

**Doctoral Thesis**  
**博士論文**

**Identification of a novel p53 target, COL17A1, that inhibits  
breast cancer cell migration and invasion**

(乳がんの浸潤転移を抑制する新規p53下流遺伝子COL17A1の同定)

ヨードスラン ワラーリー

**YODSURANG VARALEE**

# Contents

<b>ABSTRACT.....</b>	<b>1</b>
<b>INTRODUCTION.....</b>	<b>2</b>
<b>CHAPTER I IDENTIFICATION OF COL17A1 AS A NOVEL P53 TARGET .....</b>	<b>11</b>
SUMMARY .....	11
MATERIAL AND METHODS .....	12
1. Cell lines and treatments.....	12
2. Microarray.....	12
3. TCGA analysis .....	13
4. Gene Ontology (GO) enrichment analysis.....	13
5. RNA extraction and quantitative PCR (qPCR).....	13
6. Western blot analysis .....	14
7. Immunohistochemistry and immunocytochemistry .....	14
8. Antibodies .....	15
9. Animal experiments.....	15
10. Gene reporter assay.....	15
11. Chromatin immunoprecipitation (ChIP) assay.....	16
RESULTS .....	16
1. Screening for p53 downstream targets.....	16
2. Correlation of COL17A1 and p53 status .....	22
3. The p53-dependent induction of COL17A1 in human cells and mouse tissues .....	24
4. COL17A1 is a direct target of p53.....	28
DISCUSSION.....	31
<b>CHAPTER II THE ROLE OF COL17 IN BREAST CARCINOGENESIS.....</b>	<b>32</b>
SUMMARY .....	32

MATERIAL AND METHODS .....	33
1. <i>Colony formation assay</i> .....	33
2. <i>Stable cell line generation</i> .....	33
3. <i>Cell proliferation analysis (ATP assay)</i> .....	34
4. <i>Scratch assay</i> .....	34
5. <i>Invasion assay</i> .....	35
6. <i>TCGA and METABRIC analysis</i> .....	35
7. <i>DNA demethylation</i> .....	36
RESULTS .....	37
1. <i>The role of COL17A1 in cancer cell growth</i> .....	37
2. <i>The role of COL17 in cancer cell migration and invasion</i> .....	42
3. <i>The relation of COL17 and the epithelial to mesenchymal transition (EMT)</i> .....	45
4. <i>The COL17A1 expression in cancer patients</i> .....	47
5. <i>The mechanism that regulates the COL17A1 expression</i> .....	51
DISCUSSION.....	56
<b>CONCLUSION .....</b>	<b>59</b>
<b>REFERENCES.....</b>	<b>61</b>
<b>APPENDIX.....</b>	<b>67</b>
<b>ACKNOWLEDGEMENT.....</b>	<b>68</b>

## List of Figures

Figure 1.1 A schematic of breast cancer progression <sup>2</sup> .....	3
Figure 1.2 The p53 malfunctions in human cancers <sup>9</sup> .....	5
Figure 1.3 A. Pathway analysis of p53-repressed target genes <sup>17</sup> .....	6
Figure 1.3 B. Pathway analysis of p53-induced target genes <sup>17</sup> .....	7
Figure 1.4 Components of hemidesmosome and COL17 localization <sup>29</sup> .....	9
Figure 1.5 The COL17 structure and ectodomain shedding <sup>19</sup> .....	10
Figure 2.1 Screening for p53 downstream targets.....	18
Figure 2.2 MCF10A microarray results of the 17 p53 target gene candidates.....	20
Figure 2.3 TCGA results of the 17 p53 target gene candidates.....	21
Figure 2.4 Correlation of COL17A1 and p53 status.....	23
Figure 2.5 The p53-dependent induction of COL17A1 in human cells.....	25
Figure 2.6 The p53-dependent induction of COL17A1 in mouse tissues.....	27
Figure 2.7 Identification of the p53 binding sequences by Reporter assay.....	29
Figure 2.8 Identification of the p53 binding sequences by ChIP assay.....	30
Figure 3.1 The <i>COL17A1</i> over-expression in cancer cell growth.....	37
Figure 3.2 The <i>COL17A1</i> knockdown in cancer cell growth.....	39
Figure 3.3 The MDA-MB-231 stable cell lines.....	40
Figure 3.4 ATP assay of the stable cell lines.....	41
Figure 3.5 Migration assay (scratch assay) of the stable cell lines.....	43
Figure 3.6 Invasion assay of the stable cell lines.....	44
Figure 3.7 The relation of COL17 and EMT.....	46
Figure 3.8 <i>COL17A1</i> expression in multiple cancers.....	48
Figure 3.9 <i>COL17A1</i> depletion is associated with tumor progression and poor prognosis.....	49
Figure 3.10 <i>COL17A1</i> is an independent prognostic factor.....	50
Figure 3.11 Mutation in <i>COL17A1</i> .....	51

Figure 3.12 Copy number variation in <i>COL17A1</i> .....	52
Figure 3.13 <i>COL17A1</i> DNA methylation in breast cell lines.....	53
Figure 3.14 <i>COL17A1</i> DNA methylation in patients .....	54
Figure 3.15 The regulation of <i>COL17A1</i> expression in luminal B subtype.....	55
Figure 4. The regulation of breast cancer metastasis by the p53-COL17 pathway.....	60

## List of Tables

Table 1. List of 17 p53 target gene candidates from the screen .....	19
Table A1: Sequences of DNA and RNA oligonucleotides .....	67

## Abstract

This thesis contains two main chapters. The first chapter is to identify p53 downstream target genes using two transcriptome datasets, including cDNA microarrays of MCF10A breast epithelial cells with wild-type *p53* (*p53*<sup>+/+</sup>) or *p53*-null background (*p53*<sup>-/-</sup>), and RNA sequence analysis of breast invasive carcinoma. We identified 209 genes that were up-regulated following the induction of p53 by doxorubicin (Adriamycin<sup>®</sup>, ADR) in wild-type cells. Among these genes, 17 genes were also up-regulated in breast invasive carcinoma tissues with wild-type *p53* compared to those with mutant *p53*. The Gene Ontology analysis of 17-gene set identified epidermis development and ectoderm development, which *COL17A1* participates, as the significantly up-regulated pathways by wild-type *p53*. The *COL17A1* expressions increased in a p53-dependent manner in human breast cells and mouse mammary tissues. Reporter assay and ChIP assay identified intronic p53-binding sequences in the *COL17A1* gene.

The second chapter is to elucidate the role of *COL17A1* gene product (COL17) in breast carcinogenesis. COL17 showed positive results in inhibition of breast cancer metastasis, yet the relation to cancer cell growth was not clear. The MDA-MB-231 cells that genetically over-express COL17 exhibited reduced migration and invasion *in vitro*. Similarly, *COL17A1* expression was decreased in metastatic tumors compared to primary tumors and normal tissues, even from the same patients. Moreover, high *COL17A1* expression was associated with longer survival of patients with invasive breast carcinoma. In conclusion, we revealed that *COL17A1* is a novel p53 transcriptional target in breast tissues that inhibits cell migration and invasion and is associated with better prognosis.

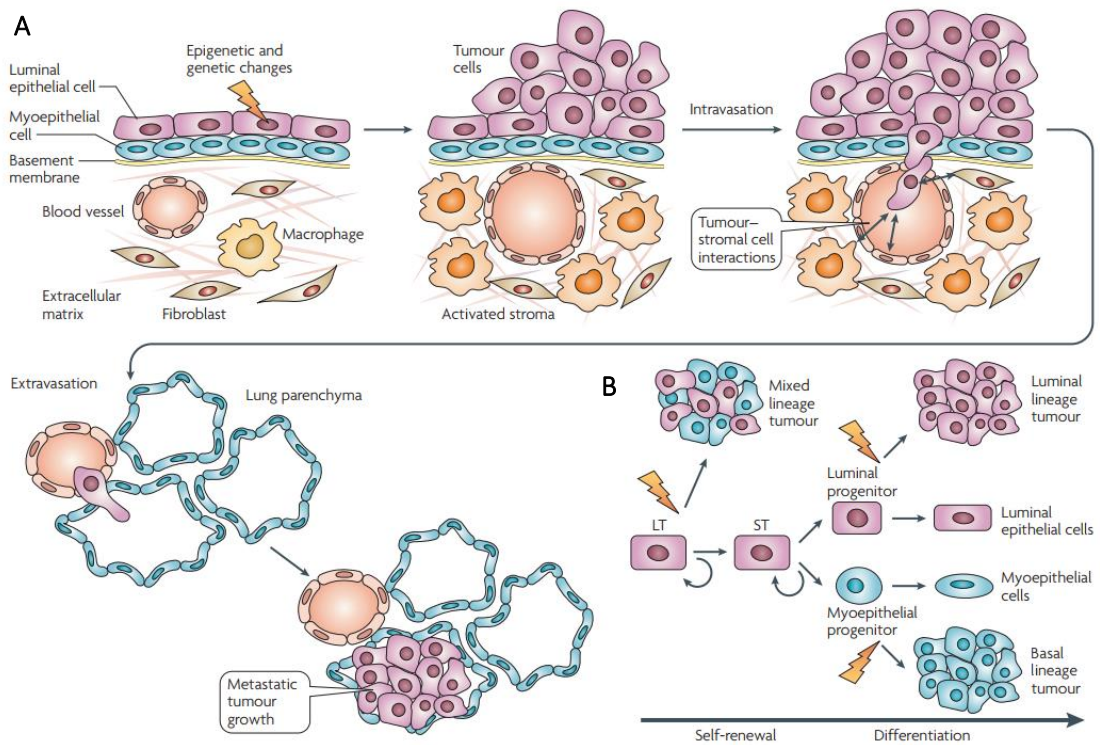
# Introduction

## Breast cancer

Breast cancer is the most frequently diagnosed cancer in women and leading cause of cancer death in worldwide<sup>1</sup>. The normal breast terminal ductal lobular unit (TDLU) contains lobules and ducts that consist of a bi-layered epithelium of luminal and myoepithelial cells (Figure 1.1)<sup>2</sup>. The subtypes of breast cancer are classified by microscope analysis of tissue (as in situ or non-invasive and invasive), by receptor status (as hormone-positive, hormone-negative, HER2 positive, triple-positive, and triple-negative), and by molecular features (as luminal A, luminal B, HER2 type, and basal type or triple-negative)<sup>3</sup>. The prognosis of invasive breast cancer is strongly influenced by the stage of the disease (stage I to IV) – that is the extent or spread of the cancer when it is first diagnosed. However, nearly 30% of women initially diagnosed with early-stage disease will ultimately develop metastatic lesions, often months or even years later<sup>4</sup>.

Typical sites of metastatic relapse for breast tumor are bone, lungs, liver, and brain<sup>5</sup>. Prognostic markers, a characteristic of a patient or tumor at the time of diagnosis, are used to estimate the risk for developing metastases or the chance of the disease recurring in the absence of therapy<sup>6</sup>. Several genes, proteins markers, and pathways involved in metastasis were identified<sup>6-8</sup>. Nowadays, the risk of developing breast cancer can be predicted by genetic tests of *BRCA1*, *BRCA2*, and *PALB2*. New biomarkers are urgently needed to estimate the risk for metastases that might enable oncologists to begin tailoring treatment strategies to individual patients<sup>6</sup>.





**Figure 1.1 A schematic of breast cancer progression<sup>2</sup>**

**A.** The transformation of breast epithelial cells to give rise to metastatic breast cancer is a combination of epigenetic and genetic changes and aberrant interactions within the microenvironment. **B.** Researchers proposed that cancer cells with stem cell-like characteristics (called ‘cancer stem cells’ or ‘tumour-initiating cells’) drive breast cancer initiation, progression and recurrence. This hypothesis is depicted as epigenetic and genetic alterations that occur in different stem or progenitor cells, including the long term (LT), short term (ST) and luminal or basal (myoepithelial) progenitors, and give rise to different subtypes of tumours that consist of different cell types

## **Tumor suppressor p53**

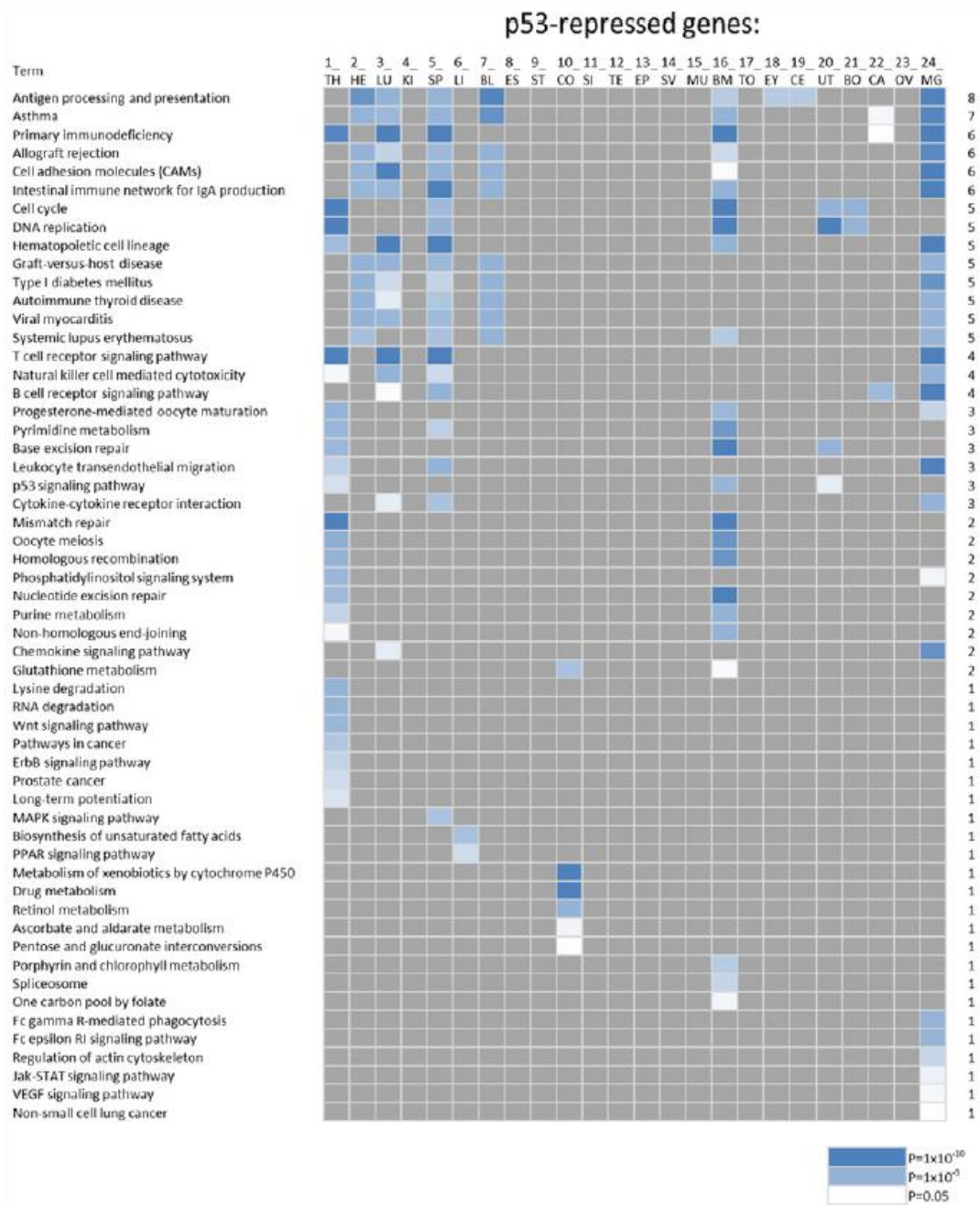
The tumor suppressor p53 functions as a transcription factor and contributes to human cancers in different ways. The loss of p53 function by several mechanisms, including missense mutation, multiplication of the MDM2 gene, and mislocalization of p53 to the cytoplasm, were found in various types of cancers (Figure 1.2)<sup>9</sup>. Somatic mutations in the *p53* gene are one of the most frequent alterations in human cancers, and germline mutations causes predisposition to early-onset cancers including breast carcinomas, sarcomas, brain tumors, and adrenal cortical carcinomas, defining the Li-Fraumeni (LFS) and Li-Fraumeni-like (LFL) syndromes<sup>10</sup>.

The mutation frequency of the tumor suppressor *p53* in breast cancer is relatively low compared to other solid tumors<sup>11</sup>; however, this mutation is the second most frequent genetic alteration observed in 30-35% of breast cancer cases<sup>12,13</sup>. A mutation in *p53* is also associated with aggressive subtypes, i.e., 12% in luminal A, 32% in luminal B, 75% in HER2, and 84% in triple-negative tumors<sup>12</sup>. Previous studies have shown that *p53* mutation is an independent marker of poor prognosis in breast cancers<sup>10</sup> and is also associated with the response to specific treatment regimens in breast cancer<sup>14</sup>. Most of *p53* somatic mutations are in the DNA binding domain and most of them are missense, leading to abrogate the regulation to its target genes<sup>10</sup>. Thus, the dysregulation of p53 targets via p53 inactivation may relate to the poor prognosis of breast cancer patients. We have previously identified several p53 targets and elucidated the molecular mechanism by which p53 regulates apoptosis, the cell cycle, senescence, iron metabolism, and post-translational modifications (Figure 1.3 A and B)<sup>15-17</sup>. However, the role of p53 in breast carcinogenesis has not been fully elucidated. Thus, the identification of p53 targets in breast tissues is important to understand the pathogenesis of breast cancer.

Mechanism of inactivating p53	Typical tumours	Effect of inactivation
Amino-acid-changing mutation in the DNA-binding domain	Colon, breast, lung, bladder, brain, pancreas, stomach, oesophagus and many others	Prevents p53 from binding to specific DNA sequences and activating the adjacent genes
Deletion of the carboxy-terminal domain	Occasional tumours at many different sites	Prevents the formation of tetramers of p53
Multiplication of the MDM2 gene in the genome	Sarcomas, brain	Extra MDM2 stimulates the degradation of p53
Viral infection	Cervix, liver, lymphomas	Products of viral oncogenes bind to and inactivate p53 in the cell, in some cases stimulating p53 degradation
Deletion of the p14 <sup>ARF</sup> gene	Breast, brain, lung and others, especially when p53 itself is not mutated	Failure to inhibit MDM2 and keep p53 degradation under control
Mislocalization of p53 to the cytoplasm, outside the nucleus	Breast, neuroblastomas	Lack of p53 function (p53 functions only in the nucleus)

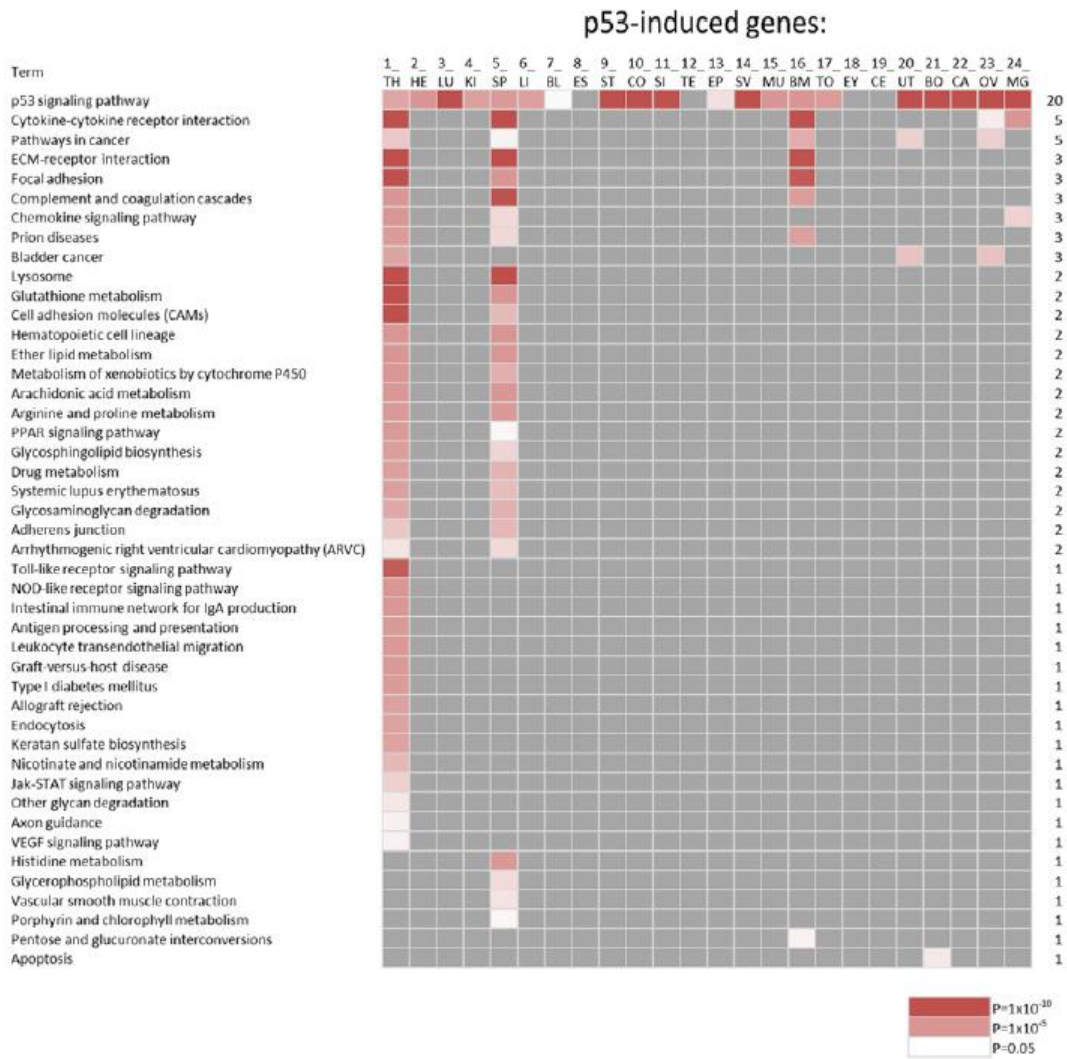
**Figure 1.2 The p53 malfunctions in human cancers<sup>9</sup>**

Three columns shows six mechanisms associated with p53 inactivation, the tumors typically associated with each type of inactivation, and the cellular effect of each type of inactivation.



