The effect of p53 on function of TFAP2C in breast cancer: detailed analysis of regulation of MUC1 gene

Yingyue Li
University of Iowa

Copyright 2012 Yingyue Li

This dissertation is available at Iowa Research Online: http://ir.uiowa.edu/etd/3338

Recommended Citation
http://ir.uiowa.edu/etd/3338.

Follow this and additional works at: http://ir.uiowa.edu/etd
Part of the Biochemistry Commons
THE EFFECT OF P53 ON FUNCTION OF TFAP2C IN BREAST CANCER:
DETAILED ANALYSIS OF REGULATION OF MUC1 GENE

By
Yingyue Li

A thesis submitted in partial fulfillment
of the requirements for the Master of
Science degree in Biochemistry
in the Graduate College of
The University of Iowa

July 2012

Thesis Supervisor: Professor Ronald J. Weigel
CERTIFICATE OF APPROVAL

MASTER'S THESIS

This is to certify that the Master's thesis of

Yingyue Li

has been approved by the Examining Committee for the thesis requirement for the Master of Science degree in Biochemistry at the July 2012 graduation.

Thesis Committee: __________________________

David H. Price, Thesis Chair

______________________________

Lori L. Wallrath

______________________________

Miles Pufall
To my beloved family and all my dear friends.
Relationship are all there is. Everything in the universe only exists because it is in relationship to everything else. Nothing exists in isolation. We have to stop pretending we are individuals that can go it alone.

Margaret Wheatley
ACKNOWLEDGMENTS

It is my pleasure to thank those who made this thesis possible.

Above all, I owe my deepest and sincerest gratitude to my PI, Dr. Ronald J. Weigel. Without his guidance and support toward my study and research, I could not achieve anything. Also, I value a lot of his understanding, encouragement and help to my career path.

I appreciate the advice and comments from the committee members. I benefit a lot from their insightful comments and logical arguments.

I give great gratitude to my rotation committee members, Dr. Ronald J. Weigel, Dr. David H. Price and Dr. Daniel L. Weeks. They polished me with great efforts on experiment design, trouble shooting, and writing and presentation skills.

Moreover, I am grateful to the help of all the lab members, especially George Woodfield, Dr. Maria Bogachek and Dr. Mikhail V Kulak. They taught me the lab techniques with great patience and helped me to initiate the data analysis.

Additionally, I would like to thank Dr. Thomas B. Bair and Dr. Tiandao Li. They performed the ChIP-seq data analysis that made my project possible.

It is my honor to thank Dr. Frederick Domann and his lab assistant Matthew Fitzgerald. Without their generous help with the DNA methylation assay and DNA accessibility assay, I could not complete the project in limited time.

Especially, I would like to give my special thanks to my family and my boyfriend. It is the encouragement from my parents and the accompany and support from my boyfriend that drives me to overcome difficulties, achieve excellence and pursue my dream.
ABSTRACT

Transcription factor AP2C (TFAP2C) is believed to be involved in breast cancer carcinogenesis. However, the molecular mechanisms of regulating its trans-activation activity are not well understood. One of the potential mechanisms is through p53-mediated regulation. Using ChIP-seq analysis to map the TFAP2C occupancy across the genome, we found that the introduction of p53 to HCT 116 p53 -/- colon cancer cell line significantly augments TFAP2C occupancy on the promoter regions of a group of genes. Of these, six genes were further investigated. First, TFAP2C binding sites were identified in the center of ChIP-seq peaks on the promoters of the six genes and these were verified by gel shift assays. One of these genes, MUC1, was then determined to be activated by TFAP2C in MCF-7 breast cancer cell line. Subsequently, MUC1 was selected as the model target gene to elucidate the mechanism for p53-mediated enhancement of TFAP2C occupancy. We hypothesized that DNA methylation of the MUC1 promoter is altered by p53, leading to the increased TFAP2C occupancy to its TFBS on MUC1 promoter. To examine this, CpG methylation assay was performed. The result showed the DNA methylation of MUC1 promoter region remains identical with or without over-expression of p53 in HCT 116 p53 -/- cell line. From these studies, I conclude that 1) introduction of p53 augments TFAP2C binding on specific gene targets; 2) MUC1 gene is activated by TFAP2C and two TFAP2C binding sites were verified; 3) promoter DNA methylation does not explain the increased occupancy of TFAP2C on MUC1 promoter.
# TABLE OF CONTENTS

## LIST OF TABLES

vi

## LIST OF FIGURES

ix

## LIST OF ABBREVIATIONS

xii

## CHAPTER I INTRODUCTION

1

1.1. Breast cancer and classification of phenotypes of breast cancer.............. 1

1.1.1. Breast cancer .................................................................................. 1

1.1.2. Breast cancer categories .............................................................. 1

1.2. AP-2 family (with focus on TFAP2C) .............................................. 2

1.2.1. AP-2 family .................................................................................. 2

1.2.2. The activity of TFAP2C in breast carcinogenesis ......................... 3

1.2.3. The molecular mechanisms of regulating TFAP2C occupancy and its trans-activation activity .................................................. 5

1.3. p53 is a potential TFAP2C cofactor .................................................. 8

1.3.1. p53 and p53 variants .................................................................. 8

1.3.2. The link between p53 and ER-α or ERBB2 ................................. 8

1.3.3. p53 and AP-2 family .................................................................. 10

1.4. MUC1, a gene model to illuminate that p53 regulates binding and trans-activation activity of TFAP2C ................................................. 12

1.4.1. Protein character of MUC1 ......................................................... 12

1.4.2. Interactions of MUC1 with other breast cancer related molecules.. 14

1.4.3. Mucin family and AP-2 family ................................................. 15

1.5. Promoter CpG methylation- the potential mechanism to explain enhanced TFAP2C occupancy on MUC1 promoter induced by p53.... 16

1.5.1. CpG methylation and tumor ....................................................... 16

1.5.2. Why examining DNA methylation ............................................. 16

1.6. Overview of thesis ......................................................................... 17

## CHAPTER II P53-MEDIATED AUGMENTATION OF TFAP2C BINDING OCCUPANCY ON SPECIFIC GENES

20

2.1. Materials and Methods .................................................................... 20

2.1.1. Cell Lines and Culture ................................................................. 20

2.1.2. Transfection of siRNA ................................................................ 21

2.1.3. Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)... 22

2.1.4. Western Blot ............................................................................... 23

2.1.5. Chromatin Immunoprecipitation (ChIP) Assays and ChIP-seq Data Analysis ............................................................................. 23

2.1.6. Gel-Shift Assay ......................................................................... 25

2.1.7. CpG Island Methylation Analysis ............................................. 28

2.2. Results ............................................................................................ 33
2.2.1. TFAP2C is a key transcriptional regulator establishing pattern of gene expression characterizing luminal type of breast cancer........ 34
2.2.2. p53 expression augments TFAP2C binding on specific genes........ 35
2.2.3. Localization of TFAP2C binding sites ........................................ 37
2.2.4. MUC1 is a TFAP2C target in MCF-7 breast cancer cell line .......... 40
2.2.5. CpG methylation is not changed within MUC1 promoter in HCT 116 p53 -/- cells with introduction of p53 ........................................ 42

CHAPTER III DISCUSSION

3.1. Limitations of ChIP-seq research ....................................................... 45
3.2. Other ways to perform DNA methylation assay ................................. 46
3.3. Other mechanisms to explain that p53-mediated augmentation of TFAP2C binding occupancy on MUC1........................................ 47
3.3.1. p53 modifies motif preference of TFAP2C ...................................... 47
3.3.2. Altered DNA accessibility by p53 .................................................. 48
3.4. Future Directions ............................................................................. 48
3.4.1. To determine MUC1 as a TFAP2C target in other breast cancer cell lines ................................................................. 48
3.4.2. To perform the ChIP-seq (or ChIP-reChIP) with anti-p53 and anti-TFAP2C in HCT116 p53 -/- cells .................................................. 49
3.4.3. To generate a set of genes where p53 blocks TFAP2C occupancy 49

REFERENCES

APPENDIX
LIST OF TABLES

Table 1 Summary of cell lines used in this study and the categorization of breast tumor. .......................................................... 4

Table 2 Summary of cell line–specific culturing media. .................................................. 20

Table 3 Transfection of siRNAs to silence TFAP2A, TFAP2C and p53 in MCF-7 cell lines. ......................................................................................... 21

Table 4 Summary of ChIP-seq datasets used in this project........................................ 24

Table 5 Summary of cell lines used in CpG island methylation assay. ...................... 29

Table 6 PCR thermal cycles in the CpG methylation assay to amplify MUC1 promoter regions. .................................................................................... 32

Table 7 List of TFAP2C binding peaks augmented with presence of p53 and their corresponsive genes. ................................................................. 54
LIST OF FIGURES

Figure 1 Multi-directional interactions among several breast cancer-related molecules-TFAP2A, TFAP2C, p53, MUC1, p21, Estrogen Receptor, HER2 and Progesterone Receptor .......................................................... 19

Figure 2 Percentage of TFAP2C target genes in gene clusters characterizing different breast cancer phenotypes ............................................................................. 51

Figure 3 Visual displays of TFAP2C binding peaks from part of chromosome 1. TFAP2C binding pattern are identical on most genomic regions among different HCT 116 p53 -/- cells transfections .................................................................................. 52

Figure 4 Sequence logo of TFAP2C consensus motif ........................................................................ 59

Figure 5 TFAP2C binding peak is significantly augmented on gene MUC1 with introduction of p53 ................................................................................................................ 60

Figure 6 TFAP2C binding peak is significantly augmented on gene DDX17 with introduction of p53 .................................................................................................................. 61

Figure 7 TFAP2C binding peak is significantly augmented on gene MYCBP2 with introduction of p53 .................................................................................................................. 62

Figure 8 TFAP2C binding peak is significantly augmented on gene FAM98B with introduction of p53 ................................................................................................................ 63

Figure 9 TFAP2C binding peak is significantly augmented on gene HDAC4 with introduction of p53 .................................................................................................................. 64

Figure 10 TFAP2C binding peak is significantly augmented on gene CNOT3 with introduction of p53 .................................................................................................................. 65

Figure 11 TFAP2C binding peak remained identical on gene FAM83A with introduction of p53 .................................................................................................................. 66

Figure 12 TFAP2C binding peak remained identical on gene ST6GALNAC2 with introduction of p53 .................................................................................................................. 67

Figure 13 TFAP2C binding peak remained identical on gene DLX4 with introduction of p53 .................................................................................................................. 68

Figure 14 Western analysis of artificially synthesized TFAP2C and p53 protein in TnT quick coupled transcription/translation system ........................................................................... 69

Figure 15 Summarization of sequences for probes of MUC1, DDX17, MYCBP2, FAM98B, HDAC4 and CNOT3 ........................................................................................................... 70
Figure 16 Titrating TFAP2C to optimize the protein amount in gel shift assay.............. 71

Figure 17 Localizing the TFAP2C binding sequences within MUC1 promoter under the ChIP-seq peak. ................................................................. 72

Figure 18 Localizing the TFAP2C binding sequences within DDX17 promoter under the ChIP-seq peak. ................................................................. 73

Figure 19 Localizing the TFAP2C binding sequences within MYCBP2 promoter under the ChIP-seq peak. ................................................................. 74

Figure 20 Localizing the TFAP2C binding sequences within FAM98B promoter under the ChIP-seq peak. ................................................................. 75

Figure 21 Localizing the TFAP2C binding sequences within HDAC4 promoter under the ChIP-seq peak. ................................................................. 76

Figure 22 Localizing the TFAP2C binding sequences within CNOT3 promoter under the ChIP-seq peak. ................................................................. 77

Figure 23 Summery and comparison of TFAP2C binding sites within those six probes. ................................................................................................. 78

Figure 24 p53 neither binds to MUC1 probe nor augments TFAP2C occupancy to TFAP2C TFBS...................................................................................... 79

Figure 25 Western analysis of protein expression of TFAP2A, TFAP2C, p53 and MUC1 in MCF-7 cell line transfected with siRNA interfering TFAP2A, TFAP2C, p53. ................................................................. 80

Figure 26 Fold changes of mRNA levels of TFAP2A, TFAP2C, p53 and MUC1 in MCF-7 cells transfected with siRNAs targeting TFAP2A, TFAP2C and p53. ......................................................................................... 81

Figure 27 Expected schematic diagram of the presence of p53 modifying CpG methylation of MUC1 promoter and enabling TFAP2C to load................ 82

Figure 28 Ideogram of MUC1 promoter region in DNA methylation assay. ............... 83

Figure 29 TFAP2C binding peaks on MUC1 gene in different cell lines................... 84

Figure 30 Western analysis of protein expression of WT p53, TFAP2C, and MU p53 in cell lines transfected with plasmids. ......................................................... 86

Figure 31 Percentage of methylated CpG in MUC1 promoter. .................................. 87

Figure 32 Percentage of methylated CpG in MUC1 promoter. .................................. 88

Figure 33 Percentage of methylated CpG in MUC1 promoter. .................................. 89
Figure 34 TFAP2C occupancy is blocked with the introduction of p53 on RPSA promoter region.
LIST OF ABBREVIATIONS

B.C.: Breast Cancer
C.C.: Colon Cancer
TFAP2C: Transcription Factor AP2C
AP-2: activating enhancing-binding protein 2 or transcription factors activating protein-2
TFBS: transcription-factor binding sites
ER: Estrogen Receptor
HER2: Human Epidermal Growth Factor Receptor-2
PR: Progesterone Receptor
GPX1: glutathione peroxidase 1
ChIP-seq: Chromatin Immunoprecipitation followed by massively parallel sequencing
MUC1-C: C-terminal subunit of MUC1
ATCC: American Type Culture Collection
qRT-PCR: Quantitative Real-Time polymerase chain reaction
WB: Western Blot
EMSA: Electrophoretic mobility shift assay
FBS: Fetal Bovine Serum
DMEM: Dulbecco's Modified Eagle Medium
shRNA: small hairpin RNA
CHAPTER I

INTRODUCTION

1.1. Breast cancer and classification of phenotypes of breast cancer

1.1.1. Breast cancer

Breast cancer is one of the leading causes of cancer death in women worldwide. In the United States, it is estimated by the National Cancer Institute that 1 out of 8 women will develop invasive breast cancer during her lifetime\(^1\). In 2011, American Cancer Society reported that about 39,970 women and 450 men in U.S. will die of breast cancer in 2012\(^2\).

1.1.2. Breast cancer categories

Breast cancer is broken into distinct phenotypes generally based on the expression of three receptor molecules, Estrogen Receptor alpha (ER-\(\alpha\)), Progesterone Receptor (PR), and the Human Epidermal Growth Factor Receptor 2 (HER2 or ERBB2) (as summarized in Table 1)\(^3\text{-}^4\). Each phenotype represents a distinct gene expression profile with distinct behavior and recurrence and death rates. Additionally, targeted therapies exist for ER and HER2 expressing tumors. Therefore, it is critical to decipher the biological mechanism of regulating carcinogenesis-related gene expression on the molecular level in each phenotype of breast cancer.
1.2. **AP-2 family (with focus on TFAP2C)**

1.2.1. **AP-2 family**

The transcription Factor Activator Protein 2 family (AP-2 family) is retinoic acid-inducible, developmentally regulated and is composed of five members in both human and mice: TFAP2A, TFAP2B, TFAP2C, TFAP2D and TFAP2E (also named as AP-2α, AP-2β, AP-2γ, AP-2δ and AP-2ε, respectively).

Structurally, the helix-span-helix domain in the carboxyl terminus is highly conserved among all AP-2 members. Together with the adjacent basic domain, it enables the formation of dimers (homo or hetero) and DNA binding. AP-2 protein directly binds to the promoter region of specific genes by recognizing the consensus sequence 5’-GCCN3GGC-3’, and regulates the gene expression. Functionally, AP-2 family proteins are expressed during early embryonic development and the differentiation of ectoderm, but each member functions in specific spatial and temporal pattern. AP-2 family members generally regulate cell growth, differentiation, apoptosis, and carcinogenesis.

In several types of cancers, the levels of AP-2 members are aberrant as compared with normal tissue, indicating that it is critical to maintain the balance of AP-2 protein level and the ratios of different isoforms for normal cell function. For instance, the elevated AP-2 is reported in seminomatous germ cell tumors, testicular carcinoma, cervical cancer and breast cancer, but the down-regulation of AP-2 is associated with deterioration of stage I cutaneous malignant melanoma and aggressive breast cancer. Notably, in breast tumor cells, the expression of AP-2 isoforms contributes to cell proliferation, interference with AP-2 function could lead to cell apoptosis induced by chemo- and radiation.
1.2.2. The activity of TFAP2C in breast carcinogenesis

To determine the function of specific AP-2 family members in breast cancer, I first specifically investigated on the trans-activation activity of TFAP2C. In various cell models, knocking down TFAP2C expression enhances cell proliferation but reduces cell migration and invasion\textsuperscript{16}. In mouse models, over-expressing TFAP2C reduces tumor initiation, elongates the latency but promotes tumor progression\textsuperscript{16-17}. In breast cancer patients, TFAP2C regulates the carcinogenesis–related genes and is associated with worse clinical outcomes\textsuperscript{18}. Furthermore, TFAP2C is shown to activate the expression of ER-\textalpha and HER2 by binding to its responsive elements on the promoter region of \textit{ESR1} and \textit{ERBB2} genes (encoding for protein ER-\textalpha and HER2, respectively)\textsuperscript{19-21}. Therefore, it is important to understand the transcriptional mechanisms and functional activity of TFAP2C in regulating cancer-related genes.
Table 1 Summary of cell lines used in this study and the categorization of breast tumor.

