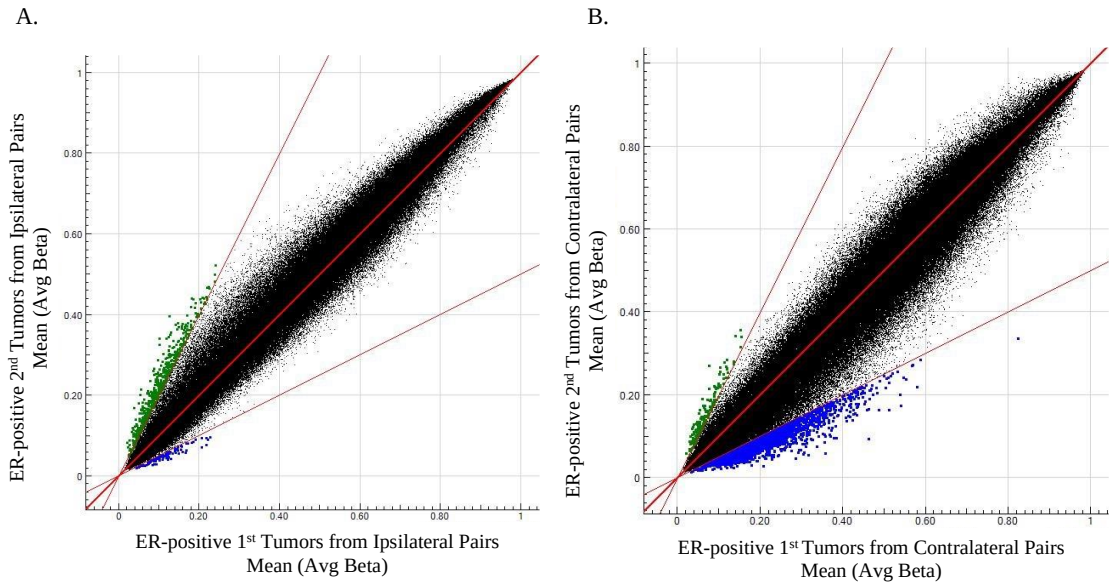


Figure 4.10. Analysis of paired tumors from Group A (ER-positive primary and ER-positive second) stratified by side of location.

8 ER-positive second tumors from ipsilateral pairs have more hypermethylated CpG sites than the first tumor (green dots). B) 6 ER-positive first tumors from contralateral pairs have more hypermethylated CpG sites than the second tumor (blue dots). Outer red lines are 2-fold change in methylation between groups; center red line equals no change in methylation between groups.

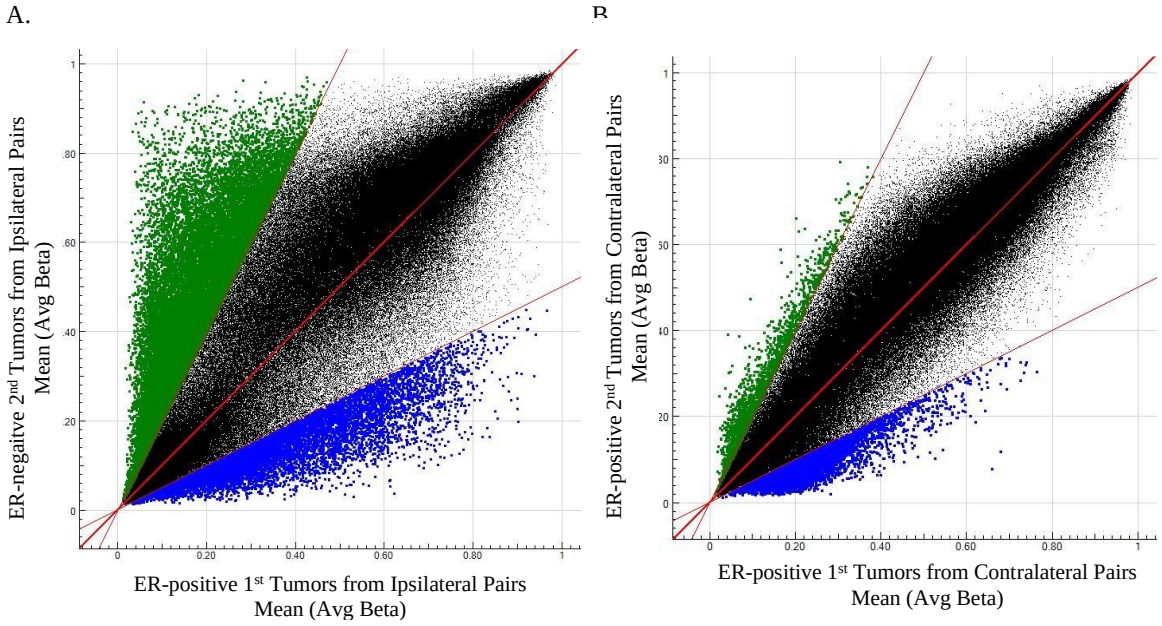


Figure 4.11. Analysis of paired tumors from Group C (ER-positive primary and ER-negative second) stratified by side of location.

2 ER-negative second tumors from ipsilateral pairs have more hypermethylated CpG sites than the first ER-positive tumor (green dots). B) 4 ER-positive first tumors from contralateral pairs have more hypermethylated CpG sites than the second ER-negative tumor (blue dots). Outer red lines are 2-fold change in methylation between groups; center red line equals no change in methylation between groups.

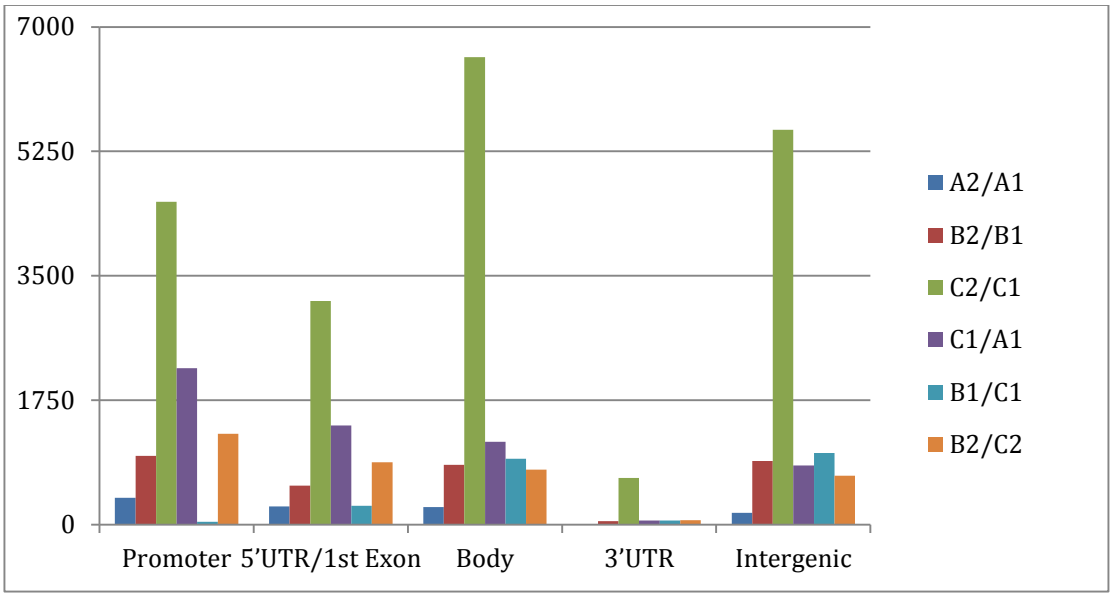


Figure 4.12 CpG site hypermethylation within gene regions.

The functional genomic location of hypermethylated CpG sites on the BeadChip was analyzed for A2 versus A1, B2 versus B1, C2 versus C1, C1 versus A1, B1 versus C1, and B2 versus C2. Promoter includes TSS200 and TSS1500 regions of the gene; Intergenic region is all undefined locations in GenomeStudio.

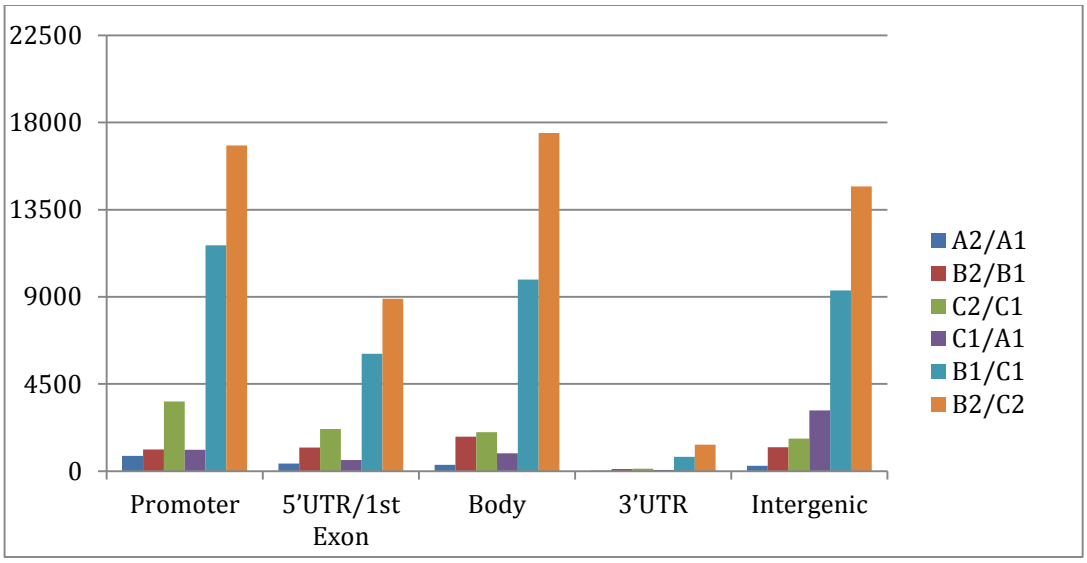
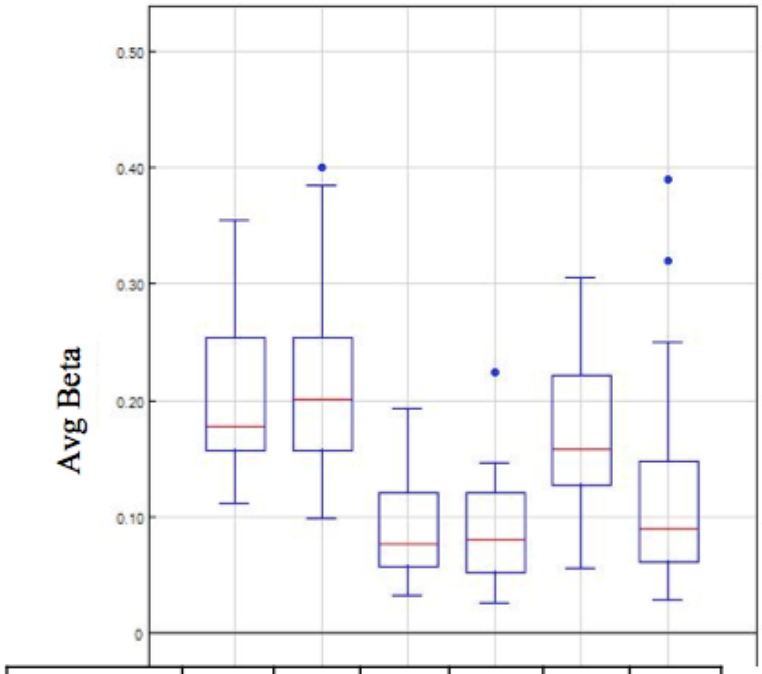


Figure 4.13 CpG site hypomethylation within gene regions.

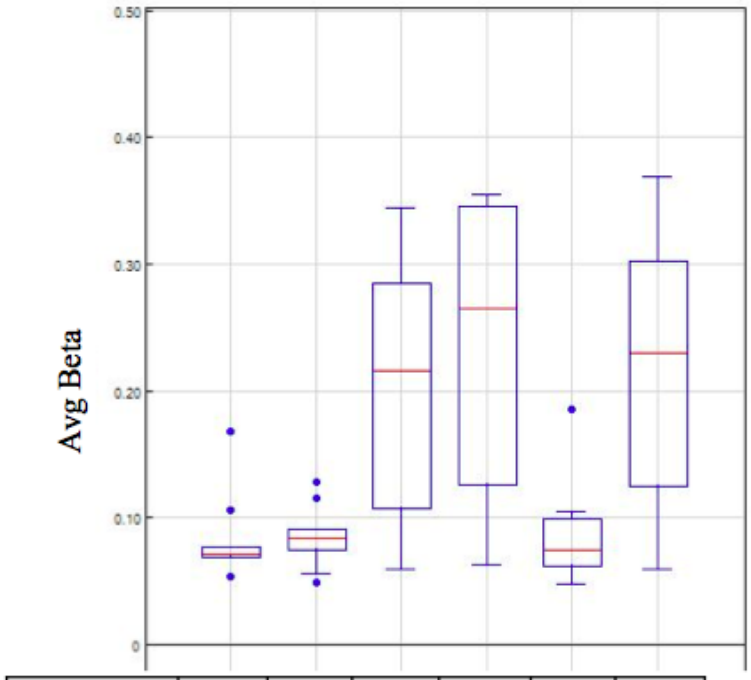
The functional genomic location of hypomethylated CpG sites on the BeadChip was analyzed for A2 versus A1, B2 versus B1, C2 versus C1, C1 versus A1, B1 versus C1, and B2 versus C2. Promoter includes TSS200 and TSS1500 regions of the gene; Intergenic region is all undefined locations in GenomeStudio.



ER-status	+	+	-	-	+	-
1° or 2°	1°	2°	1°	2°	1°	2°
Group	A1	A2	B1	B2	C1	C2

Figure 4.14. Group methylation analysis of twenty-three CpG sites previously identified as hypermethylated in ER-positive breast tumors.

Distribution analysis of 23 previously identified CpG sites that were hypermethylated in ER-positive tumors was assessed in ER-positive first tumors (A1) with ER-positive second tumors (A2), ER-negative first tumors (B1) with ER-negative second tumors (B2) and ER-positive first tumors (C1) with ER-negative second tumors (C2). Box is 25th and 75th percentiles; red line is median average beta value.



ER-status	+	+	-	-	+	-
1º or 2º	1º	2º	1º	2º	1º	2º
Group	A1	A2	B1	B2	C1	C2

Figure 4.15. Group methylation analysis of twelve CpG sites previously identified as hypermethylated in ER-negative breast tumors.

Distribution analysis of 12 previously identified CpG sites that were hypermethylated in ER-negative tumors was assessed in ER-positive first tumors (A1) with ER-positive second tumors (A2), ER-negative first tumors (B1) with ER-negative second tumors (B2) and ER-positive first tumors (C1) with ER-negative second tumors (C2). Box is 25th and 75th percentiles; red line is median average beta value.

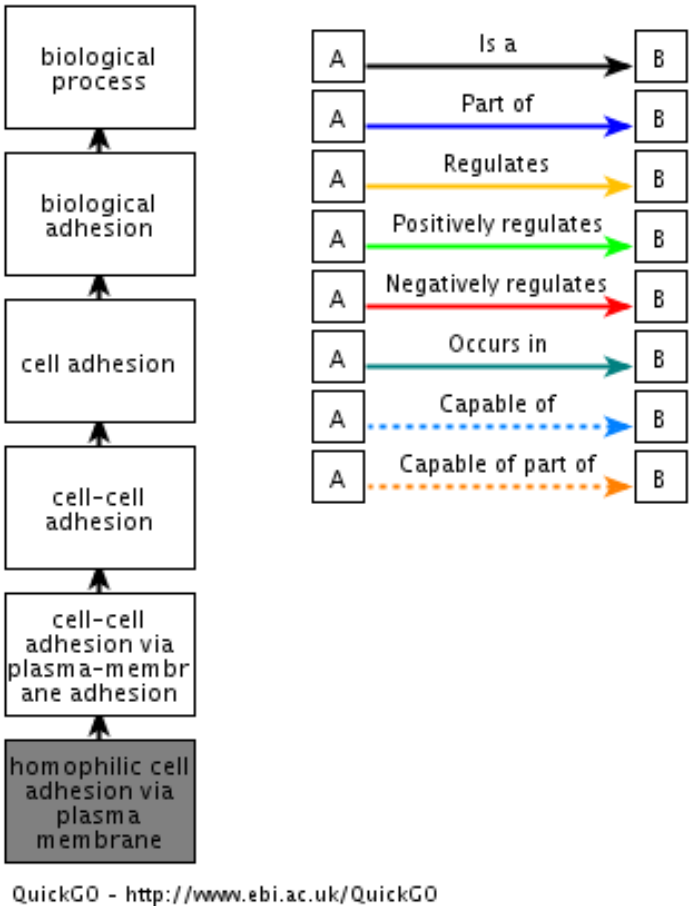


Figure 4.16. Ancestor chart for homophilic cell adhesion via plasma membrane pathway.

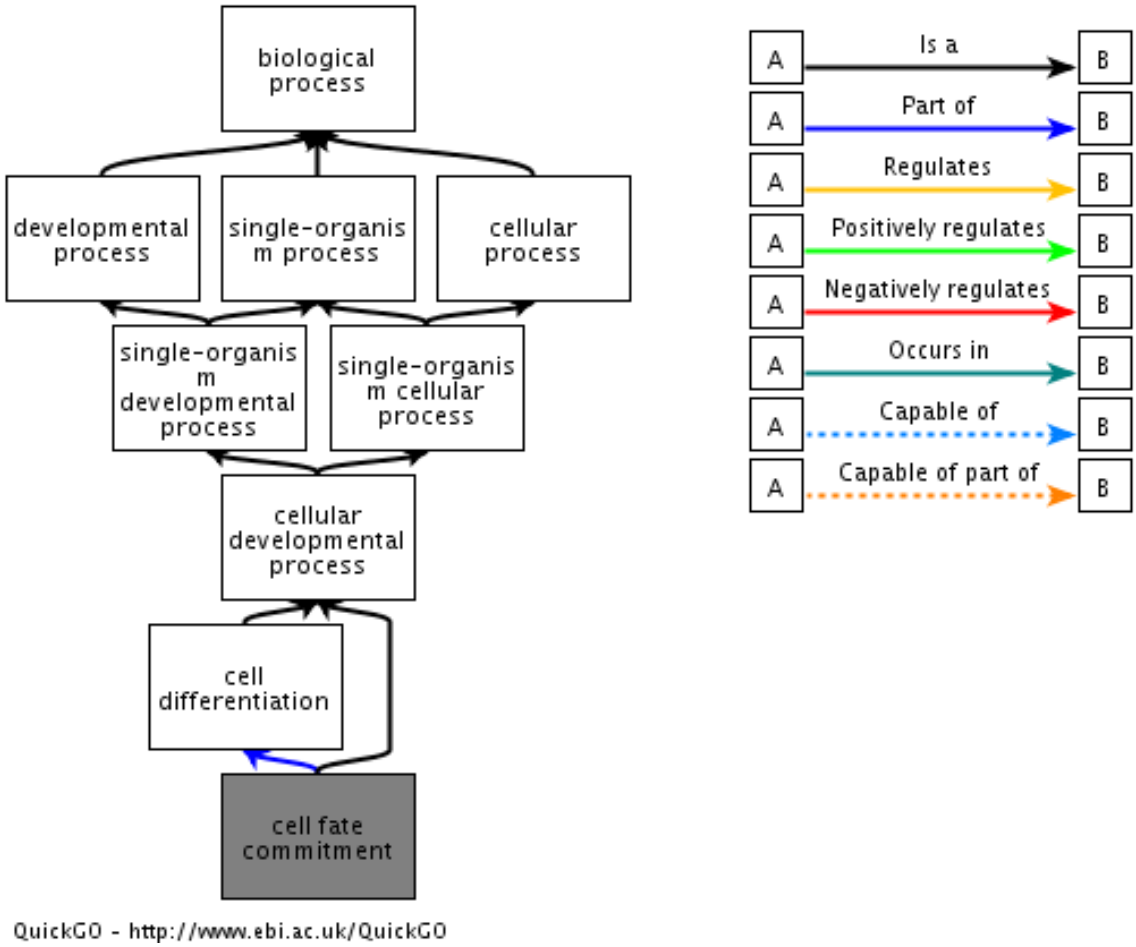


Figure 4.17. Ancestor chart for cell fate commitment pathway.

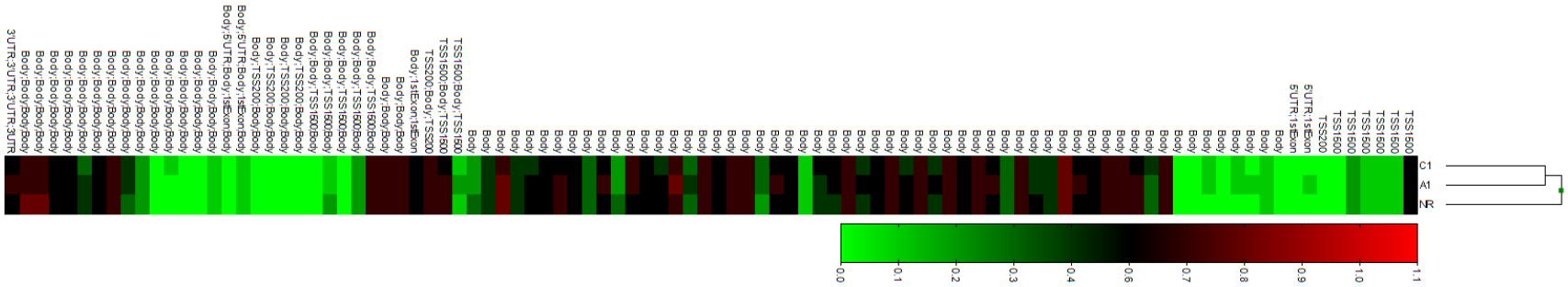


Figure 4.18. Differential RORA methylation is found in the gene body.

Of the 97 RORA CpGs included on the HM450BC, the majority of those differentially methylated were found in the body. In C1, 17 CpG sites were differentially methylated in the body region and in A1, 21 were differentially methylated in the body region.

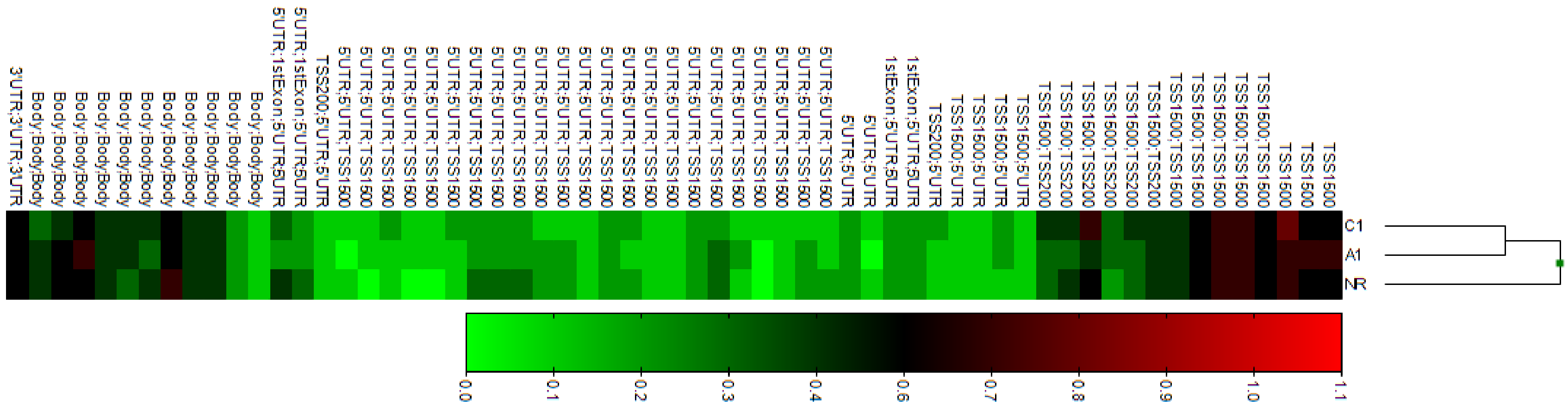


Figure 4.19. Differential methylation of GFI1 occurs most frequently in the promoter region.