**Figure 1.3 A. Pathway analysis of p53-repressed target genes<sup>17</sup>**

A pathway analysis of p53-repressed genes in 24 types of tissue. Each color represents the *P*-value as shown below. The right column shows the number of tissues used in the analysis.



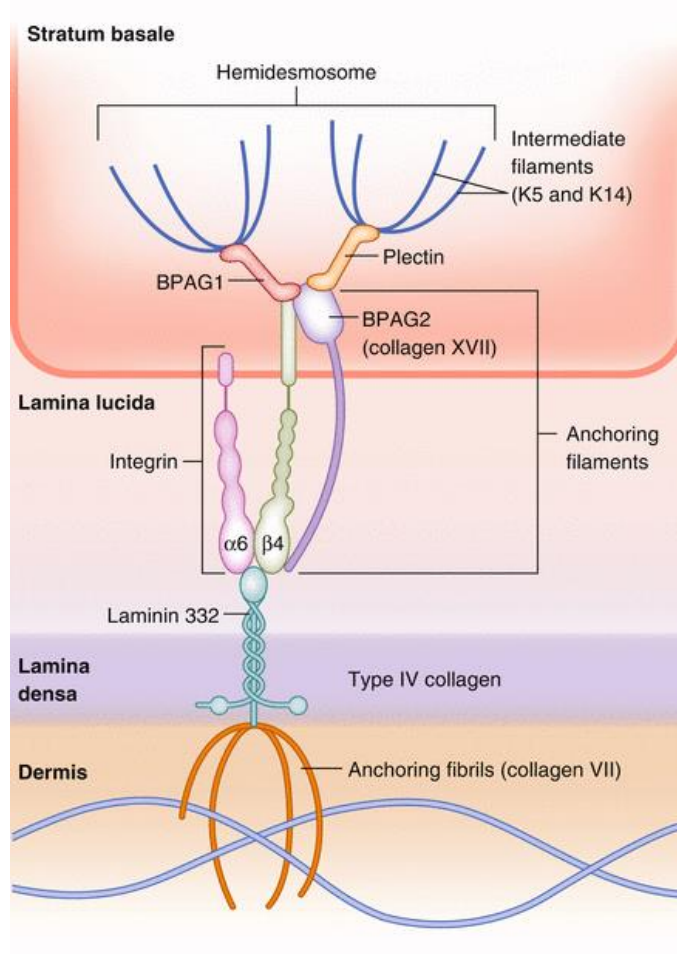
**Figure 1.3 B. Pathway analysis of p53-induced target genes<sup>17</sup>**

A pathway analysis of p53-induced genes in 24 types of tissue. Each color represents the *P*-value as shown below. The right column shows the number of tissues used in the analysis.

## Collagen XVII alpha I

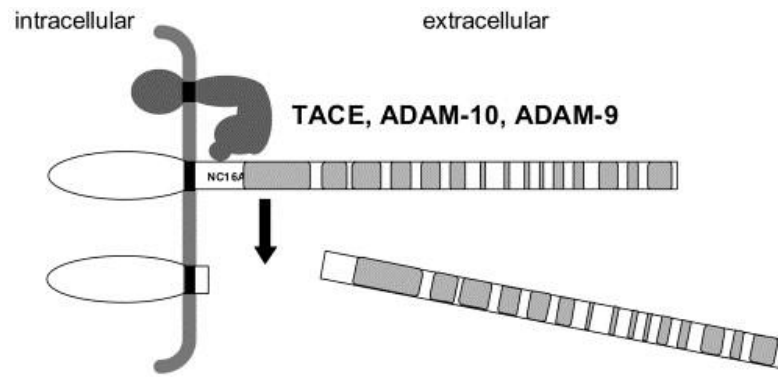
*COL17A1* encodes Collagen XVII (COL17; formerly known as BP180 or BPAG2), a 180-kDa transmembrane protein. COL17 is an essential component of type I hemidesmosomes (HDs) and functions as a cell-matrix adhesion molecule<sup>18</sup> (Figure 1.4). Each COL17 molecule is a trimer of 180 kDa alpha 1 (XVII) chains, with a globular intracellular N-terminal domain, a short transmembrane stretch, and the flexible extracellular C-terminal ectodomain (Figure 1.5). The soluble 120-kDa ectodomain, also known as LAD-1, is generated by constitutive shedding of the full-length form and has a role in regulating the function of COL17 in skin homeostasis<sup>19</sup>.

COL17 is highly expressed in tissues with a prominent epithelial component, including the mammary gland<sup>20</sup>. Autoimmunity to COL17 and mutations in *COL17A1* result in blistering skin diseases caused by a loss of attachment between the epidermis and the underlying basement membrane<sup>21,22</sup>. Type I HDs were observed in normal epithelial cells but were lost in cancer cells, including invasive breast cancer<sup>23</sup>, and pancreatic ductal epithelium<sup>24</sup>. COL17 has been previously reported as a down-regulated protein signature in stage II compared to premalignant cells and in premalignant cells compared to normal myoepithelial cells<sup>25</sup>. Although recent studies have suggested a role of COL17 in cell migration<sup>26-28</sup>, its function in breast carcinogenesis has never been investigated. Here, we revealed *COL17A1* as a novel downstream target gene of p53 that is suppressed in breast cancer tissues with a *p53* mutation.



**Figure 1.4 Components of hemidesmosome and COL17 localization<sup>29</sup>**

The boundary between the epidermis and the underlying dermis is established by deposition of a specialized layer of extracellular matrix (ECM) called the basement membrane. The stratum basale has a high density of hemidesmosomes and integrin-based adhesions that maintain cell attachment to the underlying basement membrane, which is composed of the lamina lucida and lamina densa. Hemidesmosomes connect to the basement membrane through  $\alpha6\beta4$  integrins and the transmembrane protein bullous pemphigoid antigen 2 (BPAG2; also known as collagen XVII), and are tethered to intermediate filaments by the plakin family members plectin and BPAG1.



**Figure 1.5 The COL17 structure and ectodomain shedding<sup>19</sup>**

The COL17 molecule is a trimer of three 180 kDa alpha 1 (XVII) chains, with a globular intracellular N-terminal domain of 466 amino acids, a short transmembrane domain of 23 amino acids and the extracellular C-terminal domain of 1008 amino acids in length. The ectodomain consists of 15 collagenous sequences (grey bars) intervened by 16 short non-collagenous sequences (white bars). The soluble 120-kDa ectodomain released into the extracellular space by cleavage of the COL17 within the juxtamembranous NC16A domain by membrane-anchored metalloproteinases of a disintegrin and metalloproteinase (ADAM) family (TACE, ADAM-9 and ADAM-10).



# Chapter I

## Identification of *COL17A1* as a novel p53 target

### Summary

To identify p53 targets in breast tissues, we performed transcriptome analysis using non-tumorigenic breast epithelial cell lines with or without wild-type *p53* (MCF10A *p53*<sup>+/+</sup> and MCF10A *p53*<sup>-/-</sup>, respectively). MCF10A *p53*<sup>+/+</sup> and MCF10A *p53*<sup>-/-</sup> cells were treated with 0.5 µg/ml of doxorubicin (trade name Adriamycin<sup>®</sup>, ADR). Total RNA was isolated at 12, 24, and 48 hours after ADR treatment and then subjected to cDNA microarray analysis. Next, we analyzed the mRNA expression of genes in breast invasive carcinoma cohort (BRCA), TCGA. We selected 17 p53 target gene candidates that were up-regulated after the p53 induction in MCF10A with wild-type *p53* and whose expressions were significantly high in breast cancer tissues with wild-type *p53*. We performed Gene Ontology (GO) analysis of this 17-gene set to examine the common pathways up-regulated by wild-type *p53*. The most and second most enriched pathways containing the similar set of genes; one of these is *COL17A1*, a new p53 target gene candidate whose relation to p53 had never been investigated in any detail study.

To verify the correlation of *COL17A1* and p53 status, the mRNA and protein expression of *COL17A1* were investigated after an induction or a knockdown of endogenous wild-type *p53* in human breast cells and in mouse mammary tissues. The p53-binding sequences were identified in the *COL17A1* genomic region using reporter assay and ChIP assay. To examine a specificity of p53, the binding signals were observed after *p53* or the *COL17A1* sequences had been mutated.

## Material and methods

### 1. Cell lines and treatments

The non-tumorigenic breast epithelial cell line MCF10A  $p53^{+/+}$  and its isogenic  $p53$  knockout cell line (MCF10A  $p53^{-/-}$ ) were purchased from Sigma Aldrich (St. Louis, MO, USA). HBC4 cells were a gift from Dr. Takao Yamori (Japanese Foundation for Cancer Research, Tokyo, Japan). The other cell lines were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA). Cell cultures were maintained under their depositors' recommendations. Cells were transfected with plasmids using FuGENE6 (Promega, Madison, WI, USA). siRNA oligonucleotides (see sequences in Appendix Table A1), commercially synthesized by Sigma Genosis (St Louis, MO, USA), were reverse transfected with Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's protocol. To induce  $p53$ , cells were continuously incubated with ADR for 2 hours on the day following cell plating. Based on the cytotoxic effects of the drug, MCF10A, HBL-100, and HBC4 cells were treated with 0.5, 1, and 2  $\mu\text{g}/\text{mL}$  of ADR, respectively.

### 2. Microarray

Total RNA was isolated from MCF10A  $p53^{+/+}$  or MCF10A  $p53^{-/-}$  cells at 12, 24, and 48 hours after ADR treatment, as well as from non-treated cells as described<sup>30</sup>. The integrity and purity of the RNA templates were determined using a 2100 Bioanalyzer (Agilent Technologies). Agilent's One-Color Quick Amp labeling kit (Agilent Technologies) was used to generate fluorescent cRNA using cyanine 3-labeled targets according to the manufacturer's instructions. Gene expression analysis was performed using SurePrint G3 8x60K microarray (Agilent Technologies). The fold-change after ADR in the MCF10A  $p53^{+/+}$  cells was calculated using the following equation:

$$\text{Fold} = \frac{\text{Median expression of probe in } [p53^{+/+} \text{ ADR}^+ (12, 24, 48 \text{ h})]}{\text{Maximum expression of probe in } [p53^{+/+} \text{ ADR}^- \text{ and } p53^{-/-} \text{ all}]}$$

An  $F$ -test and two-tailed Student's  $t$ -test were performed to calculate the  $P$ -value.

Genes demonstrating at least one probe satisfied the screening criteria, i.e., above 3-fold and  $P < 0.05$ , were included. The MCF10A microarray data is available from the NCBI GEO database (GSE98727).

### 3. TCGA analysis

The mRNA expression of p53 target gene candidates, the p53 mutation status, and clinical data were obtained from TCGA project by cBioPortal<sup>31,32</sup>. For the differential expression analysis, 1093 tumor tissues were categorized according to p53 mutation status and were subjected to a box plot analysis. The fold-change in wild-type p53 tumors was calculated using the following equation:

$$\text{Fold} = \frac{\text{Median expression in [p53 wild-type tumors]}}{\text{Median expression in [p53 mutant tumors]}}$$

An  $F$ -test and two-tailed Student's  $t$ -test were performed to calculate the  $P$ -value as described<sup>30</sup>. Genes that satisfied the screening criteria, i.e., above 2-fold and  $P < 0.05$ , were included.

### 4. Gene Ontology (GO) enrichment analysis

The resultant 17-gene set from the screening was submitted for Gene Ontology analyses using the DAVID Functional Annotation Tool [<https://david.ncifcrf.gov/home.jsp>]. The enrichment score for biological process was calculated using the chi-squared test.

### 5. RNA extraction and quantitative PCR (qPCR).

Total RNA was isolated from cells and tissues using RNeasy Plus Mini Kits (Qiagen) according to the manufacturer's instructions. Complementary DNA molecules were synthesized using the SuperScript III reverse transcriptase (Invitrogen). qPCR was performed using SYBR Green Master Mix and a Light Cycler 480 (Roche). The primer sequences are shown in Appendix Table A1.

## **6. Western blot analysis**

Cells were harvested by scraping and were lysed using chilled RIPA buffer (Thermo Fisher Scientific) containing 1 mM PMSF, 0.1 mM DTT, and 0.1% Protease Inhibitor Cocktail Set III (Calbiochem). The cell lysates were sonicated using a Bioruptor UCD-200 (Cosmobio, Tokyo, Japan) and centrifuged at  $15,000 \times g$  for 15 minutes at 4°C. Protein concentrations were measured using a BCA™ protein assay (Thermo Fisher Scientific). For media extract, cell culture media were changed to antibiotic-free media 12 hours before harvest. Proteins were precipitated in chilled acetone, incubated for 1 hour at -80°C, and centrifuged at  $15,000 \times g$  for 15 minutes at 4°C. The cell lysates and precipitated protein from the media were added to SDS sample buffer (Bio-Rad), boiled for 5 minutes, and separated using SDS-PAGE. The proteins were transferred onto nitrocellulose membranes, which were subsequently blocked in 5% milk. The membranes were incubated with primary antibodies according to the manufacturers' protocols. The membranes were incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies for 1 hour at room temperature. The immunoblots were developed using Amersham™ ECL (GE Healthcare), and images were acquired using the luminescent image analyzer LAS-4000 mini (Fujifilm, GE Healthcare). The western blot signal representing protein level was obtained by image quantification using ImageJ software.

## **7. Immunohistochemistry and immunocytochemistry**

For immunocytochemistry, cells were plated onto glass coverslips, fixed with 4% paraformaldehyde, and permeabilized in 0.2% Triton-X in PBS. The samples were blocked with 3% BSA, stained with anti-Collagen XVII, incubated with HRP-conjugated secondary antibodies, and counterstained with DAPI. The mounted coverslips were visualized using a confocal microscope (Olympus FluoView FV1000). For immunohistochemistry, paraffin sections of mouse mammary tissues were stained using anti-Collagen XVII according to the manufacturer's protocol. For visualization, the sections were incubated with HRP-labeled

polymer anti-rabbit (Dako) and DAB (Dako) was used as a chromogen. Then, the samples were counterstained with hematoxylin.

## 8. Antibodies

Anti-Collagen XVII (ab184996, Abcam) was used for all human and mouse experiments. Anti-actin (AC15, Sigma-Aldrich), anti-p53 (OP43, Merck Millipore), and anti-p53-S15P (9284, Cell Signaling) were used in the western blot analyses.

## 9. Animal experiments

The *p53* knockout C57BL/6J mice were provided by RIKEN BioResource Center (Ibaraki, Japan)<sup>33</sup>. The mice were maintained under specific pathogen-free conditions and were handled according to the Guidelines for Animal Experiments of the University of Tokyo. The mouse genotypes were confirmed by PCR analyses. Primer sequences are shown in Appendix Table A1. The *p53* wild-type and knockout female mice at 10 weeks of age were exposed to 10 Gy of X-rays using an X-ray irradiation system (MBR-1520R-3, Hitachi). At 24 hours after irradiation, the mice were sacrificed for mammary tissue extraction.

## 10. Gene reporter assay

A DNA fragment that includes the potential p53 binding sites of *COL17A1* was amplified and subcloned into the pGL4.24[*luc2P*/minP] vector (Promega). To create a mutant vector, point mutations were introduced at the 4<sup>th</sup> and 14<sup>th</sup> nucleotides (C to T mutations) and the 7<sup>th</sup> and 17<sup>th</sup> nucleotides (G to T mutations) within the consensus p53 binding site using site-directed mutagenesis. Reporter assays were performed using the Dual-Glo Luciferase Assay System (Promega) according to the manufacturer's protocol. The H1299 cells were co-transfected with the analyzed constructs and the control vector (pcDNA3.1), wild-type p53 (p53), or mutant p53 (p53R175H). The primer sequences for amplification and site-directed mutagenesis are shown in Appendix Table A1.

## 11. Chromatin immunoprecipitation (ChIP) assay

ChIP assays were performed using EZ-Magna ChIP G Chromatin Immunoprecipitation Kits (Merck Millipore, Darmstadt, Germany) according to the manufacturer's protocol. Briefly, HBC4 cells with or without ADR treatment were cross-linked with 1% formaldehyde for 10 minutes, washed with PBS, and lysed in nuclear lysis buffer. The lysate was then sonicated using Bioruptor UCD-200 (Cosmo Bio, Tokyo, Japan) to shear the DNA into fragments of approximately 200-1000 bp. The supernatant from  $1 \times 10^6$  cells was used for each immunoprecipitation with anti-p53 antibody (OP140, Merck Millipore) or mouse IgG (SC-2025, Santa Cruz). Before immunoprecipitation, 1% of the supernatant was removed as "input". Column-purified DNA was quantified by qPCR using primers for p53 binding site in *WAF1* and *COL17A1* BS primers (Appendix Table A1).