A. Classification of different phenotypes of breast cancer by hormone receptor status.

<table>
<thead>
<tr>
<th>Phenotypes</th>
<th>ER-α Status</th>
<th>HER2 Status</th>
<th>PR Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminal A</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Luminal B</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>HER2-positive</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Basal or Triple Negative</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

B. Categorization of the predictable outcomes of drug treatment in breast tumor phenotypes based on status of ER-α and HER2.

<table>
<thead>
<tr>
<th>ER-α Status</th>
<th>HER2 Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Drugs</th>
<th>Tamoxifen</th>
<th>Herceptin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respond</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>Not Respond</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

C. Cell lines involved in this thesis project.

<table>
<thead>
<tr>
<th>Cell Lines</th>
<th>Cancer Type</th>
<th>Phenotypes</th>
<th>ER-α</th>
<th>HER2</th>
<th>p53</th>
<th>MUC1 Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCF-7</td>
<td>B.C.</td>
<td>Luminal A</td>
<td>+</td>
<td>-</td>
<td>Wild Type</td>
<td>Medium</td>
</tr>
<tr>
<td>BT-474</td>
<td>B.C.</td>
<td>Luminal B</td>
<td>+</td>
<td>+</td>
<td>Mutant (E285K)</td>
<td>Low</td>
</tr>
<tr>
<td>MDA-MB-453</td>
<td>B.C.</td>
<td>HER2</td>
<td>-</td>
<td>+</td>
<td>Mutant (30 bp Deletion in exon 11)</td>
<td>Extremely Low</td>
</tr>
<tr>
<td>SKBR-3</td>
<td>B.C.</td>
<td>Luminal B</td>
<td>-</td>
<td>+</td>
<td>Mutant (R175H)</td>
<td>Medium</td>
</tr>
<tr>
<td>HCT 116</td>
<td>C.C.</td>
<td>Luminal B</td>
<td>/</td>
<td>/</td>
<td>Absence of p53</td>
<td>/</td>
</tr>
<tr>
<td>HCT 116</td>
<td>C.C.</td>
<td></td>
<td>/</td>
<td>/</td>
<td>Wild Type</td>
<td>/</td>
</tr>
</tbody>
</table>

Note: Herceptin is monoclonal antibody trastuzumab®

B.C.: Breast Cancer; C.C.: Colon Cancer; PR: Progesterone
1.2.3. The molecular mechanisms of regulating TFAP2C occupancy and its trans-activation activity

Given the significance of TFAP2C in breast cancer carcinogenesis, the molecular basis determining its binding target genes and trans-activation activity in various breast cancer cell lines must be determined. In spite of the great efforts in the field to understand the molecular mechanism, answers to the following questions are fully understood: 1) how are the binding activity and trans-activation activity of TFAP2C regulated? 2) why is TFAP2C able to activate different genes in various phenotypes? Previous research has revealed several molecular mechanisms, such as protein modification, cofactor interaction, and the epigenetic modification of TFAP2C target genes.

1.2.3.1. Protein modification

Protein modification is one of the ways to modulate the activity of AP-2 protein. Garcia et al. demonstrated that the PKA-induced phosphorylation stimulates AP-2 activity on APOE promoter in vivo\(^22\). Protein casein kinase 2 is also shown to phosphorylate TFAP2A, resulting in its increased stability and elevated transcriptional activity\(^23\). Moreover, Eloranta et al. identified that sumolation residing in the DNA binding and dimerization domain of AP-2 suppresses its trans-activation activity\(^24\). Unpublished data in our lab also suggests sumolation is one of the mechanisms to explain the functional specificity of TFAP2A and TFAP2C.

1.2.3.2. Interaction with cofactors

Several cofactors alter the functional activity of TFAP2C or other AP-2 members. Mitchell et al. reported that SV40 Large T antigen blocks DNA binding of AP-2 by direct protein-protein interaction in vitro\(^25\). In particular to TFAP2C, Wwox tumor suppressor
physically binds to the proline-rich domain of TFAP2C, triggers the latter to export from
the nucleus to the cytoplasm and blocks the trans-activation activity of TFAP2C in
several human cell lines\textsuperscript{26}. Ku70 and Ku80 interact with both TFAP2A and TFAP2C,
leading to their enhanced trans-activation activity and upregulation of HER2 in breast
cancer cell lines\textsuperscript{27}.

For other AP-2 isoforms, similar findings have been published. Zou \textit{et al.} reports
the over-expression of Aurora-A, a serine/threonine protein kinase, reduces the stability
of TFAP2A, and therefore suppresses its transcription activity by physical interaction\textsuperscript{28}.
Oncoprotein DEK, as a co-activator, enhances the DNA binding activity of TFAP2A \textit{in
vitro} and stimulates trans-activation activity of TFAP2A in human cells (brain HTB-14
and liver HB-8065)\textsuperscript{29}. Sivak \textit{et al.} found that the transcription factor Pax6 tethers on
TFAP2A at \textit{gelB} promoter region and cooperatively controls \textit{gelB} expression as well as
eye development\textsuperscript{30}. Another example is GAS41, as a transcriptional co-activator for
TFAP2B, stimulates the transcription activity of TFAP2B actively to regulate the
artificial reporter gene. In addition, the critical residues for the protein-protein interaction
were identified\textsuperscript{31}. There are several papers reporting that various proteins interact with
AP-2 isoforms non-specifically and modify the transcriptional activity of AP-2 on certain
target gene. Those AP-2 interacting proteins includes but not limited to: potassium
channel tetramerization domain-containing 1 (as a suppressor of AP-2)\textsuperscript{32}, transcription
factor Yin Yang 1\textsuperscript{33}, mitochondrial cytochrome c\textsuperscript{34}, polyADP-ribose polymerase\textsuperscript{35}, p53\textsuperscript{36},
and p53 together with YB-1\textsuperscript{37}. 
1.2.3.3. Epigenetic modification of AP-2 target genes

Besides the protein modification and cofactor interaction, epigenetic chromatin structure is correlated to AP-2 activity on specific target genes. In particular to TFAP2C, our lab has identified that epigenetic modification within the promoters of two AP-2 target genes (*ESR1* encoding ER-α, and *GPX1* encoding glutathione peroxidase) determines their availability for TFAP2C occupancy and is associated with protein expression. For instance, TFAP2C binds to the promoter region of the *ESR1* gene and enhances the expression of ER-α in ER-α positive breast cancer, but not in ER-α negative breast cancer. In this case, the discrepancy of TFAP2C occupancy between cell lines is explained by DNA methylation and histone acetylation status on the promoter region, in which the hypermethylation and deacetylation severely decrease the TFAP2C occupancy. Another example is *GPX1* gene. The regulatory effect of TFAP2C on the expression of glutathione peroxidase is also controlled epigenetically. In the breast cancer cell lines whose *GPX1* transcription is not responsive to the over-expression of TFAP2C, the CpG islands encompassing the TFAP2C binding site within the *GPX1* promoter is hypermethylated. The association is explained by that TFAP2C losses its ability to occupy this region (shown by Chromatin Immunoprecipitation followed by parallel sequencing dataset), leading to the down-regulation of *GPX1* expression (unpublished data by M. V. Kulak and R. Weigel).

Overall, the three mechanisms described above regulate the trans-activation activity of TFAP2C on different levels. In this project, we will determine the novel cofactor potentially mediating TFAP2C activity.
1.3. **p53 is a potential TFAP2C cofactor**

1.3.1. **p53 and p53 variants**

p53 is a prevalent tumor suppressor with the functional capacity to regulate many genes involved in cell cycle arrest, DNA repair, senescence, apoptosis, angiogenesis or changes in metabolism\textsuperscript{38-39}. Structurally, p53 is a DNA-binding protein containing an amino-terminal transcription activation domain, a DNA-binding and tetramerization domain and carboxyl terminal regulatory domains\textsuperscript{40}. p53 mutations are found in 20\%-35\% of breast tumors. Generally, these p53 mutants are more stable and are elevated in multiple types of tumors compared to the level in normal tissue\textsuperscript{38}. However, the molecular mechanisms behind the carcinogenetic effect of p53 mutants are still not well understood\textsuperscript{40-41}.

1.3.2. **The link between p53 and ER-\(\alpha\) or ERBB2**

p53, ER, and ERBB2, together with TFAP2C, are all breast cancer-related transcriptional factors. There are regulations and interactions in multi-directions between them (the network was summarized in Figure 1). Understanding the complex interaction of those molecules in this network is critical to begin the thesis project.

1.3.2.1. **p53 regulates ER-\(\alpha\) expression**

Shirley \textit{et al.} found that in MCF-7 cell line (ER-positive), p53 regulates ER expression by the transcriptional control of the \textit{ESR1} promoter, and by forming a transcription multifactor complex with CARM1, CBP, c-Jun, RNA polymerase II and Sp1 on the \textit{ESR1} promoter\textsuperscript{42}. Fuchs-Young \textit{et al.} used mouse models to demonstrate that the transcriptional regulation of p53 on ER expression and function is operational in
mammary tumorigenesis in vivo\textsuperscript{43}. Furthermore, Contant et al. suggests that inactivating p53 mutations lead to different transcriptional changes in ER negative cell lines and therefore may carry different prognostic value in ER positive and ER negative cell lines\textsuperscript{44}. Rasti et al. suggested that DNA methylation of ESR1 promoter by overexpressed p53 mutations may account for silencing ER expression in ER-negative cell line\textsuperscript{45}.

1.3.2.2. **ER-α inhibits p53**

Konduri et al. recently discovered the interaction of ER-α and p53 on the reverse direction. In stem cell-containing murine mammospheres, ER-α binds to p53 and inhibits the latter from activating p21 gene (encoding for cyclin kinase inhibitor, which is a key cell cycle regulator)\textsuperscript{46}. They also suggested that the wild type p53 is important to determine positive therapeutic response\textsuperscript{46}. This, together with the finding of the regulation of p53 on ER, suggests a potential role for the ER-p53 interaction in mammary carcinogenesis.

1.3.2.3. **Coexistence of p53 mutant and high HER2**

In breast cancer cell lines or patient samples, the presence of p53 mutants is highly frequently associated with the over-expression of HER2\textsuperscript{47-49}. Moreover, breast cancer patients who possess this coexistence of p53 mutation and high HER2 are suspected for bad prognosis and poor survival\textsuperscript{50-51}. Though the molecular mechanism behind the phenomena is still uncovered, we believe the interaction between p53 and those molecules is worthy to be investigated to help us understand breast cancer carcinogenesis.

Overall, considering that both ER-α and HER2 are TFAP2C target genes and are known to be predictive for the clinical phenotypes of breast cancer (as stated in 1.1.2 and
1.2.2), we believe it deserves efforts to investigate on the potential interaction of p53 and TFAP2C.

1.3.3. **p53 and AP-2 family**

1.3.3.1. **AP-2 alters p53 stability and p53 trans-activation activity**

1.3.3.1.1. TFAP2A decreases p53 stability and increases p53 trans-activation activity on p21 (two conflicting mechanisms)

The first finding of the interaction between AP-2 and p53 was revealed using yeast two-hybrid assay by Weigel lab in 2002\(^{52}\). They demonstrate that radiolabeled p53 interacts GST-TFAP2A and GST-TFAP2C *in vitro* and that the interaction of p53 and TFAP2A is confirmed by gel-shift following immunoprecipitation from MCF7 nuclear extract by TFAP2A-specific antibody. They then substantiated the expression of TFAP2A increased p53-mediated trans-activation on both an artificial construct and cyclin kinase inhibitor p21\(^{\text{WAF1/CIP1}}\) (a key cell cycle regulator), both were p53 responsive. However, the R175H p53 mutant, which lost its direct DNA-binding ability, was unable to activate the reporter. TFAP2A, requiring the wild type p53, not only activates p21 expression but also induces cell cycle arrest associated with *p21*’s function in HCT 116 cell lines\(^{52}\). Later in 2006, Stabach *et al.* determined the region of p53 (aa305-375) interacting with TFAP2A and identified that the over-expression of AP-2 led to destabilization of p53. Moreover, upregulation of TFAP2A level significantly reduced the half-life of p53 and p21 expression was thereby decreased\(^{53}\).
1.3.3.1.2. TFAP2B stabilizes p53 and enhances p53 trans-activation activity

Similar to TFAP2A, Hu et al. recently demonstrated that TFAP2B enhances the stability of p53 protein, increases p53 trans-activation activity on CRYAB gene encoding for αB-crystalline. This indicates that TFAP2B acts as a coactivator of p53.

1.3.3.2. p53 regulates TFAP2A and TFAP2C expression

Besides TFAP2A altering the stability and trans-activation activity of p53, Li et al. revealed that both TFAP2A and TFAP2C are transcriptional targets of wild type p53. They showed over-expression of p53 induces upregulation of TFAP2A and TFAP2C at mRNA and protein levels, whereas the R175H mutant, which cannot bind to DNA, had no such effect. They also localized the p53-binding site on AP-2 promoter and found that p53 facilitates AP-2 transcription by opening chromatin structure.

1.3.3.3. Co-activation of p53 and AP-2 on target genes

Modugno et al. first published that the interaction between p53 and AP-2 family members can regulate AP-2 family activity. They found that both wild type and mutant p53 (can not trans-activate target gene) repress the trans-activation activity of AP-2 on RPSA gene (encoding for laminin receptor precursor) by associating physically in the ovarian carcinoma cell lines. Hannay et al. substantiated that the reconstitution of wild-type p53 suppresses Rad51 promoter activity (a gene encoding for Rad51 protein that assists in repair of DNA double strand breaks) through increasing AP-2 occupancy to cis-element within the Rad51 promoter in human soft tissue sarcoma cell. Similarly, the binding sites of AP-2, p53 and YB-1 sit near each other within the -40 bp to -10 bp upstream of MMP2 gene encoding gelatinase A. They also showed that, in hepatoma cells, AP-2 and p53 in combination greatly increased MMP2 promoter activity. The
activity is further enlarged when YB-1 is included. Moreover, p53 and AP-2 were shown to activate KAI1 by binding to promoter directly but no protein-protein interaction was referred to within this paper.

Taken these data together, we propose that there is a feedback loop between AP-2 and p53. It is possible that the functional interaction of p53 and TFAP2C contributes to mechanisms of carcinogenesis and progression of cancer. Specifically, I am interested in the mechanism of the alteration of TFAP2C binding profile mediated by p53.

1.4. **MUC1, a gene model to illuminate that p53 regulates binding and trans-activation activity of TFAP2C**

*MUC1* was selected as the model gene to investigate the impact of p53 on enhancing TFAP2C occupancy and transcriptional activity.

1.4.1. **Protein character of MUC1**

MUC1 (or MUCIN1, Mucin-1) is a member of the transmembrane mucin family. Mucins are heavily glycosylated proteins with high-molecular weight. Some mucins are membrane-associated, localizing on the apical surface of epithelial cells, and others are secreted, forming a gel-like local microenvironment. It is hypothesized that mucins protects epithelial cells against the damage from environment factors such as extremes of pH, various chemicals, dehydration, and hypoxia. Mucins function as physical barriers to protect the cell surface from harsh environment, such as in liver, pancreas, gall bladder and other organs with mucosal surfaces. In response to extracellular stimuli, mucins might act as cell-surface receptors and sensors and conduct signals. Presumably, the aberrant amount of mucins on malignant cell surface creates a relatively stable local
microenvironment, helping the tumor cell survive in the harsh conditions during inappropriate invasive and metastatic growth\textsuperscript{60}.

*\textit{MUC1*} gene encodes protein MUC1. Generally, MUC1 is a membrane-bound protein whereas some isoforms of MUC1 are secreted\textsuperscript{62}. A specific proteolytic cleavage during post-translational processing cleaves MUC1 into two subunits, MUC1-C and MUC1-N. The two subunits form a tight, non-covalent hetero-dimer, remaining associated at cell surface\textsuperscript{62}. MUC1-C subunit also translocates to the nucleus and mitochondria\textsuperscript{63}.

Under normal conditions, this complex is located only at the apical border of the secretory epithelial cells. When those cells lose their polarity, MUC1 complex will translocate to the entire cell surface, leading to the loss of tight junction, cell-cell interaction, and the enhanced cell proliferation and migration\textsuperscript{62,64}. Also, tumor cells might utilize anti-adhesive effect of MUC1 to detach from the tumor mass and surrounding stroma and to initiate invasion\textsuperscript{60}. Additionally, MUC1 can facilitate the attachment of the breast tumor cells to distant metastatic sites\textsuperscript{65}.

Notably, MUC1 is over-expressed in almost 90\% of breast cancer\textsuperscript{66}. It is one of the several prevalent markers for stem cancer cells and disseminated tumor cells (DTC)\textsuperscript{67}. Furthermore, the over-expression of MUC1 is detectable in the serum of breast cancer patients and is associated with a poor prognosis. Given its great biomedical and clinical signature, MUC1 is believed to be a highly attractive target of clinical breast cancer studies. For instance, it is a target to antibody detection assay for prognosis, a promising vaccine target for immunotherapy, and a marker inducing gene expression in accordance to tumor characteristics and to clinical outcome prediction\textsuperscript{68-71}. MUC1 can alter the
transcription of a list of genes and the expression pattern of these MUC1-induced genes is in accordance with deterioration of breast and lung cancer survival.