A1 is more frequently hypomethylated in the promoter region (TSS200 and TSS1500) as compared to NR with 10 hypomethylated CpG sites out of 44 CpG sites. C1 is more frequently hypermethylated in the promoter as compared to NR with 11 out of 44 CpG sites being hypermethylated.

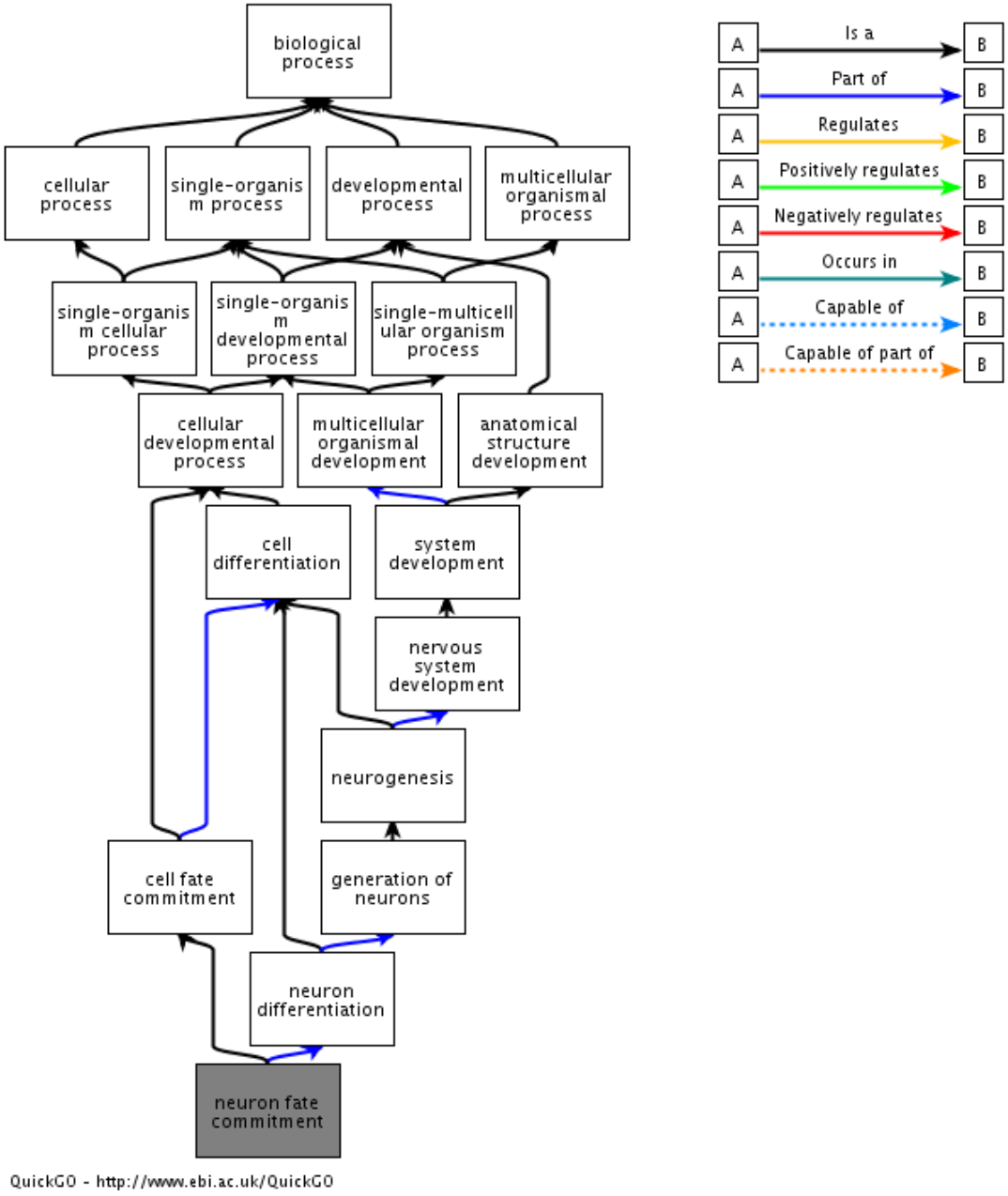


Figure 4.20. Ancestor chart for neuron fate commitment.

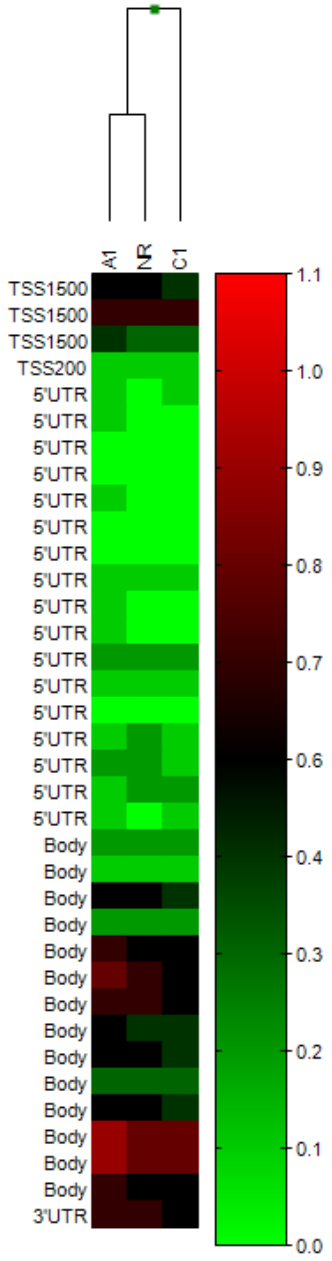


Figure 4.21. SATB2 is differentially methylated between A1, C1 and NR.

Clustering analysis shows that methylation of SATB2 is most similar between A1 and NR. Hypermethylation occurs more frequently in A1 and NR than in C1, where hypomethylation occurs more frequently. A1 and C1 have 14 and 13 dmCpG sites respectively, as compared to NR.

APPENDICES

APPENDIX A

HYPER AND HYPOMETHYLATED PATHWAYS SHARED BY TMX2-11 AND TMX2-28

Hypermethylated*		
Pathway	*Genes*	*p value*
sensory perception of smell	OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, OR4S2, OR52L1, OR8K3, OR5R1, OR5M1, OR5M3, OR4M1, OR5M9, OR10AG1, OR4D6, OR1S2, OR5M8, CNGA2, OR1S1, OR5H14, OR8J1, OR52K1, OR5H15, OR8J3, OR13F1, OR2D3, OR2L8, OR2B3, OR8U8, OR2L13, OR1G1, OR6A2, OR2T33, OR5AS1, OR10C1, OR10W1, OR5P3, OR2AG1, OR7G1, OR5P2, OR5M10, OR52E4, OR6C75, OR5AC2, OR2M4, OR2M5, OR12D2, OR52M1, OR12D3, OR4A15, OR11H6, OR4A16, OR10A6, OR52A1, OR5H2, OR4K17, OR11A1, OR51Q1, OR9A4, OR5W2, OR13A1, OR14J1, OR5AN1, OR5I1, OR52B2, OR52B4, OR2W1, OR2H1, GRM8, OR5V1, GRM7, OR6Q1, OR5D14, OR5D16, OR8H3, OR5D18, OR1L3, OR5A1, OR5A2, OR9Q1, OR9Q2, OR13C3, OR9I1, OR10G9, OR5E1P, OR5B12, OR5AP2, OR13C8, OR5B17, OR9G9, OR1B1, OR5K4, OR9G1, OR9G4, OR5B2, OR5T2, OR5B3, OR5T1, OR4X1, OR13D1, OR10Q1, OR8H2	2.13E-34
cell surface receptor linked signal transduction	ADCY5, GPR123, LPHN2, HTR1B, S1PR1, GRIN2B, OR52L1, OR4S2, GAB1, ROS1, ADAM9, OR5R1, ATRNL1, OR1S2, OR4D6, OR1S1, VEGFC, GRB10, PPBP, OR52K1, F2, HTR6, OR13F1, ROR1, ROR2, ADAMTS1, OR2D3, OR2B3, GNAI1,	4.28E-33

OR8U8, AKAP12, EPHB1, GPR141,
SORCS3, OR1G1, ITGBL1, HCRTR1,
OR6A2, BAI3, FCER1A, GUCY2F,
BMP2, OR5P3, OR2AG1, TCF7, OR7G1,
OR5P2, SMAD5, MAML2, TAX1BP3,
KCNK2, BTLA, OR12D2, OR52M1,
OR12D3, SFRP2, NOTCH4, PTENP1,
FGF5, OR10A6, OR52A1, FGF9,
OR11A1, OR51Q1, SIRPB1, EDNRB,
OR9A4, OR5W2, OR13A1, NMUR2,
SPG21, OR5AN1, OR52B2, OR52B4,
OR2W1, GRM3, CCND1, GRM2,
DACT1, CHRDL1, GRM8, OR2H1,
GRM7, OR5V1, WNT9B, NPPB, IL12B,
RAPGEFL1, CXCL5, APC2, OR5D14,
GPR63, OR5D16, CXCL2, FPR1,
OR5D18, FPR2, OR5A1, OR5A2,
OR9Q1, OR9Q2, OR13C3, DGKB,
PTK2B, OR5E1P, OR5AP2, OR5B12,
NPFFR2, OR13C8, OR5B17, OR1B1,
DGKK, DGKI, GRIA4, PTGFR, OR5T2,
OR5B2, OR5B3, OR5T1, OR13D1,
MC4R, OR10Q1, IFT52, BAMBI,
WNT7A, HTR2A, BMP10, OR5L1,
OR2J2, OR1J4, OR4D10, OR52D1,
MARCO, OR8K1, WNT2, TIAM1,
OR8K3, INSR, IRS2, OR5M1, BAIAP2,
OR5M3, FGF22, DLL1, SOCS5, OR4M1,
OR5M9, FGF20, OR10AG1, OR5M8,
NCAM1, GABRR3, OR5H14, OR8J1,
OR8J3, OR5H15, GPR50, AKAP4,
C3AR1, OR2L8, PKHD1, OR2L13,
ITGB6, OR2T33, OR5AS1, OR10C1,
ENTPD1, DEFB1, OR10W1, GABRE,
OR5M10, GABRA4, GABRA3,
GABRA6, OR52E4, OR6C75, OR2M4,
OR5AC2, OR2M5, OR4A15, OR11H6,
OR4A16, CD274, ADRA1B, ADRA1A,
MERTK, OR5H2, LEPR, PREX2,
OR4K17, TAAR8, TAAR9, IAPP,
RSPO3, HEY2, RSPO2, OR14J1, NRG1,
GABRG1, OR5I1, GABRG3, CCDC88C,
ARID5B, FSHR, ADRB2, CHRM2,
OR6Q1, GAP43, GPRC5C, OR8H3,
APH1B, GPR6, TAC1, OXTR, OR1L3,
PF4, GPRC5B, MSX2, APLNR, HRH1,

	OR10G9, OR9I1, PTN, APC, OR9G9, PTPRD, FLT1, GNAO1, OR5K4, OR9G1, OR9G4, RGS13, OR4X1, OR8H2, BMPR1B, ADAMDEC1	
neurological system process	RP1, SYT5, OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, WNT2, OR8K1, CTTNBP2, HTR1B, BDNF, S1PR1, OR4S2, GRIN2B, OR52L1, OR8K3, CHRNA7, OR5R1, KCND2, OR5M1, TRPA1, SIX3, OR5M3, OR4M1, OR5M9, OR10AG1, OR4D6, OR1S2, OR1S1, CNGA2, CTNNA2, OR5M8, GABRR3, LRAT, CAMK4, OR5H14, OR52K1, OR8J1, OR8J3, OR5H15, HTR6, OR13F1, OR2D3, OR2L8, OR2B3, OR8U8, ASZ1, RIMS1, OR2L13, OR1G1, KCNMB2, HCRTR1, CRB1, OR6A2, OR5AS1, OR2T33, OR10C1, OR10W1, GUCY2F, OR2AG1, OR5P3, OR5P2, OR7G1, OR5M10, GABRA3, GABRA6, OR52E4, OR6C75, NR4A3, OR2M4, OR5AC2, OR2M5, FOXP2, OR12D2, OR12D3, OR52M1, OR4A15, OR11H6, OR4A16, ADRA1B, CHRND, CACNA1E, RIT2, MERTK, CACNA1C, PTENP1, OR10A6, OR52A1, OR5H2, OR4K17, OR11A1, OR51Q1, MBP, SLC1A4, KCNQ5, OR9A4, WDR36, OR5W2, IAPP, SLC24A2, OR13A1, NMUR2, CNTNAP2, OR14J1, IMPDH1, KCNQ1, USH2A, OR5AN1, GABRG1, OR52B2, OR5I1, GABRG3, OR52B4, RAX, NRXN3, NRXN1, OR2W1, GRM3, EYA1, GRM2, GRM8, OR2H1, SBF2, GRM7, OR5V1, PLLP, IL12B, OR6Q1, OAT, ABLIM1, OR5D14, OR8H3, OR5D16, TAC1, OR5D18, OXTR, OR1L3, COL2A1, OR5A1, OR5A2, OR9Q1, ESPN, OR9Q2, OR13C3, HRH1, OR10G9, OR9I1, OR5E1P, OR5AP2, OR5B12, PTN, SCNN1A, NEFL, OR13C8, SCNN1D, OR5B17, OR1B1, OR9G9, OR5K4, DLGAP2, OR9G1, TBX1, AFF2, OR9G4, GRIA4, OR5T2, OR5B2, OR5B3, OR5T1, OR4X1, GRIA2, GRIA1, OR13D1, OR10Q1, OR8H2, PBX3, APBB1,	1.34E-32

	WNT7A, HTR2A	
sensory perception of chemical stimulus	OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, OR4S2, OR52L1, OR8K3, OR5R1, OR5M1, TRPA1, OR5M3, OR4M1, OR5M9, OR10AG1, OR4D6, OR1S2, OR5M8, CNGA2, OR1S1, OR5H14, OR8J1, OR52K1, OR5H15, OR8J3, OR13F1, OR2D3, OR2L8, OR2B3, OR8U8, OR2L13, OR1G1, OR6A2, OR2T33, OR5AS1, OR10C1, OR10W1, OR5P3, OR2AG1, OR7G1, OR5P2, OR5M10, OR52E4, OR6C75, OR5AC2, OR2M4, OR2M5, OR12D2, OR52M1, OR12D3, OR4A15, OR11H6, OR4A16, OR10A6, OR52A1, OR5H2, OR4K17, OR11A1, OR51Q1, OR9A4, OR5W2, OR13A1, OR14J1, OR5AN1, OR5I1, OR52B2, OR52B4, OR2W1, OR2H1, GRM8, OR5V1, GRM7, OR6Q1, OR5D14, OR5D16, OR8H3, OR5D18, OR1L3, OR5A1, OR5A2, OR9Q1, OR9Q2, OR13C3, OR9I1, OR10G9, OR5E1P, OR5AP2, OR5B12, OR13C8, SCNN1A, SCNN1D, OR5B17, OR9G9, OR1B1, OR5K4, OR9G1, OR9G4, OR5B2, OR5T2, OR5B3, OR5T1, OR4X1, OR13D1, OR10Q1, OR8H2	1.72E-32
G-protein coupled receptor protein signaling pathway	ADCY5, GPR123, OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, LPHN2, HTR1B, S1PR1, OR4S2, OR52L1, OR8K3, INSR, OR5R1, ATRNL1, OR5M1, OR5M3, OR4M1, OR5M9, OR10AG1, OR4D6, OR1S2, OR1S1, OR5M8, GABRR3, PPBP, OR5H14, OR52K1, OR8J1, OR8J3, OR5H15, HTR6, OR13F1, GPR50, OR2D3, OR2L8, C3AR1, OR2B3, GNAI1, OR8U8, PKHD1, AKAP12, OR2L13, GPR141, OR1G1, SORCS3, HCRTR1, OR6A2, OR5AS1, OR2T33, BAI3, OR10C1, ENTPD1, DEFB1, OR10W1, GABRE, OR2AG1, OR5P3, OR5P2, OR7G1, GABRA4, OR5M10,	5.72E-30

	GABRA3, GABRA6, OR52E4, OR6C75, KCNK2, OR2M4, OR5AC2, OR2M5, OR12D2, OR12D3, OR52M1, OR11H6, OR4A15, OR4A16, ADRA1B, ADRA1A, OR10A6, OR52A1, OR5H2, PREX2, OR4K17, OR11A1, OR51Q1, TAAR8, TAAR9, EDNRB, OR9A4, OR5W2, IAPP, OR13A1, NMUR2, OR14J1, OR5AN1, GABRG1, OR52B2, OR5I1, GABRG3, OR52B4, FSHR, OR2W1, GRM3, ADRB2, GRM2, GRM8, CHRM2, OR2H1, GRM7, OR5V1, OR6Q1, RAPGEFL1, GAP43, GPRC5C, CXCL5, OR5D14, GPR63, OR8H3, OR5D16, CXCL2, FPR1, GPR6, TAC1, OR5D18, OXTR, PF4, OR1L3, FPR2, GPRC5B, OR5A1, OR5A2, OR9Q1, OR9Q2, APLNR, OR13C3, HRH1, DGKB, OR10G9, OR9I1, OR5E1P, OR5AP2, NPFFR2, OR5B12, OR13C8, OR5B17, OR1B1, OR9G9, GNAO1, OR5K4, DGKK, OR9G1, OR9G4, DGKI, PTGFR, OR5T2, OR5B2, RGS13, OR5B3, OR5T1, OR4X1, OR13D1, MC4R, OR10Q1, OR8H2, HTR2A	
cognition	RP1, OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, BDNF, OR4S2, GRIN2B, OR52L1, OR8K3, CHRNA7, OR5R1, OR5M1, TRPA1, SIX3, OR5M3, OR4M1, OR5M9, OR10AG1, OR4D6, OR1S2, OR1S1, CNGA2, OR5M8, LRAT, OR5H14, OR52K1, OR8J1, OR8J3, OR5H15, OR13F1, OR2D3, OR2L8, OR2B3, OR8U8, RIMS1, OR2L13, OR1G1, CRB1, OR6A2, OR2T33, OR5AS1, OR10C1, OR10W1, GUCY2F, OR5P3, OR2AG1, OR5P2, OR7G1, OR5M10, OR52E4, OR6C75, OR5AC2, OR2M4, OR2M5, FOXP2, OR12D2, OR52M1, OR12D3, OR11H6, OR4A15, OR4A16, ADRA1B, MERTK, CACNA1C, PTENP1, OR10A6, OR52A1, OR5H2, OR4K17, OR11A1, OR51Q1, SLC1A4, OR9A4, WDR36, OR5W2, IAPP, SLC24A2, OR13A1, OR14J1,	2.95E-27

	IMPDH1, KCNQ1, USH2A, OR5AN1, OR52B2, OR5I1, OR52B4, RAX, OR2W1, EYA1, GRM8, OR2H1, OR5V1, GRM7, IL12B, OR6Q1, OAT, ABLIM1, OR5D14, OR8H3, OR5D16, OXTR, TAC1, OR5D18, OR1L3, COL2A1, OR5A1, OR5A2, OR9Q1, ESPN, OR9Q2, OR13C3, OR10G9, OR9I1, OR5E1P, OR5AP2, OR5B12, PTN, OR13C8, SCNN1A, SCNN1D, OR5B17, OR1B1, OR9G9, OR5K4, OR9G1, AFF2, TBX1, OR9G4, OR5T2, OR5B2, OR5B3, OR5T1, OR4X1, GRIA1, OR13D1, OR10Q1, OR8H2, APBB1, HTR2A	
sensory perception	RP1, OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, OR4S2, GRIN2B, OR52L1, OR8K3, OR5R1, OR5M1, TRPA1, SIX3, OR5M3, OR4M1, OR5M9, OR10AG1, OR4D6, OR1S2, OR1S1, CNGA2, OR5M8, LRAT, OR5H14, OR52K1, OR8J1, OR8J3, OR5H15, OR13F1, OR2D3, OR2L8, OR2B3, OR8U8, RIMS1, OR2L13, OR1G1, CRB1, OR6A2, OR2T33, OR5AS1, OR10C1, OR10W1, GUCY2F, OR5P3, OR2AG1, OR5P2, OR7G1, OR5M10, OR52E4, OR6C75, OR5AC2, OR2M4, OR2M5, OR12D2, OR52M1, OR12D3, OR11H6, OR4A15, OR4A16, MERTK, OR10A6, OR52A1, OR5H2, OR4K17, OR11A1, OR51Q1, OR9A4, WDR36, OR5W2, IAPP, SLC24A2, OR13A1, OR14J1, IMPDH1, KCNQ1, USH2A, OR5AN1, OR52B2, OR5I1, OR52B4, RAX, OR2W1, EYA1, GRM8, OR2H1, OR5V1, GRM7, IL12B, OR6Q1, OAT, ABLIM1, OR5D14, OR8H3, OR5D16, OR5D18, TAC1, COL2A1, OR1L3, OR5A1, OR5A2, OR9Q1, ESPN, OR9Q2, OR13C3, OR9I1, OR10G9, OR5E1P, OR5AP2, OR5B12, OR13C8, SCNN1A, SCNN1D, OR5B17, OR9G9, OR1B1, OR5K4, OR9G1, TBX1, OR9G4, OR5T2, OR5B2, OR5B3, OR5T1, OR4X1, OR13D1, OR10Q1, OR8H2, HTR2A	6.98E-26

ion transport	SLC9A9, SLC9A6, SCN3A, GABRB1, SLC26A2, KCNIP4, WNT2, MARCO, KCNQ5, CTTNBP2, ATP2B4, GRIN2B, SLC24A3, SLC24A2, NMUR2, SLCO1C1, ANO2, ANO5, CHRNA7, ANO4, MCOLN2, KCNQ1, OCA2, GABRG1, GABRG3, SVOP, KCND2, ATP4B, TRPA1, CACNG3, CNGA2, GABRR3, ATP2C2, ATP2C1, RYR3, F2, CLIC6, RYR2, KCNH8, SCN11A, PLLP, KCNH5, SLC38A4, SLC39A12, KCNA4, ASZ1, KCNMB2, KCNS3, KCNS2, SLCO1A2, CYP27B1, SCN9A, SCNN1A, SCNN1D, GABRQ, GABRE, TRPC4, SLC8A1, SLC12A1, GABRA4, GABRA3, KCNB2, GABRA6, ATP11B, GRIA4, CACNA2D3, KCNK2, KCNJ4, SLCO1B3, SLC17A3, GRIA2, KCNJ9, SLC17A4, GRIA1, SLC5A8, CACNA1H, SCN4B, CHRND, HEPH, CACNA1E, KCTD16, CACNA1C, CLCN7	5.82E-08
cell-cell signaling	FGF5, SYT5, FGF9, FGF14, ILDR2, GDNF, MBP, WNT2, SLC1A4, KCNQ5, CTTNBP2, HTR1B, BDNF, GRIN2B, IAPP, NMUR2, LTB, GABRG1, GABRG3, KCND2, NRXN3, SIX3, NRXN1, FGF20, CTNNA2, GABRR3, GRM3, GRB10, GRM2, CAMK4, GRM8, GRM7, HTR6, WNT9B, GPR50, CXCL5, ASZ1, TAC1, OXTR, RIMS1, EPHB1, HCRTR1, IL17A, HRH1, CRB1, IFNA7, FCER1A, BMP2, TRHDE, GABRA3, DLGAP2, GABRA6, GRIA4, TNFSF9, GRIA2, GRIA1, SALL1, MC4R, ADRA1B, ADRA1A, CACNA1E, RIT2, LRP2, MERTK, CACNA1C, WNT7A, HTR2A	4.29E-07
transmission of nerve impulse	SYT5, ASZ1, TAC1, OXTR, RIMS1, KCNMB2, MBP, WNT2, SLC1A4, CTTNBP2, KCNQ5, HCRTR1, HRH1, HTR1B, S1PR1, GRIN2B, NMUR2, CNTNAP2, GABRG1, GABRG3, KCND2, DLGAP2, NRXN3, GABRA3,	2.94E-06