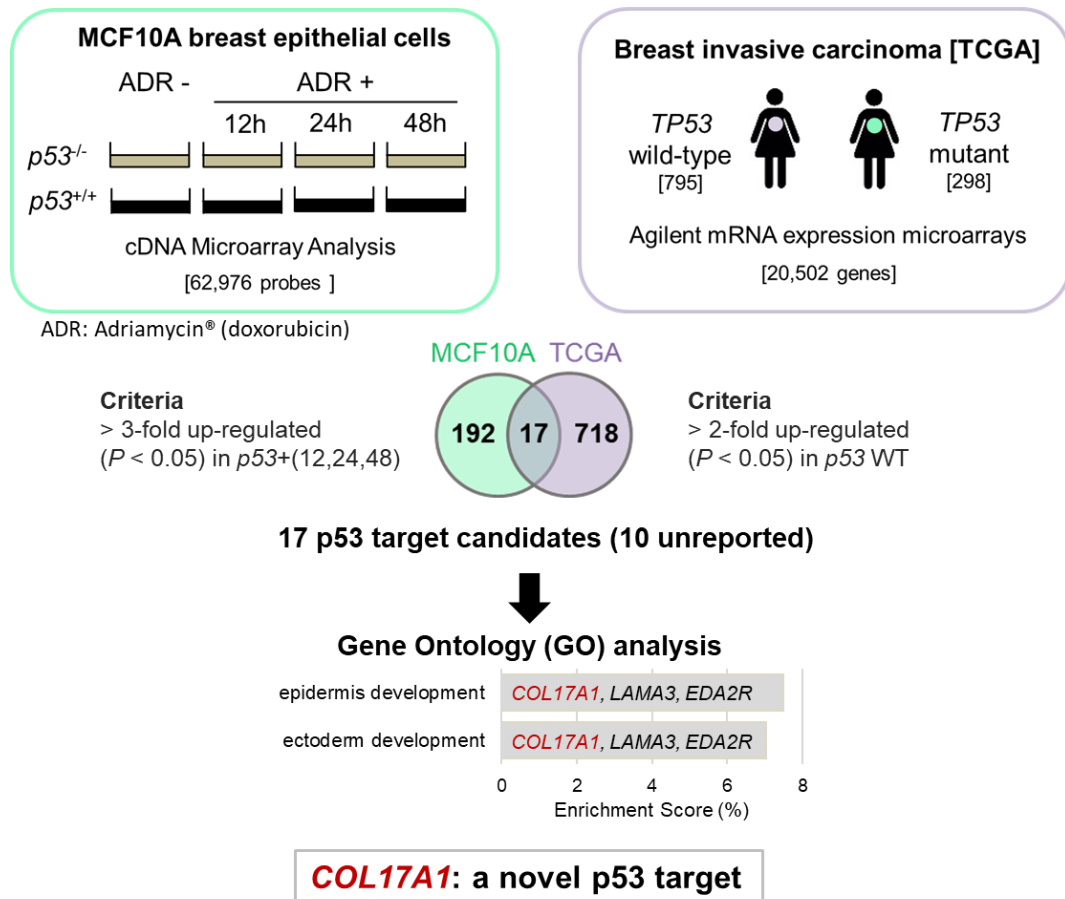
## Results

### 1. Screening for p53 downstream targets

In the first screening, we identified list of 209 genes with fold inductions that were up-regulated more than 3-fold ( $P < 0.05$ ) at 12, 24, or 48 hours after ADR treatment in MCF10A *p53*<sup>+/+</sup> cells compared to MCF10A *p53*<sup>+/+</sup> cells without ADR treatment or MCF10A *p53*<sup>-/-</sup> cells at any timepoint (Figure 2.1). Second, we used the data obtained from a breast invasive carcinoma cohort (BRCA), TCGA. We selected 735 genes that were up-regulated more than two-fold in breast cancer tissues with wild-type *p53* compared to those with mutant *p53* ( $P < 0.05$ ). The 17 overlapping genes in the two analyses, including 7 reported p53 target genes, is shown in Table 1, and those results are displayed in Figure 2.2 and 2.3. Among them, 16 genes except *SYTL2* exhibited significant association even after Bonferroni correction ( $P < 0.05/209$ ) using the number of genes those showed significant association in the first screening (MCF10A *p53*<sup>+/+</sup> and MCF10A *p53*<sup>-/-</sup> cells), indicating low possibility of false positive associations.

Therefore, those set of genes are likely to be regulated by p53 both *in vitro* and *in vivo*.

The overlapping genes from the two analyses were further investigated for enriched biological process using Gene Ontology (GO) enrichment analysis (Figure 2.1). Epidermis development and ectoderm development were the most and second most significant GO terms, respectively. *COL17A1*, *LAMA3*, and *EDA2R* were included in both terms. *EDA2R* (ectodysplasin A2 receptor, XEDAR) is a known p53 target in several cell types including cells in the breast<sup>34,35</sup>. *LAMA3* encodes a subunit of laminin-5 (Laminin-332) whose expression is altered by mutant p53 in a breast cell line<sup>36</sup>. A positive correlation between p53 level and the transcription level of *COL17A1* has been suggested by high-throughput analyses<sup>37,38</sup>, although direct regulation of *COL17A1* by p53 was not investigated previously.



### Figure 2.1 Screening for p53 downstream targets

The p53 targets were screened using MCF10A microarray analysis and invasive breast carcinoma (TCGA). A Venn diagram shows the number of genes that satisfied the inclusion criteria from each analysis and the 17 overlapping genes. GO biological processes of the 17-gene set with *P* < 0.05 and enrichment score > 5% are shown in the bar chart with indicated participating genes.



**Table 1. List of 17 p53 target gene candidates from the screen**

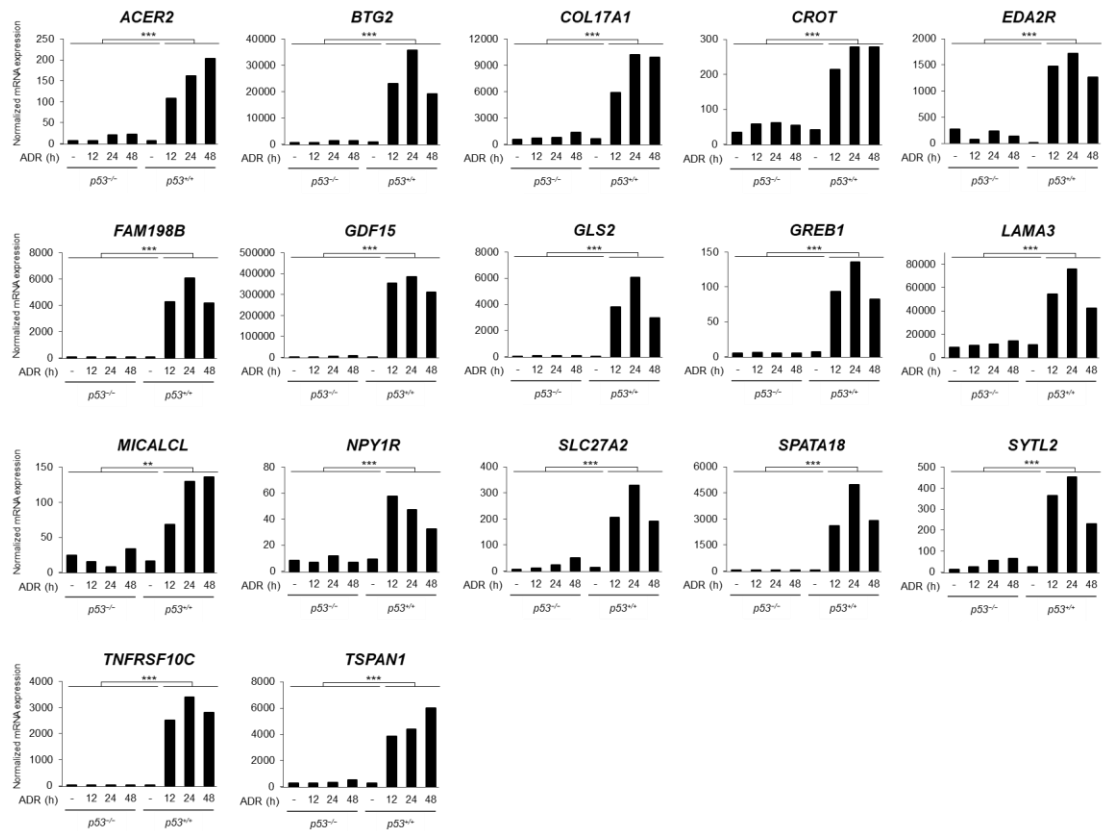
Gene	Description	MCF10A		TCGA		Accession No.	Reference PMID*
		FC <sup>1</sup>	P-value <sup>2</sup>	FC <sup>3</sup>	P-value <sup>b</sup>		
<i>ACER2</i> *	alkaline ceramidase 2	7	0.0320	2	1.83E-27	NM_001010887	26943039
<i>BTG2</i> *	BTG anti-proliferation factor 2	17	0.0382	3	3.47E-44	NM_006763	11814693, 8944033
<i>COL17A1</i>	collagen type XVII alpha 1 chain	7	0.0284	3	1.44E-04	NM_000494	
<i>CROT</i>	carnitine O-octanoyltransferase	4	1.80E-05	2	5.56E-06	NM_001243745	
<i>EDA2R</i> *	ectodysplasin A2 receptor	5	2.21E-05	2	1.51E-46	NM_001242310	19543321
<i>FAM198B</i>	family with sequence similarity 198 member B	78	0.0154	2	2.36E-14	NM_016613	
<i>GDF15</i> *	growth differentiation factor 15	52	0.0036	3	1.35E-12	NM_004864	17276395
<i>GLS2</i> *	glutaminase 2	51	0.0438	3	3.41E-21	NM_013267	20351271
<i>GREB1</i>	growth regulation by estrogen in breast cancer 1	13	0.0260	8	8.25E-23	NM_014668	
<i>LAMA3</i>	laminin subunit alpha 3	4	0.0405	3	4.88E-08	NM_198129	
<i>MICALCL</i>	MICAL C-terminal like	4	0.0457	2	3.88E-10	NM_032867	
<i>NPY1R</i>	neuropeptide Y receptor Y1	4	0.0346	5	2.34E-13	NM_000909	
<i>SLC27A2</i>	solute carrier family 27 member 2	4	0.0342	3	3.55E-09	NM_003645	
<i>SPATA18</i> *	spermatogenesis associated 18	45	0.0437	3	1.52E-37	NM_145263	21300779
<i>SYTL2</i>	synaptotagmin like 2	5	0.0337	3	0.0042	NM_032943	
<i>TNFRSF 10C</i> *	TNF receptor superfamily member 10c ( <i>TRID</i> , <i>TRAIL-R3</i> )	162	0.0080	2	6.71E-23	NM_003841	10435597
<i>TSPAN1</i>	tetraspanin 1	8	0.0203	4	5.57E-09	NM_005727	

<sup>1</sup> Fold-change, the up-regulated fold expression after ADR in p53 wild-type cells (see calculation in Material and methods).

<sup>2</sup> Two-tailed Student's *t*-test *P*-value.

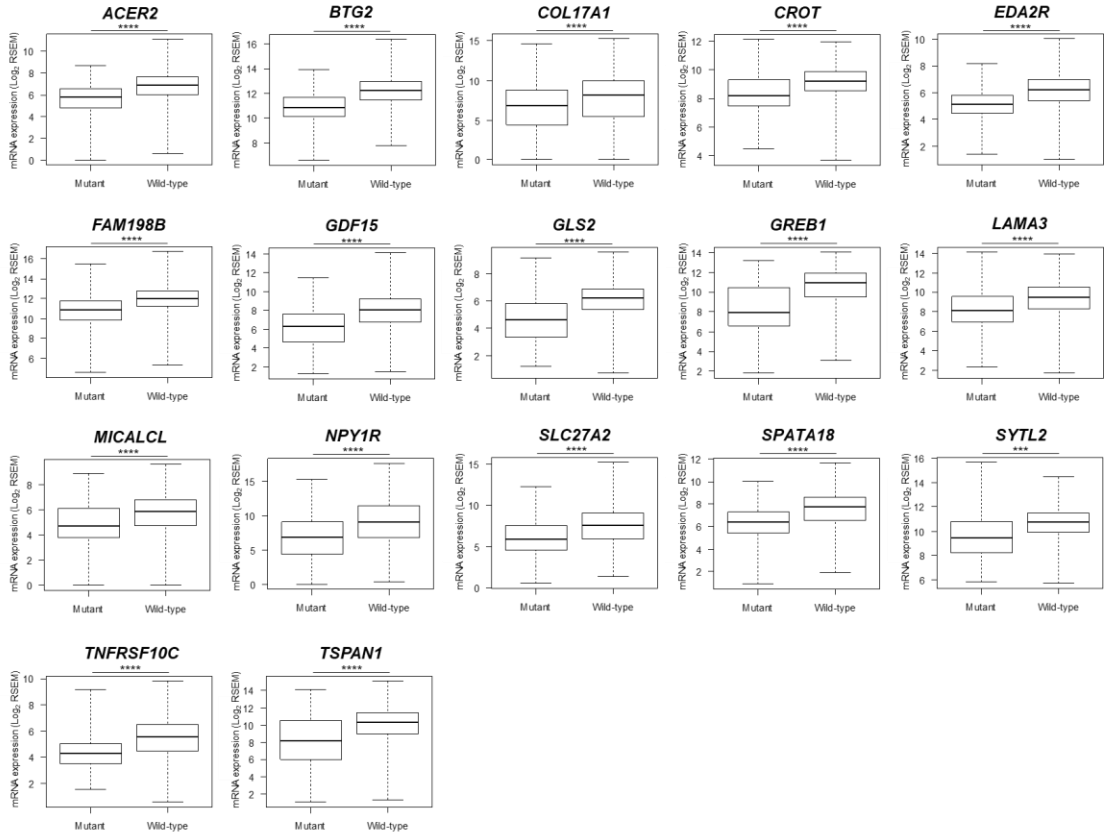
<sup>3</sup> Fold-change, the increased fold expression in wild-type compared to mutant p53 tumors.

\* Reported p53 downstream target gene and the reported reference study.



**Figure 2.2 MCF10A microarray results of the 17 p53 target gene candidates**

Normalized expression of 17 genes' criteria-satisfied probes obtained from the cDNA microarray of MCF10A cells bearing *p53* wild-type (*p53*<sup>+/+</sup>) or knockout (*p53*<sup>-/-</sup>), with or without ADR; times (hours) indicate the period after ADR treatment. Two-tailed Student's *t*-test; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001.

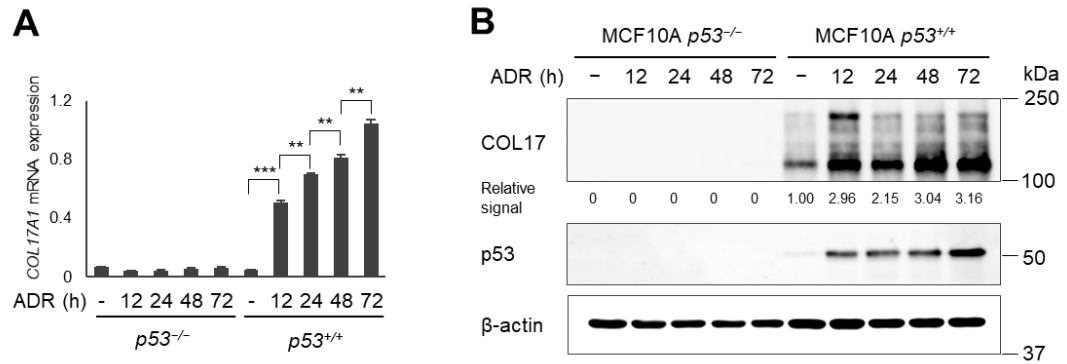


**Figure 2.3 TCGA results of the 17 p53 target gene candidates**

Differential mRNA expression of 17 genes in *p53* wild-type and mutant tumors. Number of tissues; 298 *p53* mutant tumors, and 795 *p53* wild-type tumors. Two-tailed Student's *t*-test; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.00024 (0.05/209).

## **2. Correlation of COL17A1 and p53 status**

We verified the induction of mRNA and protein expression of COL17A1 in MCF10A *p53<sup>+/+</sup>* cells at several points of time after ADR treatment (Figure 2.4). The p53 levels were gradually induced in accordance with time after ADR treatment (Figure 2.4 B). Similarly, the *COL17A1* levels increased with time from 12 to 72 hours after ADR treatment in a p53-dependent manner (Figure 2.4 A). The COL17 protein was constitutively cut and excreted as the ectodomain which can be detected in culture media as described in the Introduction. The COL17 protein level was not 100% correlated with the upregulated RNA level (Figure 2.4 B) because the whole cell lysates did not contain the excreted ectodomain; however, the detected COL17 levels increased with time from 24 to 72 hours. These results supported the screening results that COL17A1 expression was regulated by p53.



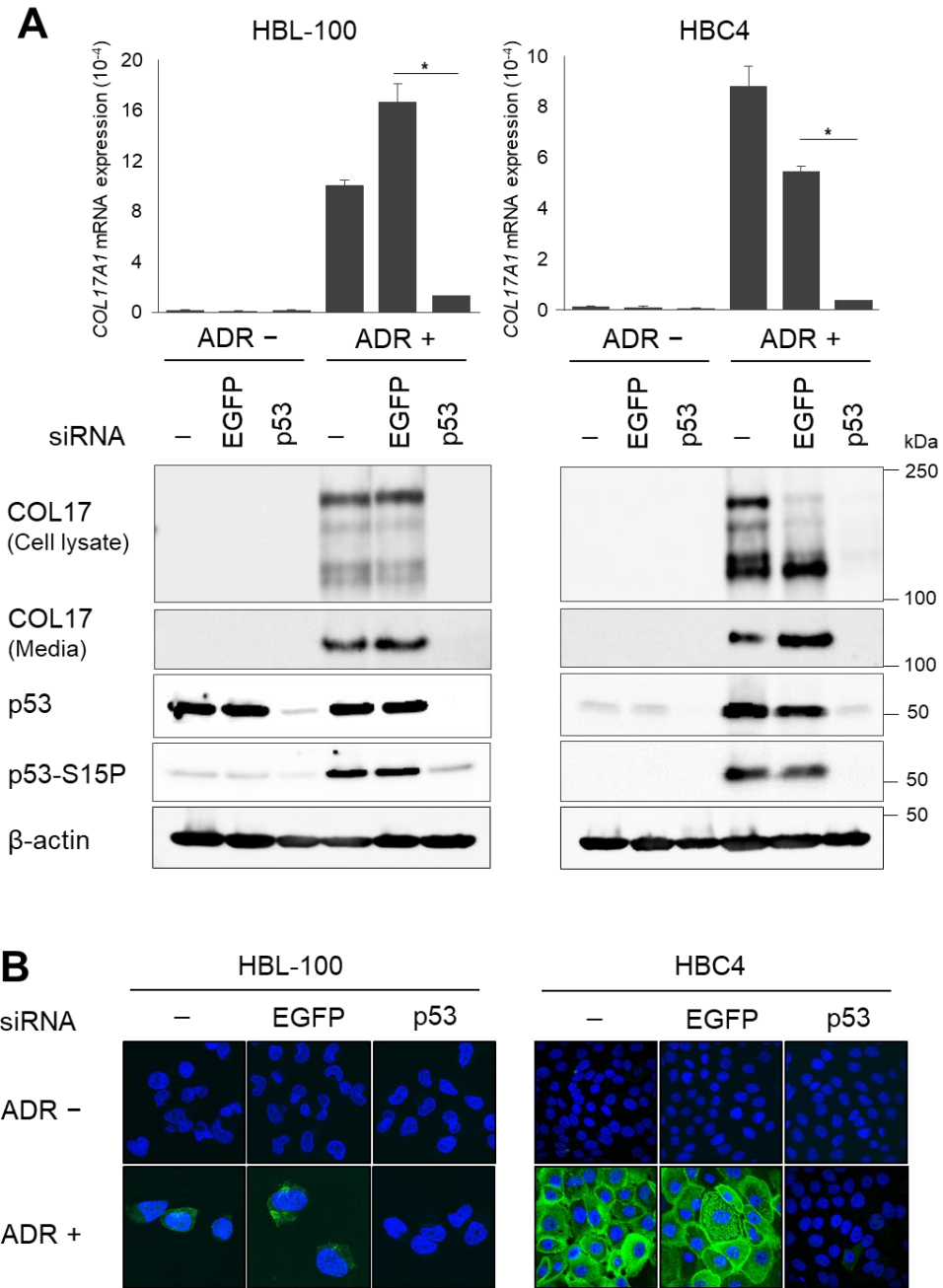
**Figure 2.4 Correlation of COL17A1 and *p53* status**

**A.** The MCF10A cells were treated with 0.5  $\mu\text{g/ml}$  of ADR for 2 hours. Total RNA was isolated at 12, 24, 48 and 72 hours after ADR treatment. The qPCR result shows the relative mRNA expression of *COL17A1* normalized by *B2M* in MCF10A cells bearing *p53* wild-type (*p53<sup>+/+</sup>*) or knockout (*p53<sup>-/-</sup>*), with or without ADR; times (hours) indicate period after ADR treatment.