1.4.2. Interactions of MUC1 with other breast cancer related molecules

1.4.2.1. The interaction of MUC1 and ER-α.

Using TRANSFAC and TSSG, Zaretsky et al. predicted numerous cis-elements for all potential transcription factors within a 2872 bp fragment of MUC1 promoter. They confirmed that the ER-α binds to its responsive element in the MUC1 promoter in vitro. Also, it was demonstrated that MUC1/SEC responds positively to estrogen, possibly by the integration of ER-α to the ER-α binding site on MUC1 promoter in human breast cancer cell lines (ER-α positive). Interestingly, another isoform, MUC1/TM fails to respond to estrogen treatment. Further studies revealed that MUC1 is also capable of modulating the activity of ER-α. Wei et al. identified the physical interaction between MUC1-C and ER-α. Moreover, MUC1-ER-α complex binds to estrogen-responsive elements on promoter regions, enhances occupancy and activity of the promoters and consequently elevates the survival of breast cancer cells.

1.4.2.2. The interaction of MUC1 and p53

MUC1-C directly binds to the regulatory domain of p53 in the nucleus and the interaction becomes stronger in response to DNA damage. Interestingly, at the p53-responsive element of the p21 gene, MUC1 is associated with p53 and able to coactivate p21 transcription. Therefore, the association of MUC1 with p53 in cancer activates the cell cycle arrest and survival response to DNA damage. This indicates that human
tumors with over-expressed MUC1 might have a survival advantage to genotoxic stress. Additionally, it has been uncovered that MUC1-C represses p53 expression by enhancing the recruitment of two repressors targeting to p53 promoters, Kruppel-like factor 4 and histone deacetylases\textsuperscript{76}.

On the other hand, p53 directly targets to two p53 elements on \textit{MUC2} promoter and therefore activates the \textit{MUC2} gene under exposure to various forms of stress in various cell lines, including human colon cell line (DLD-1), breast cancer cell line (MCF-7), and the lung cancer cell line (A427)\textsuperscript{77}.

\textbf{1.4.2.3. \textit{MUC1} is regulated by other factors}

It was shown that progesterone receptor (PR) and proinflammatory cytokines (TNF\(\alpha\) and interferon \(\gamma\))-activated factors cooperatively regulate \textit{MUC1} gene expression in human uterine epithelial cell line\textsuperscript{78}. In a breast cancer cell line, GATA3, NFkappaB, p65, and STAT1\(\alpha\) were demonstrated to regulate \textit{MUC1} expression acting as mediator (GATA3), or transcription factors (NFkappaB, p65, and STAT1\(\alpha\)), respectively\textsuperscript{79-80}.

\textbf{1.4.3. Mucin family and AP-2 family}

Previous findings have reported a role of the AP-2 family on regulating the expression of the mucin family. In human pancreatic tissue, TFAP2A suppresses MUC4 expression and the corresponding responsive element on \textit{MUC4} promoter has been identified. Moreover, over-expression of TFAP2A represses MUC4 expression and thereby significantly decreases cell proliferation and invasion\textsuperscript{81}. Another mucin family member, MUC8, was shown to be induced by treatment with phorbol 12-myristate 13-acetate (PMA) in an airway epithelial cell line where TFAP2A is involved in this
These findings agree with our hypothesis and led us to investigate the potential role of TFAP2C on the activating MUC1 transcription.

1.5. **Promoter CpG methylation-the potential mechanism to explain enhanced TFAP2C occupancy on MUC1 promoter induced by p53**

1.5.1. **CpG methylation and tumor**

CpG methylation refers to the addition of a methyl group to the 5 position of the pyrimidine ring of cytosine within a CpG dinucleotide. CpG methylation within the promoter region epigenetically regulates gene transcription as a marker for silencing in general. In many human tumors, aberrant DNA methylation patterns have been identified, including CpG island promoter hypermethylation of tumor suppressor genes and global genomic hypomethylation.

1.5.2. **Why examining DNA methylation**

We hypothesize that DNA methylation of the MUC1 promoter is altered by p53, therefore TFAP2C occupancy at its responsive element on MUC1 promoter is increased. The reasons supporting our hypothesis are listed as follows. First, CpG methylation at the promoter blocks TFAP2C binding. As described in 2.3.3, CpG methylation of ESR1 and GPX1 promoter are negatively associated with TFAP2C occupancy and protein expression. Additionally, one CpG dinucleotide is nine bases away from one of the two predicted AP-2 binding motifs in MUC1 promoter (Figure 17). Therefore, we wondered whether CpG methylation occurs near the AP-2 regulatory region of the MUC1 promoter.
Second, \textit{MUC1} is defined as an epigenetically controlled gene. MUC1 expression level varies dramatically among different breast cancer cell lines\textsuperscript{86}. Yamada \textit{et al.} found that the epigenetic changes, including CpG methylation and histone H3 lysine 9 modifications on \textit{MUC1} promoter explained the low transcription level of \textit{MUC1} in MUC1-negative breast cancer cell lines\textsuperscript{87}. Therefore, we propose that the status of CpG methylation defines the MUC1 expression.

1.6. Overview of thesis

Transcription factor AP2C (TFAP2C) is believed to be involved in breast cancer carcinogenesis, but the molecular mechanisms of regulating its trans-activation activity are not well understood. Within this thesis project, I first analyzed ChIP-seq datasets and found that the introduction of p53 to HCT 116 p53 \textendash/- cell line significantly augments TFAP2C occupancy on promoter regions of a group of genes (associated with 115 peaks). Of these, six genes were further investigated, \textit{MUC1}, \textit{DDX17}, \textit{MYCBP2}, \textit{FAM98B}, \textit{HDAC4} and \textit{CNOT3}. First, TFAP2C binding sites were identified in the center of ChIP-seq peaks on the promoters of the six genes and these sites were verified by gel shift assays. One of these genes, \textit{MUC1}, was then determined to be activated by TFAP2C in MCF-7 cell line. Subsequently, \textit{MUC1} was selected as the model target gene to elucidate the mechanism for p53-mediated enhancement of TFAP2C occupancy. We hypothesized that DNA methylation of the \textit{MUC1} promoter is altered by p53, and therefore \textit{MUC1}’s binding site occupancy for TFAP2C is increased. To examine this, CpG methylation assay was performed, but the result showed the DNA methylation of \textit{MUC1} promoter region remains identical with or without over-expression of p53 in HCT 116 p53 \textendash/- cell line. From these studies I conclude that 1) introduction of p53 augments TFAP2C binding
on specific gene targets; 2) *MUC1* gene is activated by TFAP2C and a TFAP2C binding sites was verified; 3) promoter DNA methylation does not explain the increased binding site occupancy of *MUC1* for TFAP2C.

To sum up, this study shows that p53 alters TFAP2C occupancy on certain target genes and advances our understanding of the interaction between the two. This study may possibly reveal the broad long-term objectives – studying the significance of TFAP2C as a key indicator on the molecular classification of breast cancer. This project eventually may contribute to provide vital insights to predict outcomes of breast cancer drug-therapy by translational medical methods.
Figure 1 Multi-directional interactions among several breast cancer-related molecules- TFAP2A, TFAP2C, p53, MUC1, p21, Estrogen Receptor, HER2 and Progesterone Receptor.

ER: Estrogen Receptor. PR: Progesterone Receptor. Solid red line indicates that protein suppresses the promoter activity of its target gene following the direction of the arrow. Solid green line indicates protein activates the promoter activity of its target gene. Dashed red line indicates protein suppresses its target protein. Dashed green line indicates protein activates its target protein.
CHAPTER II

P53-MEDIATED AUGMENTATION OF TFAP2C BINDING OCCUPANCY ON SPECIFIC GENES

2.1. Materials and Methods

2.1.1 Cell lines and culture

All cell lines were obtained from American Type Culture Collection (ATCC, Rockville, MD), except that HCT 116 p53-/- colon cell line was a generous gift from The Vogelstein lab. Cells were cultured at 37°C with 5% CO2, except that MDA-MB-453 cell line was cultured without CO2. Subculture and preservation were performed following ATCC protocol. The basic media used to culture cells are listed in Table 2, with supplementary 10% Fetal Bovine Serum.

Table 2 Summary of cell line–specific culturing media.

<table>
<thead>
<tr>
<th>Cell Lines</th>
<th>Basic Media</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCF-7</td>
<td>DMEM</td>
</tr>
<tr>
<td>BT-474</td>
<td>RPMI 1640</td>
</tr>
<tr>
<td>MDA-MB-453</td>
<td>Leibovitz’s L-15</td>
</tr>
<tr>
<td>SKBR-3</td>
<td>McCoy’s 5a</td>
</tr>
<tr>
<td>BT-20</td>
<td>DMEMF12</td>
</tr>
<tr>
<td>HCT 116 P53-/-</td>
<td>DMEM</td>
</tr>
<tr>
<td>HCT 116 P53+/+</td>
<td>DMEM</td>
</tr>
</tbody>
</table>

Note: DMEM: Dulbecco’s Modified Eagle Medium

All media is from Gibco® Life Technologies
2.1.2. Transfection of siRNA

siRNA was utilized to silencing transcription of TFAP2A, TFAP2C, p53. siRNAs for Non-Targeting, TFAP2A, TFAP2C and p53 were obtained from DharmaFECT Transfection Reagents (Thermo Scientific). SiTFAP2A: siGENOME siRNA human TFAP2A, D-006348-03; siTFAP2C: siGENOME siRNA human TTFAP2C, D-005238-01; siNon-Targeting: siGENOME Control siRNA Non-Targeting siRNA #2, D-001210-02-30; sip53: siGENOME SMART pool, Human TP53, M-003329-03-0005.

Transfecting siRNA into MCF-7 and BT-474 breast cancer cell lines was performed using Lipofectamine™ RNAiMAX (Invitrogen) following the protocol “Transfecting Stealth™ RNAi or siRNA into MCF7 cells, using Lipofectamine™ RNAiMAX”. http://tools.invitrogen.com/downloads/MCF7_RNAiMAX.pdf (Table 3). 30 pmoles of siRNA targeting each molecule was transfected per well (962 mm²) in the 6-well tissue culture dishes (Falcon) when the cells reached to 20% to 30% confluence (0.3X10⁶ cells per well). Two controls of siNon-Targeting at different dose (30 pmole or 60 pmole per well) was used as normalization.

Table 3 Transfection of siRNAs to silence TFAP2A, TFAP2C and p53 in MCF-7 cell lines.

<table>
<thead>
<tr>
<th>siRNA Transfection Conditions</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>NT</td>
<td>TFAP2A</td>
<td>TFAP2C</td>
<td>p53</td>
<td>TFAP2A+p53</td>
<td>TFAP2C+p53</td>
<td></td>
</tr>
</tbody>
</table>

Note:

NT: Non-Targeting
96 hours post-transfection with siRNA, mRNA and protein were harvested using QIAGEN RNeasy kit (QIAGEN) and RIPA Lysis Buffer (Millipore), respectively, according to the manufacturer. All transfections were carried out in technical replicate and biological triplicate.

2.1.3. Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

For RT-PCR analysis of individual gene, total RNA was harvested using RNeasy® Mini Kit (QIAGEN®). Reverse transcription was performed using High-Capacity cDNA Reverse Transcription Kits (Applied Biosystems, Life Technologies).

Quantitative RT-PCR was conducted using cDNA as template, following the protocol of Taqman® Fast Universal PCR Master Mix (2x) (No AmpErase UNG®, Applied Biosystems, Life Technologies). Probes and primers were purchased from Taqman® Gene Expression Assays (VIC® / MGB Probe, Applied Biosystems, Life Technologies), listed as follows. MYCBP2 Assay: Hs00209335_m1. FAM98B Assay: Hs00543341_m1. DDX17 Assay: Hs00428757_m1. HDAC4 Assay: HS01041648_m1. CNOT3 Assay: HS00248115_m1. CD44 Assay: HS01075861_m1. Mucin1 Assay: Hs00159357_m1. TFAP2A Assay: Hs01039413_m1. TFAP2C Assay: Hs00231476_m1. TP53 Assay: Hs01034249_m1. FAM83A Assay: Hs00999384_m1. ST6GALNAC2 Assay: Hs01032565_m1. DLX4 Assay: Hs00231080_m1. Endogenous control HuB2M Beta-2-Microglobulin (Assay: 4326319E), and Eukaryotic 18S rRNA (Assay4319413E) (VIC® / MGB Probe, Primer Limited, Applied Biosystems, Life Technologies), were used for normalization of target gene expression.
The machines used to perform RT-PCR were either StepOne Plus™ Real-Time PCR Systems (Applied Biosystems) or ABI PRISM 7900 sequence detection system (Applied Biosystems) in the DNA Core Facility at the University of Iowa. Cycling conditions were: 10 min at 95 °C followed by 40 repeats of the cycle: 95 °C for 60 s, 60 °C for 20 s.

Quantitative RT-PCR results were expressed as mean values with standard errors generated from technical triplicate and biological triplicate. Statistical analysis was performed combining two methods, StepOne software V2.2 (provided by Applied Biosystems), or following the User Bulletin #2 (with ABI Prism 7700 Sequence Detection System) [http://dna-9.int-med.uiowa.edu/files/Compar_Anal_Bulletin2.pdf](http://dna-9.int-med.uiowa.edu/files/Compar_Anal_Bulletin2.pdf). Student-T test was performed to examine the significant difference.

### 2.1.4. Western Blot

Protein concentrations were measured using Pierce® BCA Protein Assay (Thermo Scientific). Western Blots were carried out according to the lab standard protocol, using the following antibodies: Mucin 1 (C-Mu1) antibody: sc-53376, p53 (DO-1) antibody: sc-126, p53 (DO-2) antibody: sc-53394, AP-2α (3B5) antibody: sc-12726, AP-2γ (6E4/4) antibody: sc-12762, AP-2γ (V-18) antibody: sc-31935 and GAPDH (6C5) antibody: sc-32233 (Santa Cruz); MUC1 isoform 7 antibody: 1797-1 and AP-2α (EPR2688-2, Epitomics).

### 2.1.5. Chromatin Immunoprecipitation (ChIP)

**Assays and ChIP-seq Data Analysis**

To assess the genomic map of TFAP2C binding targets in different cell lines, Chromatin Immunoprecipitation followed by massively parallel sequencing (ChIP-seq)
was performed following the Myers lab protocol (ChIP-seq Protocol, v042211.1 and v042211.2). TFAP2C occupancy were identified using ChIP-seq in four different breast cancer cell lines (MCF-7, BT-474, MDA-MB-453 and SKBR-3), as well as in colorectal cancer cell line HCT 116 p53 -/- with different expression profile of p53 and TFAP2C (over-expressed TFAP2C, overexpressed TFAP2C and wild-type p53, over-expressed TFAP2C with EEqQ p53 mutant). EEqQ p53 mutant was created in the lab, which remains the transcriptional activity but the ability to directly interact with TFAP2C protein is completely eliminated. ChIP procedure and data analysis were described in previous published paper. ChIP pulled down with IgG antibody was performed as negative control in each case. Here list the data sets of TFAP2C ChIP-seq involved in the thesis project (Table 4):

Table 4 Summary of ChIP-seq datasets used in this project.

<table>
<thead>
<tr>
<th>Cell Lines</th>
<th>Transfect Conditions</th>
<th>ChIP-seq Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 MCF-7</td>
<td>N/A</td>
<td>MCF7/END/AP2C MM6E4/4 (72)</td>
</tr>
<tr>
<td>2 BT-474</td>
<td>N/A</td>
<td>BT474/END/AP2C (164)</td>
</tr>
<tr>
<td>3 MDA-MB-453</td>
<td>N/A</td>
<td>MDAMB453/END/AP2C (164)</td>
</tr>
<tr>
<td>4 SKBR-3</td>
<td>N/A</td>
<td>SKBR3/END/AP2C (164)</td>
</tr>
<tr>
<td>5 HCT 116 p53</td>
<td>TFAP2C</td>
<td>HCT116-/-AP2C/AP2C MM6E4/4 (99)</td>
</tr>
<tr>
<td>6 HCT 116 p53</td>
<td>TFAP2C + p53</td>
<td>HCT116-/-p53AP2C/AP2C (99)</td>
</tr>
<tr>
<td>7 HCT 116 p53</td>
<td>TFAP2C + EEqQ p53 mutant</td>
<td>HCT116-/-EqEqp53AP2C/AP2C (99)</td>
</tr>
</tbody>
</table>

Note: EEqQ p53 mutant cannot interact with TFAP2C.
To compare TFAP2C occupancy with different expression profile of p53 in HCT 116 cell lines, Tiandao Li and Thomas B. Bair analyzed the ChIP-seq datasets Number 5 and 6 (Table 4.) utilizing three individual peak prediction programs (Partek motif prediction, Findpeaks (v4.0.12) (http://sourceforge.net/apps/mediawiki/vancouvershortr/index.php) and cisGenome (v 1.2)\textsuperscript{89}, as described in previous paper\textsuperscript{59}. The criteria defining a TFAP2C binding peak is 1) the peak is unambiguously associated with a known gene if the peak was within 20 kb prior to the transcription start site, within 5’ flaking region to or in one of the first four introns with exons excluded; 2) the TFAP2C binding peak height is increased at least two-fold with presence of p53, normalized to the nearby background. To minimize the ambiguity of mapping short monotags to the genome and inability to distinguish the true ChIP enrichment from amplified noise, I cooperated with two programmers, generated peak pools overlapped in these two lists and eventually verified the 114 real binding loci with high confidence one by one. The presence of p53 significantly increases TFAP2C to bind to those targets.