	GABRA6, GRIA4, NRXN1, CTNNA2, GABRR3, GRM3, GRM2, CAMK4, GRIA2, GRIA1, SBF2, GRM8, GRM7, HTR6, PLLP, CACNA1E, RIT2, CACNA1C, WNT7A, HTR2A	
synaptic transmission	SYT5, ASZ1, TAC1, OXTR, RIMS1, MBP, WNT2, SLC1A4, CTTNBP2, KCNQ5, HCRTR1, HTR1B, HRH1, GRIN2B, GABRG1, GABRG3, KCND2, NRXN3, GABRA3, DLGAP2, GABRA6, GRIA4, NRXN1, CTNNA2, GABRR3, GRM3, GRM2, CAMK4, GRIA2, GRM8, GRIA1, GRM7, HTR6, CACNA1E, RIT2, CACNA1C, WNT7A, HTR2A	1.13E-05
neuron differentiation	TUBB2B, HELT, RORA, PRKG1, GDNF, RTN1, EFHD1, BDNF, S1PR1, CNTNAP2, ROBO2, UNC5C, ROBO3, NRXN3, STMN2, BAIAP2, NRXN1, FGF20, CTNNA2, FARP2, SLITRK2, BTG4, RELN, CNTN4, GAP43, DCC, PKHD1, SOX5, BRSK1, EPHB1, ALDH1A2, CRB1, PTK2B, TNR, GFI1, OLFM3, DCX, NEFL, APC, GNAO1, BHLHE22, PTPRZ1, NTNG1, LMX1A, PBX3, BMPR1B, APBB1, WNT7A, PTENP1	1.83E-05
metal ion transport	SLC9A9, SLC9A6, SCN3A, KCNIP4, KCNQ5, ATP2B4, GRIN2B, SLC24A3, NMUR2, SLC24A2, CHRNA7, KCNQ1, KCND2, ATP4B, TRPA1, CACNG3, ATP2C2, ATP2C1, RYR3, F2, RYR2, SCN11A, KCNH8, KCNH5, SLC38A4, SLC39A12, KCNA4, KCNMB2, KCNS3, CYP27B1, KCNS2, SCN9A, SCNN1A, SCNN1D, SLC8A1, TRPC4, SLC12A1, KCNB2, CACNA2D3, KCNK2, KCNJ4, KCNJ9, SLC17A3, SLC17A4, SLC5A8, CACNA1H, SCN4B, CACNA1E, HEPH, KCTD16, CACNA1C	2.07E-05
behavior	HELT, LEPR, GDNF, BSX, EDNRB, HTR1B, BDNF, GRIN2B, NMUR2, SLC24A2, ROBO2, CHRNA7, ROBO3,	5.32E-05

	GRM2, PPBP, GRM7, RELN, RASD2, CCL1, C3AR1, CXCL5, CXCL2, FPR1, ASTN1, TAC1, OXTR, PF4, FPR2, ESPN, PLCL1, HCRTR1, SCN9A, PTN, GFI1, DEFB1, GNAO1, TBX1, AFF2, NR4A3, FOXP2, CXCL17, GRIA1, NLGN4X, ADRA1B, MC4R, CACNA1C, PBX3, APBB1, PTENP1, HTR2A	
cell motion	NEURL, ZEB2, PRKG1, GDNF, DSTN, WNT2, CTTNBP2, EDNRB, BDNF, DAB1, VNN1, ROBO2, ROBO3, UNC5C, NR2F2, NRXN3, ARID5B, NRXN1, CTNNA2, ELMO1, VEGFC, RELN, CNTN4, IL12B, GAP43, AKAP4, DCC, VIM, FPR1, ASTN1, ASZ1, PF4, CDH2, FPR2, EPHB1, PTK2B, TNR, CLASP2, DCX, APC, FN1, FLT1, TBX1, LMX1A, ETS1, BMPR1B, APBB1, SELE, PTENP1	1.41E-04
regulation of system process	KCNE1L, BMP10, LZTS1, TAC1, OXTR, CDH2, GDNF, KCNMB2, EDNRB, HTR1B, BDNF, GRIN2B, SLC24A2, NMUR2, LTB, KCNQ1, SLC8A1, GNAO1, FLT1, KCNB2, GRIA4, LAMA2, GRM3, ADRB2, GRIA2, CHRM2, GRM8, ADRA1B, CACNA1H, NPPB, RYR2, CNTN4, PBX3, CACNA1C, HTR2A	2.78E-04
cell adhesion	DLC1, COL21A1, MYBPC2, NELL1, SDC3, REG3A, ARHGAP6, DAB1, S1PR1, DGCR6, CNTNAP2, VNN1, ROBO2, ADAM9, CNTNAP5, NRXN3, SIGLEC12, NRXN1, BTBD9, CTNNA2, NCAM1, CD84, ARVCF, ATP2C1, LSAMP, ROR2, RELN, SGCE, CNTN4, CHL1, PARVA, DCHS2, CLDN17, PLXNC1, PKHD1, TNC, ASTN1, COL2A1, FPR2, CDH2, APLP1, ITGBL1, IGSF11, LAMB4, CDH7, COL6A6, FAT3, PTK2B, FAT4, TNR, HEPN1, ITGB6, DEFB118, COL8A1, ENTPD1, FN1, APC, TNXB, LAMA2, COL19A1, NLGN4X, COL29A1, BMPR1B, SELE	3.77E-04

biological adhesion	DLC1, COL21A1, MYBPC2, NELL1, SDC3, REG3A, ARHGAP6, DAB1, S1PR1, DGCR6, CNTNAP2, VNN1, ROBO2, ADAM9, CNTNAP5, NRXN3, SIGLEC12, NRXN1, BTBD9, CTNNA2, NCAM1, CD84, ARVCF, ATP2C1, LSAMP, ROR2, RELN, SGCE, CNTN4, CHL1, PARVA, DCHS2, CLDN17, PLXNC1, PKHD1, TNC, ASTN1, COL2A1, FPR2, CDH2, APLP1, ITGBL1, IGSF11, LAMB4, CDH7, COL6A6, FAT3, PTK2B, FAT4, TNR, HEPN1, ITGB6, DEFB118, COL8A1, ENTPD1, FN1, APC, TNXB, LAMA2, COL19A1, NLGN4X, COL29A1, BMPR1B, SELE	3.91E-04
neuron projection development	DCC, PKHD1, PRKG1, GDNF, EPHB1, EFHD1, BDNF, PTK2B, TNR, ROBO2, UNC5C, ROBO3, DCX, NEFL, APC, GNAO1, NRXN3, PTPRZ1, BAIAP2, NTNG1, NRXN1, LMX1A, CTNNA2, SLITRK2, RELN, CNTN4, BMPR1B, APBB1, GAP43, PTENP1	4.75E-04
calcium ion transport	TRPC4, SLC8A1, TRPA1, CACNG3, CACNA2D3, ATP2C2, ATP2B4, CYP27B1, GRIN2B, SLC24A3, ATP2C1, RYR3, SLC24A2, NMUR2, F2, CACNA1H, RYR2, CHRNA7, CACNA1E, CACNA1C	5.98E-04

Hypomethylated**

Pathway	*Genes*	*p value*
fear response	BDNF, CCK, GRIK2	0.007256045
cell morphogenesis involved in differentiation	SEMA5A, BDNF, CCK, KAL1, C22ORF28, DCLK1	0.025362986
neuron development	SEMA5A, BDNF, CCK, KAL1, RORB, DCLK1, NTM	0.027953923
multicellular organismal response to	BDNF, CCK, GRIK2	0.028802466

stress		
neuron differentiation	SEMA5A, BDNF, CCK, KAL1, RORB, SMARCA1, DCLK1, NTM	0.029533905

APPENDIX B

HYPERMETHYLATED PATHWAYS SHARED BY TMX2-11 AND TMX2-28 IN THE PROMOTER

Pathway	*Genes*	*p value*
sensory perception of smell	OR10A6, OR5H2, OR5L1, OR11A1, OR2J2, OR1J4, OR51Q1, OR4D10, OR52D1, OR8K1, OR9A4, OR4S2, OR5W2, OR13A1, OR8K3, OR5AN1, OR52B2, OR5I1, OR52B4, OR5M1, OR4M1, OR4D6, OR1S2, OR1S1, OR5M8, CNGA2, OR5H14, OR52K1, GRM8, OR5V1, OR5H15, OR2D3, OR2L8, OR2B3, OR8U8, OR8H3, OR1L3, OR5A2, OR2L13, OR9Q1, OR1G1, OR9Q2, OR13C3, OR9I1, OR10G9, OR5E1P, OR5AP2, OR2T33, OR5B12, OR5AS1, OR10C1, OR10W1, OR5B17, OR9G9, OR5P3, OR2AG1, OR7G1, OR5P2, OR5K4, OR9G1, OR6C75, OR2M4, OR5B3, OR2M5, OR4X1, OR52M1, OR11H6, OR13D1, OR4A16, OR8H2	6.50E-32
sensory perception of chemical stimulus	OR10A6, OR5H2, OR5L1, OR11A1, OR2J2, OR51Q1, OR1J4, OR4D10, OR52D1, OR8K1, OR9A4, OR4S2, OR5W2, OR13A1, OR8K3, OR5AN1, OR52B2, OR5I1, OR52B4, OR5M1, OR4M1, OR4D6, OR1S2, OR1S1, OR5M8, CNGA2, OR5H14, OR52K1, GRM8, OR5V1, OR5H15, OR2D3, OR2L8, OR2B3, OR8U8, OR8H3, OR1L3, OR5A2, OR2L13, OR9Q1, OR1G1, OR9Q2, OR13C3, OR9I1, OR10G9, OR5E1P, OR5AP2, OR2T33, OR5B12, OR5AS1, OR10C1, SCNN1A, SCNN1D, OR10W1, OR5B17, OR9G9, OR5P3, OR2AG1, OR7G1, OR5P2, OR5K4, OR9G1, OR6C75, OR2M4, OR5B3, OR2M5, OR4X1, OR52M1, OR11H6, OR13D1, OR4A16, OR8H2	8.96E-31
G-protein coupled receptor protein signaling pathway	OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, HTR1B, S1PR1, OR4S2, OR8K3, OR5M1, ATRNL1, OR4M1, OR4D6, OR1S2, OR5M8, OR1S1, GABRR3, PPBP, OR5H14, OR52K1, HTR6, OR5H15, OR2D3, C3AR1, OR2L8, OR2B3, OR8U8, AKAP12, GPR141, OR2L13, OR1G1, OR2T33, OR5AS1, OR10C1, DEFB1, OR10W1, GABRE, OR5P3, OR2AG1, OR7G1, OR5P2, GABRA3, GABRA6, OR6C75, OR2M4, OR2M5, OR52M1, OR11H6, OR4A16, OR10A6, OR5H2, OR11A1, OR51Q1, EDNRB, TAAR9, OR9A4, OR5W2, OR13A1, OR5AN1, GABRG1, OR5I1, OR52B2, OR52B4, FSHR, GRM3,	5.19E-26

	CHRM2, GRM8, OR5V1, GAP43, CXCL5, GPR63, CXCL2, OR8H3, GPR6, FPR1, TAC1, OR1L3, PF4, FPR2, OR5A2, OR9Q1, OR9Q2, HRH1, OR13C3, DGKB, OR9I1, OR10G9, OR5E1P, OR5B12, OR5AP2, OR5B17, OR9G9, GNAO1, OR5K4, OR9G1, DGKK, OR5B3, RGS13, OR4X1, OR13D1, OR8H2	
neurological system process	OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, CTTNBP2, BDNF, HTR1B, S1PR1, OR4S2, OR8K3, OR5M1, SIX3, OR4M1, OR4D6, OR1S2, OR5M8, CNGA2, OR1S1, GABRR3, OR5H14, OR52K1, HTR6, OR5H15, OR2D3, OR2L8, OR2B3, OR8U8, ASZ1, RIMS1, OR2L13, KCNMB2, OR1G1, CRB1, OR2T33, OR5AS1, OR10C1, OR10W1, OR5P3, OR2AG1, OR7G1, OR5P2, GABRA3, GABRA6, OR6C75, OR2M4, OR2M5, OR52M1, OR11H6, OR4A16, CHRND, PTENP1, OR10A6, OR5H2, OR11A1, OR51Q1, SLC1A4, OR9A4, WDR36, OR5W2, OR13A1, CNTNAP2, KCNQ1, OR5AN1, GABRG1, OR5I1, OR52B2, OR52B4, NRXN1, GRM3, GRM8, OR5V1, OR8H3, TAC1, OR1L3, OR5A2, OR9Q1, OR9Q2, HRH1, OR13C3, OR9I1, OR10G9, OR5E1P, OR5B12, OR5AP2, SCNN1A, SCNN1D, OR5B17, OR9G9, OR5K4, DLGAP2, OR9G1, GRIA4, OR5B3, OR4X1, GRIA1, OR13D1, OR8H2, APBB1	2.31E-22
sensory perception	OR10A6, OR5H2, OR5L1, OR2J2, OR11A1, OR51Q1, OR1J4, OR4D10, OR52D1, OR8K1, OR9A4, WDR36, OR4S2, OR5W2, OR13A1, OR8K3, KCNQ1, OR5AN1, OR52B2, OR5I1, OR52B4, OR5M1, SIX3, OR4M1, OR4D6, OR1S2, OR1S1, OR5M8, CNGA2, OR5H14, OR52K1, GRM8, OR5V1, OR5H15, OR2D3, OR2L8, OR2B3, OR8U8, OR8H3, TAC1, OR1L3, RIMS1, OR5A2, OR2L13, OR9Q1, OR1G1, OR9Q2, OR13C3, CRB1, OR9I1, OR10G9, OR5E1P, OR5AP2, OR2T33, OR5B12, OR5AS1, OR10C1, SCNN1A, SCNN1D, OR10W1, OR5B17, OR9G9, OR5P3, OR2AG1, OR7G1, OR5P2, OR5K4, OR9G1, OR6C75, OR2M4, OR5B3, OR2M5, OR4X1, OR52M1, OR11H6, OR13D1, OR4A16, OR8H2	6.40E-21
cognition	OR10A6, OR5H2, OR5L1, OR2J2, OR11A1, OR51Q1, OR1J4, OR4D10, OR52D1, OR8K1, SLC1A4, BDNF, OR9A4, WDR36, OR4S2, OR5W2, OR13A1, OR8K3, KCNQ1, OR5AN1, OR52B2, OR5I1, OR52B4, OR5M1, SIX3, OR4M1, OR4D6, OR1S2, OR1S1, OR5M8, CNGA2, OR5H14, GRM8, OR52K1, OR5V1,	6.69E-21

	OR5H15, OR2D3, OR2L8, OR2B3, OR8U8, OR8H3, TAC1, OR1L3, OR5A2, RIMS1, OR2L13, OR9Q1, OR1G1, OR9Q2, OR13C3, CRB1, OR9I1, OR10G9, OR5E1P, OR5AP2, OR2T33, OR5B12, OR5AS1, OR10C1, SCNN1A, SCNN1D, OR10W1, OR5B17, OR9G9, OR5P3, OR2AG1, OR7G1, OR5P2, OR5K4, OR9G1, OR6C75, OR2M4, OR5B3, OR2M5, OR4X1, OR52M1, GRIA1, OR11H6, OR13D1, OR4A16, OR8H2, APBB1, PTENP1	
cell surface receptor linked signal transduction	BMP10, OR5L1, OR2J2, OR1J4, OR4D10, OR52D1, OR8K1, MARCO, HTR1B, S1PR1, OR4S2, OR8K3, IRS2, OR5M1, ATRNL1, FGF22, SOCS5, OR4M1, OR4D6, OR1S2, OR1S1, OR5M8, GABRR3, VEGFC, GRB10, PPBP, OR5H14, OR52K1, HTR6, F2, OR5H15, ADAMTS1, OR2D3, C3AR1, OR2L8, OR2B3, OR8U8, AKAP12, GPR141, OR2L13, OR1G1, OR2T33, OR5AS1, OR10C1, DEFB1, OR10W1, GABRE, BMP2, OR5P3, OR2AG1, TCF7, OR7G1, OR5P2, GABRA3, GABRA6, SMAD5, OR6C75, TAX1BP3, OR2M4, OR2M5, OR52M1, SFRP2, OR11H6, OR4A16, PTENP1, OR10A6, OR5H2, OR11A1, OR51Q1, TAAR9, EDNRB, OR9A4, OR5W2, OR13A1, OR5AN1, GABRG1, OR52B2, OR5I1, OR52B4, FSHR, GRM3, DACT1, CHRDL1, GRM8, CHRM2, OR5V1, GAP43, CXCL5, APC2, GPR63, OR8H3, CXCL2, GPR6, FPR1, TAC1, PF4, OR1L3, FPR2, OR5A2, OR9Q1, OR9Q2, HRH1, OR13C3, DGKB, OR9I1, OR10G9, PTK2B, OR5E1P, OR5AP2, OR5B12, APC, OR5B17, OR9G9, GNAO1, OR5K4, OR9G1, DGKK, GRIA4, RGS13, OR5B3, OR4X1, OR13D1, OR8H2, IFT52	2.52E-20
defense response to bacterium	DEFB121, PPBP, DEFB125, DEFB118, DEFB128, DEFB116, DEFB115, DEFB1, CTSG, DMBT1, DEFB119	0.001357161
gamma-aminobutyric acid signaling pathway	GABRG1, GABRE, GABRR3, GABRA3, GABRA6	0.002200012
ion transport	SLC9A6, GABRB1, SLC39A12, ASZ1, SLC26A2, KCNIP4, KCNMB2, KCNS3, MARCO, CTTNBP2, SLC24A3, SCN9A, SLCO1C1, ANO5, MCOLN2, ANO4, KCNQ1, SCNN1A, SCNN1D, GABRG1, GABRE, SLC8A1, SVOP, GABRA3, GABRA6, GRIA4, CNGA2, GABRR3, KCNJ4, ATP2C2, KCNJ9, GRIA1, F2, SLC5A8, SCN4B, CHRND, HEPH	0.002297548
transmission of nerve impulse	GABRG1, DLGAP2, GABRA3, GABRA6, ASZ1, TAC1, NRXN1, GRIA4, RIMS1, KCNMB2, SLC1A4,	0.0025710 8

	CTTNBP2, GABRR3, GRM3, HRH1, HTR1B, S1PR1, GRM8, GRIA1, HTR6, CNTNAP2	
regulation of cell migration	BMP10, IRS2, ONECUT1, TAC1, LAMA2, VEGFC, S1PR1, SERPINE2, PTK2B, UNC5C, HDAC9, PTENP1, APC	0.003347862
behavior	CCL1, C3AR1, HELT, CXCL5, CXCL2, FPR1, TAC1, PF4, FPR2, EDNRB, PLCL1, BDNF, HTR1B, SCN9A, ROBO2, GFI1, DEFB1, GNAO1, CXCL17, PPBP, GRIA1, NLGN4X, APBB1, PTENP1, RASD2	0.004200034
synaptic transmission	GABRG1, DLGAP2, GABRA3, GABRA6, ASZ1, TAC1, NRXN1, GRIA4, RIMS1, SLC1A4, CTTNBP2, GABRR3, HRH1, HTR1B, GRM3, GRM8, GRIA1, HTR6	0.005334357
chemotaxis	CCL1, C3AR1, EDNRB, CXCL17, PPBP, CXCL5, CXCL2, FPR1, ROBO2, PF4, FPR2, DEFB1	0.006194279
taxis	CCL1, C3AR1, EDNRB, CXCL17, PPBP, CXCL5, CXCL2, FPR1, ROBO2, PF4, FPR2, DEFB1	0.006194279
cell-cell signaling	CXCL5, ASZ1, TAC1, RIMS1, SLC1A4, CTTNBP2, HTR1B, HRH1, BDNF, IL17A, IFNA7, CRB1, GABRG1, BMP2, TRHDE, DLGAP2, GABRA3, GABRA6, SIX3, GRIA4, NRXN1, GABRR3, GRB10, GRM3, GRM8, GRIA1, SALL1, HTR6, LRP2	0.007381335
regulation of locomotion	BMP10, IRS2, ONECUT1, TAC1, LAMA2, VEGFC, S1PR1, SERPINE2, PTK2B, UNC5C, HDAC9, PTENP1, APC	0.009085341
response to drug	GABRE, NES, SLC8A1, GNAO1, ASZ1, ADIPOQ, CTTNBP2, CYP7B1, BDNF, PTK2B, SLC18A1, LRP2, PTENP1, APC	0.009175427
regulation of cell motion	BMP10, IRS2, ONECUT1, TAC1, LAMA2, VEGFC, S1PR1, SERPINE2, PTK2B, UNC5C, HDAC9, PTENP1, APC	0.009446985