**B.** Western blot analyses of COL17 and p53 in the whole cell lysates of MCF10A cells treated as described in Figure A.  $\beta$ -actin is shown as a loading control. The relative signal of western blot represents protein level of combined 180-kDa and 120-kDa COL17 quantified by ImageJ and normalized by  $\beta$ -actin. Two-tailed Student's *t*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

### 3. The p53-dependent induction of COL17A1 in human cells and mouse tissues

We further investigated the relationship between COL17A1 and p53 using human breast cancer cell lines with wild-type *p53*. ADR treatment in HBC4 cells induced an accumulation of p53 protein (Figure 2.5 A). Although p53 was not increased in HBL-100 cells (Figure 2.5 A), most likely due to a high basal expression of p53 as previously described<sup>39</sup>, phosphorylation of p53 at Ser15, a marker of p53 activation<sup>40</sup>, was clearly induced by ADR (Figure 2.5 A). Next, we treated these cells with siRNA against p53 (sip53). Consistent with p53 expression, COL17A1 mRNA and protein expression levels were significantly induced by ADR, and this induction was diminished by sip53 in both cell types (Figure 2.5 A and B). We detected both the 180-kDa full-length form of COL17 and its 120-kDa extracellular C-terminal domain (ectodomain) in the cell lysates, whereas only the 120-kDa form was detected in the media (Figure 2.5 A). The ectodomain was generated by constitutive shedding of the full-length form and was released from the cell surface into media<sup>41</sup>. The notably decreased 180-kDa form, with a concomitant increased 120-kD ectodomain, was observed in HBC4 cells treated with ADR and siEGFP (Figure 2.5 A), as observed in keratinocytes<sup>42</sup>. This increased ectodomain shedding can be activated by proinflammatory cytokines<sup>43</sup> which have been reported to be non-targeted induced in some siRNA-treated mammalian systems<sup>44</sup>.



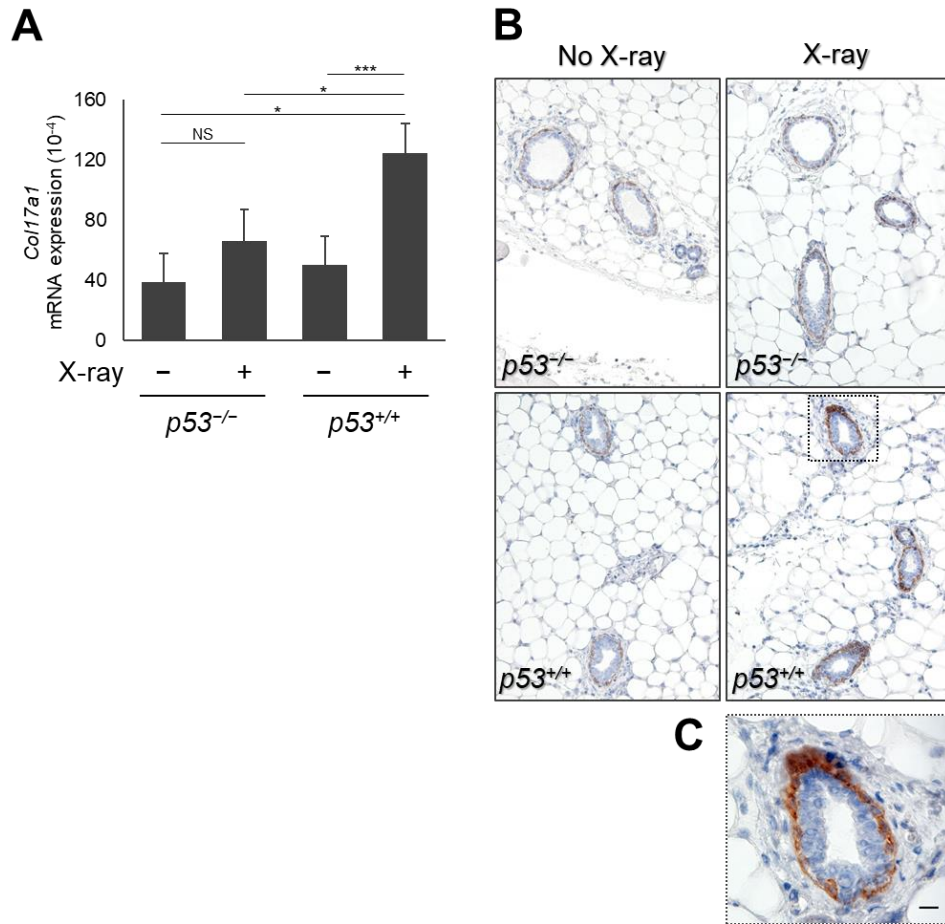
**Figure 2.5 The p53-dependent induction of COL17A1 in human cells**

**A.** The qPCR and western blot results of COL17A1 in *p53* wild-type breast cell lines, HBC4 and HBL-100, transfected with siRNA of either EGFP or p53, with or without doxorubicin (ADR) treatment. The cells were treated with ADR for 2 hours. ADR concentrations were 1  $\mu\text{g}/\text{mL}$  for HBL-100 and 2  $\mu\text{g}/\text{mL}$  for HBC4. Total RNA and protein were isolated at 48 hours after ADR treatment. siRNA of EGFP was used as a control. Relative mRNA expression of *COL17A1* was normalized to *ACTB*;  $n = 3$ . Immunoblotting results of cell lysates with

antibodies against COL17 (anti-Collagen XVII), p53, and p53-S15P (p53 phosphorylated at serine 15). Cell media were blotted with anti-Collagen XVII.  $\beta$ -actin is shown as a loading control. **B.** Immunocytochemistry of HBL-100 (40 $\times$  magnification) and HBC4 (20 $\times$  magnification) cells treated as described in Figure A. The cells were stained with DAPI (blue) and anti-Collagen XVII (green). Two-tailed Student's *t*-test; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001.

In the mouse experiment, 10-week-old *p53*<sup>+/+</sup> and *p53*<sup>-/-</sup> mice were exposed to 10 Gy X-ray irradiation. Twenty-four hours after irradiation, mammary tissues were collected for qPCR analysis and immunohistochemistry. The COL17A1 mRNA and protein levels were significantly induced by X-ray in *p53*<sup>+/+</sup> but not *p53*<sup>-/-</sup> mice (Figure 2.6 A and B). In the mouse mammary gland, elevated COL17 protein levels were detected in myoepithelial cells surrounding secretory luminal cells (Figure 2.6 C) regarding its role as a cell-matrix adhesion molecule. These data demonstrated that COL17A1 expression depends on p53 and is induced by ADR in human breast cell lines and by X-ray in mouse mammary tissues.



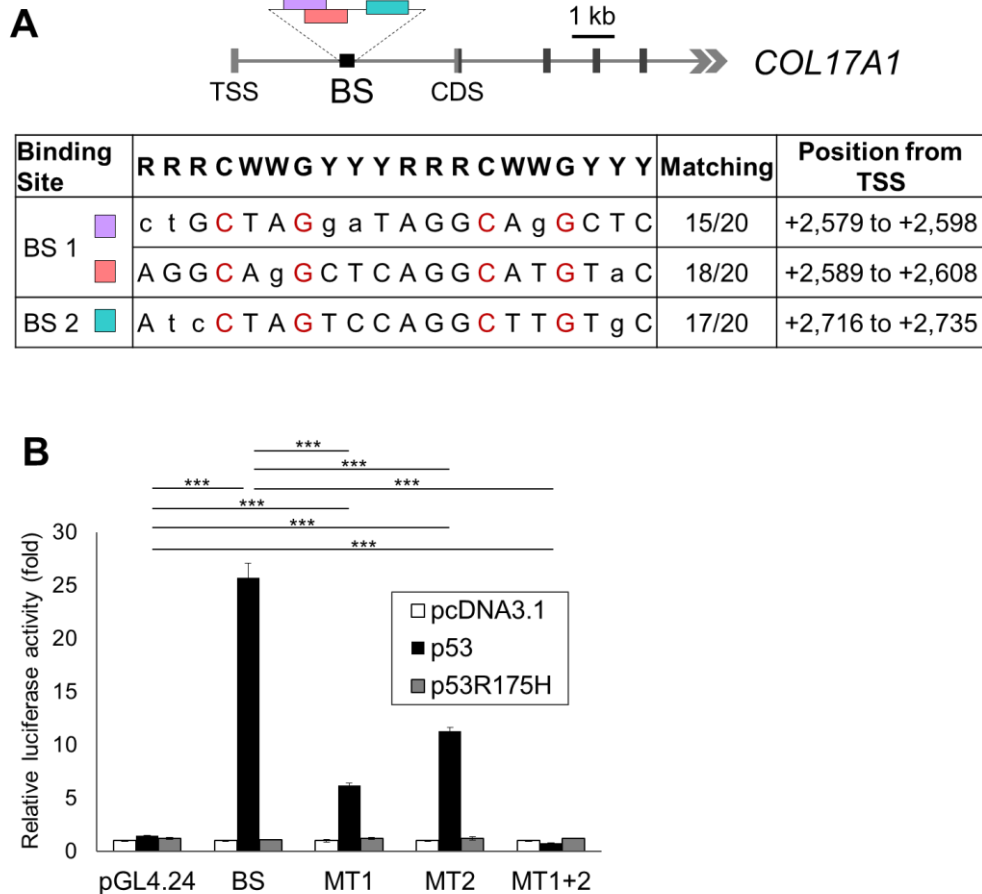


**Figure 2.6 The p53-dependent induction of COL17A1 in mouse tissues**

A. Relative mRNA expression of *Col17a1* normalized to *Gapdh* in mouse mammary tissues of *p53* knockout (*p53*<sup>-/-</sup>) or *p53* wild-type (*p53*<sup>+/+</sup>) mice; with or without X-ray irradiation;  $n = 3$ . The mice were sacrificed at 24 hours after X-ray. D. Immunohistochemistry results of mammary tissues staining with anti-Collagen XVII;  $n = 3$  mice per group. Images were obtained at 10 $\times$  magnification. B. The inset of Figure C, *p53*<sup>+/+</sup> mice with X-ray at 40 $\times$  magnification. Scale bar, 20  $\mu$ m. Two-tailed Student's *t*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; NS,  $P \geq 0.05$ .

#### **4. *COL17A1* is a direct target of p53**

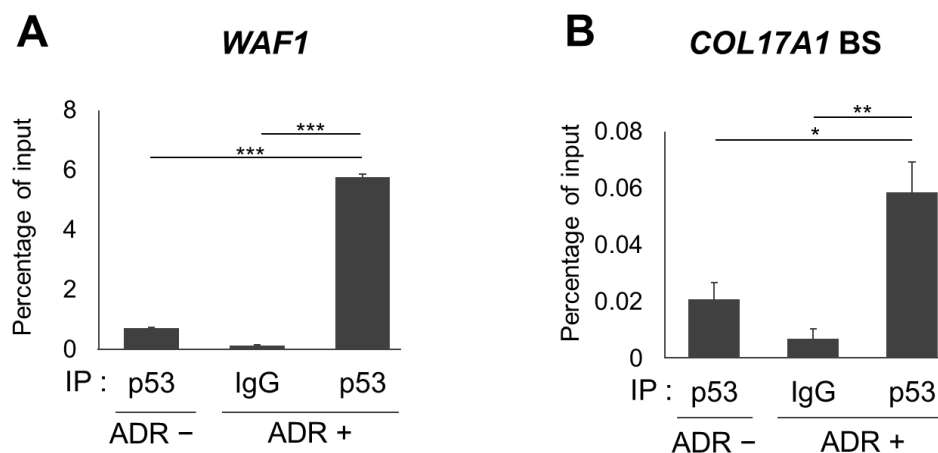
To examine whether *COL17A1* is a direct target of p53, we surveyed the genomic sequence of the *COL17A1* gene and found two putative p53 binding sites, BS1 and BS2, within the first intron (Figure 2.7 A). The DNA fragment was amplified using primers covering both BS1 and BS2 regions and then cloned into pGL4.24 vectors (BS) to evaluate the p53-dependent transcriptional activity using a reporter assay. The luciferase activity of the BS construct was markedly enhanced by co-transfection with wild-type p53 but not with mutant p53, an Arg-to-His substitution at p53 amino acid 175 (p53R175H) (Figure 2.7 B). Next, we introduced mutations at each binding site, resulting in mutations of BS1 (MT1) and BS2 (MT2) (Figure 2.7 A). MT1 and MT2 significantly reduced the luciferase activity compared to that of BS, although a high signal was still detected due to the existence of another non-mutated BS. The luciferase activity of MT2 when co-transfected with p53 was superior to that of MT1 (Figure 2.7 B), suggesting that p53 binds strongly to BS1, most likely because BS1 contains long binding sequences (30 bases) and matched highly to the consensus p53 binding sequence (Figure 2.7 A). Moreover, after two binding sites had been mutated (MT1+2), the luciferase activity was entirely drop to the control level (Figure 2.7 B), suggesting the complete deletion of p53 binding.



**Figure 2.7 Identification of the p53 binding sequences by Reporter assay**

**A.** Genomic structure of the human *COL17A1* gene indicates potential p53 binding sites (color boxes) in the BS construct (black box). Light gray box, untranslated region; dark gray box, translated region; TSS, transcription start site; CDS, coding sequence. The table shows the BS sequences aligned with the consensus p53 binding sequence (20 bases): R, purine; W, A or T; Y, pyrimidine. Identical nucleotides to the consensus are written in capital letters with the number of matching (n/20). The red-labeled nucleotides were mutated to thymine to examine the specificity of BS after mutagenesis. BS1 was mutated to MT1; BS2 was mutated to MT2; MT1+2 is double mutation of BS1 and BS2. **B.** Luciferase assay of the BS construct, the mutated BS construct (MT1, MT2, MT1+2), and the empty vector (pGL4.24) using H1299 cells co-transfected with control vector (pcDNA3.1), wild-type p53 (p53), or mutant p53 (p53R175H). Luciferase activity is indicated relative to the activity of the control vector (pcDNA3.1);  $n = 3$ . Two-tailed Student's  $t$ -test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

To further confirm whether p53 can bind to this DNA segment, we performed a chromatin immunoprecipitation (ChIP) assay using HBC4 cells with or without ADR treatment. After precipitation with an anti-p53 antibody, a DNA fragment containing p53 binding sites was quantified by qPCR. The p53 binding site in *WAF1*, a p53 target gene, was examined as a positive control region (Figure 2.8 A). The qPCR result using *COL17A1* BS primers indicated that the endogenous wild-type p53 significantly binds stronger to *COL17A1* in ADR-treated cells compared to non-treated cells (Figure 2.8 B). Taken together, our results suggest that p53 directly regulates *COL17A1* through p53-binding sites located in the first intron.



**Figure 2.8 Identification of the p53 binding sequences by ChIP assay**

ChIP assay of non-treated or ADR-treated HBC4 cells of which DNA–protein complexes were then immunoprecipitated with the indicated antibodies followed by qPCR. Anti-mouse IgG was used as a negative control. The graphs show qPCR results indicating the amount of genomic fragments containing p53-binding sequence in *WAF1* (Figure A), or *COL17A1* using BS primers (Figure B). Two-tailed Student’s *t*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

## Discussion

This chapter presents a p53-target screening based on human breast cell lines and breast cancer tissues. Ten genes including *COL17A1*, *CROT*, *FAM198B*, *GREB1*, *LAMA3*, *MICALCL*, *NPY1R*, *SLC27A2*, *SYTL2*, and *TSPAN1* were identified as p53 target gene candidates from this screening. GO analysis identified two enriched biological processes as significantly up-regulated by wild-type p53, epidermis development and ectoderm development, using a similar set of genes, i.e., *COL17A1*, *LAMA3*, and *EDA2R*. *LAMA3* encodes Laminin-5, an extracellular matrix protein, which binds to COL17 and functions in cell-matrix adhesion<sup>45</sup>. An alteration in the expression patterns of cell adhesion molecules, including Laminin-5, by knockdown and mutations of p53 in MCF10A cells has been previously described, but COL17 has not been examined<sup>36</sup>. Here, we provide evidence that the mRNA and protein expressions of COL17A1 are regulated by p53. The COL17 levels elevated after the induction of endogenous wild-type p53 following ADR-treatment in human breast cells or X-ray in mouse mammary glands. The results from a reporter assay and a ChIP assay indicate binding sites of p53 in intron 1, indicating *COL17A1* as a direct p53 target. In response to X-ray exposure, COL17 protein is expressed in the myoepithelial cells (basal cell) of the mammary gland, regarding its function in facilitating cell adhesion to the underlying basement membrane<sup>23,46</sup>. Consistent with our result, gene expression profiling demonstrated that *COL17A1* expression is high in human breast myoepithelial cells but low in luminal epithelial cells<sup>47</sup>.

## Chapter II

### The role of COL17 in breast carcinogenesis

#### Summary

The functional analyses of COL17 related to breast carcinogenesis were investigated. Due to its originally low expression in all breast cancer cell lines, *COL17A1* was over-expressed in four cells and the colony formation assays were performed to examine the relation to cell growth. On the contrary, an endogenous *COL17A1* was knockdown and the ADR-induced growth suppression was observed using cell proliferation analysis.

Regarding the function of COL17 as a cell-matrix adhesion molecule<sup>18</sup>, we examined its role in tumor metastasis using a highly invasive MDA-MB-231 cell line expressing mutant p53. We generated MDA-MB-231 cell stably expressing COL17 in the presence of doxycycline (Dox) and control cells using the tetracycline-regulated lentiviral expression system. These stable cell lines were used for subsequent experiments to investigate cell growth, cell migration, and cell invasion *in vitro*. The effect of the secreted form of COL17 on the adjacent cells' invasion was also investigated imitating the natural cell-matrix environment surrounding the mammary glands.

The epithelial to mesenchymal transition (EMT) plays an important role in invasion and subsequent metastasis of breast carcinoma<sup>48</sup>. Thus, we investigated whether COL17 regulates the metastasis through this mechanism. The protein level of an epithelial marker, E-cadherin, and a mesenchymal marker, Vimentin were examined in the MDA-MB-231 and MCF7 stable cells over-expressing COL17, and in HBC4 cells with *COL17A1* knockdown. Finally, we analyzed *COL17A1* expression in multiple types of cancer. The relationship between *COL17A1* expression and tumor progression was investigated in subgroup breast cancer patients. The mechanisms that control *COL17A1* expression were revealed.

## Material and methods

### 1. Colony formation assay

The entire coding sequence of *COL17A1* was amplified and cloned into the pCAGGSnHC-MCS vector. The plasmid sequence was confirmed by DNA sequence analysis. Primers for amplifying the *COL17A1* coding sequence (CDS) are shown in Appendix Table A1. Four breast cancer cells including HBC4, SK-BR3, MCF7, and T47D, were seeded on 6-well flat-bottomed microplates ( $n = 3$ ). The seeding cell number was  $1 \times 10^4$  for HBC4, and  $1 \times 10^5$  for other cells. At 24 hours after seeding, cells were transfected with pCAGGSnHC-MCS empty vector (mock) or pCAGGSnHC/*COL17A1* using FuGENE6 (Promega, Madison, WI, USA). The cells were cultured with Geneticin (0.2, 0.8, 0.5, and 0.6 mg/ml for HBC4, SKBR3, MCF7, and T47D cells, respectively) to select the positive transfected clones. After 2 to 3 weeks of drug selection, formed colonies were washed by PBS, fixed by methanol, and stained with 0.1% crystal violet for 24 h. After being washed by water and dried up, colonies were counted using ImageJ.