\textbf{2.1.6. Gel-Shift Assay}

Gel-shift assay was performed to examine: 1) The specific binding sites of TFAP2C on promoter regions of \textit{MUC1, DDX17, MYCBP2, FAM98B, HDAC4} and \textit{CNOT3}; 2) The potential interaction between p53 and TFAP2C augments TFAP2C occupancy on TFAP2C \textit{cis}-elements on those targets.
2.1.6.1. **Competitors preparation**

Oligonucleotides were ordered from IDT (Integrated DNA Technologies), and were resuspended to 50 µM. Double strand DNA competitors were created by annealing each pair of oligonucleotides in a single-step thermal cycle: heated up to 95°C and cooled down to room temperature. The sequences of oligonucleotides are listed in the Appendix and double or triple nucleotides mutations are underlined.

2.1.6.2. **Probe preparation**

Genomic DNA of HCT 116 p53 +/- (1000 ng/µL) was used as template to amplify ~100 bp fragments of the promoter region of genes *MUC1, DDX17, MYCBP2, FAM98B, HDAC4* and *CNOT3* using Platinum® PCR SuperMix (Invitrogen). PCR thermal cycle, generally, was 95°C for 30 s, annealing for 30 s, 72°C for 60 s, 4°C pause. The annealing temperature and the cycle number depend on individual target, as listed below. *MUC1*: 55°C, 35 cycles. *DDX17*: 50°C, 30 cycles. *MYCBP2*: 54°C, 30 cycles. *FAM98B*: 57°C, 30 cycles. *HDAC4*: 57°C, 30 cycles. *CNOT3*: 54°C, 30 cycles. The primers used to amplify these regions are listed as follows:

- **MUC1_F** 5’-ACAGGACCTCGACCTAGCTGGCTT-3’
- **MUC1_R** 5’-GAGGGGCAGTCTGGAACTTGCGC-GGC-3’
- **DDX17_F** 5’-TACGTAAACAGCGCCTCGGA-3’
- **DDX17_R** 5’-CGCAAAATACACCCATCGTC-3’
- **MYCBP2_F** 5’-CCCTTTTCTCTCCCCACCCCCC-3’
- **MYCBP2_R** 5’-AGAAGTGGACTACTTCAGAT-3’
- **FAM98B_F** 5’-TTATTGCCTCGGGGCTGGGA-3’
- **FAM98B_R** 5’-AAGTAGCCCGCCGGAAACTA-3’
PCR fragments were cut from the agarose gel, extracted by QIAquick® Gel Extraction Kit, (QIAGEN®) and diluted to final concentration of 6~8 ηg/ µl. Then oligonucleotides were radio-labeled by phosphorylation reaction using T4 Polynucleotide Kinase (M0201S, Biolabs). The recipe, modified based on the protocol from the manufacturer, is listed below: Consensus Oligonucleotides (6~8 ηg/ µl, 14 µl), T4 Polynucleotide Kinase 10X buffer (2.5µl), [γ-32 P] ATP (5 µl at 10 mCi/ml, 3000 Ci/µmol, PerkinElmer), and T4 Polynucleotide Kinase (3 µl, 10,000 U/µl, Biolabs). The mixture was incubated at 37°C for 30 min and then 65°C for 20 min.

Finally, unincorporated label was removed from oligonucleotide by Quick Spin Columns (G-25 Sephadex Columns for Radiolabeled DNA Purification, exclusion limit: <10 base pairs, Roche). The removal was performed as described in the protocol from the manufacturer.

2.1.6.3. Expression of TFAP2C and p53 proteins

TFAP2C and p53 proteins were expressed in vitro by using the TnT quick coupled transcription/ translation system (Promega). This experiment was performed following the protocol provided by the manufacturer. 2 µl of 0.5 µg/µl plasmids (pcDNA3.1-TFAP2C or pcDNA3.1-WTp53) or water was added as control. pcDNA3.1-WTp53 was a generous gift given by Dr. Domann lab. The TnT crude was aliquoted and saved at -80°C. To estimate the relative concentration of proteins expressed in the TnT
system, a Western Blot was performed (Figure 14). Amount of proteins is estimated compared to the 20 KDa band in the ladder.

2.1.6.4. Gel-shift assay

Gel shift experiment was conducted following the protocol of Gel Shift Assay System (Promega). The details are described as follows. Anti-TFAP2C and anti-p53 used in this assay are: AP-2γ (6E4/4): sc-12762, AP-2γ (V-18): sc-31935 and p53 (DO-2) antibody: sc-53394 (Santa Cruz), respectively. Reaction was mixed and incubated at room temperature for 30 minutes. Then, add 1µL gel loading 10Xbuffer, and load the samples to Novel® 6% DNA retardation gel (1.0 mm, 10 Well, Invitrogen™). Start running gel during the process of loading at 50V, to minimize the dissociation of the protein-DNA complex when the reaction mix was loading to the well and was exposed to 0.5% TBE buffer. Gel was run at 100V for about 1 hour until the dye reached the bottom of the gel. Subsequently, gel was taken out and placed on Biolabs filter paper, covered with plastic wrap and dried on a gel dryer. Signal was detected either by exposing the gel on X-ray film (CLASSIC X—Ray Film for Autoradiography and Chemiluminescence, RPI Research Product International Corp) or using phosphorimaging instrumentation (FLA7000, Fuji).

2.1.7. CpG Island Methylation Analysis

Allele-specific bisulfite sequencing was performed to determine the CpG methylation of the MUC1 promoter in different cell lines and various transfection conditions (Table 5). The steps are described in detail as follows.
2.1.7.1. **Plasmids transfection**

HCT 116 p53-/ cells were transfected with plasmids using Lipofectamine™ LTX Reagent (Invitrogen), following the manufacturer’s protocol “Transfecting plasmid DNA into HCT 116 cells”.


Table 5 Summary of cell lines used in CpG island methylation assay.

<table>
<thead>
<tr>
<th>Cell Lines</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HCT</td>
<td>HCT</td>
<td>TFAP2C</td>
<td>TFAP2C</td>
<td>HCT 116p53</td>
<td>MDA-MB-453</td>
</tr>
<tr>
<td></td>
<td>116</td>
<td>p53-/</td>
<td></td>
<td>+</td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p53</td>
<td>-</td>
<td></td>
<td>WT p53</td>
<td>Mu p53</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Note:

EV: Empty Vector

WT p53: wild type p53

MU p53: mutant R175H p53

N/A: Not Applicable.

24 hours post-transfection, genomic DNA and protein were harvested using DNeasy blood and tissue kit (QIAGEN) and RIPA Lysis Buffer (Millipore), respectively, according to the manufacturer. All transfections were carried out in biological replicate.
2.1.7.2. **Bisulfite modification of DNA**

Complete conversion of unmethylated cytosines to uracils was achieved by treating genomic DNA with bisulfite using EpiTect® Bisulfite Kit (QIAGEN) according to the manufacturer.

2.1.7.3. **PCR amplification**

Primers were designed using MethPrimer software\(^9^0\). (http://www.urogene.org/methprimer/index1.html). To screen pair of primers amplifying target region specifically, the EpiTect® PCR Control DNA (QIAGEN\(^®\)) were used as templates.

Since the bisulfite treated gDNA was broken into small fragments due to the C to U conversion, large regions were hard to be amplified. Therefore, to amplify the weak signal of the target part, Nested-PCR was performed using Platinum\(^®\) PCR SuperMix (Invitrogen) or Platinum\(^®\) Blue PCR SuperMix (Invitrogen). PCR thermal cycle, generally, was 35 cycles of 95°C for 30 s, annealing for 30 s, 72°C for 60 s. However, the annealing temperature depends on individual PCR experiment (Table 6). PCR conditions including the primer sets, the amplicon lengths, and the thermal cycle sets were modified after many trials. Table 6 summarized the primers eventually used for subsequent steps and the final conditions to individually amplify the three fragments within the MUC1 promoter. The three fragments of MUC1 promoter region I cloned are as follows and CpG dinucleotides are underlined and highlighted. Also, the two potential TFAP2C responsive elements are shown in bold and double-underlined.

Fragment A (-467 to -200, 293 bp)
GGGGGAGGGAGCCCAAAACTAGCACCTAGTCACCTATTATCCAGCCCTC
TATTTTCTGGCCCCCGCTCTGGCTTTAGTGGACCCCGGGAGGGGGC
GGGAAGCTGGAGTGGGAGACCTAGGGGTGGGCTTCCCGACCTTGCTGTACAGGACCT
CGACCTAGCTGGCTTTGTCCCCATCCCCACGTAGTTGTTGCCCTGAGGC
CTAAAAACTAGAGCCCAAGGGGCCCAAGTTCCAGACTGCCCTCCCCCTC
CCCGAGCCAGGGAGTGGTTGGTGAAAGGGGGAGGCCAGCTGGAGAACAAA
CGGTAGTCAGGGGGTTGAGATTAGAGCCCTTGTACCTACCCAGGAATGGTTG
GGAGGGAGGAGGAAGAGGTAGGAGGGGAGGGGGCGGGTTTTGTCACCTGC
CTAGCTCTAGCTCCACCTCTCAAGCAGCCAGCCTGCCTGAATCTGTTCTGCC
CCCTCC

Fragment O (-213 to -71, 143bp)
GGGGGAGGAGCCAGCTGGAGAACAAA CGGTAGTCAGGGGGTTGAGATTAG
AGCCCTTGTACCTACCCAGGAATGGTTGGGAGGGAGGAGGAAGAGGTAG
GAGGTAGGGGAGGGGGCGGGTTTTGTCACCTGCACCTGTACTGCTC

Fragment X (-158 to 48, 206 bp)
TTGTACCCCTACCCAGGAATGGTTGGGAGGGAGGAGGAAGAGGTAGGAGGT
AGGGAGGGGGCGGGTTTTTGTACCTGCACCTGCTACCTGCCTGGCTGTCCTGCTA
GGCCGGCGGGCGGGAGGTGGGGGAGCCGTATAAAGCGGTAGGGGCGCTG
TGCCCGCTCCACCTCTCAAGCAGCCAGGGCCTGCCTGAATCTGTTCTGC
CCCTCC
Table 6 PCR thermal cycles in the CpG methylation assay to amplify *MUC1* promoter regions.

A. Fragment A (-492 bp to -200 bp)

<table>
<thead>
<tr>
<th>Pair of Primers</th>
<th>Annealing Tm.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Run</strong></td>
<td></td>
</tr>
<tr>
<td>-492F: 5’-GGGGGAGGAGGAGTTTTAAAATTAGTA</td>
<td></td>
</tr>
<tr>
<td>-71R: 5’-AAACAAATAACAAATAACAAACCC</td>
<td></td>
</tr>
<tr>
<td>46°C</td>
<td></td>
</tr>
<tr>
<td><strong>Second Run</strong></td>
<td></td>
</tr>
<tr>
<td>-492F: 5’-GGGGGAGGAGGAGTTTTAAAATTAGTA</td>
<td></td>
</tr>
<tr>
<td>-200R: 5’-CAACTAACCTCCCCCTTTCCAC</td>
<td></td>
</tr>
<tr>
<td>50°C</td>
<td></td>
</tr>
</tbody>
</table>

B. Fragment O (-213 bp to -71 bp)

<table>
<thead>
<tr>
<th>Pair of Primers</th>
<th>Annealing Tm.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Run</strong></td>
<td></td>
</tr>
<tr>
<td>-492F: 5’-GGGGGAGGAGGAGTTTTAAAATTAGTA</td>
<td></td>
</tr>
<tr>
<td>-71R: 5’-AAACAAATAACAAATAACAAACCC</td>
<td></td>
</tr>
<tr>
<td>46°C</td>
<td></td>
</tr>
<tr>
<td><strong>Second Run</strong></td>
<td></td>
</tr>
<tr>
<td>-213F: 5’-GGGGAGGTTAGTTGGAGAATAAA</td>
<td></td>
</tr>
<tr>
<td>-71R: 5’-AAACAAATAACAAATAACAAACCC</td>
<td></td>
</tr>
<tr>
<td>50°C</td>
<td></td>
</tr>
</tbody>
</table>

C. Fragment X (-158bp to 48bp)

<table>
<thead>
<tr>
<th>Pair of Primers</th>
<th>Annealing Tm.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Run</strong></td>
<td></td>
</tr>
<tr>
<td>-213F: 5’-GGGGAGGTTAGTTGGAGAATAAA</td>
<td></td>
</tr>
<tr>
<td>168R: 5’-CACCTAAAACACAAACCCTCTC</td>
<td></td>
</tr>
<tr>
<td>48°C</td>
<td></td>
</tr>
<tr>
<td><strong>Second Run</strong></td>
<td></td>
</tr>
<tr>
<td>-158F: 5’-TTGTATTTTTATTTTAGGAATGGTTGG</td>
<td></td>
</tr>
<tr>
<td>48R: 5’-AAAAAAAAACAAAAACATTCAAAACAA</td>
<td></td>
</tr>
<tr>
<td>54°C</td>
<td></td>
</tr>
</tbody>
</table>

Note:

Tm.: temperature

F: forward primer

R: reverse primer
2.1.7.4. **Ligation and transformation**

PCR products with expected size were cut from agarose gel, extracted by QIAquick® Gel Extraction Kit (QIAGEN®) and ligated to pCR®2.1 TOPO vector by TA Cloning® Kit (Invitrogen) following the manufacturer’s protocol. The ligation mixture was transfected into One Shot® TOP10 Chemically Competent E. coli (Invitrogen). With the ampicillin-resistance and White/Blue screening, single colonies were streaked out. The colony number I picked varies between 10 to over 100. Then, plasmid DNA was isolated and then screened by digested with Eco RI HF Restriction Enzyme and analyzed by gel electrophoresis confirming size of amplicon inserted. Multiple screened plasmids were sent to DNA Core Facility at the University of Iowa for sequencing. Sequence analysis and blast were then performed using Mega4 software.

**Results**

TFAP2C is a transcription factor which associates with breast cancer carcinogenesis. It not only is involved in the estrogen-signaling pathway by regulating ER-α¹⁹, but also controls genes that do not belong to the luminal-cluster genes. The overall aim of this thesis project is to decipher the mechanism of p53-mediated enhancement of TFAP2C occupancy on specific sites. To start, we first performed ChIP-seq to map TFAP2C binding genes and found that the introduction of p53 to HCT 116 p53 -/- cell line significantly augments TFAP2C occupancy on the promoter regions of a group of genes. Taking six genes from this pool as examples, we identified TFAP2C binding sites by gel-shift, which agree with published consensus motif. One of these genes, *MUC1*, was then determined to be activated by TFAP2C in MCF-7 breast cancer cells. In addition, we showed that promoter DNA methylation does not explain the increased occupancy of TFAP2C on *MUC1* promoter with the introduction of p53. In
addition, I showed that TFAP2C is a key transcriptional regulator establishing pattern of
gene expression characterizing luminal type of breast cancer.

2.2.1. **TFAP2C is a key transcriptional regulator**

**establishing pattern of gene expression**

**characterizing luminal type of breast cancer**

To determine TFAP2C transactivation activity in different phenotypes of breast
cancers, genome-wide map of TFAP2C occupancy in several representative breast cancer
cell lines (MCF-7, BT-474, MDA-MB-453 and SKBR-3) have been generated by ChIP-seq. These cell lines are believed to possess various molecular expression profile and
biological characters, and they represent unique phenotypes of breast cancer as well. The question we first addressed here is whether TFAP2C is associated with defining
gene-expression pattern for specific phenotypes of breast cancer.

To answer this question, I first examined the TFAP2C occupancy on a cluster of
genes which are categorized as markers to establish gene expression patterns for different
groups of breast cancer (basal, luminal cytokeratins, ER-associated and luminal
differentiation). Surprisingly, I identified that TFAP2C occupies on above 80% of
genesis (38 from 47) that characterizing luminal type, whereas only less than 10% of genes
(1 from 9) that characterizing basal type (Figure 2). This interesting result indicates
TFAP2C is a potential key transcriptional regulator establishing pattern of gene
expression potentially determining luminal type of breast cancer.
2.2.2. p53 expression augments TFAP2C binding on specific genes

As stated in the Introduction, the association of p53 with AP-2 is found to regulate AP-2 transcriptional activity on genes RPSA and Rad51. The Weigel lab also indicated that TFAP2A decreases p53 stability and increases p53 trans-activation activity on p21. Based on these results, we formulated the general hypothesis that p53, as a cofactor, interacts with TFAP2C and changes TFAP2C trans-activation activity on specific target genes in a combinatorial manner.