APPENDIX C

HYPERMETHYLATED PATHWAYS SHARED BY TMX2-11 AND TMX2-28 IN THE GENE BODY

Pathway	*Genes*	*p value*
ion transport	SLC9A9, KCNIP4, WNT2, KCNQ5, GRIN2B, NMUR2, SLC24A2, ANO2, CHRNA7, KCNQ1, OCA2, GABRG3, KCND2, ATP4B, TRPA1, CACNG3, CNGA2, ATP2C1, RYR3, CLIC6, RYR2, PLLP, SCN11A, KCNH8, KCNH5, SLC38A4, SLC39A12, SCN9A, SCNN1A, GABRQ, GABRE, TRPC4, SLC12A1, GABRA4, KCNB2, GABRA3, ATP11B, GRIA4, CACNA2D3, KCNK2, SLCO1B3, GRIA2, GRIA1, SLC17A4, CACNA1H, SCN4B, CACNA1E, HEPH, KCTD16, CACNA1C, CLCN7	7.35E-09
metal ion transport	SLC9A9, SLC38A4, SLC39A12, KCNIP4, KCNQ5, GRIN2B, SLC24A2, NMUR2, SCN9A, CHRNA7, KCNQ1, SCNN1A, TRPC4, SLC12A1, KCND2, ATP4B, KCNB2, TRPA1, CACNG3, CACNA2D3, KCNK2, ATP2C1, SLC17A4, RYR3, CACNA1H, SCN4B, RYR2, HEPH, SCN11A, CACNA1E, KCNH8, KCTD16, CACNA1C, KCNH5	4.77E-07
cell adhesion	DLC1, DCHS2, PLXNC1, PKHD1, NELL1, ASTN1, COL2A1, FPR2, CDH2, SDC3, APLP1, ITGBL1, REG3A, IGSF11, ARHGAP6, COL6A6, FAT3, DGCR6, VNN1, CNTNAP2, COL8A1, ENTPD1, ADAM9, NRXN3, CNTNAP5, NRXN1, BTBD9, CTNNA2, NCAM1, LAMA2, COL19A1, ATP2C1, LSAMP, NLGN4X, ROR2, RELN, SGCE, CNTN4, COL29A1, SELE, CHL1, PARVA	2.95E-06
biological adhesion	DLC1, DCHS2, PLXNC1, PKHD1, NELL1, ASTN1, COL2A1, FPR2, CDH2, SDC3, APLP1, ITGBL1, REG3A, IGSF11, ARHGAP6, COL6A6, FAT3, DGCR6, VNN1, CNTNAP2, COL8A1, ENTPD1, ADAM9, NRXN3, CNTNAP5, NRXN1, BTBD9, CTNNA2, NCAM1, LAMA2, COL19A1, ATP2C1, LSAMP, NLGN4X, ROR2, RELN, SGCE, CNTN4, COL29A1, SELE, CHL1, PARVA	3.02E-06
cell-cell signaling	FGF9, FGF14, ILDR2, RIMS1, EPHB1, WNT2, KCNQ5, BDNF, GRIN2B, NMUR2, LTB, BMP2, GABRG3, KCND2, TRHDE, NRXN3, GABRA3, GRIA4, NRXN1, FGF20, TNFSF9, CTNNA2, CAMK4, GRIA2, GRM8, GRIA1, GRM7, ADRA1B, WNT9B, ADRA1A, GPR50,	1.78E-05

	CACNA1E, RIT2, LRP2, CACNA1C, WNT7A	
cation transport	SLC9A9, SLC38A4, SLC39A12, KCNIP4, KCNQ5, GRIN2B, SLC24A2, NMUR2, SCN9A, CHRNA7, KCNQ1, SCNN1A, TRPC4, SLC12A1, KCND2, ATP4B, KCNB2, TRPA1, CACNG3, CACNA2D3, KCNK2, ATP2C1, SLC17A4, RYR3, CACNA1H, SCN4B, RYR2, HEPH, SCN11A, CACNA1E, KCNH8, KCTD16, CACNA1C, KCNH5	1.96E-05
multicellular organismal response to stress	EDNRB, BDNF, GRIN2B, NMUR2, GRM7, TRPA1, SCN9A, RELN	7.55E-05
transmission of nerve impulse	GABRG3, KCND2, GABRA3, NRXN3, NRXN1, GRIA4, RIMS1, CTNNA2, WNT2, KCNQ5, CAMK4, GRIA2, GRIN2B, GRM8, SBF2, GRIA1, NMUR2, GRM7, CNTNAP2, PLLP, CACNA1E, RIT2, CACNA1C, WNT7A	8.78E-05
neurological system process	WNT2, KCNQ5, BDNF, GRIN2B, SLC24A2, NMUR2, CNTNAP2, CHRNA7, KCNQ1, IMPDH1, USH2A, GABRG3, OR5R1, KCND2, RAX, OR5M1, NRXN3, TRPA1, OR5M3, NRXN1, OR5M9, CTNNA2, CNGA2, OR5M8, EYA1, LRAT, CAMK4, SBF2, OR2H1, GRM8, GRM7, PLLP, ABLIM1, OR8U8, COL2A1, RIMS1, OR9Q1, OR5AP2, NEFL, SCNN1A, GUCY2F, OR9G9, OR5M10, GABRA3, OR9G1, TBX1, AFF2, GRIA4, FOXP2, GRIA2, GRIA1, ADRA1B, CACNA1E, RIT2, CACNA1C, PBX3, WNT7A	1.01E-04
appendage development	MSX2, ALDH1A2, RAX, FGF9, PRRX1, COL2A1, ZBTB16, ASPH, SP8, MECOM, NR2F2, WNT7A	1.10E-04
limb development	MSX2, ALDH1A2, RAX, FGF9, PRRX1, COL2A1, ZBTB16, ASPH, SP8, MECOM, NR2F2, WNT7A	1.10E-04
calcium ion transport	TRPC4, TRPA1, CACNG3, CACNA2D3, GRIN2B, ATP2C1, RYR3, SLC24A2, NMUR2, CACNA1H, RYR2, CACNA1E, CHRNA7, CACNA1C	1.33E-04
cell surface receptor linked signal transduction	FGF9, GPR123, LEPR, ADCY5, PREX2, WNT2, EDNRB, GRIN2B, NMUR2, GAB1, RSPO2, HEY2, NRG1, INSR, ADAM9, GABRG3, OR5R1, OR5M1, CCDC88C, BAIAP2, ARID5B, DLL1, OR5M3, FGF20, OR5M9, OR5M8, NCAM1, CHRDL1, GRM8, OR2H1, GRM7, WNT9B, ROR1, GPR50, ROR2, GAP43, AKAP4, GPRC5C, OR8U8, PKHD1, GNAI1, APH1B, AKAP12, FPR2, EPHB1, OR9Q1, ITGBL1, SORCS3, MSX2, APLNR, DGKB, OR5AP2, BAI3, ENTPD1, GUCY2F, GABRE, OR9G9, PTPRD, BMP2, FLT1, GABRA4, OR5M10, GABRA3, MAML2, SMAD5, OR9G1, DGKK, DGKI, GRIA4, PTGFR, KCNK2, BTLA, NOTCH4,	1.35E-04

	ADRA1B, ADRA1A, BAMBI, WNT7A, ADAMDEC1	
response to pain	EDNRB, GRIN2B, NMUR2, TRPA1, SCN9A, RELN	1.60E-04
neuron differentiation	PKHD1, RORA, PRKG1, RTN1, EPHB1, EFHD1, ALDH1A2, BDNF, CNTNAP2, ROBO3, OLFM3, NEFL, NRXN3, STMN2, PTPRZ1, BAIAP2, NTNG1, NRXN1, LMX1A, FGF20, CTNNA2, BTG4, CNTN4, RELN, PBX3, WNT7A, GAP43	1.66E-04
muscle organ development	POU6F1, ARID5B, UTRN, MYL1, MOV10L1, TBX1, FOXP2, LAMA2, COL19A1, HLX, CACNA1H, VGLL2, ZFPM2, SGCE, NRG1, NR2F2, TCF12	2.15E-04
cell motion	ASTN1, ZEB2, CDH2, FPR2, PRKG1, EPHB1, DSTN, WNT2, EDNRB, BDNF, VNN1, CLASP2, ROBO3, NR2F2, FLT1, NRXN3, ARID5B, TBX1, NRXN1, LMX1A, ELMO1, CTNNA2, ETS1, CNTN4, RELN, SELE, GAP43, AKAP4	2.50E-04
regulation of system process	LZTS1, FLT1, KCNB2, CDH2, GRIA4, LAMA2, EDNRB, BDNF, GRIA2, GRIN2B, GRM8, NMUR2, SLC24A2, ADRA1B, CACNA1H, RYR2, CNTN4, PBX3, CACNA1C, KCNQ1, LTB	3.11E-04
di-, tri-valent inorganic cation transport	TRPC4, TRPA1, CACNG3, CACNA2D3, GRIN2B, ATP2C1, RYR3, SLC24A2, NMUR2, CACNA1H, RYR2, CHRNA7, CACNA1E, HEPH, CACNA1C	3.26E-04
synaptic transmission	GABRG3, KCND2, GABRA3, NRXN3, NRXN1, GRIA4, RIMS1, CTNNA2, WNT2, KCNQ5, CAMK4, GRIA2, GRIN2B, GRM8, GRIA1, GRM7, CACNA1E, RIT2, CACNA1C, WNT7A	5.24E-04

APPENDIX D

GENES WITH TWO OR MORE HYPERMETHYLATED CPG SITES IN THE PROMOTER IN BOTH TMX2-11 AND TMX2-28 IN ORDER OF DECREASING NUMBER OF HYPERMETHYLATED CPGS

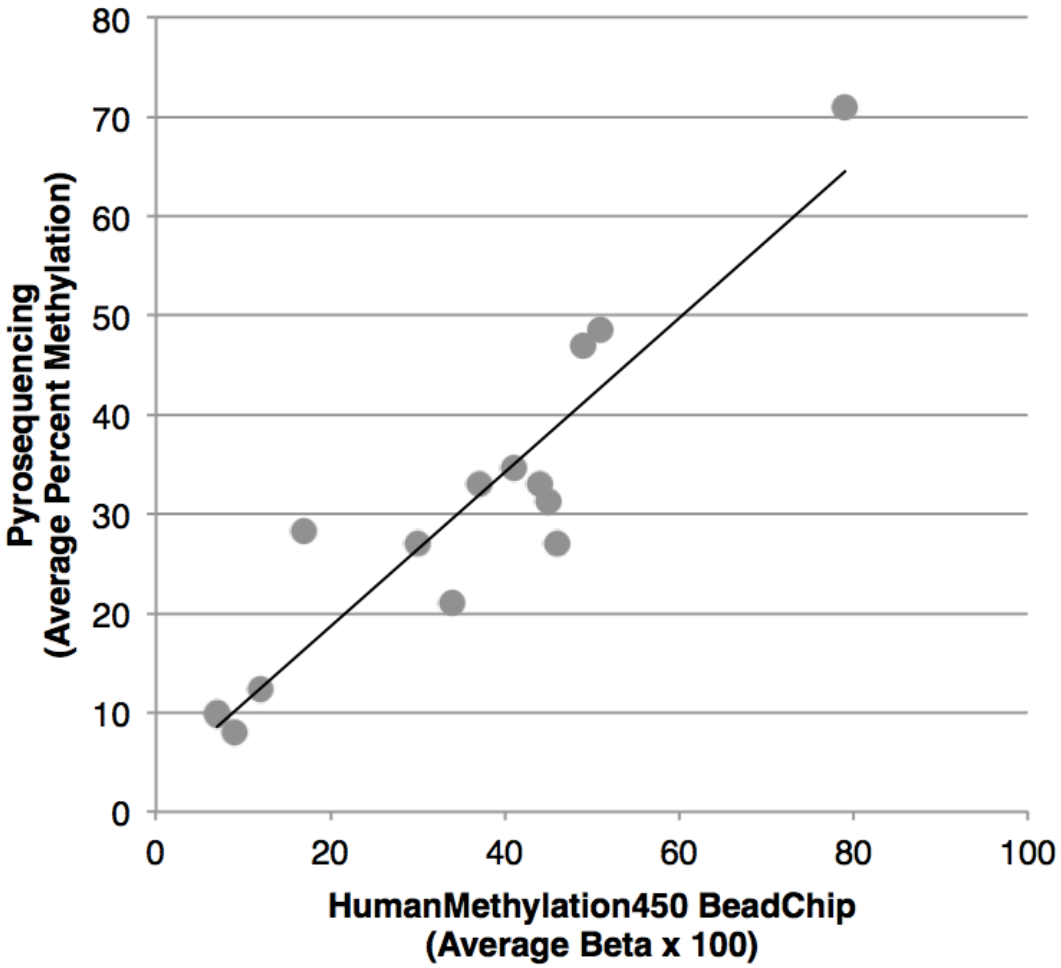
Official Gene Symbol	Number of hypermethylated CpGs	Average fold change TMX2-11	Average fold change TMX2-28
EDNRB	11	2.44	3.19
SORBS2	7	4.98	7.95
EVX2	6	3.05	3.70
ZBTB20	6	4.19	5.56
CXorf67	5	3.34	3.08
MIR568	5	4.33	5.19
ZNF350	5	4.63	3.82
BMP2	4	9.09	9.55
COL21A1	4	4.62	7.30
CSGALNACT1	4	2.88	5.56
CXCL2	4	6.35	16.98
FMO6P	4	2.24	2.52
RGS13	4	3.00	3.50
ZNF385B/MIR1258	4	5.90	9.33
ASZ1	3	2.44	2.94
C18orf20	3	2.67	3.05
CHRDL1	3	2.30	2.55
COL29A1	3	5.17	4.29
COX7B2	3	2.67	3.98
DGKK	3	2.39	2.70
ELAVL4	3	3.66	5.16
FAM55D	3	4.69	9.05
GFI1	3	4.98	7.07
HDAC9	3	2.56	3.92
HOXB9	3	2.67	3.94
LOC284688	3	3.39	3.93
LZTS1	3	3.49	4.67
MAGEA5	3	3.44	5.32
MIR452	3	4.03	2.87
MIR548I4	3	3.35	5.59
MOV10L1	3	2.32	4.45
OR2M5	3	2.63	6.26
PAK3	3	2.63	2.75
PRKCDBP	3	2.26	2.36
VANGL2	3	5.57	7.40
ZNF215	3	4.96	10.56
ZNF22	3	3.51	3.87
ACTL9	2	3.00	3.49
AKAP12	2	2.30	2.40
ALDH3A1	2	2.06	2.95
ASCL2	2	2.18	3.45

BMP15	2	3.96	4.14
C6orf64	2	3.49	4.92
CDC14C	2	4.28	4.25
CDKL2	2	2.49	2.38
CEACAM18	2	3.03	6.18
CNGA2	2	6.83	7.06
COL6A6	2	3.55	3.12
CXCL5	2	2.96	2.76
DEFB115	2	2.86	4.38
DEFB116	2	4.44	8.06
DEFB118	2	2.18	2.53
FAM151B	2	2.25	2.83
FAM19A4	2	5.55	4.79
FAM55A	2	3.18	6.89
FPR2	2	2.44	3.61
FSHR	2	2.66	2.74
GABRE	2	2.97	2.77
GAP43	2	4.87	4.37
GLYATL1	2	2.17	3.45
GRB10	2	3.40	5.56
GRM3	2	3.19	3.93
GRM8	2	2.15	3.38
GRXCR1	2	2.61	3.76
GUCY1A2	2	5.19	8.96
HCG4	2	2.60	3.36
HELT	2	4.80	3.55
HS3ST1	2	5.88	10.5
IL17A	2	2.38	2.55
IL17F	2	2.23	2.41
KCNIP4	2	2.75	3.93
KCNJ4	2	2.20	2.44
KRTAP11-1	2	2.80	4.67
KRTAP27-1	2	2.44	4.55
LOC441666	2	2.19	2.14
LOC728640	2	2.29	2.80
LRP2	2	4.87	3.04
MAGEA10	2	2.82	3.87
MAGED1	2	3.56	5.58
MAP3K15	2	6.48	5.50
MAPRE2	2	2.67	2.85
METTL11B	2	2.88	3.12
MIR105-1	2	6.28	6.03
MIR548F3	2	3.81	5.46
MIR592	2	3.42	4.53
MIR921	2	2.66	3.32
MOXD2	2	3.11	4.64
MS4A13	2	2.17	3.31
MS4A4A	2	2.06	3.14
MS4A6A	2	2.15	3.14
MTUS2	2	3.09	4.55
OR1S1	2	5.29	6.93
OR2J2	2	2.73	3.15
OR4D6	2	2.59	3.37
OR5AS1	2	2.18	4.25
OR5B12	2	11.67	17.35
OR5E1P	2	3.25	6.42
OR5H15	2	2.69	4.28
OR5P3	2	4.93	10.71
OR5V1	2	3.80	6.60

OR5W2	2	3.06	2.57
OR8H3	2	2.92	2.70
PPBP	2	2.78	4.04
PRKG2	2	2.85	3.59
REG1A	2	2.31	4.19
REG1B	2	3.44	5.86
RORA	2	3.00	2.46
SGIP1	2	2.10	3.13
SIX3	2	3.26	7.35
SLC1A4	2	7.09	9.10
SLC35F4	2	2.88	5.10
SLC46A1	2	2.16	2.22
SSRP1	2	3.29	6.83
TAC1	2	2.30	3.36
TCF12	2	2.77	4.32
TIGIT	2	2.70	2.45
UNC5D	2	2.15	2.52
ZNF396	2	4.21	4.36
ZNF615	2	2.56	4.01
ZSCAN18	2	2.45	2.23

APPENDIX E

COMPARISON OF HM450BC AND PYROSEQUENCING DATA SHOWS CORRELATION BETWEEN ANALYSIS METHODS

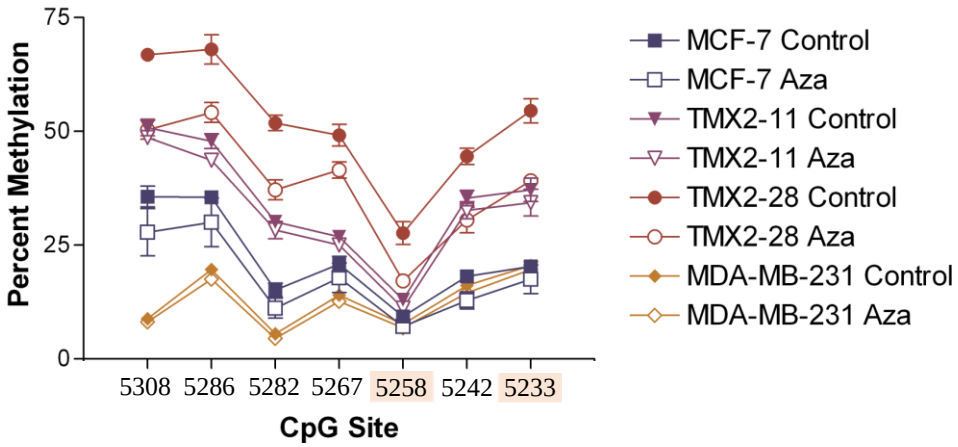


CpG site methylation data from Table 4 were plotted to determine the correlation coefficient of the HM450 BeadChip (x-axis) and pyrosequencing (y-axis) data. All four ZNF350 and the single MAGED1 CpG sites are included on the graph. A single replicate was used for HM450 BeadChip data and the average of three replicates was used for pyrosequencing data. HM450 BeadChip beta value times 100 was calculated to compare with pyrosequencing percents. Pearson coefficient (R) = 0.931, p-value =<0.0001.

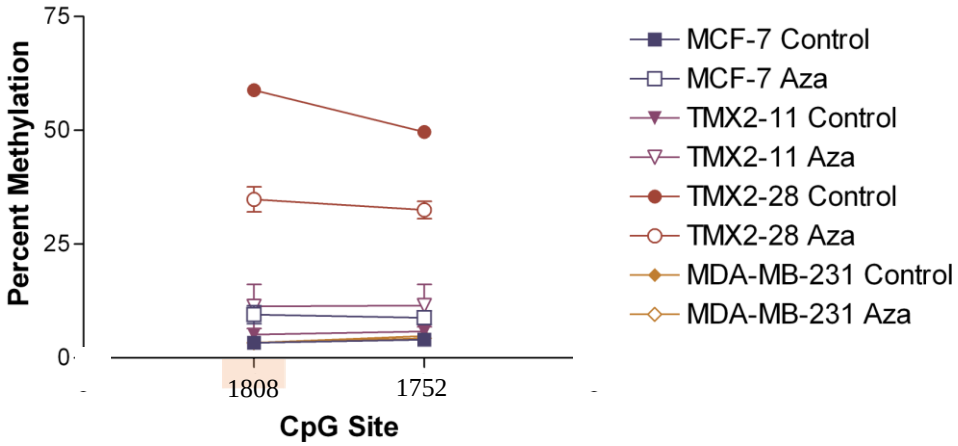
APPENDIX F

METHYLATION OF RORA AND THBS1 ACROSS CPG SITES INTERROGATED BY PYROSEQUENCING

A. RORA



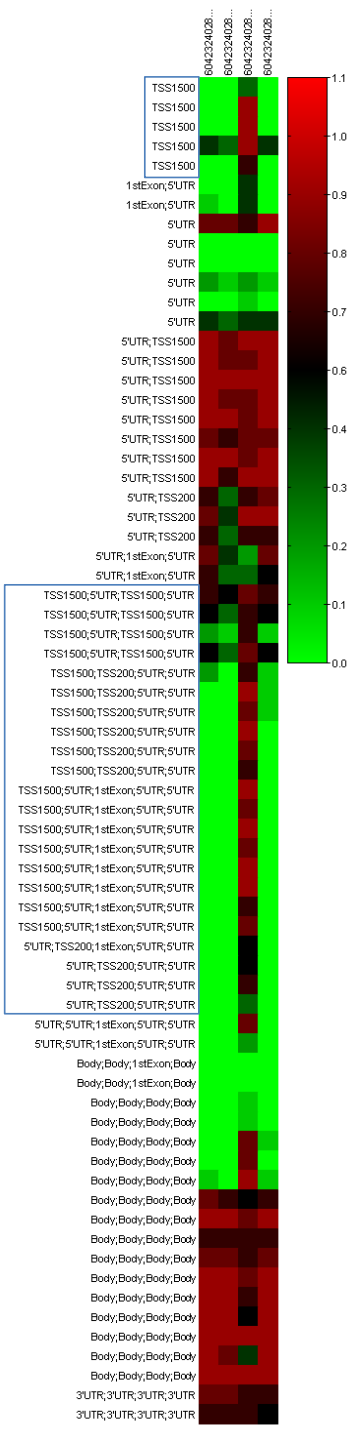
B. THBS1



A) RORA methylation decreases in all seven CpG sites analyzed by pyrosequencing. Methylation of CpG site 5258 (MAPINFO 60885258, highlighted in orange), a site interrogated by the HM450BC, decreases to levels similar to those in MCF-7 and TMX2-11. Site 5233 (MAPINFO 60885233, highlighted in orange), another site interrogated by the HM450BC, also shows a decrease in methylation to levels similar to TMX2-11. B) Two sites were analyzed by pyrosequencing in the THBS1 gene, one of which, 1808 was on the HM450BC. Pyrosequencing analysis shows that treatment with 5-Aza decreases the levels of both CpG sites interrogated, however levels were not decreased to those of MCF-7 or TMX2-11. Analysis was completed using two separate experiments conducted 9 months apart (n=6).

APPENDIX G

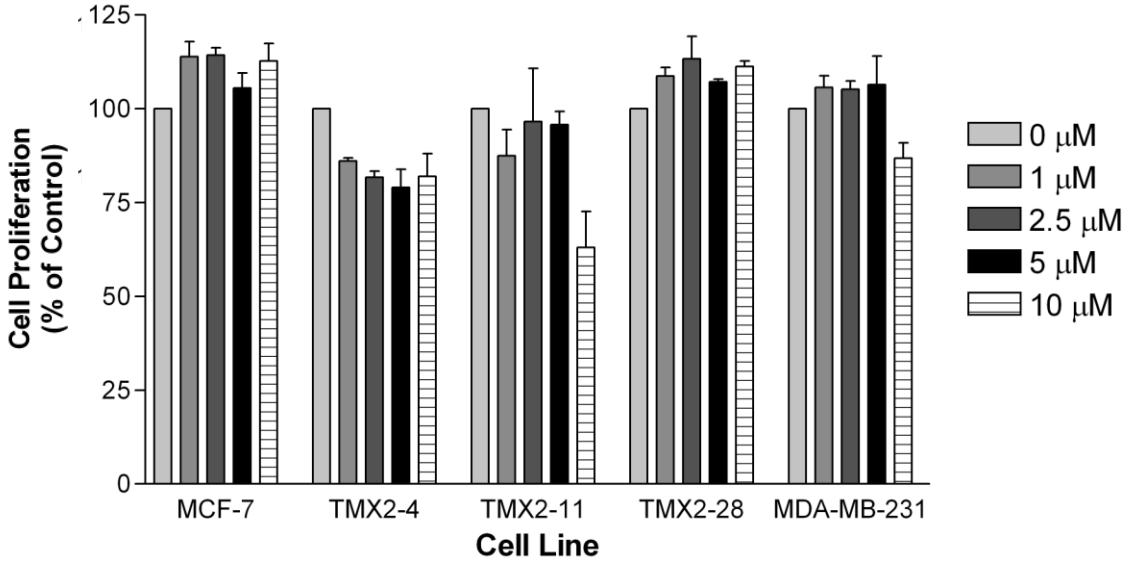
HM450BC METHYLATION OF THE ER GENE



Hypermethyation was present in 26 out of 38 CpG sites in the promoter region in TMX2-28 as compared with MCF-7 (blue boxes).