### 2. Stable cell line generation

The full-length *COL17A1* genomic DNA fragment containing Kozak sequence was cloned into the entry vector (pENTR<sup>TM</sup>3C). The primer sequences are shown in Appendix Table A1. The *COL17A1* fragment was integrated into the destination vector (pLenti6.3/TO/V5-DEST) using Gateway LR Clonase II enzyme mix and then transformed into Stbl3 cells to generate the expression clone. Human embryonic kidney cells transformed with the SV40 large T antigen (293FT) were provided in the ViraPower<sup>TM</sup> Lentiviral Expression System (Thermo Fisher Scientific). Lentiviruses containing *COL17A1* expression construct, empty vector, or *TetR* vector (pLenti3.3/TR) were produced in 293FT cells using ViraPower<sup>TM</sup> Packaging Mix according to the manufacturer's protocol. The virus-containing media were harvested and used to determine the viral titer by qPCR. The *TetR* construct was co-transduced at 10 MOI into MDA-MB-231 and MCF7 cells with either the *COL17A1* construct for COL17A1-expressing

cells or empty vector for mock cells. To select the transduced cells with viral vector containing antibiotic-resistant genes (Blasticidin in destination vector and Geneticin<sup>®</sup> in *TetR* vector), cells were continuously cultured in media containing Blasticidin (6 ug/mL for MCF7, and 10 ug/mL for MDA-MB-231) and Geneticin<sup>®</sup> (0.8 mg/mL for MCF7, and 1 mg/mL for MDA-MB-231) for 3 weeks.

### **3. Cell proliferation analysis (ATP assay)**

HBC4 cells were plated and transfected with siRNA. At 24 h after transfection, cells were treated with 2 mg/mL of ADR for 2 hours. At 48 hours after ADR treatments, cells were subjected to ATP assay. The MDA-MB-231 stable cells were treated with 1 µg/mL doxycycline (Dox) for 48 hours before cell plating. The cells ( $1.2 \times 10^6$ ) were seeded onto a 24-well plate. The cells were subjected to ATP assay at 24, 48, 72, and 96 hours after plating ( $n = 3$ ). ATP measurement assay was performed using Cell Titer-Glo Luminescent Viability Assay (Promega) according to the manufacturer's protocol. The fluorescence of the solution was measured by an ARVO X3 plate reader (Perkin Elmer, USA) according to the manufacturer's protocol.

### **4. Scratch assay**

The MDA-MB-231 stable cells were treated with 1 µg/mL Dox for 48 hours before cell plating. The cells ( $1.2 \times 10^5$ ) were seeded onto the COL1-coated 24-well plates using 1% FBS-containing culture media with or without Dox. Twenty-four hours after plating, the confluent cells were scraped using a CELL Scratcher (Iwaki, Japan), washed and replaced with media. The culture plates were marked as reference points close to the scratch. Images of the scratched area at the reference points were recorded immediately after the scratch and 24 hours later using a phase-contrast microscope with 10× magnification. The distances of the scratched area were determined and measured using ImageJ software<sup>49</sup>. The average migration distance of each well was calculated and subjected to a box plot analysis using R programming.



## 5. Invasion assay

The MDA-MB-231 stable cells were treated with 1 µg/mL Dox for 48 hours before cell plating. The cells ( $7.5 \times 10^4$ ) were seeded into the 24-well BioCoat Matrigel invasion chambers (Corning) or control inserts according to the manufacturer's protocol. The upper chamber filled with serum-free DMEM was placed in a well containing 10% FBS DMEM with the same condition of Dox, i.e., presence or absence. For the cell assay, stable cell lines were used with fresh media in both the upper chamber and the lower well. For the media assay, parental MDA-MB-231 cells were seeded into the chamber using conditioned media harvested from the stable cell culture dishes; and the chambers were placed in fresh media. After 24 hours of incubation, the cells invading to the bottom surface of the chamber membrane were fixed with 4% paraformaldehyde for 30 minutes and stained with 0.1% crystal violet for 2 h. The invading cells were counted using CELLCOUNTER software<sup>50</sup> and calculated as the percent invasion through the Matrigel relative to the control. The invasion index is displayed as the ratio of % invasion of Dox+ over those of Dox- condition.

## 6. TCGA and METABRIC analysis

The mRNA expression of *COL17A1*, allelic copy number, mutation data, methylation data, clinical data and the type of tissues of TCGA and METABRIC<sup>51</sup> project were obtained from cBioPortal<sup>31,32</sup>. In multi-cancer analysis, the ratio (fold-change) of *COL17A1* expression in normal over tumor tissue from each patient was calculated, and box plot analysis was performed using log scale of fold-change. The differential expression of *COL17A1* was analyzed in breast tissues categorized by tissue type (as normal, tumor, or metastasis), and by allelic copy number of *COL17A1* (as deep deletion, shallow deletion, diploid, gain, or amplification) using box plot analysis. The expression and promotor methylation of *COL17A1* were subgroup analyzed in different cancer subtypes and in patients with metastasized tissues. A survival analysis was performed using the log-rank test stratified by the expression level of *COL17A1* in tumors (above or below the median expression level). Cox's proportional hazards

model was used to adjust for the following variables: *p53* status (wild-type or mutant), patient menopausal state (pre- or post-menopause), and *COL17A1* expression level (above or below the median expression). Box plot and survival analyses were performed using the EZR software program<sup>52</sup>.

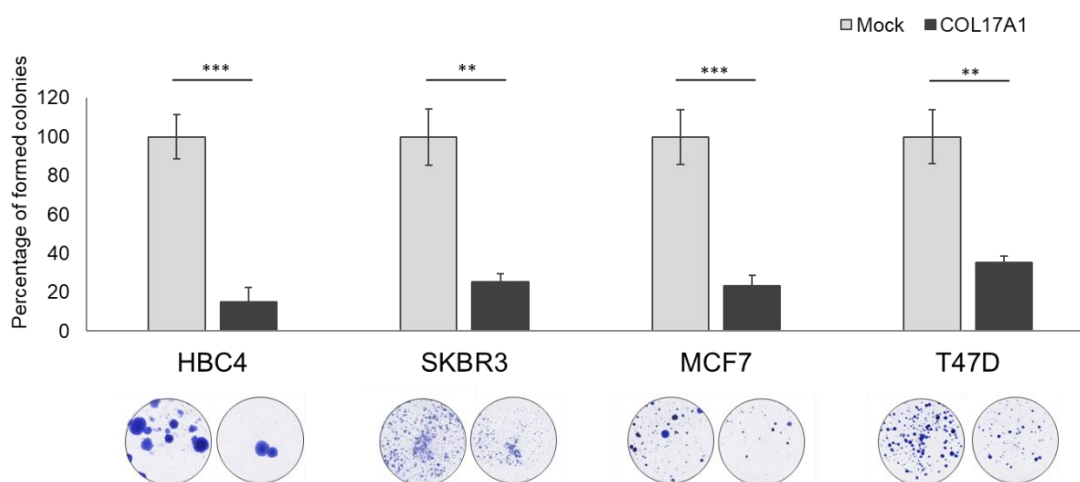
## **7. DNA demethylation**

Four breast cancer cell lines with *p53* mutant, including BT20, BT-549, HCC1395, and MDA-MB-435s, were seeded in 6-well plate, and placed at 37°C incubator overnight. The next day, the cell culture media was replaced with fresh media containing 5, 10 and 15 µM 5-aza-2'-deoxycytidine (DAC), and plates were wrapped in aluminum foil to avoid light exposure. The media were changed every 12 hours for 2 days. After 2 days, the media were replaced with fresh media without DAC. The cells were cultured for an additional 48 hours, and then extracted for total RNA. Complementary DNA were synthesized and subjected to qPCR with *COL17A1* primers and *B2M* primers (see the primer sequences in Appendix Table A1).

## Results

### 1. The role of *COL17A1* in cancer cell growth

The colony formation assays were performed in four breast cancer cells. The cells over-expressed *COL17A1* remarkably exhibited the decreased number of colonies formed to 15-35% compared with control cells (100%) (Figure 3.1).

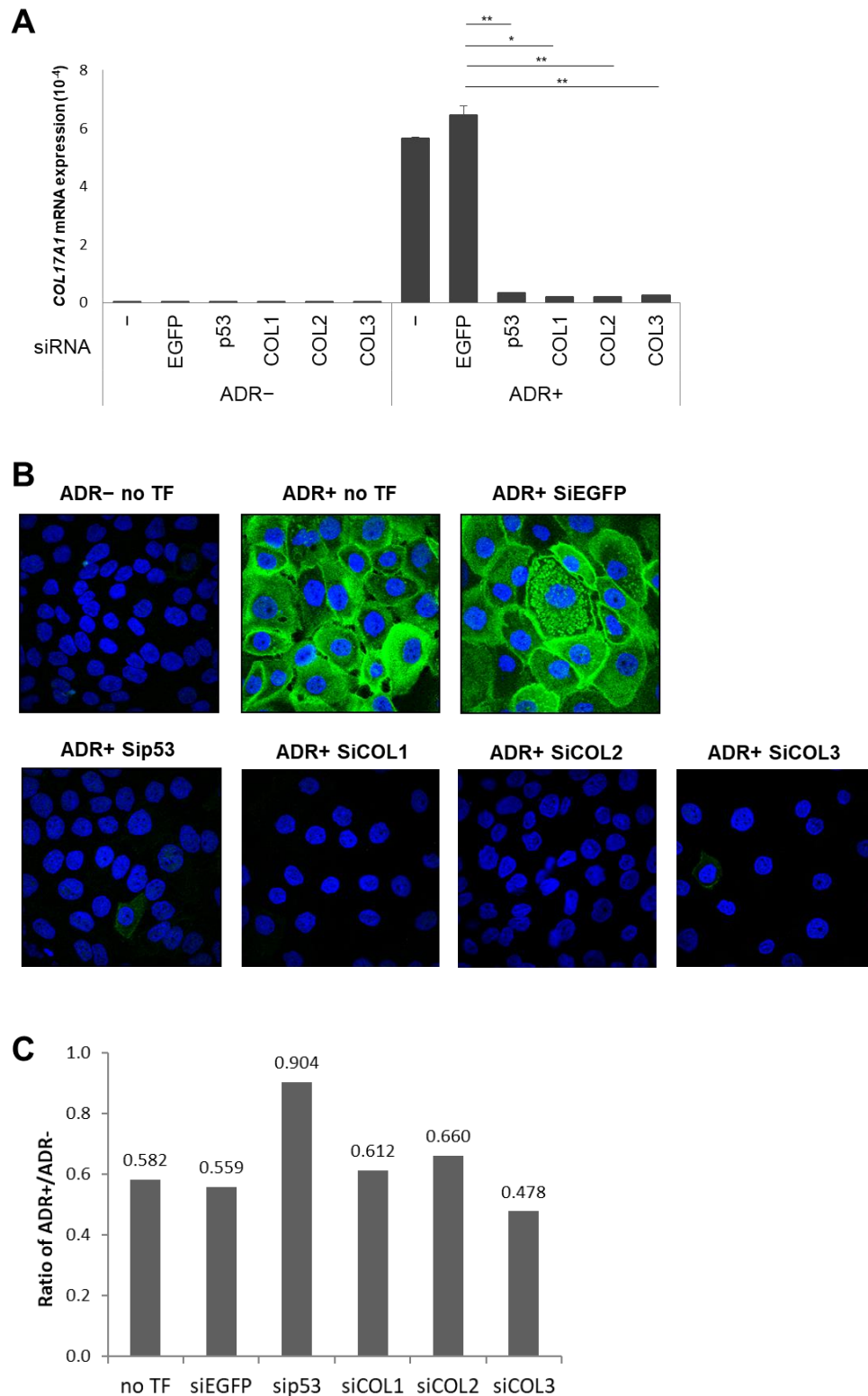


**Figure 3.1** The *COL17A1* over-expression in cancer cell growth

Colony formation assays were performed using four breast cancer cells with wild-type *p53* transfected with plasmid expressing *COL17A1* or mock. Cells were cultured in the presence of Geneticin (0.2, 0.8, 0.5, and 0.6 mg/ml for HBC4, SKBR3, MCF7, and T47D cells, respectively) for at least 2 weeks for selecting transformed colonies. The number of colonies was quantified using the Image J software. Columns, means; error bars, S.D.;  $n = 3$ . Two-tailed Student's *t*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

Next, we generated three siRNAs (siCOL1, siCOL2, and siCOL3) designed to simultaneously knockdown the *COL17A1* (see the siRNA oligonucleotides in Appendix) and confirmed the mRNA suppression in HBC4 cells with or without ADR treatments (Figure 3.2 A). The endogenous *COL17A1* expression is excessively low in non-ADR-treated HBC4, thus its expression level between non-transfected cell and *COL17A1*-knockdown cell cannot be differentiated. Though, the siRNA against *COL17A1* significantly inhibited the induction of

mRNA and protein level of *COL17A1* following ADR treatment (Figure 3.2 A and B). Then we analyzed the cell proliferation of siCOL17A1-treated HBC4 cells in the presence of ADR. The *COL17A1* knockdown did not inhibit the ADR-induced growth suppression as found in cells treated with sip53 (Figure 3.2 C).

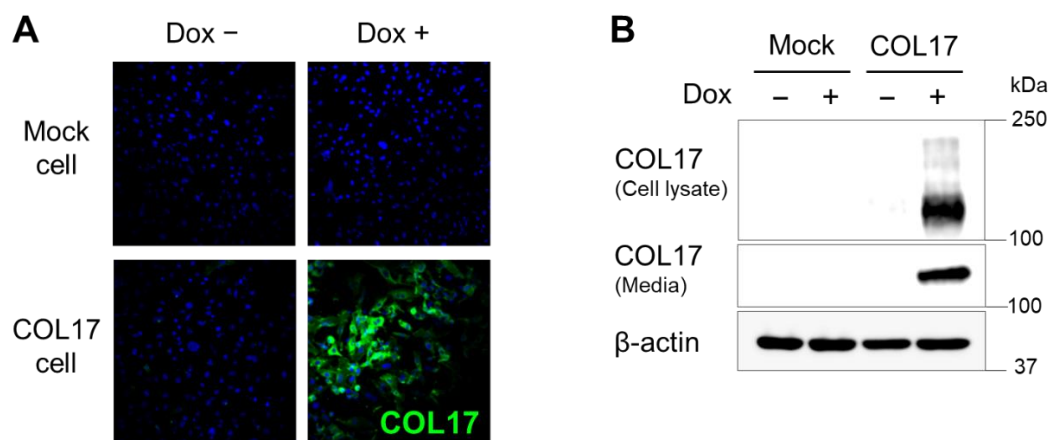


**Figure 3.2 The *COL17A1* knockdown in cancer cell growth**

HBC4 cells were plated and transfected with siRNA of either EGFP, p53, or COL17A1 (three different oligonucleotides: COL1, COL2, and COL3) with or without ADR treatment; No TF, no transfection with any siRNA. At 24 hours after plating, the cells were treated with 2 µg/mL of ADR for 2 hours. Total RNA and protein were isolated at 48 hours after ADR treatment. siRNA of EGFP was used as a control. **A.** Relative mRNA expression of *COL17A1* normalized to *ACTB*;  $n = 3$ . **B.** Immunocytochemistry of HBC4 (20× magnification) cells stained with DAPI (blue) and anti-Collagen XVII (green). **C.** Cell proliferation analysis of HBC4 cells at 48 hours after ADR treatment. Relative cell viability was calculated by dividing the average fluorescence signal of ADR-treated cells by that of untreated cells ( $n = 3$ ). Two-tailed Student's *t*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

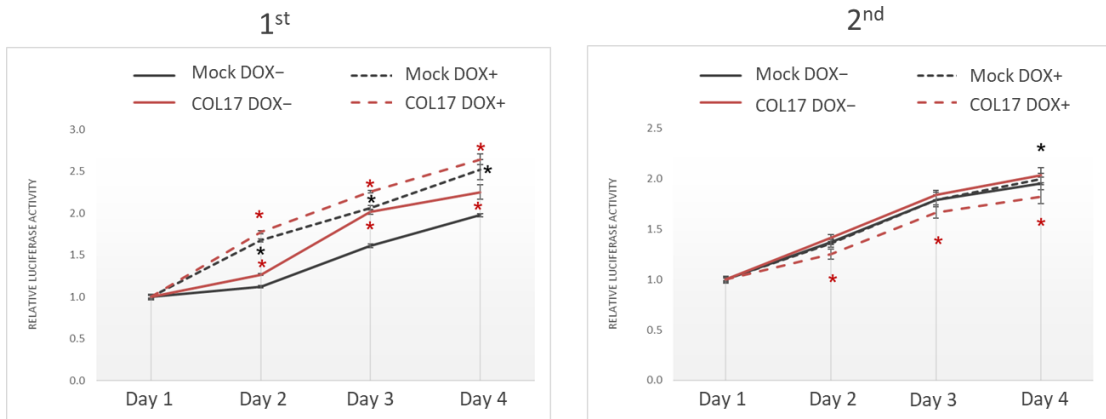
Owing to the inconsistent results from the colony formation assay and the *COL17A1* knockdown by siRNA, the relation of COL17 and cell proliferation was further examined using stable cell lines. Firstly, the MDA-MB-231 cell lines stably expressing COL17 and mock cells were generated. In the absence of Dox, the Tet repressor suppresses COL17 expression (Figure 3.3 A and B, Dox<sup>-</sup>). By contrast, after the addition of Dox into cell culture media for 48 hours, the 180-kDa form of COL17 is expressed, with the shed 120-kDa ectodomain form that is predominantly detected in the culture media (Figure 3.3 A and B, Dox<sup>+</sup>).

Cell proliferation of the stable cell lines was examined using ATP assay at 1, 2, 3, and 4 days after cell seeding (Figure 3.4). The results of two independent experiments were inconsistent. The cell proliferation of COL17-expressing cells was highest at first, but lowest at the repeat, comparing to non-expressing cells. Taken together, the effect of COL17 on cell growth is still controversial.



**Figure 3.3 The MDA-MB-231 stable cell lines**

The MDA-MB-231 cells stably expressing COL17 in the presence of doxycycline (Dox) were generated using the tetracycline-regulated lentiviral expression system. Mock cells were used as controls. Cells were not treated (Dox<sup>-</sup>) or treated with 1  $\mu$ g/mL of doxycycline (Dox<sup>+</sup>) for 48 hours before analysis. **A.** Immunocytochemistry of the stable cells stained with anti-Collagen XVII (green) and DAPI (blue). **B.** Western blot of the stable cell lysates and media extract blotting with anti-Collagen XVII.



**Figure 3.4 ATP assay of the stable cell lines**

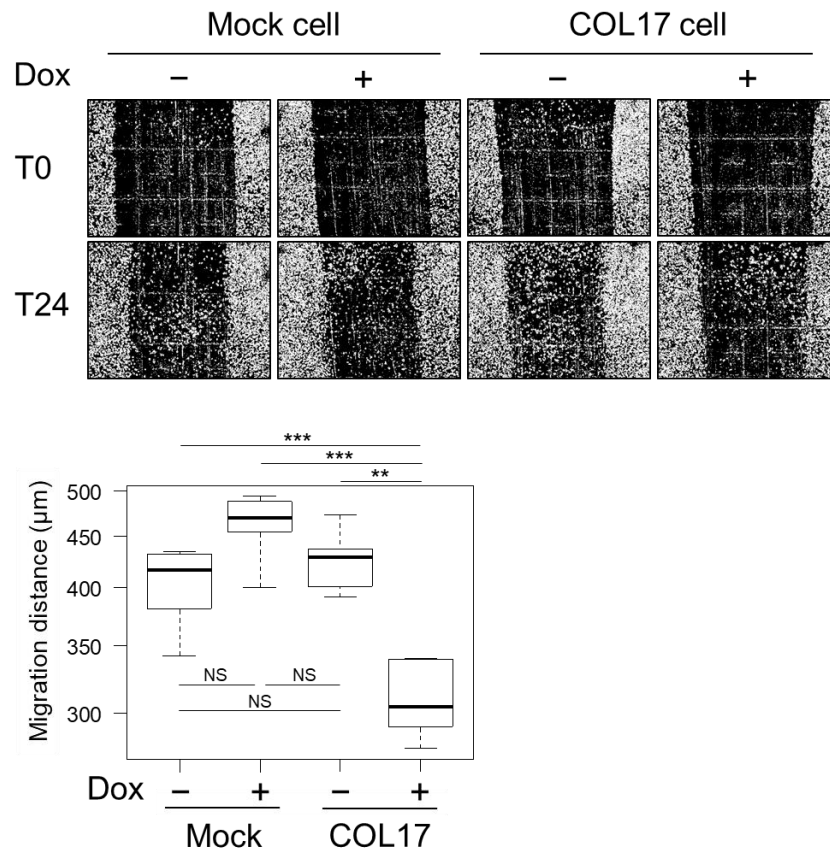
MDA-MB-231 cells stably overexpressing COL17 and mock cells were not treated (Dox-) or treated with 1  $\mu\text{g}/\text{mL}$  of doxycycline (Dox+) for 48 hours before plating. The cells were plated onto a 24-well plate and continuously cultured in media with the indicated condition (Dox- or Dox+). ATP assays were performed at 1, 2, 3, 4 days after plating ( $n = 4$ ). The experiment was performed two times separately. Two-tailed Student's  $t$ -test;  $*P < 0.05$ .