TFAP2C occupancy has been mapped by ChIP-seq in HCT 116 p53-/- colon cancer cell line (p53-negative cell line) with over-expression of TFAP2C and exogenous WT or mutant p53, as summarized in Number 5-7 in Table 4. This cell line was selected to avoid the background noise of endogenous p53 in breast cancer cell system. Discrepancy of TFAP2C occupancy in the context with or without p53 was compared. The ChIP-seq peak heights of TFAP2C at each locus after normalized to nearby background are comparable in HCT 116 p53-/- cells with or without exogenous p53 because of the following reasons. First, the total reads for the two ChIP-seq dataset are identical and enough to generate convincing peaks. Second, the overall views of TFAP2C binding for two ChIP-seq datasets are similar, but demonstrate big difference only on specific sites, as illustrated in figure 3. To identify the potential TFAP2C targets where TFAP2C occupancy is largely increased upon the presence of p53, we followed the defined criteria: 1) peak should locate within the -20 kb to the transcription start site of a known gene; and 2) this region should be found and overlapped in lists generated by
different programs. Also, biological replicate of TFAP2C ChIP-seq is currently performing to achieve more confident result.

Subsequently, 115 satisfactory peaks with 111 corresponsive genes were grouped as a candidate pool (Table 7). Among those, five genes (BCAN, GRHL3, TFAP4, ITGA3, ID1) contain two individual peaks, and two pair of genes (S100A6 and S100A5, ACAD8 and THYN1) share one peak. Motif analysis and pathway analysis on these targets were performed by Tomas B. Bair in the DNA Core Facility of the University of Iowa. The TFAP2C binding site motif was generated by the Partek software analysis of 300 bp regions centered on the peaks yielded from this study (Figure 4). It agrees with previously published TFAP2C binding motif\(^\text{39}\). However, pathway analysis fails to identify a common or particular shared theme of gene families being regulated by TFAP2C (data not shown).

Using this list as a candidate gene pool, I first selected six representative genes for further investigation: MYCBP2, FAM98B, DDX17, HDAC4, CNOT3, and MUC1. TFAP2C occupancy on the promoter regions of these genes were significantly increased with the introduction of WT p53 (Figure 5 to Figure 10). Also, other TFAP2C binding genes, FAM83A, ST6GALNAC2 and DLX4, were selected as negative controls, since the TFAP2C occupancy on those three genes remained identical in the context of over-expressed p53 (Figure 11 to Figure 13). These observations indicate that exogenous p53 increases TFAP2C occupancy on specific genes.
2.2.3. Localization of TFAP2C binding sites

Moving forward, to identify the responsive DNA elements under the ChIP-seq peak in the promoter region of these six genes and confirm the direct protein-DNA interaction, I performed Electrophoretic Mobility Shift Assays (EMSA) in vitro.

First, TFAP2C and p53 were synthesized in vitro using the TnT system and successful synthesis was confirmed by Western Blot (Figure 14). The TnT lysate and the whole protein in BT-474 cell line were used as negative control and positive control respectively. I also amplified and radiolabeled DNA fragments about 100-120 bp long under the ChIP-seq peaks of these genes, MUC1, DDX17, MYCBP2, FAM98B, HDAC4 and CNOT3 (Figure 15). To optimize the conditions, TFAP2C amount was titrated with the range from 1/300 µL to 3 µL TnT reaction mix. The band of TFAP2C-probe complex responds positively to the increase of the TFAP2C (Figure 16). Therefore, I chose 1 µL TnT reaction crude in the following gel-shift experiment, because this amount enables us to see detectable band and save the expensive TFAP2C proteins as well. Moreover, the incubation of the TFAP2C-probe complex with the anti-TFAP2C antibody leads to a super-shifted complex, arrowed as Super-shift.

To further examine the specific consensus binding sites for TFAP2C under the sequences of ChIP-seq peak region within the six promoters (MUC1, DDX17, MYCBP2, FAM98B, HDAC4 and CNOT3), competitive gel shift assays were performed using series of competitors, some of which containing two or three nucleotides mutation in the potential binding sites. When TFAP2C protein is incubated with 100 bp-long p32-radiolabeled MUC1 probe, a complex is formed and exhibits a slower migration rate, shown as a shifted band in lane 2 (Figure 17). This indicates that TFAP2C bind to MUC1
probe directly. A super complex was formed in the presence of anti-TFAP2C antibody in the reaction, suggesting that it is TFAP2C protein that binds to MUC1 probe (Figure 17, lane 10). In contrast, the TnT lysate or anti-TFAP2C antibody alone did not interact with MUC1 probe (Figure 17, lane 10 and lane 12). We then incubated hot probe, TFAP2C protein, together with cold competitors (the molecular ratio between probe and competitor is about 1:200 to 1:250), to examine that whether specific competitor is capable of competing with probe and therefore blocking the hot probe-TFAP2C shifted band. We found competitor number 3 and number 4 compete with the MUC1 probe, leading to the disappearance of hot band (Figure 17, lane 5 and lane 6). Notably, these two fragments contain TFAP2C binding consensus motif in the center. The introduction of mutated competitors (two nucleotides mutation in each of them) greatly eliminated the ability to interact with TFAP2C protein (Figure 17, lane 8 and 9), indicating the nucleotides at these positions are vital for physical binding to TFAP2C protein. In sum, we identified two TFAP2C binding sites within MUC1 probe. Also, the sequence sitting in the center of either competitor agrees with the consensus motif revealed by peak analysis program (Figure 4), further supporting that TFAP2C directly bind to these elements by protein-DNA interaction.

Similarly, TFAP2C protein interacts directly with probes of MYCBP2, DDX17, FAM98B, HDAC4, and CNOT3 (Figure 18 to 22). TFAP2C binding sites are identified in competitor 2 for DDX17 probe (Figure 18), competitor 3 for MYCBP2 probe (Figure 19), competitors 1 and 6 for FAM98B probe (Figure 20), competitors 5 and 6 for HDAC4 probe (Figure 21), and competitor 4 for CNOT3 probe (Figure 22). Again, all these confirmed competitors share the consensus TFAP2C binding motif in the center of these
sequences (Figure 23). Notably, some competitors contain an extra single nucleotide within the common motif sequence and still compete with their probes, but with differential occupancy, e.g. competitor 3 for FAM98B partly competes (Figure 20), competitor 2 for HDAC4 partly competes (Figure 21), competitor 4 for CNOT3 totally competes (Figure 22), as summarized in Figure 23.

I observed that the super shifted band was separated to several ones with different molecular-weight components along with the extended period of electrophoresis (figure 16, 18, 19, 20, 22, and 24). It might due to the different combinations of the several molecules. For example, two binding sites in one probe may attract the TFAP2C dimer individually, and/or the polyclonal anti-TFAP2C antibody recognizes and interact with several antigen on the surface of the TFAP2C protein.

Considering the absence of p53 binding motif (CATG******CATG)\textsuperscript{93} in all probes used in the gel shift experiments, I anticipate that p53 will not bind to those six DNA probes directly, but may interact with TFAP2C and increase TFAP2C binding ability to its TFBS. To test this hypothesis, I performed gel shift experiments to test the protein-DNA interaction \textit{in vitro} (Figure 24). The result demonstrated that p53 protein cannot interact with \textit{MUC1} probe directly (Figure 24, lane 90). Also, incubating p53 with TFAP2C-\textit{MUC1} complex neither yielded a super-shifted complex nor enhanced the band intensity of the TFAP2C-\textit{MUC1} complex (Figure 24, lane 4, 5, 8 and 9). A similar result was observed on other candidate genes, such as \textit{DDX17} (data not shown here). Taken these findings together, I conclude that p53 neither bind to \textit{MUC1} probe directly nor augment TFAP2C occupancy to its TFBS on \textit{MUC1} or \textit{DDX17} probe \textit{in vitro}. 
In addition, to identify other co-factors, I also incubated TFAP2C protein, *MUC1* or *DDX17* probe, with the nuclear extract from HCT 116 p53 -/- and HCT 116 p53 +/- cells in gel shift experiments. However, I did not find any promising supershift bands that indicate the presence of potential TFAP2C co-factors (data not shown).

To summarize, based on the information of TFAP2C occupancy provided by ChIP-seq datasets, I confirmed that TFAP2C protein binds to the consensus TFBSs of these six probes. However, the preliminary result does not exhibit that p53 binds to these probes directly or augment TFAP2C occupancy on these DNA fragments.

### 2.2.4. *MUC1* is a TFAP2C target in MCF-7 breast cancer cell line

In spite of the robust power of ChIP to exhibit transcription factor occupancy, we do not know the TFAP2C function on any occupied sites. It is highly possible that TFAP2C does neither activate nor repress transcription of most of the TFAP2C binding site-correlated genes.

To further determine that the six representative genes (*MYCBP2*, *FAM98B*, *DDX17*, *HDAC4*, *CNOT3* and *MUC1*) identified in HCT 116 p53 -/- colon cancer cell line are TFAP2C targets in breast cancer cells as well, I knocked down the expression of TFAP2C, TFAP2A, or p53 by siRNA in MCF-7 cells. The knockdown experiment is exhibited in Table 3. TFAP2A, TFAP2C and p53 on protein level were greatly reduced ninety-six hours post-transfection (Figure 25). Then transcription levels of interested genes (*MYCBP2*, *FAM98B*, *DDX17*, *HDAC4*, *CNOT3* and *MUC1*) were measured by quantitative Real-Time PCR. Human B2M (Beta-2-Microglobulin) was used as control.
It is consistently shown that transient elimination of TFAP2C in MCF-7 cell line greatly reduced cellular MUC1 on both mRNA (Figure 26) and protein level (Figure 25). However, MUC1 transcription does not respond to elimination of TFAP2A (Figure 26). Therefore, I conclude that MUC1 is a TFAP2C specific target in MCF-7 cell lines.

To examine the effect of p53 on the activity of TFAP2C in regulating MUC1 expression, the expression of p53 was inhibited using siRNA in MCF-7 cell line. The preliminary data suggests that silencing p53 neither reverses nor enhances the positive correlation between MUC1 and TFAP2C expression. Therefore, to explain the phenomenon that p53 augments TFAP2C occupancy on MUC1 promoter region in colon cancer cell line, it requires for additional mechanisms besides the direct protein-protein interaction between TFAP2C and p53.

Also, it is worthy to mention that it is achieved to stably knockdown TFAP2C expression by small hairpin RNA (shRNA) in MCF-7 cells line by other lab members. They showed that MUC1 protein is significantly decreased following TFAP2C level drops down dramatically, two to three weeks after shRNA procedure performed. This supports the notion that MUC1 is a TFAP2C direct target in this breast cancer cell line (data not shown).

Altogether, our results support that TFAP2C activates MUC1 gene expression in MCF-7 cell line.
2.2.5. CpG methylation is not changed within 

*MUC1* promoter in HCT 116 p53 -/- cells 

with introduction of p53 

To construe the mechanism for the enhancement of TFAP2C occupancy with the presence of p53 on *MUC1* promoter, we formulated the hypothesis that the introduction of p53 into HCT 116 p53 -/- cell line demethylates CpG dinucleotides on *MUC1* promoter region, therefore this region is more accessible for TFAP2C to load and to trans-activate *MUC1* expression (Figure 27). To test this hypothesis, I over-expressed TFAP2C and p53 in HCT 116 p53 -/- cell line (Table 5) and performed allele-specific CpG methylation assay for *MUC1* promoter covering the *MUC1* promoter region from -492 bp to 48 bp within the TFAP2C binding peak identified from ChIP-seq (Figure 28, and Figure 29). Given the correlation between the absence of TFAP2C occupancy on *MUC1* promoter (Figure 29) and the methylation of this region in MDA-MB-453 cell line, I included this cell line in the DNA methylation assay as a positive control. 

Twenty four hours post-transfection, expression of exogenous proteins were confirmed by Western Blot (Figure 30). Interestingly, the expression of wild type p53 (in transfection 3) down-regulated the over-expression of TFAP2C as compared to that when p53 is absent or when mutant p53 (R175H) is introduced. In contrast, the over-expression of wild type p53 in this cell line has little effect on the relative low amount of endogenous TFAP2C. Biological repeat confirmed this observation (Figure 30 B). I anticipated that p53 might activate the promoter activity of TFAP2C as a transcription factor or p53 might regulate TFAP2C protein level by post-translational mechanism. Moreover, MUC1 protein level is very low but detectable in each condition of HCT 116
cells, and it remained at identical level in each condition. This could be interpreted by that specific factors of the transcription machinery for *MUC1* in breast cancer cells is absent in this colon cancer cell system, therefore the transcription of MUC1 remains at basic level and does not respond to TFAP2C activation. Overall, we think this colon cancer cell system could be utilized to map the TFAP2C occupancy associated with the present of p53 on target genes but some pieces of the puzzle might be missing for activating MUC1 expression in colon cells.

To identify the DNA methylation changes in *MUC1* promoter, genomic DNA from these four transfected samples, HCT 116 p53 +/- cell line and MDA-MB-453 cell line were treated with bisulfite conversion, and followed by nested PCR, cloning and sequencing. After intense trials, I succeeded to amplify three fragments (namely fragment A, O, and X) covering the -492 bp to 48 bp of *MUC1* promoter, which containing 20 CpG dinucleotides. The map of amplicon, all CpG sites and promoter sequences are demonstrated in figure 28.

The preliminary DNA methylation result is summarized in Figure 31 to Figure 33. First, the methylation map of *MUC1* in MDA-MB-453 cell line shows that about 10 CpG dinucleotides (number 11 to 20) in the upstream of TSS are methylated at high percentage. This agrees with the published data by Yamada *et al.* indicating that this DNA methylation experiment was successfully carried out. Next, I compared the methylation map of other cell lines from condition 1 to 5. When wild type or mutant p53 is introduced in HCT 116 p53 +/- cell line, CpG sites of number 7, 10 and 11 are methylated at low percentage (10% to 20%). We think two reasons may contribute to this low percentage. One is not full conversion from unmethylated cytosine to uracil. The
other is that one allele is methylated but the other one is not. Though these variance of CpG methylation changes were observed in association with p53, it is still likely that the CpG methylation status is identical among HCT 116 p53 -/- cell line (condition 1) and HCT 116 p53 +/- cell line (condition 5) and among HCT 116 p53 -/- cells with different expression profile of p53 (condition 2, 3 and 4). Therefore, based on these preliminary data, we tend to believe that DNA methylation status of MUC1 promoter region studied in this project is not the primary mechanism to explain why the introduction of p53 augments TFAP2C binding on this promoter in HCT 116 p53 -/- colon cancer cell line.
CHAPTER III
DISCUSSION

3.1. Limitations of ChIP-seq research

In this project, I observed that TFAP2C occupancy was greatly increased on the promoter regions of *MYCBP2, FAM98B, DDX17, HDAC4, CNOT3* and *MUC1* after p53 is introduced to HCT 116 p53 -/- cells. However, transient knocking down of TFAP2C in MCF-7 cell line has almost no evident effect on transcription level of these genes, *MYCBP2, FAM98B, DDX17, HDAC4* and *CNOT3* (data not shown). This supports that in spite of robust power of ChIP-seq tool to exhibit transcription factor occupancy, even when direct protein-DNA interaction has been shown *in vitro*, I cannot conclude the TFAP2C functions on all the occupied genes in cells. It is highly possible that TFAP2C activates or represses only on small portion of its target genes.

On the other hand, another explanation is that the transcriptions of those genes require complete transcriptional machinery composed of several cooperatively-functioning factors. Since this machinery is fine-tuned, the effect of one particular factor on the expression of target gene may be hardly detected by transiently silencing it (rather than stably knocking down this factor or disrupting several factors involved in this machine). For instance, *CD44* is a TFAP2C occupied gene in MCF-7 cells. Transient knocking-down of TFAP2C has no evident effect on *CD44* transcription (or the result is not consistent among biological triplicate), whereas *CD44* is greatly increased when TFAP2C is stably repressed (data from lab member Dr. Mikhail V Kulak).

Therefore, to determine the TFAP2C function on *MYCBP2, FAM98B, DDX17, HDAC4* and *CNOT3*, it is required to conduct more biochemical or cellular experiments.
For example, the transcription level of those genes when TFAP2C is stably knocked down could be studied. However, the argument for stable knocking-down is TFAP2C may regulate specific gene expression by either direct binding to its promoter region or through indirect downstream effect.

### 3.2. Other ways to perform DNA methylation assay

Within this thesis project, I performed allele-specific bisulfite sequencing to map the CpG methylation status covering the TFAP2C binding sites in *MUC1* promoter region. Each clone represents a single allele providing allele specific information. Therefore, the convincing data demands large number of clones. In our case, the PCR to amplify target region is very challenging: 1) the desired target fragment is very sensitive to primer design and the thermal cycler setting; 2) the efficiency to achieve positive clone among sequenced samples is lower than 10%. Therefore, to get information from a large population with great statistical significance and to save labors as well, bisulfite pyro-sequencing is a potential method that could be tried to measure the CpG island map of short to medium length DNA sequences accurately. Also, bisulfite PCR followed by restriction analysis (COBRA) is another method to reveal CpG dinucleotide methylation. Additionally, a novel way named as methylation specific electrophoresis (MSE) was reported recently to map mucin gene methylation. It could be investigated in the near future as well.
3.3. Other mechanisms to explain that p53-mediated augmentation of TFAP2C binding occupancy on MUC1

To examine the mechanism of p53 functioning to alter TFAP2C binding on specific sites, I introduced ETEQ mutant p53 (six amino acids were mutated to glutamic acid or glutamine, therefore this mutant cannot interact with TFAP2C protein) into HCT 116 p53 -/- cell line, and mapped TFAP2C binding loci by ChIP-seq. I observed that this ETEQ mutant increases TFAP2C occupancy on promoter regions of MUC1, DDX17, MYCBP2, FAM98B, HDAC4 and CNOT3, showing identical patterns as compared to wild type p53. This supports that p53 augments TFAP2C binding not by p53-DNA interaction, but through other indirect mechanisms. Therefore, we hypothesized two potential mechanisms.