APPENDIX H

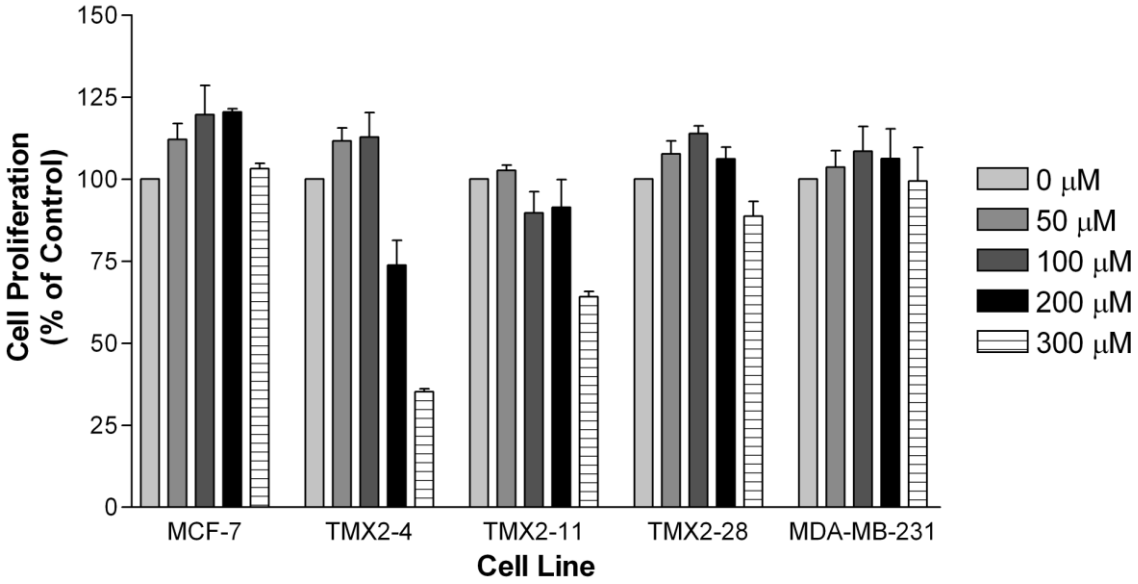
CYTOTOXCICITY ASSAY TO DETERMINE 5-AZA CONCENTRATION



Cells were treated in triplicate with 1, 2.5, 5, or 10 μM 5-Aza or a vehicle control (0 μM) for 4 days. TMX2-4 was most sensitive to the drug at low concentrations, with a 14 and 19% decrease in proliferation at 1 and 2.5 μM respectively. Consequently, a concentration of 2.5 μM was chosen for further experiments.

APPENDIX I

CYTOTOXCICITY ASSAY TO DETERMINE SAM CONCENTRATION



Cells were treated in triplicate with 50, 100, 200, or 300 μM SAM or a vehicle control (0 μM) for 6 days. TMX2-4 was most sensitive to the drug concentrations, with a 27 and 65% decrease in proliferation at 200 and 300 uM respectively. Therefore, concentration of 100 μM was chosen for further experiments.

APPENDIX J

INDIVIDUAL PATIENT CHARACTERISTICS

Patient	Group[1]	Age	Menopausal	Primary or Recurrence	Time to recurrence[2]	ER	PR	HER2	Ki67	CK5	CK14	p40	Laterality	Tumor size (mm)	Tumor Type[3]	Tumor Grade	Follow-up Treatment[6]			
																	Anti-estrogen[4]	Chemo[5]	Radiation	
A11A	A1	45	No	0		8	4	1	30	Neg	Neg	Neg	Left	.	IDC	.	.	AC+ pac	Yes	Matched pairs: ER+ to ER+
A11B	A2	46	No	1	12	8	6	0	30	Neg	Neg	Neg	Left	50	IDC	3				
A13A	A1	60	No	0		8	6	0	0	Neg	Neg	Neg	Right	.	DCIS	0	.	None	No	
A13B	A2	72	Yes	1	143	8	4	0	1	Neg	Neg	Neg	Right	12	DCIS	0				
A14A	A1	62	No	0		8	7	0	20	Neg	Neg	Neg	Left	10	ILC	1	Tam & AI	None	Yes	
A14B	A2	74	Yes	1	142	8	0	0	0	Neg	Neg	Neg	Left	30	ILC	3				
A15A	A1	59	Yes	0		8	8	0	1	Neg	Neg	Neg	Right	25	IDC	3	AI	AC+ pac	Yes	
A15B	A2	61	Yes	1	17	8	0	0	15	Neg	Neg	Neg	Left	.	DCIS	0				
A16A	A1	77	Yes	0		7	0	0	15	Neg	Neg	Neg	Right	22	IDLC	2	AI	None	No	
A16B	A2	80	Yes	1	45	8	4	0	25	Neg	Neg	Neg	Right	33	IDLC	3				
A17A	A1	77	Yes	0		8	7	0	1	Neg	Neg	Neg	Left	.	IDC	1	Tam & AI	None	No	
A17B	A2	78	Yes	1	17	8	7	0	1	Neg	Neg	Neg	Right	30	IDC	2				
A18A	A1	37	No	0		8	8	0	1	Neg	Neg	Neg	Left	12	IDC	2	AI	other	No	
A18B	A2	58	Yes	1	252	8	8	0	1	Neg	Neg	Neg	Left	15	IDC	2				
A19A	A1	43	No	0		8	4	0	0	Neg	Neg	Neg	Left	3	IDC	1	Tam & AI	None	No	
A19B	A2	49	No	1	70	8	6	0	2	Neg	Neg	Neg	Right	5	IDC*	3				
A1A	A1	47	No	0		8	8	0	0	Neg	Neg	Neg	Right	8	IDC	2	AI	other	No	
A1B	A2	59	Yes	1	148	8	6	1	3	Neg	Neg	Neg	Left	.	IDC*	.				

A20A	A1	59	Yes	0		8	3	0	0	Neg	Neg	Neg	Left	10	IDC	2	Tam & AI	None	No	
A20B	A2	68	No	1	103	8	0	0	0	Neg	Neg	Neg	Right	7	IDC	1				
A21A	A1	38	No	0		5	4	3	20	Neg	Neg	Neg	Left	15	IDC	3	Tam	None	No	
A21B	A2	40	No	1	21	4	4	3	20	Neg	Neg	Neg	Left	10	IDC	3				
A2A	A1	41	No	0		7	8	1	4	Neg	Neg	Neg	Left	15	IDC	2	Tam	AC+ pac	No	
A2B	A2	46	No	1	55	7	4	3	20	Neg	Neg	Neg	Left	13	IDC	3				
A7A	A1	84	Yes	0		8	8	0	3	Neg	Neg	Neg	Left	20	ILC	1	.	unkn	No	
A7B	A2	90	Yes	1	74	8	8	0	6	Neg	Neg	Neg	Left	22	ILC	2				
A8A	A1	66	Yes	0		8	7	0	0	Neg	Neg	Neg	Right	15	IDC	3	.	unkn	No	
A8B	A2	76	Yes	1	128	8	6	1	0	Neg	Neg	Neg	Left	55	ILC	1				
A8C		81	Yes	2	55	8	6	3	30	Neg	Neg	Neg	Right	35	ILC	3				
A9B		61	No	1	134	7	4	0	0	Neg	Neg	Neg	Left	.	IDC	3				Unmatched
A12B		39	No	1	38	8	7	0	20	Neg	Neg	Neg	.	.	IDC*	.				
A9C		62	No	2	21	0	0	0	40	Neg	Neg	Neg	Right	.	IDC	3				
A3B		65	Yes	1	52	8	6	0	4	Neg	Neg	Neg	Right	15	IDC	2				
A5B		72	Yes	1	76	8	5	1	7	Neg	Neg	Neg	Left	9	IDC	2				
B2A	B1	66	No	0		0	0	0	20	Pos	.	Occa	Left	15	IDC	3	No	AC+ doc	Yes	
B2B	B2	74	No	1	90	0	0	0	20	Pos	Pos	Neg	Left	10	IDC	3				
B3A	B1	46	No	0		0	0	0	2	Pos	.	Neg	Right	9	IDC	3	No	AC+ doc	No	
B3B	B2	48	No	1	37	0	0	0	25	Pos	Pos	Neg	Right	5	IDC	3				
B4A	B1	79	No	0		0	0	0	2	Neg	Neg	Neg	Left	38	IDC	3	No	None	Yes	
B4B	B2	80	No	1	10	0	0	0	30	Neg	Neg	Neg	Left	.	IDC	3				
B5A	B1	52	Yes	0		0	0	3	20	Neg	Neg	Neg	Right		IDC	3	No	AC+ doc	No	
B5B	B2	53	Yes	1	11	0	0	3	20	Neg	Neg	Neg	Right	.	IDC	3				
B7A	B1	64	Yes	0		0	0	0	25	Neg	Neg	Neg	Right	20	IDC	3	No	AC+ pac	Yes	
B7B	B2	66	Yes	1	22	0	0	0	35	Pos	Pos	PosF	Right	45	IDC	3				

Matched pairs: ER- to ER-

B4C		81	No	2	11	0	0	0	10	Neg	Neg	Neg	Left		IDC	3			Unmatched
B8A		52	No	0		0	0	0	35	Pos	Pos F	Occ a	Left	.	IDC	3	No	unkn	No
B9B		78	Yes	1	25	0	0	0	15	Neg	Neg	Neg	Right	.	IDC	3			
B1B		66	Yes	1	45	0	0	0	20	Neg	Pos	Neg	Left	.	IDC	.			
B5C		55	Yes	2	32	0	0	3	30	Neg	Neg	Neg	Right		IDC	3			
C2A	C1	51	No	0		8	0	3	15	Neg	Neg	Neg	Right	36	DCIS	0	Tam & AI	.	No
C2B	C2	53	No	1	17	0	0	3	2	Neg	Neg	Neg	Left	60	DCIS	0			
C3A	C1	47	No	0		8	8	0	10	Neg	Neg	Neg	Left	.	IDC	2	Tam	unkn	No
C3B	C2	65	Yes	1	216	0	0	0	30	Pos	Pos	Neg	Left	32	IDC	3			
C4A	C1	58	Yes	0		8	0	1	3	Neg	Neg	Neg	Left	30	IDC	2	.	unkn	No
C4B	C2	64	Yes	1	71	0	0	0	15	Pos	Pos	Pos	Right	20	IDC	3			
C5A	C1	65	No	0		8	0	3	5	Neg	Neg	Neg	Left	12	IDC	3	yes	other	No
C5B	C2	68	No	1	36	0	0	3	15	Neg	Neg	Neg	Right	19	IDC	3			
C6A	C1	42	No	0		7	7	0	0	Neg	Neg	Neg	Right	.	DCIS	0	yes	unkn	No
C6B	C2	55	No	1	132	0	0	0	35	Neg	Pos	Neg	Right	14	IDC	3			
C7A	C1	55	Yes	0		8	8	0	1	Neg	Neg	Neg	Left	26	IDLC	2	yes	unkn	No
C7B	C2	57	Yes	1	25	0	0	3	0	Neg	Neg	Neg	Right	43	DCIS	0			
C1B		68	Yes	1	60	0	0	3	10	Neg	Neg	Neg	Right	50	DCIS	2	Tam		
N1		46	No	0		8	8	0	2	Neg	Neg	Neg	Right	9.6	IDC	2	yes	AC+ pac	No
N10		69	Yes	0		0	0	0	50	Pos	Neg	Pos	Left	25	IDC	3	yes	other	No
N3		69	Yes	0		8	8	0	0	Neg	Neg	Neg	Right	10	ILC	2	AI	None	No
N4		44	No	0		8	0	0	0	Neg	Neg	Neg	Left	24	DCIS	0	Tam	None	No
N5		54	No	0		8	3	1	10	Neg	Neg	Neg	Right	4	ILC	1	AI	None	No
N6		51	No	0		8	3	3	3	Neg	Neg	Neg	Left	24	IDC	3	yes	AC+ pac	No
N7		55	No	0		8	0	0	1	Neg	Neg	Neg	Left	25	IDC	2	AI	other	Yes
N8		64	Yes	0		0	0	0	1	Neg	Neg	Neg	Right	25	DCIS	0	No	None	Yes

Matched pairs: ER+ to ER-

Nonrecurrents

[1]Group refers to the 3 sets of matched primary and recurrent tumors described in text

[2]Time to recurrence is in months

[3]IDC = invasive ductal carcinoma; DCIS = ductal carcinoma in situ; ILC = invasive lobular carcinoma; IDLC = invasive ductal & lobular carcinoma

[4]Tam = Tamoxifen; AI = aromatase inhibitor; yes = treatment received but type unknown

[5]AC = Adriamycin (doxorubicin) and cyclophosphamide; pac = paclitaxel; doc = docetaxel; other = one of several different combinations; unkn = treatment received but type unknown

[6]Follow-up treatment refers to treatment following the excision of the primary tumor

*Indicates lung metastases; Samples A1B, A12B, and A19B

Missing data are represented with a "."

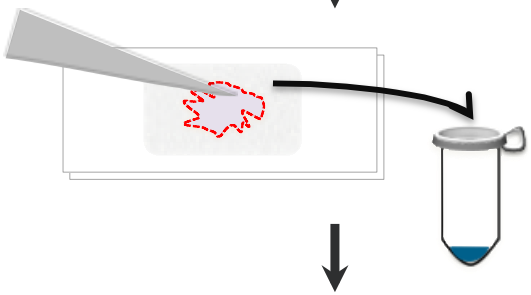
Not applicable are left blank

APPENDIX K

DIAGRAM OF TISSUE MICRODISSECTION FOR DNA PURIFICATION

5-10 µm thick FFPE tissue slice

Unstained breast tissue section on slide placed over H&E stained slide as guide for tumor



Carefully macrodissected tissue using a sterile needle and placed into microcentrifuge tube

Purified DNA using MoBio BiOstic FFPE DNA purification kit

Analyzed DNA for quality and concentration

Ethanol precipitation if needed

Sent DNA to USC for QC and HM450K analysis

APPENDIX L

35 HYPERMETHYLATED AND HYPOMETHYLATED CPG SITES AND ASSOCIATED GENES

ER-Positive Hypermethylated (23 CpG Loci)

TargetID	SYMBOL	PRODUCT
cg09068528	ACADL	acyl-Coenzyme A dehydrogenase; long chain precursor
cg00116234	ADAMTSL1	ADAMTS-like 1 isoform 2
cg00079563	ARFGAP3	ADP-ribosylation factor GTPase activating protein 3
cg08690031	CDCA7	cell division cycle associated protein 7 isoform 1
cg14998713	ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1
cg14228238	EVI1	ecotropic viral integration site 1
cg12998491	FAM78A	hypothetical protein LOC286336
cg00679738	FAM89A	hypothetical protein LOC375061
cg01657408	FLJ31951 (RNF145)	hypothetical protein LOC153830
cg18108623	FLJ34922 (SLFN11)	hypothetical protein FLJ34922
cg01561916	HAAO	3-hydroxyanthranilate 3;4-dioxygenase
cg01868782	HEY2	hairy/enhancer-of-split related with YRPW motif 2
cg12370791	HOXB9	homeo box B9
cg14188232	ITGA11	integrin; alpha 11 precursor
cg02755525	NETO2	neuropilin- and tolloid-like protein 2 precursor
cg14019317	PROX1	prospero-related homeobox 1
cg25336579	PSAT1	phosphoserine aminotransferase isoform 2
cg12717594	RECK	RECK protein precursor
cg06377278	RUNX3	runt-related transcription factor 3 isoform 2

cg15239123	SMOC1	secreted modular calcium-binding protein 1
cg00893242	SND1	staphylococcal nuclease domain containing 1
cg01186777	TNFSF9	tumor necrosis factor (ligand) superfamily; member 9
cg12874092	VIM	vimentin
ER-Negative Hypermethylated (12 CpG Loci)		
TargetID	**SYMBOL**	**PRODUCT**
cg08090772	ADHFE1	alcohol dehydrogenase; iron containing; 1
cg27650175	DAB2IP	DAB2 interacting protein isoform 1
cg14063008	DAB2IP	DAB2 interacting protein isoform 1
cg07981910	DAB2IP	DAB2 interacting protein isoform 2
cg05684891	DAB2IP	DAB2 interacting protein isoform 2
cg00720137	DYNLRB2	dynein; cytoplasmic; light polypeptide 2B
cg14874121	HSD17B4	estradiol 17 beta-dehydrogenase 4
cg07433344	HSD17B8	estradiol 17 beta-dehydrogenase 8
cg16541031	IRF7	interferon regulatory factor 7 isoform a
cg23178308	PDXK	Pyridoxal (pyridoxine, vitamin B6) kinase
cg17599586	PISD	phosphatidylserine decarboxylase
cg08108311	WNK4	WNK lysine deficient protein kinase 4

APPENDIX M

SHARED CPG SITES BETWEEN HYPERMETHYLATED AND HYPOMETHYLATED PATHWAYS

MAPINFO	MAPINFO	MAPINFO	MAPINFO	MAPINFO	MAPINFO
1000471	13889871	21814787	39969264	65494101	97507666
100075304	139017424	21814789	39969297	65601167	97657432
100075396	139144055	21814793	40204479	6565046	97657553
100259329	139208223	21835945	40254712	65670237	97657608
100259341	139283681	219646481	40428713	65670841	98279166
100318154	139396524	21984339	40428986	66049932	98480124
1004213	139538222	21985589	40825980	66135527	98620518
100441223	139539092	21985613	41086050	66135734	98620572
100784781	13957159	21985628	41086274	66189297	98850691
100881458	139589617	21985706	41086680	6622294	98850881
100882213	140261697	219857927	41086800	6622297	98881259
100882334	140306205	219867925	41131315	6622338	98945131
1008964	140306208	21988147	41267981	6622425	98945949
100913842	140306377	21988173	41268130	66597378	99283201
101111427	140345966	220025712	41511175	66878654	99357437
101380601	140346199	220025719	41624841	66878709	99382193
101419479	140346236	220406470	41786692	66878810	99439119
101419483	140346263	22122872	41818756	66878871	99439685
101570703	140346293	222436985	41818770	67048577	99470550
101618394	140346309	22263347	4198058	67134587	99470560
102090791	140346394	223175334	42139507	67350976	99470618
102090979	140346403	223900693	42445059	67351490	99531765
102107676	140381745	224903393	42543442	67360805	9957221
102158209	140566710	224903977	42633783	67410806	9957467
102414501	140566767	22518317	42982929	67410995	9957531
102568886	140700272	22539908	4298749	67837483	996280
102729375	140855547	226411247	43099350	67837689	
10276984	140855582	226736713	43142086	68165759	
102807849	140855857	226925181	43142093	68166241	
102809984	140856330	228029433	43153216	68260574	
102810101	141490067	228029452	43153223	68481049	
102810157	141490097	228225687	43188257	68481342	
102863004	141490153	228604037	43242911	68546467	
103046736	14183018	228604240	43428471	68546507	
103046930	14183064	228645045	4357579	68546579	

103046943	14183086	228645634	4357716	68547088	
103047053	14183212	228645688	43597565	688132	
103052468	142053720	228645710	43662979	68864324	
1031205	14212636	228645796	43702080	68864523	
10316358	142139002	228645859	43702277	68864529	
10316367	142494953	228646224	43864466	68864546	
103352267	142552914	228872121	43864628	68864549	
103352294	142607799	23038177	43902315	68870812	
103425389	142607939	230579435	43922174	68870892	
103600846	142682652	231762962	43971928	69240429	
103601188	142682682	23208349	43989949	69243285	
10382828	143060037	23217017	44040206	69243398	
10383164	143532735	232260305	44059003	69243486	
104153148	143743939	232395061	44258271	69243510	
104178949	143766916	232527265	44258572	69243752	
104383499	144533322	23261117	44319719	69326897	
104383715	144621971	233464194	44320173	69634372	
104383722	144629702	233464229	44331108	69634400	
104383724	1446802	233792732	44420276	69725117	
104383730	144680893	235813272	44450962	69725223	
104383798	145027948	23607951	44540794	69726397	
104584098	14516815	236227653	44763979	69760119	
10458667	145473631	237027592	44763993	69786306	
104609432	145566975	23712355	44764101	70033592	
10479966	145718405	23712364	44928967	70033624	
104850745	145878963	2375010	44928992	70033627	
104851909	145878979	2375420	44928995	70034475	
104902678	146257484	24130450	44947976	70034491	
10532886	146257862	241760025	45056008	7005733	
10534809	146258427	24229575	45056057	70313417	
105628017	1463554	242743214	45056325	70351524	
105628022	1463620	242743588	45056646	70415514	
105628027	14643925	24358236	45065052	7041565	
105663093	146614298	24358327	45097940	7041599	
105760444	146614596	24360156	45128769	7041930	
105760890	146833171	24360391	45128775	70655862	
10591633	1475742	24449668	45241222	70745466	
106331160	148033708	244893503	45331082	70745540	
106331166	148651359	2461278	45403493	70745601	
106532491	148651363	246785611	45444623	70746654	
106533997	148663426	247611502	45444767	70758965	
107005390	148663571	248020692	45496476	71122946	
107019166	148663583	248020697	45496482	71123004	

107282483	148663829	24837564	45613825	71123154	
107502023	148664322	24857719	45614290	71123196	
107713049	148664420	25019044	45614300	71123290	
107713740	148881226	25019747	45614720	71264579	
107777602	149068864	25219897	45614848	71264678	
107798913	149068974	25219914	45662262	71404443	
107799197	14924712	25219961	45670865	71497113	
107799660	14925198	25219976	45670875	71640369	
10812817	149632705	25241427	45695643	71666682	
1083002	150186289	25241462	45696455	71679959	
1083491	150211855	25242031	45696465	72052924	
108523056	150390267	25439375	4580777	72123958	
108523409	150390274	25464008	45960367	72141625	
108924366	150464124	25465219	45960834	72352968	
108924560	150464265	25465242	46317036	72388282	
10895950	150497346	25559743	46354562	72416469	
10895978	150497496	25621328	46354565	72416677	
109235928	150497663	25703528	46354630	72459953	
109294168	150497669	25864083	46414355	72596557	
109648208	150497703	25935223	46438743	72612125	
109747333	150497740	25944514	46477273	73030356	
109964033	150674972	26108391	46477288	73099917	
109964223	150674980	26108399	46507705	73145006	
109964581	150675025	26108401	46641708	73149356	
110009282	150675859	26108410	46655387	73331405	
110223967	151479894	26108412	46655394	73435229	
110228269	1523917	26172302	46655561	73449519	
110453107	152850698	26223310	46655580	73676012	
110454063	153601123	26284938	46655588	73706527	
110581872	153652247	26370552	46655790	73893202	
110582247	15413437	26370717	46669221	73972820	
110582420	15413461	2652640	4667753	73973008	
110583036	154335014	2659081	46703422	74005557	
110583377	154681128	2659305	46719761	74005562	
110583599	154797867	26625273	46742865	74005566	
11059577	154797892	26625307	46800467	74044539	
11059677	154797917	26625359	46825840	74045075	
110703888	155247706	26625766	46924050	74236251	
110704080	155664311	26626134	46932308	74236288	
110704116	155665249	2674285	46992909	74347435	
110754087	156433108	268923	47133964	74533976	
110754101	156433300	268950	4715520	74534090	
110958892	156433317	268976	47301741	74725980	