## 2. The role of COL17 in cancer cell migration and invasion

The stable cell lines were plated onto a COL1-coated container using a lower percentage of serum to minimize cell proliferation, as previously described<sup>53</sup>. Twenty-four hours after the scratch, the COL17-expressing cells exhibited significantly decreased migration compared to the non-producing cells (Figure 3.5). To strengthen this result, we performed an invasion assay using basement membrane matrix Matrigel-coated chambers to mimic the *in vivo* extracellular environment<sup>54</sup>. Cells were allowed to invade depending on their invasive potential through membrane pores to the lower chamber for 24 hours. The invading cells were fixed, stained, and quantified. The COL17-producing cell exhibited nearly 50% repressed invasive ability compared to mock cells (Figure 3.6 A), suggesting a role of COL17 in suppressing cell migration and invasion.

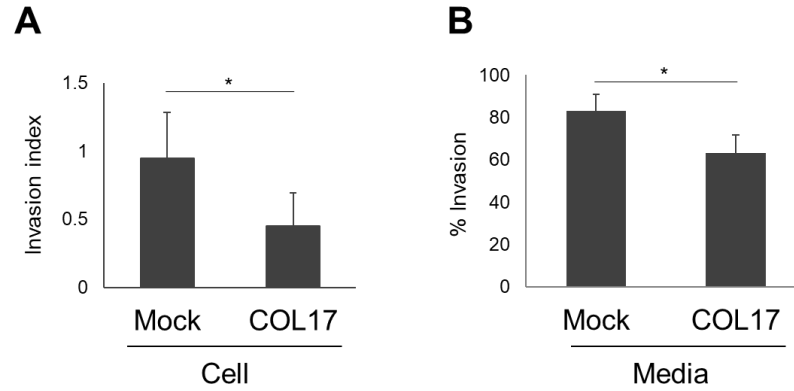
Next, we examined the effect of the secreted form of COL17 on cells imitating the natural cell-matrix environment surrounding the mammary glands. Parental MDA-MB-231 cells were plated using conditioned media harvested from the stable cell culture dishes. The cells cultured in the COL17-enriched media demonstrated a significantly lower percentage of invasion compared to the cells cultured in mock media (Figure 3.6 B), indicating that the COL17 ectodomain regulated cancer cell invasion. Our results established a role of COL17 in suppressing breast tumor migration and invasion *in vitro*.





**Figure 3.5 Migration assay (scratch assay) of the stable cell lines**

MDA-MB-231 cells stably overexpressing COL17 and mock cells were not treated (Dox-) or treated with 1 µg/mL of doxycycline (Dox+) for 48 hours before plating. Then, cells were plated onto a COL1-coated container using 1% serum-containing media. The images of scratch region were taken immediately after the scratch (T0) and 24 hours later (T24) using a phase-contrast microscope with 10× magnification;  $n = 6$  per group. Migration distances from T0 to T24 were measured at the reference points ( $n = 3$ ) using ImageJ, and the average was calculated and used for the box plot analysis. Two-tailed Student's  $t$ -test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; NS,  $P \geq 0.05$ .



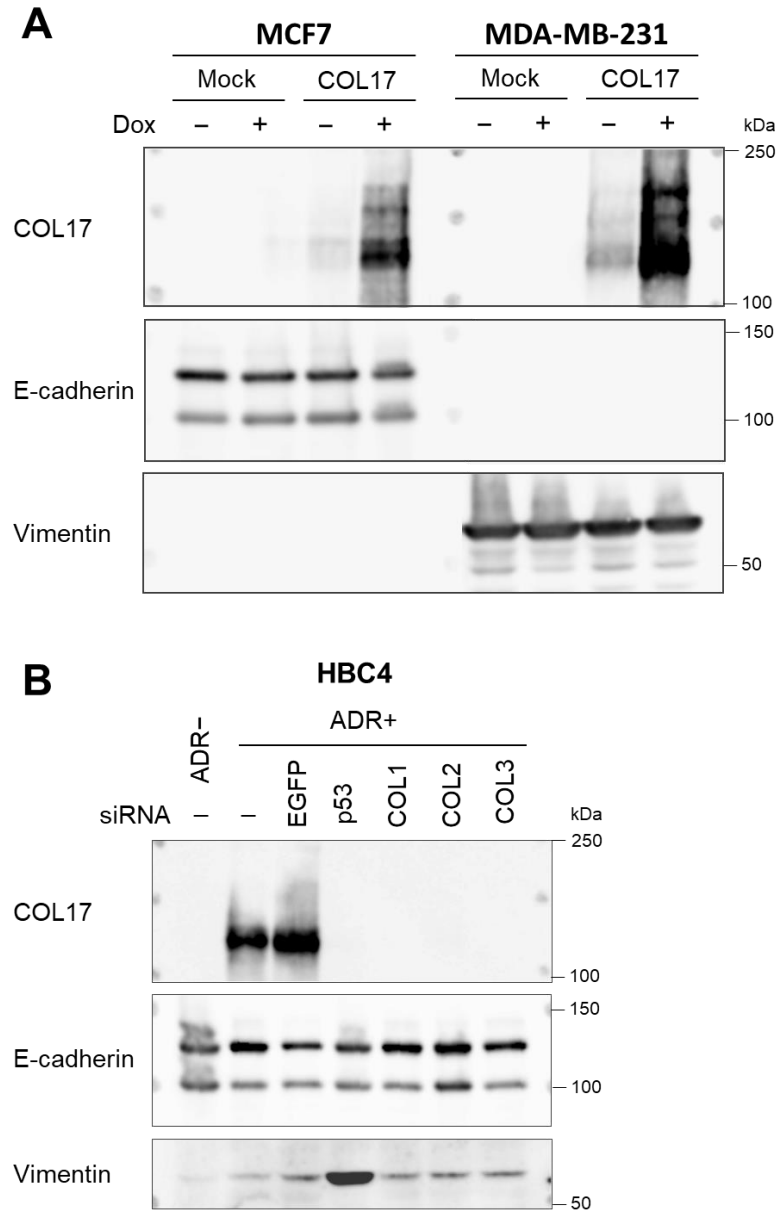
**Figure 3.6 Invasion assay of the stable cell lines**

MDA-MB-231 cells stably overexpressing COL17 and mock cells were not treated (Dox<sup>-</sup>) or treated with 1  $\mu\text{g}/\text{mL}$  of doxycycline (Dox<sup>+</sup>) for 48 hours before plating or harvesting of media.

**A.** Invasion assay of stable cells analyzed at 24 hours after plating. The number of invading cells was quantified and calculated as the % invasion through the Matrigel over control chambers. The invasion index displays a proportion of the % invasion for the Dox<sup>+</sup>/Dox<sup>-</sup> condition. **B.** Media assay using parental MDA-MB-231 cells plated with conditioned media harvested from the stable COL17 or mock cell culture dishes. The parental cells invasive ability was displayed as the % invasion through the Matrigel over the control. Two-tailed Student's *t*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

### **3. The relation of COL17 and the epithelial to mesenchymal transition (EMT)**

To investigate the relation of COL17 and EMT, EMT markers were examined in breast cancer cells with over-expression or knockdown of COL17. First, we checked EMT markers in two stable cell lines, including epithelial-like MCF7 and mesenchymal-like MDA-MB-231. The ectopic expression of COL17 did not either up-regulate the epithelial marker E-cadherin which is usually detectable in MCF7, or down-regulate the mesenchymal marker vimentin which is normally high in MDA-MB-321 (Figure 3.7 A). Next, we analyze these markers in HBC4 cells pre-treated with siRNA against *COL17A1* or *p53*. The COL17 knockdown did not affect any marker level, contrary to p53 knockdown that enhance vimentin level suggesting its role in inhibition of mesenchymal transition (Figure 3.7 B).



**Figure 3.7 The relation of COL17 and EMT**

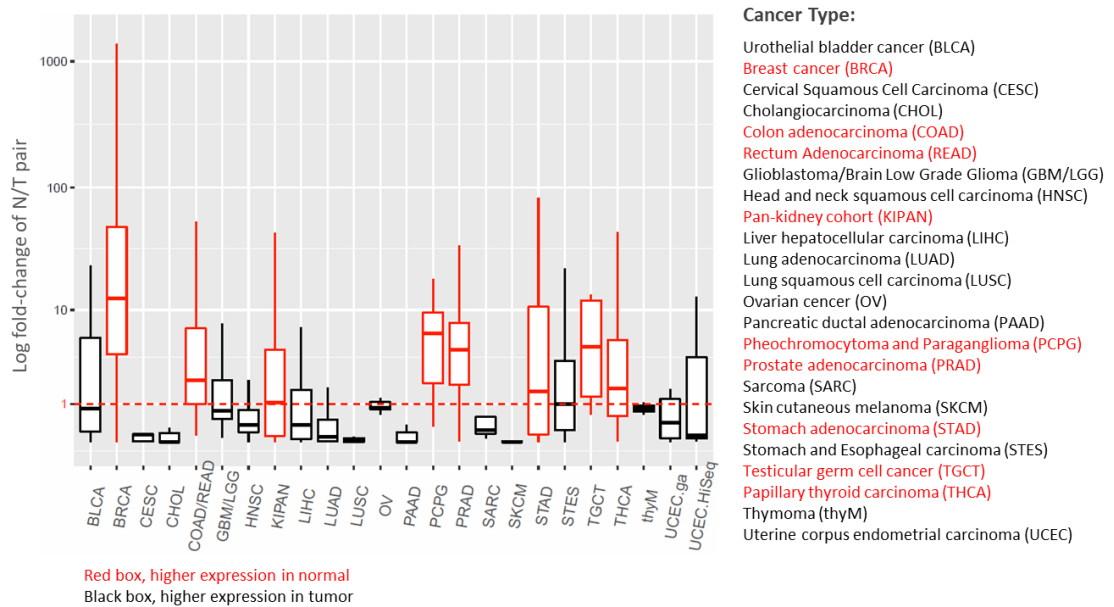
Western blot analysis using anti-Collagen XVII, E-cadherin (BD 610181), or vimentin (sc-32322) according to the manufacturers' protocols. **A.** The stable cell lines, MCF7 and MDA-MB-231, stably overexpressing COL17 and mock cells were not treated (Dox<sup>-</sup>) or treated with 1  $\mu$ g/mL of doxycycline (Dox<sup>+</sup>) for 48 hours before protein extractions. **B.** HBC4 cells were plated and transfected with siRNA of either EGFP, p53, or COL17A1 (three different oligonucleotides: COL1, COL2, and COL3) with or without ADR treatment. At 24 hours after plating, the cells were treated with 2  $\mu$ g/mL of ADR for 2 hours. The whole cell lysates were isolated at 48 hours after ADR treatment. siRNA of EGFP was used as a control.

#### 4. The *COL17A1* expression in cancer patients

The mRNA expression level of *COL17A1* was analyzed in multiple cancers (Figure 3.8). The alteration of *COL17A1* expressions in tumor versus normal tissues varied depending on cancer type, suggesting the different roles of *COL17A1* in different organs. The remarkably increased expression of *COL17A1* was detected in skin melanoma, lung carcinoma, pancreatic ductal adenocarcinoma, cholangiocarcinoma, and cervical carcinoma. On the contrary, the above 5-fold depleted expression of *COL17A1* was found in pheochromocytoma and paraganglioma, prostate adenocarcinoma, and testicular germ cell cancer. Interestingly, the *COL17A1* expression in breast cancer was extremely repressed more than 10 fold.

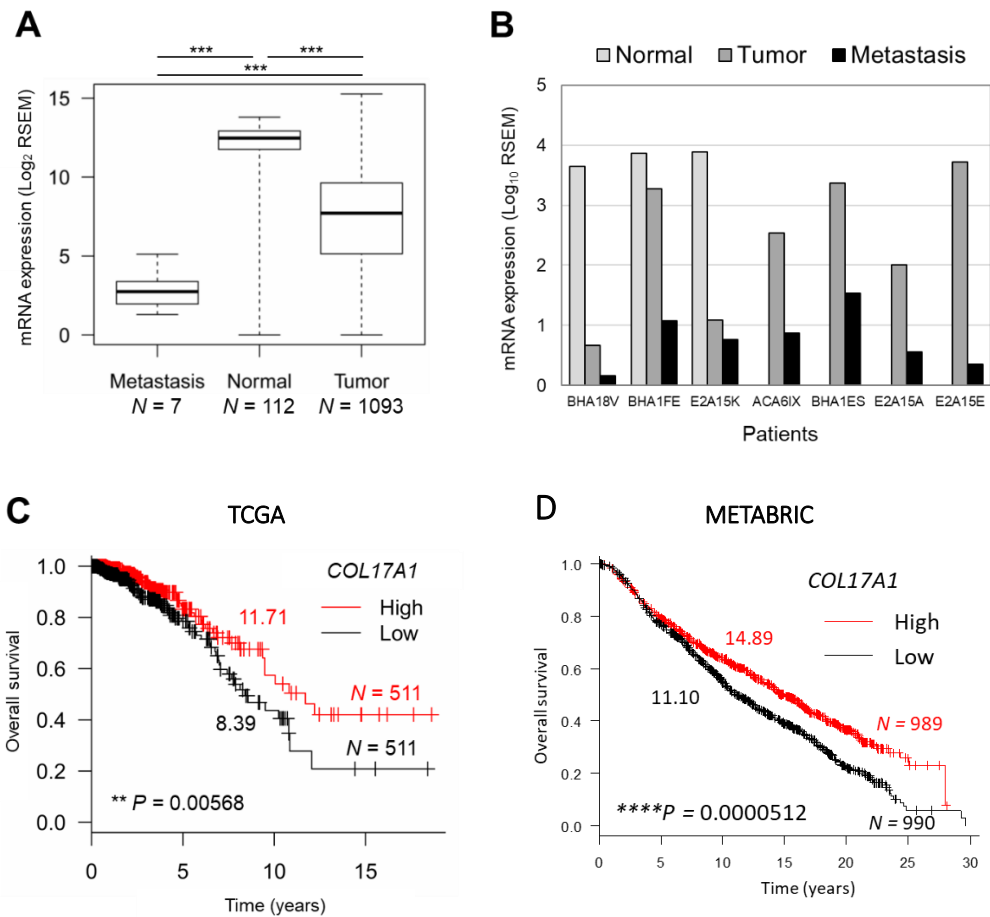
In breast cancer, the *COL17A1* expression was significantly decreased in metastatic tissues compared to primary tumors (Figure 3.9 A) and even from the same patients (Figure 3.9 B). Moreover, low *COL17A1* expression was significantly associated with a shorter median survival time among patients with breast invasive carcinoma in two dependent breast cancer studies (Figure 3.9 C and D). Indeed, the *p53* mutation status and menopausal state are significantly associated with prognosis of breast cancer patients in METABRIC (Figure 3.10 A). The varied expression level of *COL17A1* was found depending on *p53* status and menopausal state (Figure 3.10 B). To investigate whether *COL17A1* expression is an independent prognostic factor, multivariate analyses were performed with *p53* status and menopausal state as covariates, and the *COL17A1* level was still associated with overall survival of breast cancer patients (Figure 3.10 C).

In summary, we elucidated the role of *COL17A1* as a breast tumor metastasis suppressor by inhibition of cancer cell migration and invasion, resulting in a better prognosis of patients.



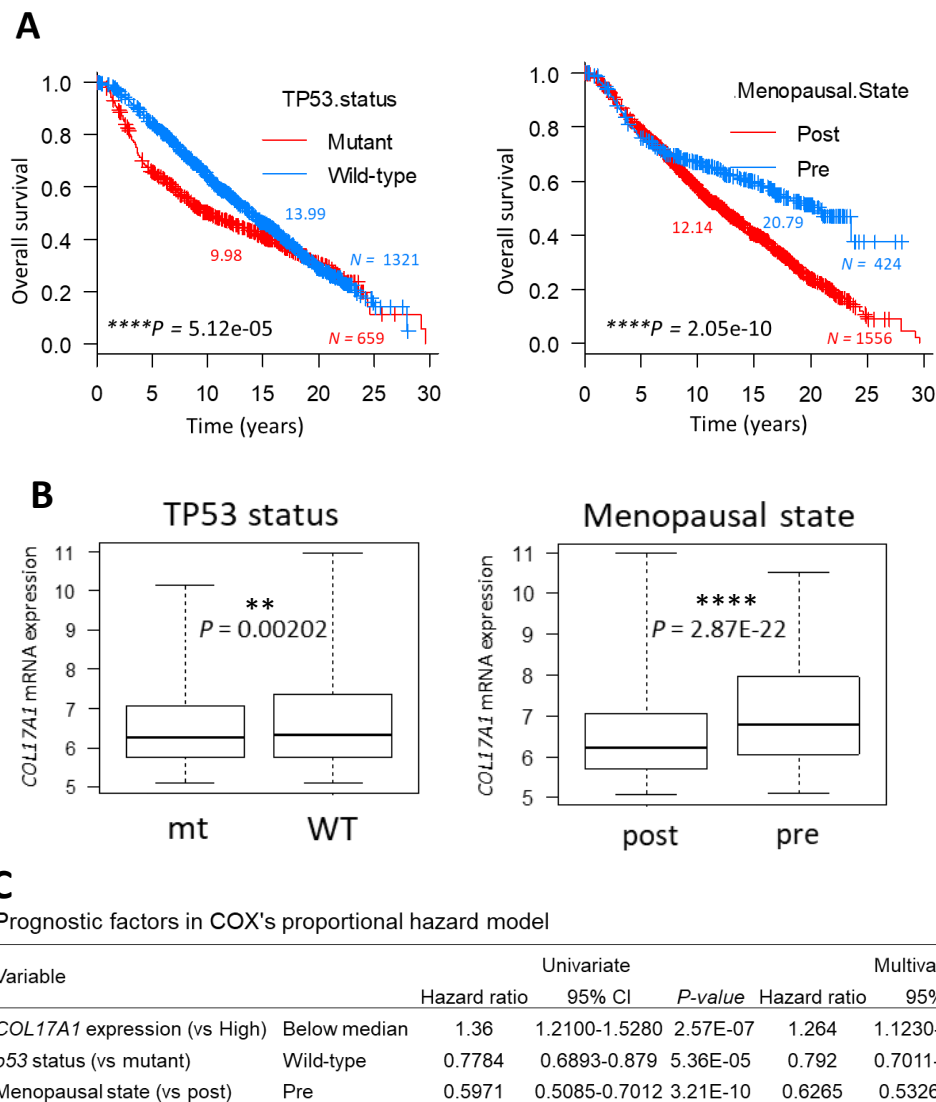
**Figure 3.8** *COL17A1* expression in multiple cancers

The mRNA expression data of multiple cancer studies were obtained from TCGA project by cBioPortal. Box plot analysis represents the ratio (fold-change in log scale) of *COL17A1* mRNA expression in normal over tumor tissue from each patient.



**Figure 3.9** *COL17A1* depletion is associated with tumor progression and poor prognosis

**A and B.** The mRNA expression of *COL17A1* in metastatic, normal, and tumor tissues from all patients (**A**) and from the seven patients with metastasis (**B**). Data were obtained from TCGA. *N*, number of tissues. Mann-Whitney *U*-test; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001. **C and D.** Kaplan-Meier survival curve of patients with high (red) and low (black) expression of *COL17A1* compared to median expression. Data were obtained from TCGA (**C**) and METABRIC (**D**). Number, median survival time (years). *N*, number of patients. *P*-value was assessed by log-rank test.