3.3.1. p53 modifies motif preference of TFAP2C

Our TFAP2C binding motif analysis based on 115 peaks showed that the common motif does contain cytosine in position 10 and adenine in position 11 at high preference (over 80%) (Figure 4 A). However, these two positions are not shown on the published motifs (Woodfield et al., 2010). A recent publication referred that the presence of cofactor (Exd) evokes latent sequence specificities between all eight Hox proteins. Furthermore, the binding site of TFAP2A is demonstrated to be flexible by Systematic Evolution of Ligands by Exponential Enrichment (SELEX, but the effect of cofactor on the modifying TFAP2A binding site preference is not tested. Therefore, it is possible that direct protein-protein interaction between p53 and TFAP2C modifies the binding specificity of the latter. To test it, SELEX could be performed. If this is the case, it would
be another mechanism for p53 to modify transcription factor function and to regulate cancer-related gene expression.

### 3.3.2. Altered DNA accessibility by p53

Another mechanism to explain the increased TFAP2C occupancy on *MUC1* is that the introduction of p53 alters chromatin structure and the DNA accessibility on the *MUC1* promoter region, therefore facilitating the recruitment of TFAP2C and subsequent transcription. Also, the exogenous p53 might change histone modification of *MUC1* promoter region, leading to the opened chromatin structure and augmentation of TFAP2C occupancy. To test this, DNA accessibility assay and ChIP with anti-histone antibodies targeting various histone modification could be performed.

### 3.4. Future Directions

#### 3.4.1. To determine *MUC1* as a TFAP2C target in other breast cancer cell lines

To examine *MUC1* is a TFAP2C target in MCF-7, BT-474, SKBR-3, but not MDA-MB-453, knocking-down experiment of TFAP2C could be performed on these cell lines, followed by qPCR and WB against both TFAP2C and MUC1. The variant p53 is expected to demonstrate different effect on altering AP2C regulation activity regarding *MUC1* expression.
3.4.2. To perform the ChIP-seq (or ChIP-reChIP) with anti-p53 and anti-TFAP2C in HCT116 p53 -/- cells

To achieve more convincing ChIP-seq dataset with lower background, we are repeating the ChIP-seq with anti-TFAP2C in HCT116 p53 -/- cells with 1) over-expression of exogenous TFAP2C; 2) over-expression of exogenous TFAP2C and wild type p53; 3) over-expression of exogenous TFAP2C and EQEQ p53 mutant. Also, to get genome-wide occupancy of p53, another lab member has recently performed ChIP-seq with anti-p53. However, it is bad ChIP-seq dataset because of high background. To solve the problem, it is required to find a better ChIP-quality anti-p53 antibody and to optimize the immunoprecipitation procedure.

Additionally, ChIP-reChIP method, in which DNA is immunoprecipitated using two different antibodies, allows us to identify multiple concurrently binding proteins on a single DNA fragment. It is a robust tool to reveal the interaction of multiple transcription factors and cofactors to regulate gene transcription corporately. Therefore, it could be applied to understand the TFAP2C-p53 genome-wide co-occupancy.

We believe that by analyzing the co-occupancy of TFAP2C and p53, more targets potentially co-regulated by TFAP2C-p53 interaction will be found. Therefore, the mechanism to reveal the interaction of TFAP2C and p53 will be better understood.

3.4.3. To generate a set of genes where p53 blocks TFAP2C occupancy

Modugno et al. reported that p53 represses the expression of laminin precursor (encoded by RPSA gene), partly by blocking TFAP2C occupy to TFBS on RPSA
promoter in ovarian carcinoma cell line\textsuperscript{57}. Interestingly, our ChIP-seq data supports this idea: the introduction of wild type p53 represses TFAP2C occupancy on \textit{RPSA} promoter, yet the presence of EQEQ p53 mutant (which cannot interact with TFAP2C) has no effect on TFAP2C binding (Figure 34). Taking the pattern we found on \textit{MUC1} promoter into account, we think the role of p53 on altering TFAP2C occupancy and function is complicated and target-dependent. Also, being inspired by the agreement of our ChIP-seq dataset with literature\textsuperscript{57}, we may generate a set of genes exhibiting the pattern that the presence of p53 blocks TFAP2C occupancy whereas EQEQ p53 mutant cannot block.

To sum, I have observed several different patterns of how introduction of p53 altering TFAP2C occupancy on various genes. In the near future, it could help us to better understand how the TFAP2C trans-activation activity is regulated and how p53-TFAP2C interaction is associated with the breast cancer carcinogenesis.
Figure 2 Percentage of TFAP2C target genes in gene clusters characterizing different breast cancer phenotypes.
Number of genes examined in core clusters: Basal: 9 genes; Luminal Cytokeratins: 14 genes; ER-associated: 19 genes, Luminal Differentiation: 14 genes
Figure 3 Visual displays of TFAP2C binding peaks from part of chromosome 1. TFAP2C binding pattern are identical on most genomic regions among different HCT 116 p53 -/- cells transfections. (ChIP-seq dataset). The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C; ChIP-seq data in middle panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and wild type p53; ChIP-seq data in bottom panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and mutant p53 (EQEQ).
Table 7 List of TFAP2C binding peaks augmented with presence of p53 and their corresponsive genes.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Chromosome</th>
<th>Peak Start</th>
<th>Peak Stop</th>
<th>Subtraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABTB2</td>
<td>chr11</td>
<td>34,241,599</td>
<td>34,241,984</td>
<td>385</td>
</tr>
<tr>
<td>ACAD8</td>
<td>chr11</td>
<td>133,628,304</td>
<td>133,628,604</td>
<td>300</td>
</tr>
<tr>
<td>ACCS</td>
<td>chr11</td>
<td>44,044,219</td>
<td>44,044,470</td>
<td>251</td>
</tr>
<tr>
<td>ADD1</td>
<td>chr4</td>
<td>2,814,996</td>
<td>2,815,396</td>
<td>400</td>
</tr>
<tr>
<td>AHDC1</td>
<td>chr1</td>
<td>27,739,509</td>
<td>27,739,712</td>
<td>203</td>
</tr>
<tr>
<td>AIG1</td>
<td>chr6</td>
<td>143,539,470</td>
<td>143,539,766</td>
<td>296</td>
</tr>
<tr>
<td>ANK2</td>
<td>chr4</td>
<td>114,496,246</td>
<td>114,496,504</td>
<td>258</td>
</tr>
<tr>
<td>ANKHD1</td>
<td>chr5</td>
<td>139,760,646</td>
<td>139,760,946</td>
<td>300</td>
</tr>
<tr>
<td>ANKMY2</td>
<td>chr7</td>
<td>16,651,767</td>
<td>16,652,067</td>
<td>300</td>
</tr>
<tr>
<td>ANKRD11</td>
<td>chr16</td>
<td>87,914,207</td>
<td>87,914,524</td>
<td>317</td>
</tr>
<tr>
<td>APEH</td>
<td>chr3</td>
<td>49,686,564</td>
<td>49,686,864</td>
<td>300</td>
</tr>
<tr>
<td>APITD1</td>
<td>chr1</td>
<td>10,410,736</td>
<td>10,410,950</td>
<td>214</td>
</tr>
<tr>
<td>ARID5A</td>
<td>chr2</td>
<td>96,566,205</td>
<td>96,566,475</td>
<td>270</td>
</tr>
<tr>
<td>ATRN</td>
<td>chr20</td>
<td>3,399,379</td>
<td>3,399,679</td>
<td>300</td>
</tr>
<tr>
<td>ATXN7L1</td>
<td>chr7</td>
<td>105,203,793</td>
<td>105,204,041</td>
<td>248</td>
</tr>
<tr>
<td>BC031953</td>
<td>chr15</td>
<td>38,985,576</td>
<td>38,985,876</td>
<td>300</td>
</tr>
<tr>
<td>BCAN</td>
<td>chr1</td>
<td>154,876,154</td>
<td>154,876,528</td>
<td>374</td>
</tr>
<tr>
<td>BCAN</td>
<td>chr1</td>
<td>154,896,602</td>
<td>154,896,854</td>
<td>252</td>
</tr>
<tr>
<td>BCL9L</td>
<td>chr11</td>
<td>118,287,846</td>
<td>118,288,146</td>
<td>300</td>
</tr>
<tr>
<td>BEGAN</td>
<td>chr14</td>
<td>100,084,113</td>
<td>100,084,413</td>
<td>300</td>
</tr>
<tr>
<td>C10orf136</td>
<td>chr10</td>
<td>43,669,624</td>
<td>43,669,833</td>
<td>209</td>
</tr>
<tr>
<td>C11orf46</td>
<td>chr11</td>
<td>30,301,006</td>
<td>30,301,290</td>
<td>284</td>
</tr>
<tr>
<td>C12orf75</td>
<td>chr12</td>
<td>104,248,253</td>
<td>104,248,562</td>
<td>309</td>
</tr>
<tr>
<td>C1orf25</td>
<td>chr1</td>
<td>183,392,536</td>
<td>183,392,836</td>
<td>300</td>
</tr>
<tr>
<td>C1orf26</td>
<td>chr1</td>
<td>183,392,536</td>
<td>183,392,836</td>
<td>300</td>
</tr>
<tr>
<td>C3orf18</td>
<td>chr3</td>
<td>50,581,351</td>
<td>50,581,651</td>
<td>300</td>
</tr>
<tr>
<td>C3orf32</td>
<td>chr3</td>
<td>8,646,145</td>
<td>8,646,489</td>
<td>344</td>
</tr>
<tr>
<td>C5orf51</td>
<td>chr5</td>
<td>41,960,922</td>
<td>41,961,199</td>
<td>277</td>
</tr>
<tr>
<td>C6orf126</td>
<td>chr6</td>
<td>35,852,395</td>
<td>35,852,695</td>
<td>300</td>
</tr>
<tr>
<td>CATSPERG</td>
<td>chr19</td>
<td>43,518,049</td>
<td>43,518,317</td>
<td>268</td>
</tr>
<tr>
<td>CCDC36</td>
<td>chr3</td>
<td>49,257,259</td>
<td>49,257,543</td>
<td>284</td>
</tr>
<tr>
<td>CCDC46</td>
<td>chr17</td>
<td>61,618,613</td>
<td>61,618,955</td>
<td>342</td>
</tr>
<tr>
<td>CCNG2</td>
<td>chr4</td>
<td>78,296,735</td>
<td>78,297,028</td>
<td>293</td>
</tr>
<tr>
<td>CCNL1</td>
<td>chr3</td>
<td>158,360,600</td>
<td>158,360,889</td>
<td>289</td>
</tr>
</tbody>
</table>
Table 7. (Continued)