110958977	156433350	27012615	47307743	74726136	
110959356	156433367	27044629	47394431	74868604	
110959647	156572731	27044685	47468154	74868670	
110960177	157180005	27045048	47518747	74868725	
110960186	157261106	27053221	47576957	74892744	
110960198	15761854	27053247	47653309	74951923	
111162296	158141570	27053250	47691301	74988951	
111217406	158365160	27053559	4771201	75139139	
111472183	158449803	27112287	4775638	75139347	
11148769	15847977	27112305	48046859	75139364	
11148932	15848067	2721819	48075568	75139470	
11149068	15848072	27224092	48086162	75139482	
111637200	15848253	27238279	48086580	75139490	
111730545	15848264	27239728	48206749	751830	
111840957	159200676	27280423	48206884	751833	
111981687	160761085	2738746	48397225	752149	
112531653	160973736	27389879	48398535	752180	
112533260	16179633	27411030	48398730	75368902	
112629604	16179660	27455004	48462306	75369055	
112832547	16180048	27717801	48462862	75369224	
112930675	162273011	27723409	48494536	75379166	
112931182	162930177	27763865	48503057	75480246	
112931194	162930188	27782628	48503163	75528813	
11349023	162930325	27782658	4868983	7554681	
114176105	162930666	27799380	4870012	75601276	
114178220	162930671	28032282	48988066	75612223	
114180814	16310359	28034797	49372180	75629039	
11430447	16375404	28034816	49378032	7566656	
114695695	1639534	28199109	49427684	75670903	
114898409	164264853	28199456	49521109	7568539	
114898440	164265004	28219613	495736	75787820	
114918693	164265012	28241425	495817	75788038	
114937784	16555379	28351906	49737236	75788212	
114938002	166039444	28352098	49812963	75788231	
115132761	166421992	28366600	49888944	75788287	
115377939	166795323	28396039	49926070	75831210	
115530150	166809971	28492265	49962507	75896720	
115630748	167350883	28503060	49962509	75896801	
115803805	167599527	2861914	49962609	75905604	
115918745	167599631	2864736	49962729	75955844	
11592958	167956581	289390	50134915	76189770	
116163765	168149401	28983117	5018805	76381687	
116381660	168149519	28983120	50194228	76381937	

116383634	1685555	28983122	5026393	76382149	
11653600	16871106	28995458	5026543	76476158	
116997308	168728076	28998195	50354998	76510579	
117317515	168728081	28998214	50355307	76633627	
117513042	168728213	29106422	50496857	76838549	
117513101	168728270	29125104	50608218	76838560	
11752089	16884364	29302016	50638648	76976057	
117537264	16884576	29336262	50717667	77160030	
117586817	16926335	29338258	50791419	77168916	
117798954	16926680	29338432	50810682	77246968	
11783294	169387139	29394775	50817213	77270312	
118030848	169940192	29450584	50828905	77459677	
118030970	170600562	29451235	50832651	77460719	
11810173	170646566	29711918	5085586	77607222	
118603742	17137306	29711922	5085597	77719326	
118603880	171572714	29711941	5085713	77719408	
118603908	172359575	29716597	50893226	77734206	
118603938	172543905	29716601	50969997	77736904	
118753705	17281327	29794657	50970086	77765305	
118899024	17282284	29855347	50982818	77776587	
118960340	17282333	29855435	51147769	78004303	
119040090	17398264	29894737	51221603	78079801	
11919376	17399405	29894903	5140001	78080068	
119227631	174219209	29894986	51568336	78080193	
119234736	174871289	29895037	51842336	78237432	
119274192	175084743	29911036	51869353	78272372	
119298355	1753623	29967531	51869589	78272378	
119298513	1753636	30018720	51869672	78272393	
119419786	175547056	30116489	51869678	78423530	
119610526	175547399	30138339	51869680	78449647	
11969958	175792176	30144547	51869687	78450357	
11970145	175792510	30206046	5207312	7851280	
119769268	175792973	30227883	52074483	78549324	
119916017	175793049	30431699	52196506	78913147	
1202468	176047485	30449081	52241186	78933407	
12038718	176986856	30484033	52419380	789928	
120514968	176987174	30606026	52419483	789930	
120628572	17699773	30670159	52456361	79151188	
120628874	177004604	30890555	52499571	79151719	
1206682	177099102	30890583	52511853	79152102	
1206693	17726965	30938073	52626476	79152112	
1206703	178017571	3109053	52626524	79234251	
1206739	178017578	31126267	52626668	79318146	

1206758	178257266	31158158	53013726	79318875	
120806383	178257292	31170822	53350498	79373381	
120969079	178257317	31193985	53564	79578343	
121412970	178257407	31360817	54155663	79724519	
121413143	178257585	31483285	54156330	79739863	
12156718	178257590	31483774	54164310	79740185	
121783544	178313656	31609347	54402704	80055067	
121993049	178450763	31609351	54411773	80186763	
121993098	179660224	31619231	54422775	80530701	
121993378	179754483	31620378	54423807	80531597	
121993696	179754521	31684836	54477	80656888	
122399344	179754529	31831428	54618544	80696445	
122854796	179754603	31831553	54671440	8103101	
122872351	179754615	31831803	54675137	81123613	
122872836	180017349	31833013	54675154	8112956	
122872838	18011922	31833016	54912286	8113573	
122872996	180123317	31835534	54966247	81187125	
122873018	180479829	31835577	55033982	81187198	
12290322	180631909	3202077	55092884	81187610	
123167276	180725907	32110523	55102849	819227	
123167507	180726035	32111001	55366610	82116365	
123167522	180726249	32226684	55370423	82116392	
123167770	180726252	3227741	5543548	82116456	
12329223	180726328	3227751	5544099	82116571	
12329242	180726349	3227981	5544169	82116596	
12329745	180979631	3232246	55443757	82167764	
12329826	18208422	32433833	55524333	82167774	
12330263	182521880	32456910	55672036	82192663	
12330332	183066337	32456912	55690381	82193449	
123379308	183155573	32457158	55865288	82193460	
123673331	18319362	32460799	55991782	82193704	
123748753	183993995	32485396	56224627	82444798	
123753198	18486241	32489203	5629371	82965088	
124124875	18487771	32489555	5629683	82965868	
124529273	185827293	32526065	5630016	83377929	
124639215	186648279	32639011	5630348	83478878	
124639221	186648285	32714023	563891	8367425	
124738871	186648544	32714038	56536599	83679643	
124782817	187025886	32907636	56609274	83776269	
124791197	187025897	32907639	56623215	83776271	
124890314	187065831	33037792	56623312	84002370	
125259369	187066279	330732	56623390	84002599	
125284203	187092	3310268	56666334	841436	

125284241	187476599	33173024	56677335	8429439	
125284245	1875011	33183713	5681323	84419202	
125284359	1876131	33216697	56833197	84419242	
125932073	18900772	33245114	56879554	84651521	
125986472	189203962	33245457	56879559	84651816	
126006749	189838686	33245471	56879662	85358477	
126069089	190539103	33351568	56915347	85430336	
126069645	191045309	33397679	56998622	85640810	
126240239	191045668	33486937	57016682	85954316	
126625864	19188414	33487058	5713100	85954331	
126626298	19189456	33590495	57194562	85954336	
126626364	19191417	33701353	57198533	860925	
126626557	192126742	33792148	57250780	86300605	
126873139	192127457	34392203	57365116	86383236	
127439539	192232077	34405681	57463991	86383252	
127440000	19281019	34442360	57836242	86383300	
127644075	19281140	34442377	57978892	86539022	
127807500	19281185	34460557	58213109	86543092	
127837931	19281197	34625543	58227267	86543519	
12809014	194208441	34629400	58358384	86658385	
12809420	19437157	34645109	58514497	8699462	
128152075	19483394	34657676	58520847	87105216	
12838028	19562810	34894648	58545728	87521180	
12838136	1958851	35395837	58609473	88070661	
1283875	1958883	35517516	58609730	88071195	
128431947	19617439	356222	58609744	8810139	
128433373	196366451	35639492	58609764	8810980	
128458240	19668147	35639632	58609770	88295472	
128458281	196729632	35639653	58619169	88324169	
128470599	1967781	35639656	5890411	88324597	
128530306	1968232	35772996	5890430	88324879	
128720884	19706186	35772999	59104825	88717134	
128796079	19706365	3585406	59189791	88731631	
128796097	19748910	3585609	59216714	88752056	
129685695	197808017	3606853	59477172	88793252	
130339733	19812592	36389575	59990120	88875836	
13043878	19970334	36531652	60383672	88875840	
130645645	19988706	36591112	6063654	88928467	
130646497	200323147	36788387	6069929	8906601	
13124215	200323768	36909413	60779652	8907213	
1312475	20044428	36920332	60780415	89080179	
13137129	20085047	36985996	60795818	89080191	
131513927	200992656	36986513	61051317	89163804	

131593259	201450560	36988217	61051341	8925752	
1316815	201450575	36989454	61051348	89323758	
131800430	201450601	37005570	61148894	89324033	
13201456	201857621	37006107	61149009	89346205	
13201650	20255992	37006611	61188431	89378460	
13201655	202611730	37096148	6121909	89748238	
13201662	202679456	37096487	6122107	89911298	
13201844	202679504	37288705	61276447	90035945	
13202437	202899877	3730151	61276678	90039805	
13202476	20346426	37366294	61335254	90039809	
132052870	20349241	37403242	61468757	90114063	
132052887	20349568	37434260	61553938	90228815	
132262353	204797566	37434950	61553954	90343174	
132312702	205424985	3751076	61554106	90343192	
132329846	205538293	3751199	61583910	90343204	
132657819	205538310	3751618	61583924	90343208	
132722045	205538427	3751739	61583979	90343220	
132722421	20626133	37764118	61584072	90455799	
132904689	20690628	37783689	61584108	9049273	
132906443	20690807	37783879	61615739	90527978	
133066054	20691126	37839684	61615913	90895083	
133110244	20691161	37839769	61703735	90967113	
133320728	20691180	37839829	61777090	90967157	
133492476	20691429	37839918	62185401	90967165	
133493171	20711985	37839923	6243880	90967894	
133493200	207308829	37839943	62477362	91643544	
133758089	20768389	37840438	62477480	91670134	
133758161	207817630	37956425	62732892	91870067	
133890532	20792243	37956434	62784802	92053021	
134067789	20792323	38071166	6280583	92053063	
134125573	20792476	38071301	6281104	920786	
134201505	20806543	38071498	6281197	92417998	
134261413	20806663	38258441	62948235	92937735	
134282207	208084415	38258451	63263824	93389246	
134282404	208084436	38258884	63285425	93520269	
134282864	208084456	38334166	63422640	93520275	
134285972	20834843	38334176	63429038	93551004	
134755679	209848535	38670804	63461803	93583253	
134755955	209921370	38670840	63506148	93583693	
134871962	21084420	38670985	637035	93616072	
134871966	21092936	38691791	637162	93966978	
134871973	21121564	38691799	637173	94136	
134901294	212780243	38691801	637175	94449797	

134974048	21280288	38794831	63792210	94473671	
134974192	2137067	3881141	63802272	94501467	
13504687	214148856	38879544	63802280	95225681	
135050702	214148892	38879699	63802491	95361208	
135090367	214148958	39196998	64188108	95361223	
135150497	214148976	39197233	64188163	95402189	
135150517	214161251	39197256	64188166	95570657	
1354522	216100	39197260	64240654	95620306	
13546145	21634945	39197417	64241013	95653942	
135476297	2163608	39197424	6439864	95691647	
135476893	21646662	39261432	6440065	95947146	
136114344	21649722	393239	64512188	96012592	
13617012	21650028	39466575	64513068	97055150	
137243382	21656374	39853885	64513156	97172480	
137244090	2167496	39893185	64839958	97173034	
137522960	217559020	39893261	65127973	97203640	
13839913	217559131	39893438	65128112	9745295	
138666040	217559207	39893520	65183745	97505764	
138666050	21769430	39957298	6546777	97506251	

APPENDIX N

SHARED GENES BETWEEN HYPERMETHYLATED AND HYPOMETHYLATED PATHWAYS

Gene	Gene	Gene	Gene	Gene	Gene
MARCH11	CNST	GRID2IP	MIR548N	RASL10A	UNC13A
SEPT5	CNTN1	GRIK5	MIR9-3	RASSF2	UNC5A
A2BP1	CNTNAP2	GRIN2A	MIXL1	RAX	UNC5C
AASS	CNTNAP5	GRIN2C	MKX	RBM20	UNCX
ABCB1	COBL	GRIN2D	MME	RBM24	UQCRC1
ABCB4	COCH	GRIN3B	MMP14	RBMS1	USP2
ABCG2	COL11A2	GSC2	MMP15	RBP4	USP4
ABCG5	COL15A1	GSG1L	MMP17	RCSD1	UST
ABLIM3	COL18A1	GSTCD	MMP9	RELL2	UVRAG
ABR	COL23A1	GSTP1	MNX1	REM2	VASH1
ABRA	COL25A1	GSX1	MOBKL2B	RFPL2	VAX1
ACADM	COL2A1	GSX2	MOXD1	RFTN1	VCAN
ACADS	COL4A1	H2AFJ	MPO	RFX4	VENTX
ACAN	COL4A2	HAAO	MPPED2	RFX7	VEPH1
ACCN1	COL4A3	HAPLN3	MRC2	RFX8	VGF
ACCN4	COL4A4	HAPLN4	MRGPRF	RGL2	VGLL2
ACE	COL5A1	HAS2	MRPL18	RGMA	VIM
ACHE	COL5A3	HAS2AS	MSI1	RGNEF	VIPR2
ACOXL	COL6A2	HAS3	MSRA	RGS14	VPS37B
ACSF2	COL7A1	HBM	MSX1	RGS17	VSIG10L
ACTN2	COL8A1	HCCA2	MT1DP	RGS20	VSTM2B
ADAM11	COL9A1	HCG4	MT1M	RGS7BP	VSTM2L
ADAM23	COL9A2	HCG4P6	MT2A	RHCG	WBSCR26
ADAM8	COLEC12	HCN1	MT3	RHOU	WDR17
ADAMTS1	COMP	HCN2	MTNR1A	RIBC1	WDR86
ADAMTS10	CORO2B	HCN3	MTP18	RIMS1	WFDC10A
ADAMTS16	COTL1	HDAC3	MUL1	RIMS2	WFDC3
ADAMTS17	COX19	HECW1	MUM1	RIMS3	WFDC9
ADAMTS19	COX6B2	HECW2	MXI1	RIPK4	WHAMM
ADAMTS2	CPAMD8	HELT	MXRA7	RNASEH2A	WHAMML1
ADAMTS3	CPEB1	HEPACAM	MYCBPAP	RNF123	WHAMML2
ADAMTS7	CPLX1	HERC5	MYEF2	RNF150	WIPF1
ADAMTS9	CPLX2	HES5	MYO1C	RNF157	WIT1
ADARB2	CPM	HEY2	MYO3A	RNF175	WNK2
ADCY1	CPNE2	HFM1	MYO7A	RNF180	WNT1
ADCY3	CPNE8	HGSNAT	MYOD1	RNF182	WNT16
ADCY5	CPT1C	HHEX	NACAD	RNF213	WNT2
ADCY8	CR1L	HHIP	NAV1	RNF219	WNT3
ADORA1	CREB3L1	HHIPL1	NAV2	RNF32	WNT3A
ADORA2B	CREG2	HIC1	NBEA	RNLS	WNT9B
ADPRH	CRHBP	HIF3A	NBL1	ROBO3	WRB
ADRA2C	CRHR1	HIST1H2AE	NCAM1	ROR1	WRN
ADRB1	CRHR2	HIST1H2AJ	NCAN	ROR2	WSCD2
AFAP1L1	CRISPLD1	HIST1H2BG	NCOA7	RORA	WT1
AGAP1	CRMP1	HIST1H2BM	NCRNA00171	RPA2	XKR5

AGAP11	CRTAC1	HIST1H4K	NDFIP2	RPGRIP1L	XKR6
AGPAT9	CRYBA2	HIST3H2A	NDRG4	RPL31	XKR7
AGPS	CRYM	HIST3H2BB	NECAB2	RPP25	ZAK
AGTR1	CSF1	HK2	NEDD9	RPRM	ZBTB17
AHDC1	CSMD2	HLA-A	NEK10	RPRML	ZBTB7B
AJAP1	CSMD3	HLA-C	NELL1	RPS4X	ZC3H12C
AK5	CTHRC1	HLA-DQB1	NEU1	RPTOR	ZCWPW1
AKAP12	CTNNA2	HLA-DRB5	NEURL3	RREB1	ZDHHC22
AKR1B1	CTNND2	HLA-F	NEUROD2	RSPH9	ZEB1
ALDH1A2	CTSC	HLA-G	NEUROG1	RSPO1	ZFP82
ALDH1A3	CTTNBP2	HLA-H	NFASC	RSPO3	ZFPM2
ALDH1L1	CUGBP2	HLA-J	NFATC1	RSPO4	ZIC2
ALDH7A1	CUX1	HLA-L	NFATC2	RTBDN	ZIC4
ALK	CUX2	HLX	NFATC4	RTN1	ZIC5
ALKBH3	CWH43	HMGCLL1	NFIC	RTN4R	ZNF134
ALOX12B	CYBA	HNF1B	NFIX	RUNX2	ZNF154
ALPL	CYGB	HOXA10	NFKBID	RXFP3	ZNF175
ALX4	CYP1B1	HOXA11	NGB	RYR3	ZNF214
AMOTL1	CYP26B1	HOXA11AS	NGEF	S100A6	ZNF215
AMPH	CYP27A1	HOXA13	NGFR	SALL1	ZNF217
ANGPTL4	CYP2U1	HOXA7	NHEJ1	SALL3	ZNF22
ANGPTL6	CYP4V2	HOXB13	NHLH2	SAMD11	ZNF232
ANK1	CYP7B1	HOXB3	NICN1	SAMD14	ZNF233
ANKLE1	CYS1	HOXB4	NID1	SAMD4A	ZNF268
ANKRD19	CYTSB	HOXC4	NID2	SAMD5	ZNF287
ANKRD24	D4S234E	HOXC8	NKAIN2	SASH1	ZNF311
ANKRD26	DAB1	HPCA	NKAIN4	SATB2	ZNF334
ANKRD34A	DACT1	HPSE	NKD2	SBF2	ZNF37A
ANKS1A	DACT2	HPSE2	NKIRAS2	SC65	ZNF382
ANO1	DBC1	HR	NKX1-2	SCAMP5	ZNF385B
ANO5	DBX2	HRH2	NKX2-1	SCARF2	ZNF385D
ANO8	DCBLD2	HRH3	NKX2-3	SCD	ZNF391
ANTXR1	DCDC2	HRK	NKX2-5	SCGB3A1	ZNF454
AOX1	DCHS2	HRNBP3	NKX2-6	SCGN	ZNF501
AP3B2	DCLK1	HS3ST1	NKX2-8	SCHIP1	ZNF513
APBB1	DDAH2	HS3ST2	NKX3-2	SCN3B	ZNF518B
APC2	DDB2	HS3ST3A1	NKX6-1	SCN4B	ZNF529
APCDD1L	DDIT4L	HS3ST3B1	NLGN1	SCRN1	ZNF542
AQP5	DDX28	HS3ST4	NLRP14	SCRT1	ZNF606
ARHGAP20	DENND3	HS3ST6	NLRP3	SCRT2	ZNF615
ARHGAP22	DEPDC7	HSD11B1L	NLRX1	SDC2	ZNF660
ARHGAP24	DFFA	HSD17B12	NMBR	SDK1	ZNF727
ARHGDIG	DGAT2	HSD17B8	NME5	SDK2	ZNF777
ARHGEF10	DGKE	HSPA2	NMNAT2	SEC14L5	ZNF836
ARHGEF17	DGKG	HSPG2	NMNAT3	SEC23B	ZNF879
ARHGEF4	DGKZ	HTR2A	NMUR1	SEL1L3	ZSCAN1
ARHGEF7	DIRAS1	HTR4	NOG	SELO	ZSCAN12
ARID5A	DISC1	HTR7	NOL4	SELV	ZSCAN18
ARL10	DKK1	HTRA3	NOS1	SEMA3B	
ARMC3	DKK2	HUNK	NOTUM	SEMA5B	
ARNT2	DLC1	HUS1	NOVA2	SEMA7A	
ARPC1B	DLL1	HYDIN	NPAS2	SENP2	
ARVCF	DLL3	ICAM5	NPAS4	4-Sep	
ASAH2B	DLX4	ICOSLG	NPPB	9-Sep	
ASAM	DLX6AS	IDUA	NPPC	SERGEF	
ASB18	DMPK	IGF2	NPR1	SERP2	

ASCL1	DMRT1	IGF2AS	NPTX1	SERPINA10	
ASS1	DMRT2	IGF2BP1	NPY5R	SERPINE2	
ATCAY	DNAJB13	IGF2R	NQO1	SFRP1	
ATF3	DNAJB6	IGFBP3	NR2E1	SFRP4	
ATP10A	DNAJC15	IGFBP5	NRG2	SFRP5	
ATP11A	DNAJC17	IGLON5	NRG3	SFRS8	
ATP1B2	DNAJC6	IGSF11	NRXN2	SFT2D3	
ATP5G1	DNER	IGSF9B	NTF3	SGCE	
ATP8A2	DNTTIP1	IL17RC	NTM	SH3BP2	
ATP8B2	DOC2A	IL17RD	NTN1	SH3PXD2B	
ATP9A	DOCK10	IL17RE	NTRK2	SHANK1	
AUTS2	DOT1L	IMMP2L	NTRK3	SHISA2	
AVP	DPF1	INPP4B	NTSR2	SHISA4	
B3GAT1	DPF3	INPP5A	NUAK1	SHISA7	
B3GAT2	DPP10	INS-IGF2	NUDT19	SHROOM3	
B3GNT7	DPP4	INSM1	NXN	SIGIRR	
B4GALNT1	DPP6	INSM2	NXPH1	SIM2	
B4GALNT3	DPY19L2P4	INTS12	NXPH2	SIRPA	
B4GALT4	DPYD	IQSEC2	NXPH3	SIX1	
BAALC	DPYSL3	IQSEC3	OCA2	SIX4	
BACH2	DRD1	IRF4	ODZ4	SKAP1	
BAHCC1	DRD4	IRF8	OGDHL	SKAP2	
BAI1	DSCAML1	IRX2	OLFM2	SLAIN1	
BAI2	DSEL	IRX4	OLIG1	SLC12A5	
BAIAP2	DST	ISL1	ONECUT1	SLC15A1	
BARHL2	DUOX2	ISL2	ONECUT2	SLC16A12	
BARX2	DUOXA2	ISLR2	ONECUT3	SLC16A3	
BASP1	DUS2L	ISM1	OPCML	SLC16A9	
BAT2	DUSP6	ISM2	OPRK1	SLC17A9	
BCAS3	DYDC1	ITGA11	OPRM1	SLC18A3	
BCAT1	DYDC2	ITGA8	OSBPL3	SLC1A2	
BCL11A	DYNC1I1	ITGB3	OSR2	SLC27A6	
BCL11B	DYNC1I2	ITIH5	OTOP1	SLC2A4	
BDNF	DYSF	ITPKA	OTX1	SLC2A9	
BEGAIN	DZIP1	ITPKB	OXR1	SLC30A3	
BEND6	EBF3	JAKMIP1	OXTR	SLC32A1	
BEST3	EBF4	JAM2	P4HA3	SLC34A1	
BHLHE22	ECE2	JAZF1	PAK7	SLC35D3	
BMP3	ECEL1	JPH2	PALLD	SLC35F1	
BMP4	EDC3	JPH3	PANX2	SLC37A1	
BMP7	EDIL3	KAAG1	PAPLN	SLC39A3	
BMP8A	EDN1	KAL1	PAPPA	SLC43A2	
BMP8B	EDNRB	KANK4	PAQR8	SLC43A3	
BMPER	EEF1D	KATNB1	PAQR9	SLC44A4	
BNC2	EFHA2	KBTBD11	PARD3	SLC47A1	
BOC	EFNA2	KBTBD8	PARD6G	SLC4A4	
BOLA2	EFNA5	KCNA1	PARP14	SLC4A8	
BOLA2B	EGFLAM	KCNA3	PARP8	SLC6A12	
BRD2	EIF3E	KCNA6	PAX1	SLC6A15	
BRDT	ELAC1	KCNB1	PAX2	SLC6A18	
BRSK2	ELAVL3	KCNB2	PAX3	SLC6A2	
BRUNOL4	ELMO1	KCNC3	PAX5	SLC6A5	
BRUNOL6	ELMOD1	KCNC4	PAX6	SLC7A14	
BSCL2	ELOVL4	KCND3	PAX7	SLC8A3	
BSN	ELP2P	KCNG1	PAX9	SLC9A3	
BTBD11	EMB	KCNH2	PCBP1	SLC9A9	