**Figure 3.10 COL17A1 is an independent prognostic factor**

Data were obtained from METABRIC. **A.** Kaplan-Meier survival curve of patients with different *p53* status and menopausal state. Pre- or post-menopause, patients age below or above 50 years, respectively. **B.** *COL17A1* mRNA expression in different *p53* status and menopausal state. **C.** Univariate and multivariate analyses using *p53* status and menopausal state as covariates. CI, confidence interval. Mann-Whitney *U*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .



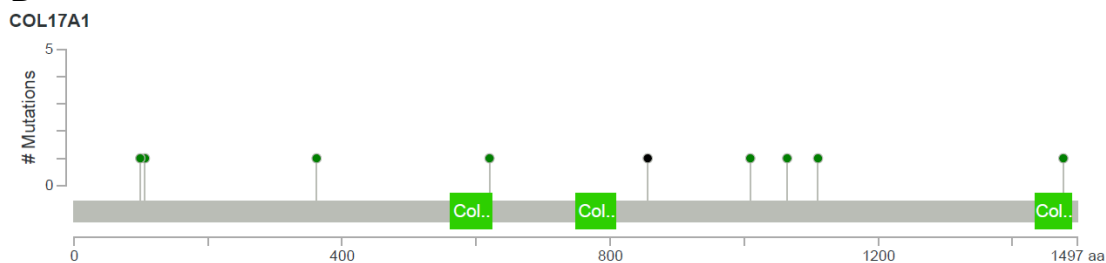
## 5. The mechanism that regulates the *COL17A1* expression

To investigate the mechanism, other than *p53* status, that regulates the decreased expression level of *COL17A1* in breast cancer, the patient data from TCGA and METABRIC were analyzed. The mutations in *COL17* were found in 9 cases (1%) only in TCGA. There was no common mutation (Figure 3.11 A and B) and no specific position (Figure 3.11 B).

### A

AA change	Type	Copy Number	Mutation Assessor
I362M	Missense	ShallowDel	Medium
S1009R	Missense	ShallowDel	Low
V1064A	Missense	Diploid	Low
P1110T	Missense	Diploid	Medium
E99Q	Missense	Diploid	Medium
R106H	Missense	Diploid	Low
K1476Q	Missense	Diploid	Low
Q856*	Nonsense	ShallowDel	
M620I	Missense	Gain	Low

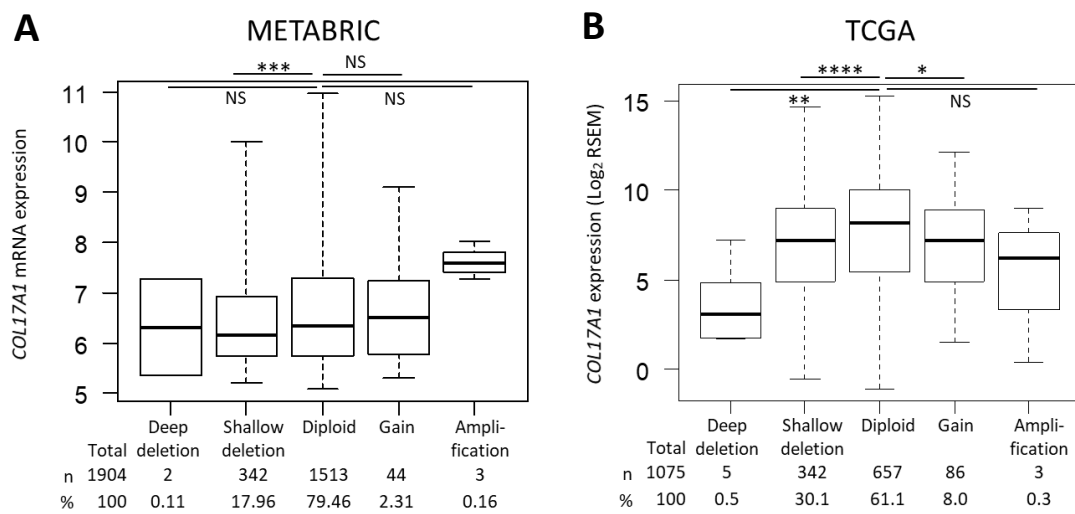
### B



**Figure 3.11 Mutation in *COL17A1***

Mutation data were obtained from TCGA by cBioportal. **A.** List of mutations found in *COL17A1*. AA change, amino acid change. **B.** Number of mutation and the position on *COL17* indicated by amino acid position.

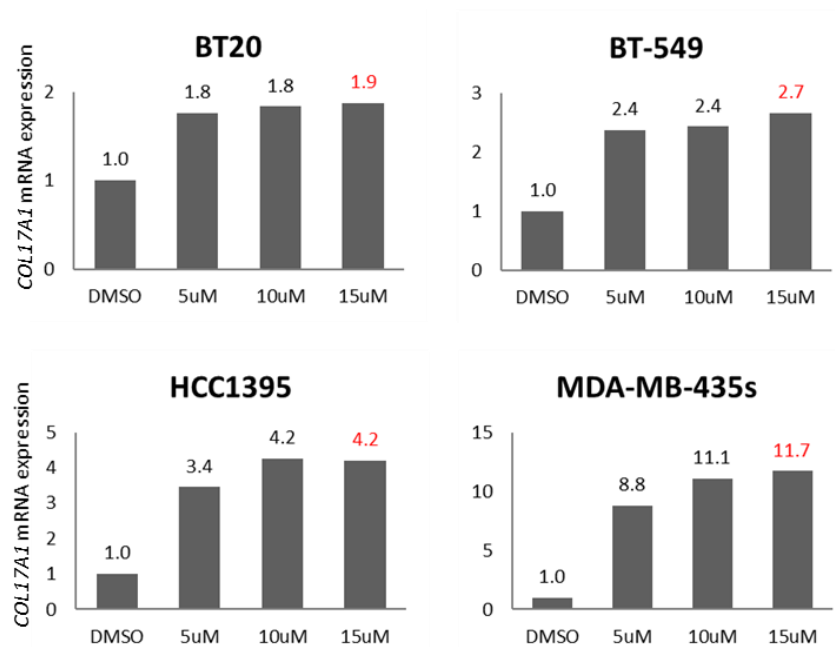
Next, the changes in allelic copy numbers of *COL17A1* were analyzed in breast cancers from TCGA and METABRIC (Figure 3.12 A and B). While most of the breast cancers (73%, average from two studies) had retained two copies of the *COL17A1* locus, 22.5% lost one or both alleles. This led to a significant reduction of *COL17A1* mRNA levels compared to diploid tumors. The remaining 4.5% of the tumors had gained extra copies of the *COL17A1* locus, however, their mRNA levels compared to those in diploid cancers were non-significantly different in METABRIC (Figure 3.12 A), and significantly lower in TCGA (Figure 3.12 B).



**Figure 3.12 Copy number variation in *COL17A1***

**A and B.** *COL17A1* allelic copy number gains and losses in relation to *COL17A1* mRNA expression level in tumor tissues. Data were obtained from METABRIC (**A**) and TCGA (**B**). *P*-values, Mann-Whitney U-test; \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001; NS, *P* > 0.05.

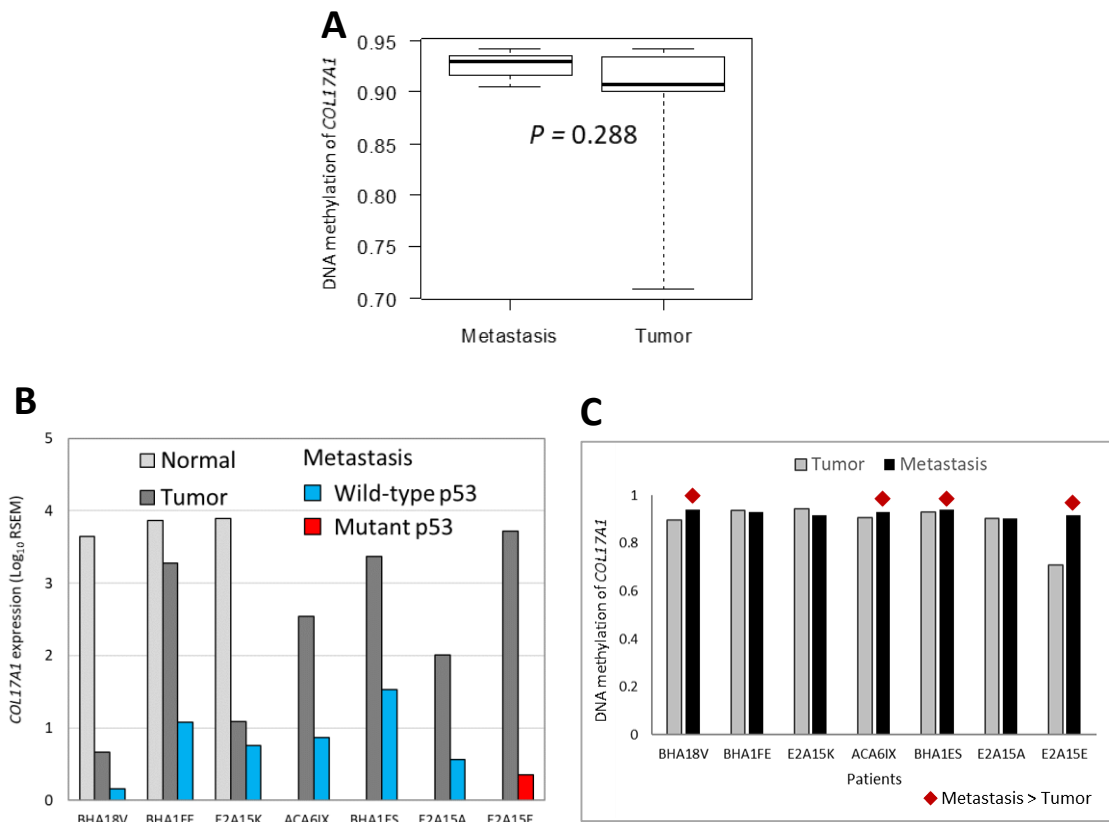
Epigenetic changes such as DNA methylation play a major role in cancer through transcriptional silencing of tumor suppressor genes<sup>55</sup>. Hence, we analyzed the mRNA expression of *COL17A1* in four breast cancer cell lines with *p53* mutant treated with demethylation drug. After demethylation, the expression levels of *COL17A1* increased in a concentration-dependent manner (Figure 3.13). Interestingly, the up-regulated *COL17A1* levels was remarkably high in MDA-MB-435s cells that had been derived from the metastatic pleural effusion (regarding the manufacturer's datasheet). Previous study revealed that the *COL17A1* promoter is indeed hypermethylated in breast tumors in TCGA, and there is a strong negative correlation between *COL17A1* promoter methylation and gene expression<sup>56</sup>.



**Figure 3.13 *COL17A1* DNA methylation in breast cell lines**

Four breast cancer cell lines with *p53* mutant were seeded in 6-well plate and placed at 37°C incubator overnight. The next day, the cell culture media was replaced with fresh media containing 5, 10 and 15  $\mu$ M 5-aza-2'-deoxycytidine (DAC) or DMSO as a control. The media was changed every 12 hours for 2 days, then replaced with fresh media without DAC, and cultured for an additional 48 hours. The bar graphs represent mRNA expression of *COL17A1* normalized with *B2M*.

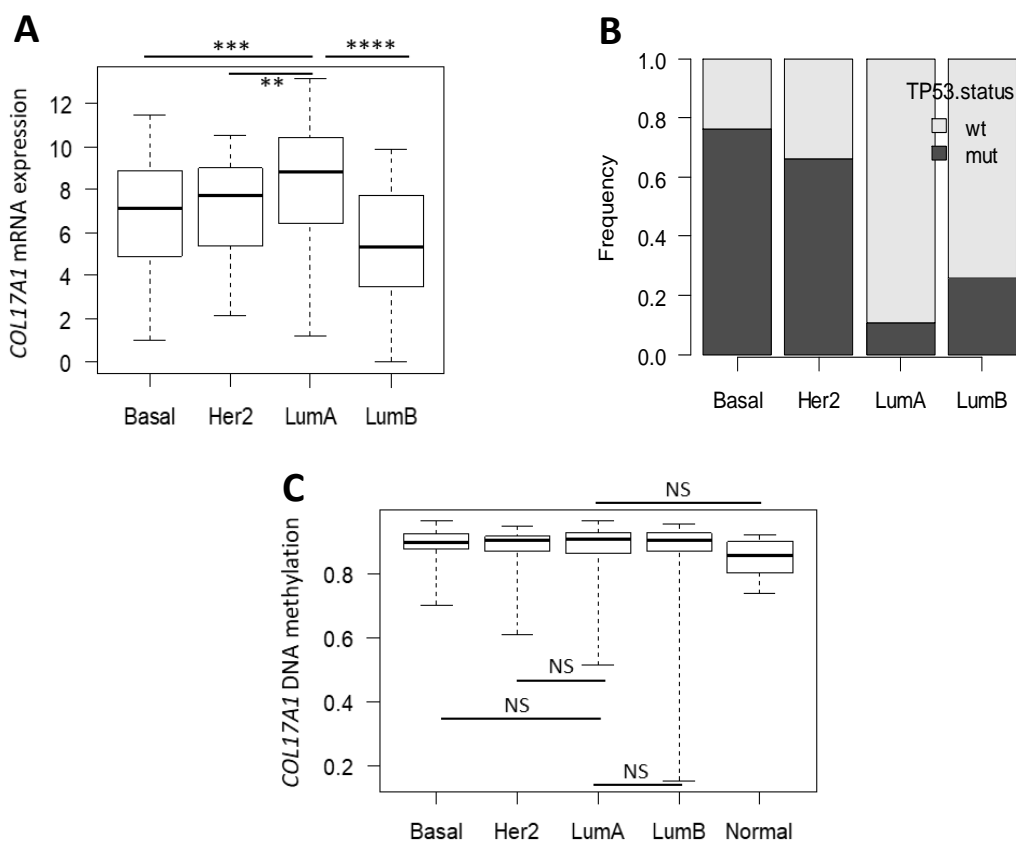
The impact of DNA methylation on *COL17A1* expression levels in breast cancers was further analyzed in patients with metastasis. The median methylation level in metastasized tissues was higher than those in tumor tissues, though this was not significant (Figure 3.14). The analysis of individual patient with metastasis revealed that *p53* mutation was found only in one patient (14.3%) (Figure 3.14 B); whereas the high methylation levels were found in four patients (57%) (Figure 3.14 C). These results suggested that DNA methylation of *COL17A1* might be the major regulator of *COL17A1* expression level in metastasis.



**Figure 3.14** *COL17A1* DNA methylation in patients

**A and C.** DNA methylation level of *COL17A1* in metastatic and tumor tissues derived from seven patients. Data were obtained from TCGA. **B.** The mRNA expression of *COL17A1* in, normal, tumor, and metastatic tissues in the seven patients. The *p53* status of metastatic tissues was labeled as blue (wild-type) or red (mutant). **C.** The DNA methylation level of *COL17A1* in tumor and metastatic tissues. Red diamonds indicate the patients that have higher DNA methylation in metastatic compared to those in tumor tissues. Mann-Whitney *U*-test *P*-value.

Lastly, the *COL17A1* expression in breast cancer subtype was analyzed together with *p53* mutation status and *COL17A1* DNA methylation. The *COL17A1* level was highest in the luminal A (Figure 3.15 A) which had the lowest frequency of *p53* mutation (Figure 3.15 B). Interestingly, the *COL17A1* level was the most down-regulated in luminal B; despite *p53* mutation frequency was not highest, and *COL17A1* DNA methylation did not alter from other subtypes (Figure 3.15 C). This result suggested that additional mechanisms regulate *COL17A1* expression in luminal B subtype.



**Figure 3.15 The regulation of *COL17A1* expression in luminal B subtype**

Data were obtained from TCGA. **A.** Differential mRNA expression of *COL17A1* in PAM50 subtype of breast cancer, including Basal-like, HER2-enriched, Luminal A, and Luminal B. **B.** Frequency of *p53* mutation status in PAM50 subtype. **C.** The DNA methylation level of *COL17A1* in PAM50 subtype and normal tissues. Mann-Whitney *U*-test; \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ ; NS,  $P > 0.05$ .

## Discussion

The loss of function and mutation of p53 not only prevent breast cancer cells from undergoing oncogene-induced senescence and apoptosis but also result in the disruption of metastasis-involved molecules<sup>57-59</sup>. Several metastasis pathways that affect, or are affected by, p53 include cell stemness, extracellular matrix maintenance, cell adhesion, and cell motility<sup>59</sup>. To metastasize, cells must invade through the basement membrane, enter the vasculature (intravasate), survive in the absence of adhesion, exit the vasculature (extravasate) and establish a new tumor in a foreign microenvironment<sup>2</sup>. The alteration and loss of cell adhesion structures are involved in an early step in epithelial cancer progression<sup>60</sup>. Our results introduce the role of *COL17A1* in inhibiting cancer cell growth by an ectopic expression of *COL17A1*; but its knockdown did not inhibit the ADR-induced growth suppression. The over-expression of COL17 protein in MDA-MB-231 stable cells was not associated with cell proliferation. Taken together, the role of *COL17A1* in cancer cell growth was still controversial.

*COL17A1* expression is particularly low in metastatic breast cancer cells<sup>25</sup>. Conversely, MDA-MB-231 cells that genetically over-express COL17, exhibited significantly decreased invasive properties that is not related to epithelial-mesenchymal transition. The released COL17 ectodomain, also known as LAD-1, was detected in cell culture media, human skin extracts, and serum obtained from patients with an autoimmune blistering disorder<sup>61</sup>. The skin blistering of this disease is caused by autoantibodies against the COL17 ectodomain, which introduces its role in maintaining dermal-epidermal cohesion<sup>62</sup>. Recent studies have clarified its function in decreased keratinocyte migration by dampening mTOR signaling<sup>26</sup>. This study revealed the role of the 120-kDa COL17 ectodomain in the suppression of breast cancer cell invasion. Indeed, myoepithelial cells act as natural tumor suppressors by secreting various molecules that have inhibitory effects on tumor cell growth, invasion and angiogenesis<sup>63</sup>. Our results propose a new mechanism supporting myoepithelial cell function as a physical barrier to prevent the invasion of tumor cells from the duct to the stroma by secreting the COL17 ectodomain.

Moreover, patients bearing high *COL17A1*-expressing tumors exhibit a longer survival time compared with those bearing low *COL17A1*-expressing tumors. *COL17A1* expression level is an independent prognosis marker for breast cancer patients. Taken together, our results implicate a role of COL17 in suppressing breast cancer cell migration and invasion, whereby a high level of *COL17A1* expression leads to a better prognosis of patients with breast invasive carcinoma.

The misexpression of *COL17A1* is cancer type specific. While the *COL17A1* expression is over-expressed in various cancer types, it is extremely suppressed in breast cancer. An analysis of metastasis breast invasive carcinoma revealed that each patient demonstrates different expression levels of *COL17A1* in different types of tissues, i.e., normal, tumor, and metastasis. The increased *COL17A1* levels following DAC treatment suggest the role of DNA methylation in regulating the *COL17A1* expression in breast cancer supporting the previous publication<sup>56</sup>. However, it is too early to conclude that DNA methylation regulates *COL17A1* expression in metastasis because of the less number of metastatic samples.

The *COL17A1* expression is remarkably suppressed in luminal B subtype, which shows more aggressive phenotype, higher proliferation, and worse outcomes in patients comparing to the luminal A. In luminal B, it seems that *p53* status is not a major regulator of *COL17A1* expression; and DNA methylation does not change between cancer subtypes. Previous study<sup>64</sup> revealed that the copy loss events (loss of heterozygosity, LOH) in *p53* is highest in luminal B cancers compared to other types. This observation agrees with our results that the mRNA and protein expression of COL17A1 are suppressed in *p53*-knockout compared to *p53* wild-type MCF10A cells without ADR treatment (Figure 2.4). These results suggest that not only *p53* mutation status, but the LOH of *p53* also plays an important role in the regulation of *COL17A1*.