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Chromosome</th>
<th>Peak Start</th>
<th>Peak Stop</th>
<th>Subtraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD320</td>
<td>chr19</td>
<td>8,279,132</td>
<td>8,279,432</td>
<td>300</td>
</tr>
<tr>
<td>CD3EAP</td>
<td>chr19</td>
<td>50,600,468</td>
<td>50,600,768</td>
<td>300</td>
</tr>
<tr>
<td>CD44</td>
<td>chr11</td>
<td>35,144,814</td>
<td>35,145,102</td>
<td>288</td>
</tr>
<tr>
<td>CFDP1</td>
<td>chr16</td>
<td>74,024,582</td>
<td>74,024,835</td>
<td>253</td>
</tr>
<tr>
<td>CLEC16A</td>
<td>chr16</td>
<td>10,993,137</td>
<td>10,993,478</td>
<td>341</td>
</tr>
<tr>
<td>CNOT3</td>
<td>chr19</td>
<td>59,332,929</td>
<td>59,333,218</td>
<td>289</td>
</tr>
<tr>
<td>COL7A1</td>
<td>chr3</td>
<td>48,607,987</td>
<td>48,608,287</td>
<td>300</td>
</tr>
<tr>
<td>CPNE6</td>
<td>chr14</td>
<td>23,620,484</td>
<td>23,620,795</td>
<td>311</td>
</tr>
<tr>
<td>DAGLA</td>
<td>chr11</td>
<td>61,221,146</td>
<td>61,221,446</td>
<td>300</td>
</tr>
<tr>
<td>DBN1</td>
<td>chr5</td>
<td>176,814,984</td>
<td>176,815,284</td>
<td>300</td>
</tr>
<tr>
<td>DDHD2</td>
<td>chr8</td>
<td>38,208,170</td>
<td>38,208,426</td>
<td>256</td>
</tr>
<tr>
<td>DDX17</td>
<td>chr22</td>
<td>37,232,219</td>
<td>37,232,555</td>
<td>336</td>
</tr>
<tr>
<td>DDX58</td>
<td>chr9</td>
<td>32,516,180</td>
<td>32,516,580</td>
<td>400</td>
</tr>
<tr>
<td>DDX59</td>
<td>chr1</td>
<td>198,905,207</td>
<td>198,905,441</td>
<td>234</td>
</tr>
<tr>
<td>DHX35</td>
<td>chr20</td>
<td>37,024,249</td>
<td>37,024,509</td>
<td>260</td>
</tr>
<tr>
<td>DRAM1</td>
<td>chr12</td>
<td>100,812,757</td>
<td>100,812,986</td>
<td>229</td>
</tr>
<tr>
<td>DSCR3</td>
<td>chr21</td>
<td>37,561,652</td>
<td>37,561,962</td>
<td>310</td>
</tr>
<tr>
<td>EHMT1</td>
<td>chr9</td>
<td>139,632,906</td>
<td>139,633,249</td>
<td>343</td>
</tr>
<tr>
<td>ELOVL5</td>
<td>chr6</td>
<td>53,299,366</td>
<td>53,299,640</td>
<td>274</td>
</tr>
<tr>
<td>EML2</td>
<td>chr19</td>
<td>50,837,167</td>
<td>50,837,474</td>
<td>307</td>
</tr>
<tr>
<td>EPB41L2</td>
<td>chr6</td>
<td>131,426,133</td>
<td>131,426,440</td>
<td>307</td>
</tr>
<tr>
<td>ERI3</td>
<td>chr1</td>
<td>44,566,114</td>
<td>44,566,442</td>
<td>328</td>
</tr>
<tr>
<td>EXOC6</td>
<td>chr10</td>
<td>94,642,107</td>
<td>94,642,405</td>
<td>298</td>
</tr>
<tr>
<td>FAM120A</td>
<td>chr9</td>
<td>95,368,522</td>
<td>95,368,785</td>
<td>263</td>
</tr>
<tr>
<td>FAM175A</td>
<td>chr4</td>
<td>84,625,124</td>
<td>84,625,373</td>
<td>249</td>
</tr>
<tr>
<td>FAM76A</td>
<td>chr1</td>
<td>27,924,624</td>
<td>27,924,929</td>
<td>305</td>
</tr>
<tr>
<td>FAM98B</td>
<td>chr15</td>
<td>36,533,439</td>
<td>36,533,751</td>
<td>312</td>
</tr>
<tr>
<td>FXYD5</td>
<td>chr19</td>
<td>40,345,685</td>
<td>40,346,020</td>
<td>335</td>
</tr>
<tr>
<td>GEM</td>
<td>chr8</td>
<td>95,343,631</td>
<td>95,343,921</td>
<td>290</td>
</tr>
<tr>
<td>GFM2</td>
<td>chr5</td>
<td>74,098,559</td>
<td>74,098,862</td>
<td>303</td>
</tr>
<tr>
<td>GLTSCR2</td>
<td>chr19</td>
<td>52,938,226</td>
<td>52,938,626</td>
<td>400</td>
</tr>
<tr>
<td>GRAMD1B</td>
<td>chr11</td>
<td>122,956,570</td>
<td>122,956,808</td>
<td>238</td>
</tr>
<tr>
<td>GRHL3</td>
<td>chr1</td>
<td>24,520,816</td>
<td>24,521,143</td>
<td>327</td>
</tr>
<tr>
<td>GRHL3</td>
<td>chr1</td>
<td>24,521,291</td>
<td>24,521,591</td>
<td>300</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Chromosome</td>
<td>Peak Start</td>
<td>Peak Stop</td>
<td>Subtraction</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HDAC4</td>
<td>chr2</td>
<td>239,988,064</td>
<td>239,988,464</td>
<td>400</td>
</tr>
<tr>
<td>HIVEP3</td>
<td>chr1</td>
<td>42,079,086</td>
<td>42,079,486</td>
<td>400</td>
</tr>
<tr>
<td>HSH2D</td>
<td>chr19</td>
<td>16,115,331</td>
<td>16,115,600</td>
<td>269</td>
</tr>
<tr>
<td>ID1</td>
<td>chr20</td>
<td>29,644,861</td>
<td>29,645,161</td>
<td>300</td>
</tr>
<tr>
<td>ID1</td>
<td>chr20</td>
<td>29,645,125</td>
<td>29,645,425</td>
<td>300</td>
</tr>
<tr>
<td>IFT20</td>
<td>chr17</td>
<td>23,686,629</td>
<td>23,686,929</td>
<td>300</td>
</tr>
<tr>
<td>INA</td>
<td>chr10</td>
<td>105,022,838</td>
<td>105,023,067</td>
<td>229</td>
</tr>
<tr>
<td>IQCD</td>
<td>chr12</td>
<td>112,137,226</td>
<td>112,137,452</td>
<td>226</td>
</tr>
<tr>
<td>ISCA1</td>
<td>chr9</td>
<td>88,087,425</td>
<td>88,087,725</td>
<td>300</td>
</tr>
<tr>
<td>ITGA3</td>
<td>chr17</td>
<td>45,494,793</td>
<td>45,495,162</td>
<td>369</td>
</tr>
<tr>
<td>ITGA3</td>
<td>chr17</td>
<td>45,497,094</td>
<td>45,497,394</td>
<td>300</td>
</tr>
<tr>
<td>ITGA6</td>
<td>chr2</td>
<td>173,015,099</td>
<td>173,015,381</td>
<td>282</td>
</tr>
<tr>
<td>JKAMP</td>
<td>chr14</td>
<td>59,020,359</td>
<td>59,020,693</td>
<td>334</td>
</tr>
<tr>
<td>KIAA0284</td>
<td>chr14</td>
<td>104,401,665</td>
<td>104,401,965</td>
<td>300</td>
</tr>
<tr>
<td>KIAA0355</td>
<td>chr19</td>
<td>39,528,600</td>
<td>39,528,926</td>
<td>326</td>
</tr>
<tr>
<td>KRT8</td>
<td>chr12</td>
<td>51,585,035</td>
<td>51,585,277</td>
<td>242</td>
</tr>
<tr>
<td>LARP1B</td>
<td>chr4</td>
<td>129,201,264</td>
<td>129,201,598</td>
<td>334</td>
</tr>
<tr>
<td>LASS6</td>
<td>chr2</td>
<td>169,020,730</td>
<td>169,021,018</td>
<td>288</td>
</tr>
<tr>
<td>LIF</td>
<td>chr22</td>
<td>28,978,443</td>
<td>28,978,743</td>
<td>300</td>
</tr>
<tr>
<td>LOC100129361</td>
<td>chr12</td>
<td>11,214,935</td>
<td>11,215,235</td>
<td>300</td>
</tr>
<tr>
<td>LOC285540</td>
<td>chr4</td>
<td>24,785,474</td>
<td>24,785,716</td>
<td>242</td>
</tr>
<tr>
<td>LOC730101</td>
<td>chr6</td>
<td>52,643,661</td>
<td>52,643,919</td>
<td>258</td>
</tr>
<tr>
<td>MAPK15</td>
<td>chr8</td>
<td>144,872,969</td>
<td>144,873,269</td>
<td>300</td>
</tr>
<tr>
<td>MCM2</td>
<td>chr3</td>
<td>128,799,661</td>
<td>128,799,965</td>
<td>304</td>
</tr>
<tr>
<td>METAP1D</td>
<td>chr2</td>
<td>172,572,931</td>
<td>172,573,197</td>
<td>266</td>
</tr>
<tr>
<td>MGAT5B</td>
<td>chr17</td>
<td>72,380,150</td>
<td>72,380,413</td>
<td>263</td>
</tr>
<tr>
<td>MIB1</td>
<td>chr18</td>
<td>17,574,554</td>
<td>17,574,872</td>
<td>318</td>
</tr>
<tr>
<td>MIOS</td>
<td>chr7</td>
<td>7,570,561</td>
<td>7,570,892</td>
<td>331</td>
</tr>
<tr>
<td>MIR639</td>
<td>chr19</td>
<td>14,501,504</td>
<td>14,501,804</td>
<td>300</td>
</tr>
<tr>
<td>MMAA</td>
<td>chr4</td>
<td>146,759,755</td>
<td>146,760,093</td>
<td>338</td>
</tr>
<tr>
<td>MOBKL2C</td>
<td>chr1</td>
<td>46,842,488</td>
<td>46,842,724</td>
<td>236</td>
</tr>
<tr>
<td>MRPL14</td>
<td>chr6</td>
<td>44,209,351</td>
<td>44,209,651</td>
<td>300</td>
</tr>
<tr>
<td>MUC1</td>
<td>chr1</td>
<td>153,429,415</td>
<td>153,429,771</td>
<td>356</td>
</tr>
<tr>
<td>MYCBP2</td>
<td>chr13</td>
<td>76,799,199</td>
<td>76,799,480</td>
<td>281</td>
</tr>
<tr>
<td>MYO16</td>
<td>chr13</td>
<td>108,662,027</td>
<td>108,662,264</td>
<td>237</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Chromosome</td>
<td>Peak Start</td>
<td>Peak Stop</td>
<td>Subtraction</td>
</tr>
<tr>
<td>---------------</td>
<td>------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MYO1E</td>
<td>chr15</td>
<td>57,356,269</td>
<td>57,356,526</td>
<td>257</td>
</tr>
<tr>
<td>NBN</td>
<td>chr8</td>
<td>91,065,496</td>
<td>91,065,796</td>
<td>300</td>
</tr>
<tr>
<td>NCOR2</td>
<td>chr12</td>
<td>123,551,232</td>
<td>123,551,501</td>
<td>269</td>
</tr>
<tr>
<td>NCRNA00183</td>
<td>chrX</td>
<td>73,080,769</td>
<td>73,081,001</td>
<td>232</td>
</tr>
<tr>
<td>NGFR</td>
<td>chr17</td>
<td>44,940,822</td>
<td>44,941,093</td>
<td>271</td>
</tr>
<tr>
<td>NLRX1</td>
<td>chr11</td>
<td>118,544,501</td>
<td>118,544,801</td>
<td>300</td>
</tr>
<tr>
<td>ORAI1</td>
<td>chr12</td>
<td>120,548,718</td>
<td>120,549,018</td>
<td>300</td>
</tr>
<tr>
<td>OSM</td>
<td>chr22</td>
<td>28,978,443</td>
<td>28,978,743</td>
<td>300</td>
</tr>
<tr>
<td>OTUB1</td>
<td>chr11</td>
<td>63,510,270</td>
<td>63,510,547</td>
<td>277</td>
</tr>
<tr>
<td>P2RY2</td>
<td>chr11</td>
<td>72,607,112</td>
<td>72,607,384</td>
<td>272</td>
</tr>
<tr>
<td>PA2G4</td>
<td>chr12</td>
<td>54,798,052</td>
<td>54,798,300</td>
<td>248</td>
</tr>
<tr>
<td>PARP4</td>
<td>chr13</td>
<td>23,984,716</td>
<td>23,985,055</td>
<td>339</td>
</tr>
<tr>
<td>PFKFB3</td>
<td>chr10</td>
<td>6,321,248</td>
<td>6,321,516</td>
<td>268</td>
</tr>
<tr>
<td>PHLDAA2</td>
<td>chr11</td>
<td>2,911,235</td>
<td>2,911,535</td>
<td>300</td>
</tr>
<tr>
<td>PIWIL2</td>
<td>chr8</td>
<td>22,266,214</td>
<td>22,266,530</td>
<td>316</td>
</tr>
<tr>
<td>PPA1</td>
<td>chr10</td>
<td>71,652,214</td>
<td>71,652,445</td>
<td>231</td>
</tr>
<tr>
<td>PQLC3</td>
<td>chr2</td>
<td>11,215,376</td>
<td>11,215,646</td>
<td>270</td>
</tr>
<tr>
<td>RBM20</td>
<td>chr10</td>
<td>112,422,128</td>
<td>112,422,469</td>
<td>341</td>
</tr>
<tr>
<td>RFC1</td>
<td>chr4</td>
<td>39,044,337</td>
<td>39,044,600</td>
<td>263</td>
</tr>
<tr>
<td>RPTOR</td>
<td>chr17</td>
<td>76,444,875</td>
<td>76,445,144</td>
<td>269</td>
</tr>
<tr>
<td>S100A5</td>
<td>chr1</td>
<td>151,775,169</td>
<td>151,775,426</td>
<td>257</td>
</tr>
<tr>
<td>S100A6</td>
<td>chr1</td>
<td>151,775,132</td>
<td>151,775,432</td>
<td>300</td>
</tr>
<tr>
<td>SFXN4</td>
<td>chr10</td>
<td>120,915,207</td>
<td>120,915,559</td>
<td>352</td>
</tr>
<tr>
<td>SH3PXD2A</td>
<td>chr10</td>
<td>105,435,911</td>
<td>105,436,184</td>
<td>273</td>
</tr>
<tr>
<td>SLC22A4</td>
<td>chr5</td>
<td>131,658,814</td>
<td>131,659,036</td>
<td>222</td>
</tr>
<tr>
<td>SNRNPB2</td>
<td>chr20</td>
<td>16,658,382</td>
<td>16,658,681</td>
<td>299</td>
</tr>
<tr>
<td>SPOP</td>
<td>chr17</td>
<td>45,109,858</td>
<td>45,110,124</td>
<td>266</td>
</tr>
<tr>
<td>SRSF1</td>
<td>chr17</td>
<td>53,439,638</td>
<td>53,440,038</td>
<td>400</td>
</tr>
<tr>
<td>ST20-MTHFS</td>
<td>chr15</td>
<td>77,929,276</td>
<td>77,929,531</td>
<td>255</td>
</tr>
<tr>
<td>STK25</td>
<td>chr2</td>
<td>242,096,538</td>
<td>242,096,838</td>
<td>300</td>
</tr>
<tr>
<td>STX16</td>
<td>chr20</td>
<td>56,659,443</td>
<td>56,659,743</td>
<td>300</td>
</tr>
<tr>
<td>TFAP4</td>
<td>chr16</td>
<td>4,243,765</td>
<td>4,244,057</td>
<td>292</td>
</tr>
<tr>
<td>TFAP4</td>
<td>chr16</td>
<td>4,261,217</td>
<td>4,261,473</td>
<td>256</td>
</tr>
<tr>
<td>TFDP2</td>
<td>chr3</td>
<td>143,141,803</td>
<td>143,142,203</td>
<td>400</td>
</tr>
</tbody>
</table>
Table 7. (Continued)

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Chromosome</th>
<th>Peak Start</th>
<th>Peak Stop</th>
<th>Subtraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>THYN1</td>
<td>chr11</td>
<td>133,628,303</td>
<td>133,628,630</td>
<td>327</td>
</tr>
<tr>
<td>TMEM131</td>
<td>chr2</td>
<td>97,853,481</td>
<td>97,853,714</td>
<td>233</td>
</tr>
<tr>
<td>TPK1</td>
<td>chr7</td>
<td>144,163,960</td>
<td>144,164,264</td>
<td>304</td>
</tr>
<tr>
<td>TRIB1</td>
<td>chr8</td>
<td>126,511,313</td>
<td>126,511,613</td>
<td>300</td>
</tr>
<tr>
<td>TRIM16</td>
<td>chr17</td>
<td>15,503,818</td>
<td>15,504,101</td>
<td>283</td>
</tr>
<tr>
<td>TSPAN14</td>
<td>chr10</td>
<td>82,221,438</td>
<td>82,221,752</td>
<td>314</td>
</tr>
<tr>
<td>TTC7A</td>
<td>chr2</td>
<td>47,067,707</td>
<td>47,067,969</td>
<td>262</td>
</tr>
<tr>
<td>VAV2</td>
<td>chr9</td>
<td>135,728,973</td>
<td>135,729,286</td>
<td>313</td>
</tr>
<tr>
<td>VGLL4</td>
<td>chr3</td>
<td>11,584,620</td>
<td>11,584,886</td>
<td>266</td>
</tr>
<tr>
<td>WDR60</td>
<td>chr7</td>
<td>158,341,877</td>
<td>158,342,142</td>
<td>265</td>
</tr>
<tr>
<td>WDR74</td>
<td>chr11</td>
<td>62,363,500</td>
<td>62,363,760</td>
<td>260</td>
</tr>
<tr>
<td>WSB2</td>
<td>chr12</td>
<td>116,974,638</td>
<td>116,974,938</td>
<td>300</td>
</tr>
<tr>
<td>YARS2</td>
<td>chr12</td>
<td>32,799,330</td>
<td>32,799,616</td>
<td>286</td>
</tr>
<tr>
<td>ZHX1-C8ORF76</td>
<td>chr8</td>
<td>124,356,192</td>
<td>124,356,531</td>
<td>339</td>
</tr>
<tr>
<td>ZNF335</td>
<td>chr20</td>
<td>44,034,252</td>
<td>44,034,552</td>
<td>300</td>
</tr>
<tr>
<td>ZNF462</td>
<td>chr9</td>
<td>108,694,865</td>
<td>108,695,101</td>
<td>236</td>
</tr>
<tr>
<td>ZNF609</td>
<td>chr15</td>
<td>62,601,223</td>
<td>62,601,464</td>
<td>241</td>
</tr>
<tr>
<td>ZNF639</td>
<td>chr3</td>
<td>180,523,057</td>
<td>180,523,357</td>
<td>300</td>
</tr>
</tbody>
</table>

Note: Subtraction=peak end position-peak start position
Figure 4 Sequence logo of TFAP2C consensus motif.
Vertical axis shows the information content of the base frequency at each position. Bits 2 indicates a perfectly conserved nucleotide, and bits 0 illustrates no specific nucleotide is preferred at this site. The horizontal axis refers to consensus base position from 1 to 11.
A. TFAP2C binding motif derived using Partek analysis of ChIP-seq data from 115 peaks listed in Table 7.
B. TFAP2C consensus motifs derived using Partek analysis of ChIP-seq data from 447 TFAP2C target genes that demonstrated >1.3-fold expression change\textsuperscript{59}.
Figure 5 TFAP2C binding peak is significantly augmented on gene MUC1 with introduction of p53. (Lab ChIP-seq dataset). The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). The TFAP2C peak on MUC1 promoter is circled. ChIP-seq data in top panel was generated from HCT 116 p53 +/- cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 +/- cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes. One isoform of MUC1 is shown as representative.
Figure 6 TFAP2C binding peak is significantly augmented on gene DDX17 with introduction of p53.
The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53−/− cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53−/− cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 7 TFAP2C binding peak is significantly augmented on gene MYCBP2 with introduction of p53. The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 8 TFAP2C binding peak is significantly augmented on gene \textit{FAM98B} with introduction of p53. The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 \textit{/-} cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 \textit{/-} cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 9 TFAP2C binding peak is significantly augmented on gene **HDAC4** with introduction of p53. The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 10 TFAP2C binding peak is significantly augmented on gene CNOT3 with introduction of p53.

The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 11 TFAP2C binding peak remained identical on gene \textit{FAM83A} with introduction of p53.

The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 \textit{--/} cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 \textit{--/} cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 12 TFAP2C binding peak remained identical on gene *ST6GALNAC2* with introduction of p53.
The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 /- cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 /- cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 13 TFAP2C binding peak remained identical on gene DLX4 with introduction of p53. The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C; ChIP-seq data in bottom panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and wild type p53. The gene and gene prediction tracks are from database of Refseq Genes.
Figure 14 Western analysis of artificially synthesized TFAP2C and p53 protein in TnT quick coupled transcription/translation system. Lane TnT lysate indicate the TnT lysate without plasmids added, used as a negative control. Panel BT-474 contains whole proteins from BT-474 cell lines. Ladder: MagicMark™ XP Western Protein Standard (20-220 kDa, Life Technologies™). p53 (DO-2) antibody: sc-53394 (Santa Cruz Biotechnology, INC). AP-2γ (V-18) Antibody: sc-31935, Santa Cruz Biotechnology, INC.
MUC1
ACAGGACCTCGAACCCTAGCTGGCTTTGTTCATCCCATCACCACGTAGTTGTGGCCTGAGGCTAACTAGAGCCCAAGGGGCCCACAGTTCCAGACTGCCCCCTC
DDX17
TACGTAACAGCGCCTCCGAATAAGCCCCAGGGGCTCCGCGGGGAACTACGCTTATTGGAAGCAGCCTTTACGTCAGTTGCGACGATGGGTTAGTTATTTGCG
MYCBP2
CCCTTTTCTCTCCCACCCCCCTCTCCGCCACCACCCTCCCGCCCTCGGGGGGTTTGGGTGGGTGGGTTGGCTGCTAAGTAGTCCACTTCT
FAM98B
TTATTGCCTCGGGGCTGGAACCTTTTGCGCCCTTTTCTCAGCTGTGCTCTCAGGCAGTTACCGGCCAGACCCGGACCGGGAACCTTGCGCGGTCGAGCTAAGTTTCC
GCCGGGGCTACTT
HDAC4
TGCGAGTGCGTCACGTGGTGGCTGCTGCGGCAGGAGGGCCGCGCGGCGCCCCTGCTCGCCTGCCCCCGCGGGCGCAGGC
CNOT3
ACGACACAGGTCTCTTGTAGTTGAGCGTTCTGCTTGGCTCGCGTGGCTGCTGCTCTGCTGAGGCAGGAGGGAGGAGGAGGAGGAAGAAAACCTGAGCCAAT

Figure 15 Summarization of sequences for probes of MUC1, DDX17, MYCBP2, FAM98B, HDAC4 and CNOT3.
The binding sites confirmed within this region are underlined.
Figure 16 Titrating TFAP2C to optimize the protein amount in gel shift assay.
A. Gel shift with MUC1 probe using TFAP2C protein.

B. Map of MUC1 probe targeted by TFAP2C with oligo competitors sequences. The top line represents the sequence of MUC1 probe. The following lines show the sequences of competitors, aligned to the probe. The binding sites are shown in red and double-underlines. Nucleotide mutations in competitor 3MU and 4MU are labeled as bold and underlined.