BTBD17	EMID2	KCNH6	PCDH21	SLCO5A1	
BTNL9	EML1	KCNH8	PCDHA1	SLFN11	
C10orf107	EMX1	KCNIP2	PCDHA10	SLIT1	
C10orf140	EN1	KCNIP3	PCDHA11	SLIT2	
C10orf25	EN2	KCNJ10	PCDHA12	SLIT3	
C10orf46	ENTPD3	KCNJ12	PCDHA13	SLITRK1	
C10orf55	EOMES	KCNJ2	PCDHA2	SLITRK5	
C10orf58	EPB41L3	KCNJ6	PCDHA3	SMARCD3	
C10orf90	EPHA10	KCNK10	PCDHA4	SMC1A	
C10orf91	EPHA3	KCNK13	PCDHA5	SMCP	
C10orf93	EPHA4	KCNK3	PCDHA6	SMO	
C11orf87	EPHA6	KCNK5	PCDHA7	SMPD3	
C11orf90	EPHA8	KCNMA1	PCDHA8	SNAP25	
C11orf92	EPHB1	KCNMB4	PCDHA9	SNAP91	
C11orf93	EPHB2	KCNN1	PCDHAC1	SNCB	
C12orf56	EPHB6	KCNN3	PCDHAC2	SNHG6	
C13orf15	EPO	KCNQ1	PCDHB9	SNTG2	
C13orf27	ERGIC1	KCNQ2	PCDHGA1	SNX22	
C13orf36	ESRRG	KCNQ3	PCDHGA10	SNX32	
C13orf38	ETFB	KCNQ4	PCDHGA11	SOCS1	
C14orf37	ETV1	KCNQ5	PCDHGA12	SOCS2	
C15orf38	ETV5	KCNS2	PCDHGA2	SORBS3	
C15orf59	EVC	KCNT1	PCDHGA3	SORCS1	
C17orf102	EVX2	KCTD1	PCDHGA4	SORCS2	
C17orf51	EXD1	KCTD10	PCDHGA5	SORCS3	
C17orf62	F7	KCTD12	PCDHGA6	SOX1	
C19orf30	FAAH2	KCTD14	PCDHGA7	SOX11	
C19orf44	FABP5	KCTD19	PCDHGA8	SOX17	
C19orf51	FAM114A2	KCTD8	PCDHGA9	SOX18	
C19orf70	FAM123C	KDM2A	PCDHGB1	SOX2OT	
C19orf76	FAM131B	KDM2B	PCDHGB2	SOX8	
C1QL1	FAM135A	KDR	PCDHGB3	SP5	
C1QL2	FAM149A	KHDC1	PCDHGB4	SP9	
C1QTNF1	FAM150B	KIAA0922	PCDHGB5	SPAG16	
C1orf115	FAM163A	KIAA1024	PCDHGB6	SPATA13	
C1orf173	FAM164A	KIAA1026	PCDHGB7	SPATA18	
C1orf190	FAM167B	KIAA1217	PCDHGC3	SPDYA	
C1orf38	FAM171A1	KIAA1239	PCOLCE2	SPG20	
C1orf70	FAM172A	KIAA1462	PCSK2	SPHK1	
C1orf86	FAM176A	KIAA1530	PDE1C	SPOCK1	
C1orf92	FAM176B	KIAA1755	PDE3B	SPOP	
C1orf94	FAM181B	KIAA1804	PDE4A	SPRED3	
C1orf95	FAM188B	KIAA1826	PDE4C	SPSB4	
C20orf103	FAM189A1	KIF1A	PDE4D	SRC	
C20orf166	FAM19A5	KIF21B	PDGFRA	SREBF1	
C20orf195	FAM20A	KIF5C	PDIA6	SRF	
C20orf200	FAM24B	KIFC3	PDK4	SRRM3	
C21orf29	FAM38B	KISS1R	PDLIM3	SRRM4	
C2orf43	FAM43A	KIT	PDLIM4	SS18L1	
C2orf88	FAM53A	KL	PDX1	SSTR1	
C2orf89	FAM60A	KLC1	PDZRN4	ST3GAL2	
C3orf55	FAM66C	KLC2	PEG10	ST5	
C3orf72	FAM69B	KLHL14	PEX14	ST6GAL1	
C4orf31	FAM69C	KLHL29	PEX5L	ST6GAL2	
C5orf38	FAM78B	KLK1	PFKFB3	ST8SIA1	
C5orf49	FAM82B	KLK13	PFKP	ST8SIA2	

C6orf145	FAM89A	KNDC1	PGCP	ST8SIA4	
C6orf174	FAR1	KRR1	PGF	STARD9	
C6orf176	FBLIM1	KY	PHACTR1	STC1	
C7orf13	FBLL1	L3MBTL3	PHF21B	STK32A	
C7orf31	FBLN5	LAMC2	PHLDA2	STK32B	
C7orf57	FBRSL1	LATS2	PHLDB3	STL	
C7orf58	FBXL15	LBH	PHLPP1	STMN2	
C8orf34	FBXL16	LBX2	PHOSPHO1	STMN3	
C8orf42	FBXO17	LBXCOR1	PHOX2B	SULF2	
C8orf45	FBXO32	LCLAT1	PI4K2A	SULT1A3	
C8orf56	FDFT1	LDHD	PIAS1	SULT1A4	
C8orf73	FER1L6	LEPR	PID1	SULT4A1	
C8orf79	FES	LEPREL1	PIK3AP1	SV2B	
C8orf84	FEZF2	LGALS3BP	PIK3CD	SVIL	
C9orf142	FGF12	LGR4	PITX2	SYCE1L	
CA10	FGF14	LGR6	PITX3	SYF2	
CA4	FGF18	LHB	PLA2G7	SYN2	
CA7	FGF19	LHFP	PLAU	SYN3	
CABP1	FGF2	LHFPL5	PLCB1	SYNE1	
CABP7	FGF3	LHX1	PLCD1	SYNGAP1	
CACHD1	FGF4	LHX2	PLCL2	SYNJ2	
CACNA1A	FGF5	LHX5	PLD5	SYNPR	
CACNA1B	FGFR1	LHX6	PLEC1	SYPL2	
CACNA1C	FGFRL1	LHX8	PLEKHA2	SYT1	
CACNA1G	FJX1	LHX9	PLEKHH2	SYT17	
CACNA1H	FKBP10	LIMD2	PLEKHH3	SYT2	
CACNA2D3	FLI1	LIMS2	PLK1S1	SYT6	
CACNG6	FLJ16779	LINGO3	PLK5P	SYT7	
CACNG8	FLJ37307	LMO1	PLLP	T	
CALR3	FLJ39739	LMO2	PLOD2	TACC2	
CALY	FLJ40330	LMX1A	PLSCR1	TAF7	
CAMK1D	FLJ41350	LOC100128811	PLTP	TAGLN3	
CAMK2B	FLJ42709	LOC100128977	PLXDC1	TAL1	
CAMK2D	FLJ43663	LOC100130015	PLXNA4	TARSL2	
CAMKV	FLNC	LOC100130148	PLXNC1	TBC1D1	
CAMTA1	FLRT2	LOC100130691	PM20D2	TBR1	
CAND2	FMNL1	LOC100130987	PNMA1	TBX1	
CANX	FMNL2	LOC100133985	PNMA2	TBX2	
CAPN2	FMOD	LOC100192426	PNMAL2	TBX20	
CAPZB	FNDC4	LOC100286793	PNPLA3	TBX21	
CASP2	FOXA2	LOC113230	PODN	TBXA2R	
CAT	FOXC2	LOC145663	POM121L2	TCEAL2	
CBLN1	FOXD1	LOC151534	POMGNT1	TCEAL3	
CBLN4	FOXD2	LOC200726	POPDC3	TCERG1L	
CBS	FOXF1	LOC202181	POU3F1	TCF15	
CCBE1	FOXF2	LOC220930	POU3F2	TCF19	
CCDC108	FOXI3	LOC222699	POU4F1	TCF7L1	
CCDC136	FOXL2	LOC255512	POU4F3	TCP1	
CCDC140	FOXN4	LOC283856	POU6F2	TESC	
CCDC3	FOXO1	LOC285696	PPAPDC1A	TFAP2B	
CCDC46	FOXP1	LOC285780	PPARG	TFPI2	
CCDC48	FOXP4	LOC285830	PPFIA3	TFPT	
CCDC63	FOXQ1	LOC285954	PPIE	TGFB1I1	
CCDC68	FREM2	LOC348840	PPM1E	TICAM2	
CCHCR1	FREM3	LOC387763	PPP1R14A	TIGD5	
CCK	FRMPD4	LOC399815	PPP1R14C	TLE4	

CCKBR	FRZB	LOC399959	PPP1R16B	TLL1	
CCNA1	FSCN1	LOC400043	PPP1R1B	TLX1NB	
CCND2	FSTL5	LOC400940	PPP1R3G	TLX3	
CCNJL	FTO	LOC402778	PPP2R1B	TM6SF1	
CCR6	FUZ	LOC441177	PPP2R2B	TMBIM6	
CD248	FXYD7	LOC645323	PPP2R2C	TMED7-TICAM2	
CD34	FYN	LOC729991-MEF2B	PPYR1	TMEFF2	
CD40	FZD1	LOC84856	PRDM1	TMEM130	
CD7	FZD10	LOX	PRDM13	TMEM132B	
CD8A	FZD2	LOXL2	PRDM14	TMEM132E	
CD97	FZD7	LPAR3	PRDM16	TMEM151B	
CDC10L	FZD8	LPPR5	PRDM8	TMEM154	
CDC14B	G0S2	LRAT	PREX2	TMEM163	
CDC42EP5	GABARAPL2	LRFN2	PRHOXNB	TMEM17	
CDCA7	GABRA4	LRFN5	PRICKLE1	TMEM171	
CDCA7L	GABRA5	LRIT1	PRKAG2	TMEM176A	
CDH13	GABRB2	LRP11	PRKAR1B	TMEM176B	
CDH18	GAD1	LRRC10B	PRKCG	TMEM178	
CDH22	GAD2	LRRC20	PRKCH	TMEM179	
CDH23	GAL3ST2	LRRC32	PRKCI	TMEM196	
CDH6	GALNT12	LRRC33	PRKCQ	TMEM20	
CDH7	GALNT14	LRRC36	PRMT8	TMEM200B	
CDK5R2	GALNT9	LRRC3B	PRNP	TMEM229A	
CDK6	GALNTL1	LRRC4C	PROC	TMEM45B	
CDX2	GALR1	LRRC8D	PROK2	TMEM55A	
CDYL	GALR2	LRRIQ1	PROM1	TMEM57	
CEBPA	GAP43	LRRTM1	PROX1	TMEM90A	
CELSR3	GAS7	LSAMP	PRPF31	TMEM90B	
CEND1	GATA4	LSM5	PRR18	TMIE	
CENPT	GATA5	LTBP2	PRR5	TMTC1	
CENPV	GATM	LTBP4	PRR5-ARHGAP8	TNFAIP8	
CERK	GDF6	LY75	PRSS12	TNS3	
CERKL	GDNF	MACF1	PRSS27	TNXB	
CETP	GDPD5	MAD1L1	PRTFDC1	TOLLIP	
CH25H	GEFT	MAL	PRUNE2	TOX2	
CHAT	GEMIN4	MAN1C1	PSD	TP73	
CHD5	GFI1	MAP1B	PSMA1	TPK1	
CHD7	GFRA1	MAP1LC3B2	PSMA7	TPM1	
CHGA	GFRA2	MAP2K6	PTCH1	TPM4	
CHGB	GGN	MAP3K3	PTF1A	TPST1	
CHODL	GIMAP2	MAP4K4	PTGER1	TRAF6	
CHP	GIMAP7	MAP6	PTGFR	TRAIP	
CHPF	GIPC3	MAPK15	PTH1R	TRAK1	
CHRM2	GIYD1	MAPK4	PTH2	TRANK1	
CHRNA3	GIYD2	MAPKAP1	PTPN21	TRIB2	
CHRNA4	GJB2	MAPT	PTPRE	TRIL	
CHRNA7	GJB6	MARCKS	PTPRK	TRIM15	
CHRNB2	GJC1	MAST3	PTPRM	TRIM17	
CHRNB4	GLB1L2	MAT2B	PTPRN2	TRIM26	
CHST11	GLDN	MATK	PTPRS	TRIM58	
CHST12	GLIS3	MATN2	PTPRT	TRIM65	
CHST13	GLP1R	MATN4	PTRF	TRIM7	
CHST6	GMEB1	MCC	PTTG1IP	TRIM9	

CHST8	GNA15	MCF2L	PTX3	TRNP1	
CIB2	GNAL	MCF2L2	PURG	TRPA1	
CIDEA	GNAO1	ME1	PYY	TRPC3	
CITED4	GNAS	ME3	QSOX1	TRPC6	
CKB	GNG11	MECOM	QTRT1	TSC22D3	
CLCF1	GNG4	MECP2	RAB1B	TSC22D4	
CLCNKB	GOLSYN	MEF2B	RAB34	TSHZ2	
CLDN10	GPATCH4	MEGF10	RAB39	TSNAX-DISC1	
CLEC2L	GPC6	MEGF11	RAB3B	TSPAN19	
CLEC4G	GPM6A	MEOX1	RAB3C	TSPY4	
CLIC5	GPM6B	MEX3B	RAB9B	TTBK1	
CLIP4	GPR12	MFSD1	RAD21L1	TTC15	
CLSTN2	GPR120	MFSD2B	RAD51AP2	TTC21B	
CLTC	GPR123	MFSD4	RADIL	TTC39C	
CLTCL1	GPR139	MGAT3	RAI14	TTC9B	
CLU	GPR150	MGAT4C	RAP1GAP	TTYH1	
CLYBL	GPR158	MGAT5B	RAPGEFL1	TUB	
CMPK2	GPR177	MGC45800	RARRES1	TUBB2B	
CMTM3	GPRC5B	MGC87042	RASA3	TUBB8	
CMTM7	GPRIN2	MGMT	RASAL1	TULP1	
CNGA3	GPRIN3	MIAT	RASAL2	TUSC3	
CNGB1	GREB1	MIR1253	RASD1	TWIST1	
CNIH2	GREM2	MIR1256	RASD2	TXNRD1	
CNR1	GRIA2	MIR1258	RASGRF1	UBASH3B	
CNR2	GRID1	MIR375	RASGRP1	UCP1	
CNRIP1	GRID2	MIR548G	RASGRP2	ULBP3	

APPENDIX O

PATHWAYS SHARED BY A1, C1, AND NR GROUPS

Hypermethylated
In both A1 vs. NR and C1 vs. NR
cell fate commitment (GO:0045165)
cell-cell adhesion via plasma-membrane adhesion molecules (GO:0098742)
central nervous system neuron differentiation (GO:0021953)
homophilic cell adhesion via plasma membrane adhesion molecules (GO:0007156)
neuron fate commitment (GO:0048663)
In both C1 vs. NR and C1 vs. A1
cell differentiation in spinal cord (GO:0021515)
cell fate determination (GO:0001709)
dorsal spinal cord development (GO:0021516)
forebrain neuron differentiation (GO:0021879)
neural retina development (GO:0003407)
Hypomethylated
In both A1 vs. NR and C1 vs. N1
autonomic nervous system development (GO:0048483)
cell differentiation in spinal cord (GO:0021515)
embryonic eye morphogenesis (GO:0048048)
forebrain regionalization (GO:0021871)
neuron fate commitment (GO:0048663)
neuron fate specification (GO:0048665)
regulation of cardiac muscle cell proliferation (GO:0060043)
regulation of heart growth (GO:0060420)
ventral spinal cord development (GO:0021517)
In both C1 vs. NR and C1 vs. A1
neuron fate commitment (GO:0048663)
positive regulation of cardiac muscle tissue development (GO:0055025)
regulation of cardiac muscle cell differentiation (GO:2000725)
regulation of cardiac muscle tissue development (GO:0055024)
In both A1 vs. NR and C1 vs. A1
ear morphogenesis (GO:0042471)
inner ear morphogenesis (GO:0042472)
neuron fate commitment (GO:0048663)
In all 3
neuron fate commitment (GO:0048663)

BIBLIOGRAPHY

1. American Cancer Society: **Breast Cancer Facts and Figures 2015.** 2015.

2. [http://www.pathophys.org/breast-cancer/]

3. Visvader J: **Keeping abreast of the mammary epithelial heirarchy and breast tumorigenesis.** Genes Dev. 2009, **23**:2563.

4. Dimri G, Band H, Band V: **Mammary epithelial cell transformation: insights from cell culture and mouse models.** Breast Cancer Res. 2005, **7**(4):171-179.

5. Gudjonsson T, Villadsen R, Nielsen HL, Rønnov-Jessen L, Bissell MJ, Petersen OW: **Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties.** Genes Dev. 2002, **16**(6):693-706.

6. Prat A, Perou CM: **Deconstructing the molecular portraits of breast cancer.** Mol Oncol. 2011, **5**(1):5-23.

7. Inwald EC, Klinkhammer-Schalke M, Hofstädter F, Zeman F, Koller M, Gerstenhauer M, Ortmann O: **Ki-67 is a prognostic parameter in breast cancer patients: results of a large population-based cohort of a cancer registry.** Breast Cancer Res Treat. 2013, **139**(2):539-552.

8. National Cancer Institute: **Understanding Estrogen Receptors/SERMs.** 2010.

9. Burger HG: **Androgen production in women.** Fertil Steril. 2002, **77**(Suppl 4):S3-S5.

10. Fabian CJ: **The what, why and how of aromatase inhibitors: hormonal agents for treatment and prevention of breast cancer.** Int J Clin Pract. 2007, **61**(12):2051-2063.

11. Dowsett M, Howell A: **Breast cancer: aromatase inhibitors take on tamoxifen.** Nat Med. 2002, **8**(12):1341-1344.

12. Liang J, Shang Y: **Estrogen and Cancer.** Annu Rev Physiol. 2013, **75**:225.

13. [http://www.rcsb.org/pdb/101/motm.do?momID=45]

14. Hervouet E, Cartron P, Jouvenot M, Delage-Mourroux R: **Epigenetic regulation of estrogen signaling in breast cancer.** Epigenetics 2013, **8**(3):237.

15. Petersen OW, Høyer PE, van Deurs B: **Frequency and Distribution of Estrogen Receptor-positive Cells in Normal, Nonlactating Human Breast Tissue.** Cancer Res 1987, **47**:5748.

16. Haldosén LA, Zhao C, Dahlman-Wright K: **Estrogen receptor beta in breast cancer.** Mol Cell Endocrinol. 2014, **382**(1):665-672.

17. Rody A, Holtrich U, Solbach C, Kourtis K, von Minckwitz G, Engels K, Kissler S, Gätje R, Karn T, Kaufmann M: **Methylation of estrogen receptor beta promoter correlates with loss of ER-beta expression in mammary carcinoma and is an early indication marker in premalignant lesions.** Endocr Relat Cancer 2005, **12**(4):903-916.

18. Zhao C, Lam EW, Sunters A, Enmark E, De Bella MT, Coombes RC, Gustafsson JA, Dahlman-Wright K: **Expression of estrogen receptor beta isoforms in normal breast epithelial cells and breast cancer: regulation by methylation.** Oncogene 2003, **22**(48):7600-7606.

19. Mohammed H, Russell IA, Stark R, Rueda OM, Hickey TE, Tarulli GA, Serandour AA, Birrell SN, Bruna A, Saadi A, Menon S, Hadfield J, Pugh M, Raj GV, Brown GD, D'Santos C, Robinson JL, Silva G, Launchbury R, Perou CM, Stingl J, Caldas C, Tilley WD, Carroll JS: **Progesterone receptor modulates ERα action in breast cancer.** Nature 2015, **523**(7560):313-317.

20. [http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-treating-general-info]

21. [http://www.nccn.org/patients/guidelines/stage_i_ii_breast/index.html]

22. [http://ww5.komen.org/BreastCancer/TheChemotherapyDrugs.html]

23. Gradishar WJ, Anderson BO, Blair SL, Burstein HJ, Cyr A, Elias AD, Farrar WB, Forero A, Hermes Giordano S, Goldstein LJ, Hayes DF, Hudis CA, Isakoff SJ, MD P, Ljung BE, Marcom PK, Mayer IA, McCormick B, Miller RS, Pegram M, Pierce LJ, Reed EC, Salerno KE: **Breast Cancer Version 3.2014.** JNCCN 2014, **12**(4):542.

24. Cole MP, Jones CT, Todd ID: **A new anti-oestrogenic agent in late breast cancer. An early clinicial appraisal of ICI46474.** Br J Cancer 1971, **25**(2):270.

25. Ring A, Dowsett M: **Mechanisms of Tamoxifen resistance.** Endocr Relat Cancer 2004, **11**:643.

26. Early Breast Cancer Trialists' Collaborative Group (EBCTCG): **Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials.** Lancet 2005, **365**(9472):1687-1717.

27. Mo Z, Liu M, Yang F, Luo H, Li Z, Tu G, Yang G: **GPR30 as an initiator of tamoxifen resistance in hormone-dependent breast cancer.** Breast Cancer Res. 2013, **15**:R114.

28. Cook KL, Shajahan AN, Clarke R: **Autophagy and endocrine resistance in breast cancer.** Expert Rev Anticancer Ther 2011, **11**(8):1283.

29. Badia E, Oliva J, Balaguer P, Cavailles V: **Tamoxifen Resistance and Epigenetic Modifications in Breast Cancer Cell Lines.** Curr Med Chem 2007, **14**(28):3035-3043.