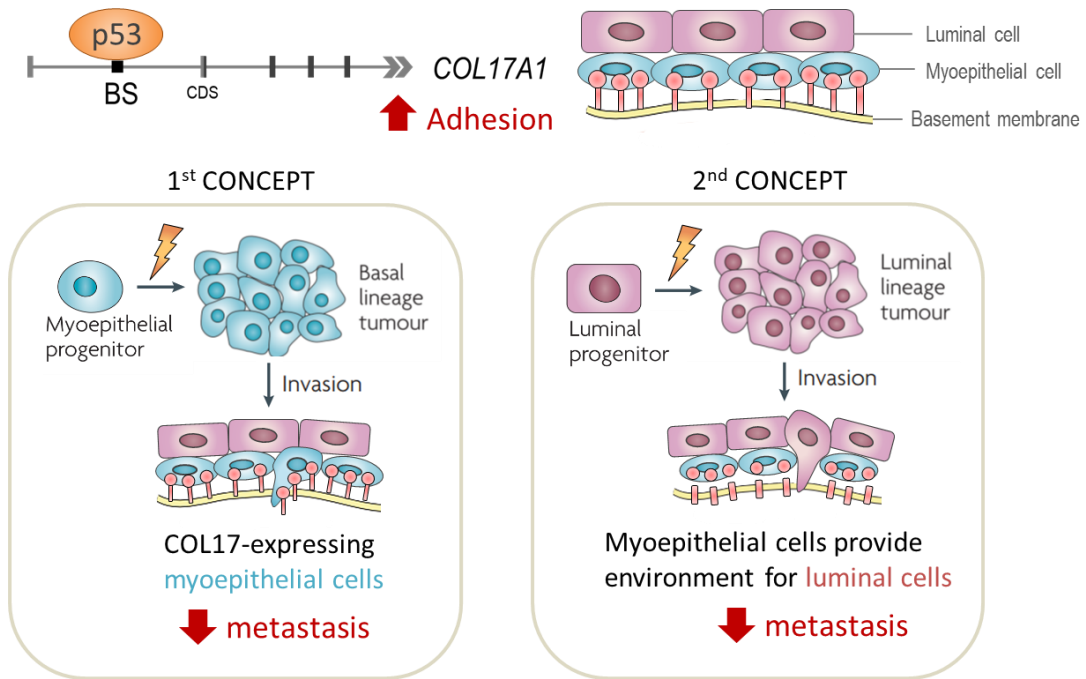
Concerning the COL17 location that is limited only the myoepithelial cells, but not the luminal cells, the cell lines used for COL17 analysis should be clarified the cell type. MCF10A cells are widely used as non-malignant breast epithelial cells to study breast cancer tumorigenesis. The cell type of MCF10A is still not clear because they exhibit a myoepithelial-

like phenotype but express luminal markers in 2D culture<sup>65</sup>. In the analysis of patients, the data taken from public database are also needed to be pointed out. The absolute *COL17A1* levels vary widely between various tumor types, as well as between various normal tissues. The different ratios between the numbers of myoepithelial and non-myoepithelial cells in different tissue types could partly account for that. In addition, the *COL17A1* levels would be largely affected by the number of normal cells within the tumor samples because the tumor content (percentage of tumor cells within samples) was unclarified.



## Conclusion

In this study, *COL17A1* was identified as a p53 transcriptional target gene. Wild-type p53 suppresses breast cancer invasion in pre-metastatic steps by up-regulating the expression of COL17, resulting in the promotion of the adhesion between the basement membranes and myoepithelial cells (Figure 4). The COL17-mediated adhesion prevents the primary tumor cells from escaping into the surrounding breast tissues. By contrast, p53 loss or mutation diminished COL17 function, resulting in less adhesive basement membranes, which cause it to eventually break down and allow tumor cells to escape into the bloodstream to form metastasis in the secondary organ, thereby contributing to a worse prognosis of patients. Although the molecules involved in the p53-COL17 signaling pathway are essential to be elucidated, our findings indicated the regulation of breast cancer metastasis by the p53-COL17 pathway and potential roles of COL17 as a prognostic biomarker.



**Figure 4. The regulation of breast cancer metastasis by the p53-COL17 pathway**

Wild-type p53 binds to *COL17A1* and up-regulates its expression, resulting in the promotion of the adhesion between the myoepithelial cells and underlying basement membranes. First concept, myoepithelial cell is a tumor progenitor. The cells over-expressing COL17 exhibit less metastasis. Second concept, myoepithelial cell is a metastasis barrier that excrete COL17 ectodomain which provide the environment preventing the metastasis of the luminal-lineage tumor. This figure is partly modified from *Nat Rev Cancer* 7, 659-72 (2007).

## References

1. Ferlay J, S.I., Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray, F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. *Lyon, France: International Agency for Research on Cancer* (2013).
2. Vargo-Gogola, T. & Rosen, J.M. Modelling breast cancer: one size does not fit all. *Nat Rev Cancer* **7**, 659-72 (2007).
3. Society, A.C. Breast Cancer Facts & Figures 2015-2016. *Atlanta: American Cancer Society, Inc* (2015).
4. Redig, A.J. & McAllister, S.S. Breast cancer as a systemic disease: a view of metastasis. *J Intern Med* **274**, 113-26 (2013).
5. Nguyen, D.X., Bos, P.D. & Massague, J. Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer* **9**, 274-84 (2009).
6. Weigelt, B., Peterse, J.L. & van 't Veer, L.J. Breast cancer metastasis: markers and models. *Nat Rev Cancer* **5**, 591-602 (2005).
7. Chuang, H.Y., Lee, E., Liu, Y.T., Lee, D. & Ideker, T. Network-based classification of breast cancer metastasis. *Mol Syst Biol* **3**, 140 (2007).
8. Jin, X. & Mu, P. Targeting Breast Cancer Metastasis. *Breast Cancer (Auckl)* **9**, 23-34 (2015).
9. Vogelstein, B., Lane, D. & Levine, A.J. Surfing the p53 network. *Nature* **408**, 307-10 (2000).
10. Olivier, M., Hollstein, M. & Hainaut, P. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol* **2**, a001008 (2010).
11. Gasco, M., Shami, S. & Crook, T. The p53 pathway in breast cancer. *Breast Cancer Res* **4**, 70-6 (2002).
12. Network, T.C.G.A. Comprehensive molecular portraits of human breast tumours.

- Nature* **490**, 61-70 (2012).
13. Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* **7**, 11479 (2016).
  14. Bertheau, P. *et al.* p53 in breast cancer subtypes and new insights into response to chemotherapy. *Breast* **22 Suppl 2**, S27-9 (2013).
  15. Mori, J. *et al.* Cystatin C as a p53-inducible apoptotic mediator that regulates cathepsin L activity. *Cancer Sci* **107**, 298-306 (2016).
  16. Koguchi, T., Tanikawa, C., Mori, J., Kojima, Y. & Matsuda, K. Regulation of myo-inositol biosynthesis by p53-ISYNA1 pathway. *Int J Oncol* **48**, 2415-24 (2016).
  17. Chizu Tanikawa, Y.-z.Z., Ryuta Yamamoto, Yusuke Tsuda, Masami Tanaka, Yuki Funauchi, Jinichi Mori, Seiya Imoto, Rui Yamaguchi, Yusuke Nakamura, Satoru Miyano, Hidewaki Nakagawa, Koichi Matsuda. The Transcriptional Landscape of p53 Signalling Pathway. *EBioMedicine (In Press)*(2017).
  18. Borradori, L. & Sonnenberg, A. Structure and function of hemidesmosomes: more than simple adhesion complexes. *J Invest Dermatol* **112**, 411-8 (1999).
  19. Franzke, C.W., Tasanen, K., Schumann, H. & Bruckner-Tuderman, L. Collagenous transmembrane proteins: collagen XVII as a prototype. *Matrix Biol* **22**, 299-309 (2003).
  20. Aho, S. & Uitto, J. 180-kD bullous pemphigoid antigen/type XVII collagen: tissue-specific expression and molecular interactions with keratin 18. *J Cell Biochem* **72**, 356-67 (1999).
  21. McGrath, J.A. *et al.* Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa. *Nat Genet* **11**, 83-6 (1995).
  22. Schumann, H. *et al.* Three novel homozygous point mutations and a new polymorphism in the COL17A1 gene: relation to biological and clinical phenotypes of junctional epidermolysis bullosa. *Am J Hum Genet* **60**, 1344-53 (1997).
  23. Bergstraesser, L.M., Srinivasan, G., Jones, J.C., Stahl, S. & Weitzman, S.A. Expression of hemidesmosomes and component proteins is lost by invasive breast cancer cells. *Am*

- J Pathol* **147**, 1823-39 (1995).
24. Laval, S. *et al.* Dual roles of hemidesmosomal proteins in the pancreatic epithelium: the phosphoinositide 3-kinase decides. *Oncogene* **33**, 1934-44 (2014).
  25. Geiger, T., Madden, S.F., Gallagher, W.M., Cox, J. & Mann, M. Proteomic portrait of human breast cancer progression identifies novel prognostic markers. *Cancer Res* **72**, 2428-39 (2012).
  26. Jackow, J., Loffek, S., Nystrom, A., Bruckner-Tuderman, L. & Franzke, C.W. Collagen XVII Shedding Suppresses Re-Epithelialization by Directing Keratinocyte Migration and Dampening mTOR Signaling. *J Invest Dermatol* **136**, 1031-41 (2016).
  27. Loffek, S. *et al.* Transmembrane collagen XVII modulates integrin dependent keratinocyte migration via PI3K/Rac1 signaling. *PLoS One* **9**, e87263 (2014).
  28. Qiao, H. *et al.* Collagen XVII participates in keratinocyte adhesion to collagen IV, and in p38MAPK-dependent migration and cell signaling. *J Invest Dermatol* **129**, 2288-95 (2009).
  29. Components and structure of the basement membrane zone (dermal-epidermal junction). (Department of Dermatology, University of Pennsylvania, Penn Presbyterian Medical Center Medical Arts Building, Philadelphia, PA, USA, 2016).
  30. Miyamoto, T., Tanikawa, C. & Yodsurang, V.Y.-z.Z., Seiya Imoto, Rui Yamaguchi, Satoru Miyano, Hidewaki Nakagawa, Koichi Matsuda. Identification of a p53-repressed gene module in breast cancer cells. *Oncotarget* (**Tentative**)(2017).
  31. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* **2**, 401-4 (2012).
  32. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, p11 (2013).
  33. Tsukada, T. *et al.* Enhanced proliferative potential in culture of cells from p53-deficient mice. *Oncogene* **8**, 3313-22 (1993).
  34. Brosh, R. *et al.* p53-dependent transcriptional regulation of EDA2R and its involvement in chemotherapy-induced hair loss. *FEBS Lett* **584**, 2473-7 (2010).

35. Tanikawa, C. *et al.* XEDAR as a putative colorectal tumor suppressor that mediates p53-regulated anoikis pathway. *Oncogene* **28**, 3081-92 (2009).
36. Zhang, Y., Yan, W. & Chen, X. Mutant p53 disrupts MCF-10A cell polarity in three-dimensional culture via epithelial-to-mesenchymal transitions. *J Biol Chem* **286**, 16218-28 (2011).
37. Allen, M.A. *et al.* Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms. *Elife* **3**, e02200 (2014).
38. Janky, R. *et al.* iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput Biol* **10**, e1003731 (2014).
39. Raman, V. *et al.* Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* **405**, 974-8 (2000).
40. Loughery, J., Cox, M., Smith, L.M. & Meek, D.W. Critical role for p53-serine 15 phosphorylation in stimulating transactivation at p53-responsive promoters. *Nucleic Acids Res* **42**, 7666-80 (2014).
41. Hirako, Y. *et al.* Cleavage of BP180, a 180-kDa bullous pemphigoid antigen, yields a 120-kDa collagenous extracellular polypeptide. *J Biol Chem* **273**, 9711-7 (1998).
42. Franzke, C.W. *et al.* Transmembrane collagen XVII, an epithelial adhesion protein, is shed from the cell surface by ADAMs. *Embo j* **21**, 5026-35 (2002).
43. Kariko, K., Bhuyan, P., Capodici, J. & Weissman, D. Small interfering RNAs mediate sequence-independent gene suppression and induce immune activation by signaling through toll-like receptor 3. *J Immunol* **172**, 6545-9 (2004).
44. Franzke, C.W., Bruckner-Tuderman, L. & Blobel, C.P. Shedding of collagen XVII/BP180 in skin depends on both ADAM10 and ADAM9. *J Biol Chem* **284**, 23386-96 (2009).
45. Van den Bergh, F., Eliason, S.L. & Giudice, G.J. Type XVII collagen (BP180) can function as a cell-matrix adhesion molecule via binding to laminin 332. *Matrix Biol* **30**, 100-8 (2011).
46. Lakhani, S.R. & O'Hare, M.J. The mammary myoepithelial cell--Cinderella or ugly

- sister? *Breast Cancer Res* **3**, 1-4 (2001).
47. Perou, C.M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-52 (2000).
  48. Felipe Lima, J., Nofech-Mozes, S., Bayani, J. & Bartlett, J.M. EMT in Breast Carcinoma-A Review. *J Clin Med* **5**(2016).
  49. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-5 (2012).
  50. Li, X., Yang, H., Huang, H. & Zhu, T. CELLCOUNTER: novel open-source software for counting cell migration and invasion in vitro. *Biomed Res Int* **2014**, 863564 (2014).
  51. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-52 (2012).
  52. Kanda, Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. *Bone Marrow Transplant* **48**, 452-8 (2013).
  53. Liang, C.C., Park, A.Y. & Guan, J.L. In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. *Nat Protoc* **2**, 329-33 (2007).
  54. Hall, D.M. & Brooks, S.A. In vitro invasion assay using matrigel: a reconstituted basement membrane preparation. *Methods Mol Biol* **1070**, 1-11 (2014).
  55. Baylin, S.B. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol* **2** **Suppl 1**, S4-11 (2005).
  56. Thangavelu, P.U., Krenacs, T., Dray, E. & Duijf, P.H. In epithelial cancers, aberrant COL17A1 promoter methylation predicts its misexpression and increased invasion. *Clin Epigenetics* **8**, 120 (2016).
  57. Arjonen, A. *et al.* Mutant p53-associated myosin-X upregulation promotes breast cancer invasion and metastasis. *J Clin Invest* **124**, 1069-82 (2014).
  58. Chander, H. *et al.* Toca-1 is suppressed by p53 to limit breast cancer cell invasion and tumor metastasis. *Breast Cancer Res* **16**, 3413 (2014).
  59. Powell, E., Piwnica-Worms, D. & Piwnica-Worms, H. Contribution of p53 to metastasis. *Cancer Discov* **4**, 405-14 (2014).

60. Davis, T.L., Goldman, A.J. & Cress, A.E. SUPPRESSION AND ALTERATION OF ADHESION STRUCTURES IN HUMAN EPITHELIAL CANCER PROGRESSION. in *Cell Adhesion and Cytoskeletal Molecules in Metastasis* (eds. Cress, A.E. & Nagle, R.B.) 19-46 (Springer Netherlands, Dordrecht, 2006).
61. Marinkovich, M.P., Taylor, T.B., Keene, D.R., Burgeson, R.E. & Zone, J.J. LAD-1, the linear IgA bullous dermatosis autoantigen, is a novel 120-kDa anchoring filament protein synthesized by epidermal cells. *J Invest Dermatol* **106**, 734-8 (1996).
62. Nishie, W. *et al.* Ectodomain shedding generates Neopeptides on collagen XVII, the major autoantigen for bullous pemphigoid. *J Immunol* **185**, 4938-47 (2010).
63. Pandey, P.R., Saidou, J. & Watabe, K. Role of myoepithelial cells in breast tumor progression. *Front Biosci (Landmark Ed)* **15**, 226-36 (2010).
64. Creighton, C.J. The molecular profile of luminal B breast cancer. *Biologics* **6**, 289-97 (2012).
65. Qu, Y. *et al.* Evaluation of MCF10A as a Reliable Model for Normal Human Mammary Epithelial Cells. *PLoS One* **10**, e0131285 (2015).



## Appendix

**Table A1: Sequences of DNA and RNA oligonucleotides**

<b>siRNA oligonucleotides</b>	<b>Sense</b>	<b>Antisense</b>
siEGFP	GCAGCACGACUUCUUAAGT	CUUGAAGAAGUCGUGCUGC
sip53	GACUCCAGUGGUAUUCUACTT	GUAGAUUACCACUGGAGUCTT
siCOL1	CCACAAGACUUACAUCUUTT	AAGGAUGUAAGUCUUGUGGTT
siCOL1	GCCAGGAGAUUCAGCAGUATT	UACUGCUGAAUCUCCUGGCTT
siCOL3	CCAUCUCUUCUGAAGACAUTT	AUGUCUUCAGAAGAGAUGGTT

<b>Primers</b>	<b>Forward</b>	<b>Reverse</b>
<b>Quantitative PCR</b>		
Human <i>COL17A1</i>	CTGACTTTGCTGGAGATCTGG	TAGGCCATCCCTTGCAGTAG
Human <i>B2M</i>	TTCTGGCCTGGAGGCTATC	TCAGGAAATTTGACTTTCCATTC
Human <i>ACTB</i>	CCCTGGAGAAGAGCTACGAG	TGAAGGTAGTTTCGTGGATGC
Mouse <i>Col17a1</i>	GAAAGGAGACAAAGGTGACCA	CGGCTTGATGGCAATACTTC
Mouse <i>Gapdh</i>	AATGTGTCCGTCGTGGATCTGA	GATGCCTGCTTCACCACCTTCT
<b>Mouse genotyping</b>		
Mouse <i>p53</i>	GTTATGCATCCATACAGTACA	CCGCAGGATTTACAGACACC
<b>Gene reporter assay</b>		
<i>COL17A1</i> BS	TGGAAGACGAACACACTGGT	AGGTACGTGTTGGGAGACTG
MT1	GTCTCAGGTATTTACCTGGGCAG GAAACGTTC	TACCTATCATAACAGGATTCACCA GTGTGTTTCGTCTTC
MT2	AGGTTTTTGCCCCACGC	GGAATAAGATCTTCCCATGGGAC
<b>ChIP assay</b>		
p53 BS in <i>WAF1</i>	CTGGACTGGGCACTCTTGTC	CTCCTACCATCCCCTTCCTC
<i>COL17A1</i> BS	TGGAAGACGAACACACTGGT	AAATGGAGAGTGATGGCGTG
<b>Colony formation assay</b>		
<i>COL17A1</i> CDS	AAAGCGGCCGCGCCGATGGAT GTAACCAAGAAAAACAAAC	AAAGTCGACCGGCTTGACAGCA ATACTTC
<b>Stable cell line generation</b>		
<i>COL17A1</i> genomic DNA	GCCGCGATGGATGTAACCAAGAA AAACAAAC	TCACGGCTTGACAGCAATACTTC

## Acknowledgement

I would like to express my deepest gratitude to my thesis advisor, Professor Koichi Matsuda, for the valuable advice, useful guidance, endless support, patience, and encouragement throughout my study. I would like to express my sincere thanks to all co-authors of my publication, Chizu Tanikawa, Takafumi Miyamoto, Makoto Hirata, and Paulisally Hau Yi Lo for the great ideas, teaching, correction, and wonderful suggestions. I especially thank to Professor Sumio Sugano, Associate Professor Susumu Goyama, Associate Professor Tsuneo Ikenoue, and Associate Professor Hiroaki Uchida for helpful discussions and recommendations for my thesis.

My gratitude and appreciation are also expressed to all researchers in my lab, Satomi Takahashi, Misato Oshima, Akane Sei, and Satoyo Oda for the technical assistances in mouse genotyping, DNA sequencing, and providing laboratory materials. I also thank my tutor, Dr. Jinichi Mori for his supports, and Liu Yuyu for the helpful ImageJ programming. I especially thank to all my seniors, friends, and juniors for taking care of laboratory mice and supporting equipment in parts of my thesis. I would like to thank the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan for the academic scholarship.

Finally, this thesis and this degree are dedicated to my well-beloved family.