Figure 17 Localizing the TFAP2C binding sequences within MUC1 promoter under the ChIP-seq peak.
Figure 18 Localizing the TFAP2C binding sequences within DDX17 promoter under the ChIP-seq peak.

A. Gel shift with DDX17 probe using TFAP2C protein.
B. Map of DDX17 probe targeted by TFAP2C with oligo competitors sequences. The top line represents the sequence of DDX17 probe. The following lines show the sequences of competitors, aligned to the probe. The binding site is shown in red and double-underlined.
Figure 19 Localizing the TFAP2C binding sequences within *MYCBP2* promoter under the ChIP-seq peak.

A. Gel shift with *MYCBP2* probe using TFAP2C protein.

B. Map of *MYCBP2* probe targeted by TFAP2C with oligo competitors sequences. The top line represents the sequence of *MYCBP2* probe. The following lines show the sequences of competitors, aligned to the probe. The binding site is shown in red and double-underlines. Nucleotide mutations in competitor 3MU are labeled as bold and underlined.
Figure 20 Localizing the TFAP2C binding sequences within *FAM98B* promoter under the ChIP-seq peak.

A. Gel shift with *FAM98B* probe using TFAP2C protein.

B. Map of *FAM98B* probe targeted by TFAP2C with oligo competitors sequences. The top line represents the sequence of *FAM98B* probe. The following lines show the sequences of competitors, aligned to the probe. The binding sites are shown in red and underlined. Nucleotide mutations in competitor 1MU, 3MU and 6MU are labeled as bold and underlined.
Figure 21 Localizing the TFAP2C binding sequences within HDAC4 promoter under the ChIP-seq peak.
A. Gel shift with HDAC4 probe using TFAP2C protein.
B. Map of HDAC4 probe targeted by TFAP2C with oligo competitors sequences.
Figure 22 Localizing the TFAP2C binding sequences within CNOT3 promoter under the ChIP-seq peak.
A. Gel shift with CNOT3 probe using TFAP2C protein.
B. Map of CNOT3 probe targeted by TFAP2C with oligo competitors sequences.
A. List of competitors that fully compete with probes:

- **MUC1**
  - GCCCTGAGGC
  - GCCCAGGGGC
- **DDX17**
  - GCCCGAGGGGC
- **MYCBP2**
  - GCCTCAGGGGC
- **FAM98B**
  - GCCTCGGGGC
  - CCGGCCGGGC
- **HDAC4**
  - GCCGTGGGC
  - CCCCCGGGGC
- **CNOT3**
  - GCCCTGAGGC

B. List of competitors that partly compete with probes:

- **FAM98B**
  - GCTCTCAGGC
- **HDAC4**
  - GCACGGAGGC

C. 

Figure 23. Summary and comparison of TFAP2C binding sites within those six probes. A. List of TFAP2C binding sites within each probe found in gel shift experiments. Conserved nucleotides are shown in bold. B. List of TFAP2C potential weak binding sites within probe FAM98B and HDAC4. Conserved nucleotides are shown in bold. C. TFAP2C binding motif derived using analysis of ChIP-seq data from 114 peaks listed in Table 7.
Figure 24 p53 neither binds to MUC1 probe nor augments TFAP2C occupancy to TFAP2C TFBS.
Figure 25 Western analysis of protein expression of TFAP2A, TFAP2C, p53 and MUC1 in MCF-7 cell line transfected with siRNA interfering TFAP2A, TFAP2C, p53.

M: molecular marker; NT: non-targeting siRNA. MUC1 (isoform 7, Epitomics); AP-2α (EPR2688-2, Epitomics).
Figure 26: Fold changes of mRNA levels of TFAP2A, TFAP2C, p53, and MUC1 in MCF-7 cells transfected with siRNAs targeting TFAP2A, TFAP2C, and p53. Value = mean ± SD. Mean values with standard errors (standard deviation) are generated from biological triplicate. * indicates that P < 0.01. siNT: siNon-Targeting.
Figure 27 Expected schematic diagram of the presence of p53 modifying CpG methylation of MUC1 promoter and enabling TFAP2C to load. Open dot indicates unmethylated CpG sites, and the filled black dot indicates methylated CpG.

A. With absence of p53, CpGs in MUC1 promoter is methylated, TFAP2C hardly load to its TFBS and to activate MUC1 transcription. (HCT 116 p53 -/- cell line).

B. With presence of p53, CpGs in MUC1 promoter is not methylated, TFAP2C occupancy to its TFBS is significantly increased and MUC1 transcription is greatly elevated. (HCT 116 p53 -/- cell line with over-expression of p53).
Nucleotide numbers are given accounting the transcription start site as +1 (arrow).

A. Ideogram of MUC1 promoter covering -492bp to +48 bp. Two TFAP2C binding sites are shown in blue bars. Each dark blue circle indicates one CpG dinucleotides. The CpG are numbered from 1 to 20. TSS: Transcription Start Site.

B. Three DNA fragments cloned in the MUC1 promoter. Fragment A (-492 to -200) is shown in green; fragment O (-213 to -71) is shown in red; fragment X (-158 to 48) is shown in blue.

Figure 28 Ideogram of MUC1 promoter region in DNA methylation assay.
Figure 29 TFAP2C binding peaks on \textit{MUC1} gene in different cell lines. The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in the top four panels were generated from breast cancer cell line, MCF-7, BT474, MDA-MB-453, and SKBR-3, respectively. ChIP-seq data in fifth panel was generated from HCT 116 p53 \(-/-\) cells with over-expressed TFAP2C. ChIP-seq data in sixth panel was generated from HCT 116 p53 \(-/-\) cells with over-expressed TFAP2C and wild type p53. ChIP-seq data in sixth panel was generated from HCT 116 p53 \(-/-\) cells with over-expressed TFAP2C and mutant p53 (EQEQ). The gene and gene prediction tracks are from database of Refseq Genes.
Figure 30 Western analysis of protein expression of WT p53, TFAP2C, and MU p53 in cell lines transfected with plasmids.
Figure 31 Percentage of methylated CpG in MUC1 promoter. 
Each row of squares represents the cytosine methylation pattern obtained from individual clone of the MUC1 promoter. Each column represents one CpG dinucleotide, numbering from 1 to 20 from left to right. Open squares indicate unmethylated CpG sites, and the filled black squares indicate methylated CpG. Fragment O in condition 2 failed to be amplified, therefore no data is available. Fragment A (-492 to -200) is shown in green; fragment O (-213 to -71) is shown in red; fragment X (-158 to 48) is shown in blue. 1. HCT116 p53/- cells transfected with empty vector. 2. HCT116 p53/- cells with TFAP2C over-expressed. 3. HCT116 p53/- cells with TFAP2C and p53 over-expressed. 4. HCT116 p53/- cells with TFAP2C and R175H mutant p53 over-expressed. 5. HCT116 p53 +/- cells. 6. MDA-MB-453 cells.
Figure 32 Percentage of methylated CpG in *MUC1* promoter. Numbers 1 to 20 on X-axis represent 20 individual CpG dinucleotides. Y-axis indicates the percentage methylation of each CpG. Fragment O in condition 2 failed to be amplified, therefore no data is available. 1. HCT116 p53 -/- E.V. 2. HCT116 p53 -/- TFAP2C 3. HCT116 p53 -/- TFAP2C + WTp53 4. HCT116 p53 -/- TFAP2C + MUtp53 5. HCT116 p53 +/- 6. MDA-MB-453
Numbers 1 to 20 on the left line represent 20 individual CpG dinucleotides. Open circle indicates unmethylated CpG sites, and the filled black circle indicates methylated CpG. Fragment O in condition 2 failed to be amplified, therefore no data is available. Numbers 1 to 6 on left line indicate the following cell lines. 1. HCT116 p53-/- cells transfected with empty vector. 2. HCT116 p53 -/- cells with TFAP2C overexpressed. 3. HCT116 p53 -/- cells with TFAP2C and p53 overexpressed. 4. HCT116 p53-/- cells with TFAP2C and R175H mutant p53 overexpressed. 5. HCT116 p53 +/- cells. 6. MDA-MB-453 cells.

<table>
<thead>
<tr>
<th>Cpg</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>11'</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of clones</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>2</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>50%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>3</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>25%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>4</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>10%</td>
<td>0%</td>
<td>0%</td>
<td>10%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>5</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>6</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>17%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

Figure 33 Percentage of methylated CpG in MUC1 promoter.
Figure 34 TFAP2C occupancy is blocked with the introduction of p53 on RPSA promoter region. (ChIP-seq dataset). The arrows under the chromosome ideogram demonstrate the location of this region within the chromosome. Vertical axis indicates TFAP2C occupancy normalized for the total number of reads obtained (counts per 1 million reads). ChIP-seq data in top panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C; ChIP-seq data in middle panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and wild type p53; ChIP-seq data in bottom panel was generated from HCT 116 p53 -/- cells with over-expressed TFAP2C and EQEQ mutant p53. The gene and gene prediction tracks are from database of Refseq Genes.
REFERENCES


59. Woodfield, G. W.; Chen, Y.; Bair, T. B.; Domann, F. E.; Weigel, R. J., Identification of primary gene targets of TFAP2C in hormone responsive breast carcinoma cells. Genes, chromosomes & cancer 2010, 49 (10), 948-62.


79. Abba, M. C.; Nunez, M. I.; Colussi, A. G.; Croce, M. V.; Segal-Eiras, A.; Aldaz, C. M., GATA3 protein as a MUC1 transcriptional regulator in breast cancer cells. Breast cancer research : BCR 2006, 8 (6), R64.


99.


APPENDIX

The sequences of competitor oligonucleotides in EMSA:

MUC1_1F 5'-ACAGGACCTCGACCTAGCTGGCTT-3'
MUC1_1R 5'-AAGCCAGCTAGGTCAGGCTCCTGT-3'
MUC1_2F 5'-TGTTCCCCATCCCCACGTTAGTTG-3'
MUC1_2R 5'-CAACTAACGTGGGGATGGGGAACA-3'
MUC1_3F 5'-TAGTTGTTGCCCTGAGGCTAA-3'
MUC1_3R 5'-GTTTTAGCCTCAGGGCAAACA-3'
MUC1_5F 5'-AAGAGAGCGCGGCCCAGA-3'
MUC1_5R 5'-ACTGGGGGCCCTGGGCTCTAGT-3'
MUC1_6F 5'-GCCCCAAGTTCCAGAGT-3'
MUC1_6R 5'-GAGGGGCAGTCTGGA-3'
MUC1_3MUF 5'-TAGTTGTTTCTGAGGCTAA-3'
MUC1_3MUR 5'-GTTTTAGCCTCAGAAACA-3'
MUC1_5MUF 5'-AAGAGAGCGCGGCCCAG-3'
MUC1_5MUR 5'-ACTGGGGGCCCTGGGCTCTAGT-3'

DDX17_1F 5'-TCACGTAAACAGCGCTRACG-3'
DDX17_1R 5'-TCCGACCTGAATGTATTACG-3'
DDX17_2F 5'-ATAGCCGAGGGCCTGCCGG-3'
DDX17_2R 5'-CCGCGGAGCCCTCGGGA-3'
DDX17_3F 5'-GGAATCGCCATTTGTCTTGG-3'
DDX17_3R 5'-CCAAGAAATACCGGTAC-3'
DDX17_4F 5'-AAGCCTTTACGTTGC-3'
DDX17_4R  5’-GCAACTGACGTAAGATGCTT-3’
DDX17_5F  5’-GACGATGGGTATTTGCG-3’
DDX17_5R  5’-CGCAAAATACACCCATCGTC-3’
DDX17_2longF  5’-ATAGCCCGAGGCTCCGCGGGAACCTACGGGTATTCTTGG-3’
DDX17_2longR  5’-CCAAGAATACCCGTAGTTCCCCCGGAGCCCTCGGGCTAT-3’
MYCBP2_1F  5’-CCCTTTTTCCTCCCAACC-3’
MYCBP2_1R  5’-GGGGGTGGGGAGGAAAAGGG-3’
MYCBP2_2F  5’-TCTCCGCCACCACTCGCAGG-3’
MYCBP2_2R  5’-GGAGGGAGTGGTGGCGGAGA-3’
MYCBP2_3F  5’-CCTCCGCCGCGGCTTTGG-3’
MYCBP2_3R  5’-CCAAAGCCCCGAGGGCAGGG-3’
MYCBP2_4F  5’-GCCTCGGGCTTTGGGTTTG-3’
MYCBP2_4R  5’-CAAACCCCAAGCCCGAGG-3’
MYCBP2_5F  5’-GGGCTGGTGTTGGGTTGG-3’
MYCBP2_5R  5’-TCCGCAACCAACCAGGCCC-3’
MYCBP2_6F  5’-ATCTGAAGTACTGTCCACTCAGAT-3’
MYCBP2_6R  5’-AGAAGTGGACTACTTCAGAT-3’
MYCBP2_3MUF  5’-CCTCGGGGCTTTGGGTTTG-3’
MYCBP2_3MUR  5’-CCAAAGCCCCGAAACGGGAGG-3’
FAM98B_1F 5'-TTATTGCCTCGGGGCTGGGAA-3'  
FAM98B_1R 5'-TTCCCAGGCGGAGGCAATAA-3'  
FAM98B_2F 5'-CCTTTGCGCCCTTCTCAGC-3'  
FAM98B_2R 5'-GCTGAAGAAGGGCGCAAGG-3'  
FAM98B_3F 5'-TGTGCTCTCAGGCAGTTACC-3'  
FAM98B_3R 5'-GGTAACTGCTGAGGACCA-3'  
FAM98B_4F 5'-GGCCAGACCCGGACCGGAAC-3'  
FAM98B_4R 5'-GTTCCGGTCCGGGTCTGGCC-3'  
FAM98B_5F 5'-CTTGGCGCGGTGACGCTAGT-3'  
FAM98B_5R 5'-ACTAGCGTCACCACCAGCCAA-3'  
FAM98B_6F 5'-TAGTTTCCGGCGGGCTACTT-3'  
FAM98B_6R 5'-AAAGTAGCCCGCCCGAAACTA-3'  
FAM98B_1MUF 5'-TTATTGCCTCGGCCGCTGGGAA-3'  
FAM98B_1MUR 5'-TTCCCAGCCGCGAGGCAATAA-3'  
FAM98B_3MUF 5'-TGTGCTCTCATTGAGTTACC-3'  
FAM98B_3MUR 5'-GGTAACTGAAATGAGGACACA-3'  
FAM98B_6MUF 5'-TAGTTTCCGGCTCTACTT-3'  
FAM98B_6MUR 5'-AAAGTAGAAAGCCCGAAACTA-3'  
HDAC4_1F 5'-TGGCAGTGGGTACGTGAGTGGCTGC-3'  
HDAC4_1R 5'-CGCAGCAACCTGACCCACTGCCA-3'  
HDAC4_2F 5'-TCGGCAGGCGCCAGCAGGGCGGCCC-3'  
HDAC4_2R 5'-GGGCGCGCCCTGCCTGGCCGTGCA-3'  
HDAC4_3F 5'-CGCAGGCGGCGGGCTCAGT-3'
HDAC4_3R  5'-AGTGACCGGGCCGCCCTGCG-3'
HDAC4_4F  5'-AGTGCCGAGGCCGCTGGGC-3'
HDAC4_4R  5'-GCCACGGCCCTGCGCCTG-3'
HDAC4_5F  5'-AGGCGGCGTGGGCCTCGCCC-3'
HDAC4_5R  5'-GGGCGAGGCCCACGGCGCCT-3'
HDAC4_6F  5'-CCTGCCCGCGGGAACGCG-3'
HDAC4_6R  5'-CGCGTTCCCGGGAACGCGG-3'
HDAC4_3MU F  5'-CGCAGGGCCGGTTGCTGACT-3'
HDAC4_3MUR  5'-AGTGACCGGAAAGGCCCTGCG-3'
HDAC4_5MU F  5'-AGGCGTTGGGCTCGCCC-3'
HDAC4_5MUR  5'-GGGCGAGGCCCACACGCT-3'
HDAC4_6MU F  5'-CCTGCCCGGCTTCGGGAACCGG-3'
HDAC4_6MUR  5'-CGCGTTCCGAAAGCGGGGCAGG-3'
CNOT3_1F  5'-ACGACACAGGTCCTTGATT-3'
CNOT3_1R  5'-AATCAAGAGACCTGTGCTG-3'
CNOT3_2F  5'-GACGTTCCGGGTCTGCGCT-3'
CNOT3_2R  5'-AGCGTTCCGGGTCTGCGCT-3'
CNOT3_3F  5'-GGCGTGTTGTGCCCTGAGG-3'
CNOT3_3R  5'-GGGAGGAGGAGGAGGAGCGG-3'
CNOT3_4F  5'-GTTGTGCCCTGAGGCGGGAG-3'
CNOT3_4R  5'-CTCCCGCCTCAGGGCAACGCT-3'
CNOT3_5F  5'-GGGAGGAGGAGGAGGAGCG-3'
CNOT3_5R  5'-CCGCTCCTCCTCCTCCTCCC-3'
CNOT3_6F  5’-GGAGGAAAACCTGAGCCAAT-3’
CNOT3_6R  5’-ATTGGCTCAGGTTTTTCTCC-3’
CNOT3_4MUF  5’-GTTGTGTGTTTGAGGGGGGAG-3’
CNOT3_4MUR  5’-CTCCGCCCTCAAAGCACAAC-3’