30. Fan M, Yan PS, Hartman-Frey C, Chen L, Paik H, Oyer SL, Salisbury JD, Cheng AS, Li L, Abbosh PH, Huang TH, Nephew KP: **Diverse gene expression and DNA methylation profiles correlate with differential adaptation of breast cancer cells to the antiestrogens tamoxifen and fulvestrant.** Cancer Res 2006, **66**(24):11954-11966.

31. Williams KE, Anderton DL, Lee MP, Pentecost BT, Arcaro KF: **High-density array analysis of DNA methylation in Tamoxifen-resistant breast cancer cell lines.** Epigenetics 2014, **9**(2):1.

32. Clarke R, Liu MC, Bouker KB, Gu Z, Lee RY, Zhu Y, Skaar TC, Gomez B, O'Brien K, Wang Y, Hilakivi-Clarke LA: **Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling.** Oncogene 2003, **22**(47):7316-7339.

33. Ross JS, Fletcher JA, Bloom KJ, Linette GP, Stec J, Symmans WF, Pusztai L, Hortobagyi GN: **Targeted therapy in breast cancer: the HER-2/neu gene and protein.** Mol Cell Proteomics 2004, **3**(4):379-398.

34. Eccleston A, DeWitt N, Gunter C, Marte B, Nath P: **Epigenetics.** Nature 2007, **447**:395.

35. Weinhold B: **Epigenetics.** Environ Health Prospectives 2006, **114**(3):A160.

36. [http://epi.grants.cancer.gov/epigen.html]

37. Adams RL, Lindsay H: **What is hemimethylated DNA?** FEBS Lett. 1993, **320**(3):243-245.

38. Das PM, Thor AD, Edgerton SM, Barry SK, Chen DF, Jones FE: **Reactivation of epigenetically silenced HER4/ERBB4 results in apoptosis of breast tumor cells.** Oncogene 2010, **29**(37):5214-5219.

39. Widschwendter M, Jones PA: **DNA methylation and breast carcinogenesis.** Oncogene 2002, **21**(35):5462-5482.

40. Handel AE, Ebers GC, Ramagopalan SV: **Epigenetics: molecular mechanisms and implications for disease.** Trends in Mol Medicine 2009, **16**(1):7.

41. Taberlay PC, Jones PA: **DNA methylation and cancer.** Prog Drug Res 2011, **67**:1.

42. Baylin S: **DNA methylation and gene silencing in cancer.** Nat Clin Practice Oncology 2005, **2**:S4.

43. Varier RA, Timmers HT: **Histone lysine methylation and demethylation pathways in cancer..** Biochim Biophys Acta. 2011, **1815**(1):75-89.

44. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G: **Gene body methylation can alter gene expression and is a therapeutic target in cancer.** Cancer Cell. 2014, **26**(4):577-90.

45. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, Turecki G, Delaney A, Varhol R, Thiessen N, Shchors K, Heine VM, Rowitch DH, Xing X, Fiore C, Schillebeeckx M, Jones SJ, Haussler D, Marra MA, Hirst M, Wang T, Costello JF: **Conserved role of intragenic DNA methylation in regulating alternative promoters.** Nature 2014, **466**(7303):253-257.

46. Yang X, Yan L, Davidson NE: **DNA methylation in breast cancer.** Endocr Relat Cancer 2001, **8**(2):115-127.

47. Gozgit JM, Pentecost BT, Marconi SA, Otis CN, Wu C, Arcaro KF: **Use of an aggressive MCF-7 cell line variant, TMX2-28, to study cell invasion in breast cancer.** Mol Cancer Res 2006, **4**(12):905-913.

48. Fagan-Solis KD, Smith-Schneider S, Pentecost BT, Bentley BA, Otis CN, Gierthy JF, Arcaro KF: **The RhoA pathway mediates MMP-2 and MMP-9-independent invasive behavior in a triple-negative breast cancer cell line.** J Cell Biochem 2013, **114**(6):1385.

49. Fasco MJ, Amin A, Pentecost BT, Yang Y, Gierthy JF: **Phenotypic changes in MCF-7 cells during prolonged exposure to tamoxifen.** Mol Cell Endocrinol 2003, **206**(1-2):33-47.

50. Gozgit JM, Pentecost BT, Marconi SA, Ricketts-Loriaux R, Otis CN, Arcaro KF: **PLD1 is overexpressed in an ER-negative MCF-7 cell line variant and a subset of phospho-Akt-negative breast carcinomas.** Br J Cancer 2007, **97**(6):809-817.

51. Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, Esteller M: **Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome.** Epigenetics 2011, **6**(6):692-702.

52. Herman JG, Graff JR, Myöhänen S, Nelkin BD, Baylin SB: **Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands.** Proc Natl Acad Sci. 1996, **93**(18):9821-9826.

53. Clark C, Palta P, Joyce CJ, Scott C, Grundberg E, Deloukas P, Palotie A, Coffey AJ: **A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip(®) for methylome profiling.** PloS one 2012, **7**(11):e50233.

54. Kandimalla R, van Tilborg AA, Zwarthoff EC: **DNA methylation-based biomarkers in bladder cancer.** Nat Rev Urol. 2013, **10**(8):327-335.

55. [http://cnx.org/content/m44536/latest/]

56. Musgrove EA, Sutherland RL: **Biological determinants of endocrine resistance in breast cancer.** Nat Rev Cancer 2009, **9**(9):631-643.

57. Zhou C, Zhong Q, Rhodes LV, Townley I, Bratton MR, Zhang Q, Martin EC, Elliott S, Collins-Burow BM, Burow ME, Wang G: **Proteomic analysis of acquired Tamoxifen resistance in MCF-7 cells reveals expression signatures associated with enhanced migration.** Br Can Res 2012, **14**(R45).

58. Badia E, Duchesne M, Semlali A, Fuentes M, Giamarchi C, Richard-Foy H, Nicolas J, Pons M: **Long-Term Hydroxytamoxifen Treatment of an MCF-7-derived Breast Cancer Cell Line Irreversibly Inhibits the Expression of Estrogenic Genes through Chromatin Remodeling.** Cancer Research 2000, **60**(15):4130-4138.

59. Garcia V, Domínguez G, García JM, Silva J, Peña C, Silva JM, Carcereny E, Menendez J, España P, Bonilla F: **Altered expression of the ZBRK1 gene in human breast carcinomas.** J Pathol 2004, **202**(2):224-232.

60. Hansen KD, Aryee M: **minfi: Analyze Illumina's 450k methylation arrays.** 2012.

61. Stone A, Valdés-Mora F, Gee JMW, Farrow L, McClelland RA, Fiegl H, Dutkowski C, McCloy RA, Sutherland RL, Musgrove EA, Nicholson RI: **Tamoxifen-Induced Epigenetic Silencing of Oestrogen-Regulated Genes in Anti-Hormone Resistant Breast Cancer.** PLoS ONE 2012, **7**(7):e40466.

62. Rao X, Evans J, Chae H, Pilrose J, Kim S, Yan P, Huang R, Lai H, Lin H, Liu Y, Miller D, Rhee J, Huang Y, Gu F, Gray JW, Huang T, Nephew KP: **CpG island shore methylation regulates caveolin-1 expression in breast cancer.** Oncogene 2013, **32**(38):4519-4528.

63. Hill VK, Ricketts C, Bieche I, Vacher S, Gentle D, Lewis C, Maher ER, Latif F: **Genome-Wide DNA Methylation Profiling of CpG Islands in Breast Cancer Identifies Novel Genes Associated with Tumorigenicity.** Cancer Res 2011, **71**(8):2988-2999.

64. Krausz C, Sandoval J, Sayols S, Chianese C, Giachini C, Heyn H, Esteller M: **Novel Insights into DNA Methylation Features in Spermatozoa: Stability and Peculiarities.** PLoS ONE 2012, **7**(10):e44479.

65. Du Q, Zhang Y, Tian X, Li Y, Fang W: **MAGE-D1 inhibits proliferation, migration and invasion of human breast cancer cells.** Oncology Reports 2009, **22**(3):659.

66. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK: **On the presence and role of human gene-body DNA methylation.** Oncotarget 2012, **3**(4):462.

67. Hellman A, Chess A: **Gene Body-Specific Methylation on the Active X Chromosome.** Science 2007, **315**(5815):1141-1143.

68. Fournel M, Sapieha P, Beaulieu N, Besterman JM, MacLeod AR: **Down-regulation of Human DNA-(Cytosine-5) Methyltransferase Induces Cell Cycle Regulators p16 ink4A and p21WAF/Cip1 by Distinct Mechanisms.** J Biol Chem 1999, **274**(34):24250-24256.

69. Zhu WG, Dai Z, Ding H, Srinivasan K, Hall J, Duan W, Villalona-Calero MA, Plass C, Otterson GA: **Increased expression of unmethylated CDKN2D by 5-aza-2-deoxycytidine in human lung cancer cells.** Oncogene 2001, **20**(53):7787-7796.

70. Liang G, Gonzales FA, Jones PA, Orntoft TF, Thykjaer T: **Analysis of Gene Induction in Human Fibroblasts and Bladder Cancer Cells Exposed to the Methylation Inhibitor 5-Aza-2'-deoxycytidine.** Cancer Res 2002, **62**(4):961-966.

71. Schmelz K, Sattler N, Wagner M, Lubbert M, Dorken B, Tamm I: **Induction of gene expression by 5-Aza-2'-deoxycytidine in acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) but not epithelial cells by DNA-methylation-dependent and -independent mechanisms.** Leukemia 2004, **19**(1):103-111.

72. Han H, Cortez CC, Yang X, Nichols PW, Jones PA, Liang G: **DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter.** Hum Mol Genet 2011, **20**(22):4299-4310.

73. Sasaki A, Hinck L, Watanabe K: **RumMAGE-D the Members: Structure and Function of a New Adaptor Family of MAGE-D Proteins.** Journal of Receptors and Signal Transduction 2005, **25**(3):181.

74. Wu Y, Alvarez M, Slamon DJ, Koeffler P, Vadgama JV: **Caspase 8 and maspin are downregulated in breast cancer cells due to CpG site promoter methylation.** BMC Cancer 2010, **10**:32.

75. Wozniak RJ, Klimecki WT, Lau SS, Feinstein Y, Futscher BW: **5-Aza-2'-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone**

H3 K9 di-methylation levels are linked to tumor suppressor gene reactivation. Oncogene 2007, **26**(1):77-90.

76. Mirza S, Sharma G, Pandya P, Ralhan R: **Demethylating agent 5-aza-2-deoxycytidine enhances susceptibility of breast cancer cells to anticancer agents.** Mol Cell Biochem 2010, **342**(1-2):101-109.

77. Fenaux P, Mufti GJ, Hellström-Lindberg E, Santini V, Gattermann N, Germing U, Sanz G, List AF, Gore S, Seymour JF, Dombret H, Backstrom J, Zimmerman L, McKenzie D, Beach CL, Silverman LR: **Azacitidine prolongs overall survival compared with conventional care regimens in elderly patients with low bone marrow blast count acute myeloid leukemia.** Journal of Clinical Oncology 2010, **28**(4):562.

78. Maurillo L, Venditti A, Spagnoli A, Gaidano G, Ferrero D, Oliva E, Lunghi M, D'Arco AM, Levis A, Pastore D, Di Renzo N, Santagostino A, Pavone V, Buccisano F, Musto P: **Azacitidine for the Treatment of Patients With Acute Myeloid Leukemia.** Cancer 2012, **118**:1014.

79. Toyota M, Kopecky KJ, Toyota, M-O., Jair, K-W., Willman CL, and Issa J-P.J.: **Methylation profiling in acute myeloid leukemia.** Blood 2001, **97**:2823.

80. Szyf M: **Epigenetics, DNA methylation, and chromatin modifying drugs.** Annu Rev Pharmacol Toxicol 2009, **49**:243-263.

81. Kelly TK, De Carvalho DD, Jones PA: **Epigenetic modifications as therapeutic targets.** Nat Biotechnol 2010, **28**(10):1069-1078.

82. Pakneshan P, Szyf M, Farias-Eisner R, Rabbani SA: **Reversal of the hypomethylation status of urokinase (uPA) promoter blocks breast cancer growth and metastasis.** J Biol Chem 2004, **279**(30):31735-31744.

83. Luo J, Li YN, Wang F, Zhang WM, Geng X: **S-adenosylmethionine inhibits the growth of cancer cells by reversing the hypomethylation status of c-myc and H-ras in human gastric cancer and colon cancer.** Int J Biol Sci 2010, **6**(7):784-795.

84. Liu W, Guan M, Hu T, Gu X, Lu Y: **Re-Expression of AKAP12 Inhibits Progression and Metastasis Potential of Colorectal Carcinoma In Vivo and In Vitro.** PLoS One 2011, **6**(8):24015.

85. Wang D, Huang P, Zhu B, Sun L, Huang Q, Wang J: **Induction of estrogen receptor α-36 expression by bone morphogenetic protein 2 in breast cancer cell lines.** Mol Med Rep. 2012, **6**(3):591.

86. Gara SK, Grumati P, Urciuolo A, Bonaldo P, Kobbe B, Koch M, Paulsson M, Wagener R: **Three novel collagen VI chains with high homology to the alpha3 chain.** J Biol Chem 2008, **283**(16):10658-10670.

87. Sasaki Y, Minamiya Y, Takahashi N, Nakagawa T, Katayose Y, Ito A, Saito H, Motoyama S, Ogawa J: **REG1A Expression is an Independent Factor Predictive of Poor Prognosis in Patients with Breast Cancer.** Annals of Surgical Oncology 2008, **15**(11):3244.

88. Jin Z, Olaru A, Yang J, Sato F, Cheng Y, Kan T, Mori Y, Mantzur C, Paun B, Hamilton JP, Ito T, Wang S, David S, Agarwal R, Beer DG, Abraham JM, Meltzer SJ: **Hypermethylation of Tachykinin-1 Is a Potential Biomarker in Human Esophageal Cancer.** Clinical Cancer Research 2007, **13**(21):6293.

89. Pang JB, Dobrovic A, Fox SB: **DNA methylation in ductal carcinoma *in situ* of the breast.** Breast Cancer Research 2013, **15**:206.

90. Du P, Tang F, Qiu Y, Dong F: **GFI1 is repressed by p53 and inhibits DNA damage-induced apoptosis.** PloS one 2013, **4**(9):e73542.

91. Yuan Z, Peng L, Radhakrishnan R, Seto E: **Histone Deacetylase 9 (HDAC9) Regulates the Functions of the ATDC (TRIM29) Protein.** Journal of Biological Chemistry 2010, **285**:39329.

92. Lee CC, Chen WS, Chen CC, Chen LL, Lin YS, Fan CS, Huang TS: **TCF12 protein functions as transcriptional repressor of E-cadherin, and its overexpression is correlated with metastasis of colorectal cancer.** Journal of Biological Chemistry 2012, **287**(4):2798.

93. Wang XX, Zhu Z, Su D, Lei T, Wu X, Fan Y, Li X, Zhao J, Fu L, Dong JT, Fu L: **Down-regulation of leucine zipper putative tumor suppressor 1 is associated with poor prognosis, increased cell motility and invasion, and epithelial-to-mesenchymal transition characteristics in human breast carcinoma.** Hum Pathol. 2011, **42**(10):1410.

94. Xiong G, Wang C, Evers BM, Zhou BP, Xu R: **RORα suppresses breast tumor invasion by inducing SEMA3F expression.** Cancer Res. 2012, **72**(7):1728.

95. Winkler DD, Muthurajan UM, Hieb AR, Luger K: **Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events.** Journal of Biological Chemistry 2011, **286**(48):41883.

96. Savage K, Leung S, Todd SK, Brown LA, Jones RL, Robertson D, James M, Parry S, Rodrigues Pinilla SM, Huntsman D, Reis-Filho J.S.: **Distribution and significance of caveolin 2 expression in normal breast and invasive breast cancer: an**

immunofluorescence and immunohistochemical analysis. Breast Cancer Res.Treat. 2008, **110**(2):245.

97. Mohammed H, D'Santos C, Serandour AA, Ali HR, Brown GD, Atkins A, Rueda OM, Holmes KA, Theodorou V, Robinson JL, Zwart W, Saadi A, Ross-Innes CS, Chin SF, Menon S, Stingl J, Palmieri C, Caldas C, Carroll JS: **Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor.** Cell Rep. 2013, **3**(2):342.

98. Barcellos-Hoff MH, Akhurst RJ: **Transforming growth factor-beta in breast cancer: too much, too late.** Breast Cancer Research 2009, **11**(1):202.

99. Lindner DJ, Wu Y, Haney R, Jmmacobs BS, Fruehauf JP, Tuthill R, Borden EC: **Thrombospondin-1 expression in melanoma is blocked by methylation and targeted reversal by 5-Aza-deoxycytidine suppresses angiogenesis.** Matrix Biol. 2013, **32**(2):123.

100. Chauvet C, Vanhoutteghem A, Duhem C, Saint-Auret G, Bois-Joyeux B, Djian P, Staels B, Danan JL: **Control of gene expression by the retinoic acid-related orphan receptor alpha in HepG2 human hepatoma cells.** PloS one 2011, **6**(7):e22545.

101. Zhu Y, McAvoy S, Kuhn R, Smith DI: **RORA, a large common fragile site gene, is involved in cellular stress response.** Oncogene 2006, **25**(20):2901.

102. Hinshelwood RA, Huschtscha LI, Melki J, Stirzaker C, Abdipranoto A, Vissel B, Ravasi T, Wells CA, Hume DA, Reddel RR, Clark SJ: **Concordant epigenetic silencing of transforming growth factor-beta signaling pathway genes occurs early in breast carcinogenesis.** Cancer Res. 2007, **67**(24):11517.

103. Hoque MO, Prencipe M, Poeta ML, Barbano R, Valori VM, Copetti M, Gallo AP, Brait M, Maiello E, Apicella A, Rossiello R, Zito F, Stefania T, Paradiso A, Carella M, Dallapiccola B, Murgo R, Carosi I, Bisceglia M, Fazio VM, Sidransky D, Parrella P: **Changes in CpG islands promoter methylation patterns during ductal breast carcinoma progression.** Cancer Epidemiol Biomarkers Prev 2009, **18**(10):2694-2700.

104. Jin H, Pi J, Huang X, Huang F, Shao W, Li S, Chen Y, Cai J: **BMP2 promotes migration and invasion of breast cancer cells via cytoskeletal reorganization and adhesion decrease: an AFM investigation.** Appl Microbiol Biotechnol. 2012, **93**(4):1715.

105. Alarmo EL, Kallioniemi A: **Bone morphogenetic proteins in breast cancer: dual role in tumourigenesis?** Endocr Relat Cancer 2010, **17**(2):R123.

106. Al-Akoum M, Dodin S, Akoum A: **Synergistic cytotoxic effects of tamoxifen and black cohosh on MCF-7 and MDA-MB-231 human breast cancer cells: an in vitro study.** Can J Physiol Pharmacol 2007, **85**(11):1153.

107. Johnson MD, Westley BR, May FEB: **Oestrogenic activity of tamoxifen and its metabolites on gene regulation and cell proliferation in MCF-7 breast cancer cells.** Br J Cancer 1989, **59**:727.

108. Detich N, Hamm S, Just G, Knox JD, Szyf M: **The Methyl Donor S-Adenosylmethionine Inhibits Active Demethylation of DNA.** Journal of Biological Chemistry 2003, **278**(23):20812-20820.

109. Mato JM, Martínez-Chantar ML, Lu SC: **S-adenosylmethionine metabolism and liver disease.** Ann Hepatol. 2013, **12**(2):183.

110. Mischoulon D, Price LH, Carpenter LL, Tyrka AR, Papakostas GI, Baer L, Dording CM, Clain AJ, Durham K, Walker R, Ludington E, Fava M: **A double-blind, randomized, placebo-controlled clinical trial of S-adenosyl-L-methionine (SAMe) versus escitalopram in major depressive disorder.** J Clin Psychiatry 2014, **75**(4):370.

111. Huang KT, Mikeska T, Li J, Takano EA, Millar EK, Graham PH, Boyle SE, Campbell IG, Speed TP, Dobrovic A, Fox SB:
Assessment of DNA methylation profiling and copy number variation as indications of clonal relationship in ipsilateral and contralateral breast cancers to distinguish recurrent breast cancer from a second primary tumour. BMC cancer 2015, **15**:669.

112. Bediaga NG, Acha-Sagredo A, Guerra I, Viguri A, Albaina C, Ruiz Diaz I, Rezola R, Alberdi MJ, Dopazo J, Montaner D, de Renobales M, Fernandez AF, Field JK, Fraga MF, Liloglou T, de Pancorbo MM: **DNA methylation epigenotypes in breast cancer molecular subtypes.** Breast Cancer Res 2010, **12**(5):R77.

113. Figueroa JD, Yang H, Garcia-Closas M, Davis S, Meltzer P, Lissowska J, Horne HN, Sherman ME, Lee M: **Integrated analysis of DNA methylation, immunohistochemistry and mRNA expression, data identifies a methylation expression index (MEI) robustly associated with survival of ER-positive breast cancer patients..** Breast Cancer Res Treat. 2015, **150**(2):457-466.

114. Fackler MJ, Umbricht CB, Williams D, Argani P, Cruz LA, Merino VF, Teo WW, Zhang Z, Huang P, Visvananthan K, Marks J, Ethier S, Gray JW, Wolff AC, Cope LM, Sukumar S: **
Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of re- currence.** Cancer Res. 2011, **71**(19):6195-6207.

115. Fang F, Turcan S, Rimner A, Kaufman A, Giri D, Morris LG, Shen R, Seshan V, Mo Q, Heguy A, Baylin SB, Ahuja N, Viale A, Massague J, Norton L, Vahdat LT, Moynahan ME, Chan TA: **Breast cancer methylomes establish an epigenomic foundation for metastasis.** Sci Transl Med. 2011, **3**(75):75ra25.

116. Naeem H, Wong NC, Chatterton Z, Hong MK, Pedersen JS, Corcoran NM, Hovens CM, Macintyre G:

Reducing the risk of false discovery enabling identification of biologically significant genome-wide methylationstatus using the HumanMethylation450 array. BMC Genomics 2014, **15**:51.

117. Anastas JN, Moon RT:
WNT signalling pathways as therapeutic targets in cancer. Nat Rev Cancer 2013, **13**(1):11-26.

118. Yao D, Dai C, Peng S: **
Mechanism of the Mesenchymal–Epithelial Transition and Its Relationship with Metastatic Tumor Formation** Mol.Cancer.Res. 2011, **9**(12):1608-1620.

119. Alarmo EL, Kallioniemi A:
Bone morphogenetic proteins in breast cancer: dual role in tumourigenesis? Endocr Relat Cancer 2010, **17**(2):R123-139.

120. Nagaishi M, Kim YH, Mittelbronn M, Giangaspero F, Paulus W, Brokinkel B, Vital A, Tanaka Y, Nakazato Y, Legras-Lachuer C, Lachuer J, Ohgaki H: **Amplification of the STOML3, FREM2,**
and LHFP genes is associated with mesenchymal differentiation ingliosarcoma. Am J Pathol. 2012, **180**(5):1816-1823.

121. Fisher B, Jeong JH, Dignam J, Anderson S, Mamounas E, Wickerham DL, Wolmark N: **Findings from recent National Surgical Adjuvant Breast and Bowel Project adjuvant studies in stage I breast cancer.** J Natl Cancer Inst Monogr. 2001, (30):62-66.

122. Szyf M: **DNA methylation signatures for breast cancer classification and prognosis.** Genome Med. 2012, **4**(3):26.