The Role of Tumor Suppressor DEAR1 in the Acquisition of Mammary Stem/Progenitor Cell Properties

Uyen Le

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations

Part of the Biology Commons, Laboratory and Basic Science Research Commons, Medicine and Health Sciences Commons, and the Molecular Genetics Commons

Recommended Citation
Le, Uyen, "The Role of Tumor Suppressor DEAR1 in the Acquisition of Mammary Stem/Progenitor Cell Properties" (2018). UT GSBS Dissertations and Theses (Open Access). 904. https://digitalcommons.library.tmc.edu/utgsbs_dissertations/904

This Dissertation (PhD) is brought to you for free and open access by the Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in UT GSBS Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact nha.huynh@library.tmc.edu.
THE ROLE OF TUMOR SUPPRESSOR *DEAR1* IN THE ACQUISITION OF MAMMARY STEM/PROGENITOR CELL PROPERTIES

by

Uyen Quynh Nguyen Le, B.S.

APPROVED:

__________________________________________
Ann Killary, Ph.D.
Advisory Professor

__________________________________________
Subrata Sen, Ph.D.

__________________________________________
Ralf Krahe, Ph.D.

__________________________________________
Sadhan Majumder, Ph.D.

__________________________________________
Kenneth Tsai, M.D./Ph.D.

APPROVED:

__________________________________________
Dean, The University of Texas
MD Anderson Cancer Center UTHHealth Graduate School of Biomedical Sciences
DEFINING THE ROLE OF TUMOR SUPPRESSOR DEAR1 IN THE ACQUISITION OF MAMMARY STEM/PROGENITOR CELL PROPERTIES

A

DISSERTATION

Presented to the Faculty of

The University of Texas

M.D. Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Uyen Quynh Nguyen Le, B.S.

Houston, Texas

December 2018
DEDICATION

For the three special women in my life:

My mother, who has showered me with unconditional love and taught me loyalty and forgiveness;

My advisor, who taught me to find the silver lining in every dark cloud;

And my grandmother, who taught me that kindness and pure hearts exist.
ACKNOWLEDGEMENTS

First and foremost, I want to thank my family with all of my heart. Without them, I would not be who I am or where I am today. Without them, I would not have traveled or experienced other cultures. Without them, I would not have found success. You have supported me since birth and continue to do so as I figure my way through the next phase of my life. Simply saying “thank you” does not seem enough to show my appreciation and gratitude. I love you all very much, forever and always.

Next, I want to thank my lab family – Dr. Ann Killary, Dr. Nanyue Chen, Dr. Seetharamen Balasenthil, and Dr. Jacquelyn Reuther. Thank you for the support, advice and encouragement that you have shown me all these years. You have taught me how optimism, perseverance and persistence are just as important in the lab as it is in life. I especially want to thank my advisor for teaching me to appreciate the small wins in life, that striving for my personal best is more worthwhile than trying to be the best, and for being my part-time therapist (free of charge).

I would like to acknowledge my advisory committee members – Dr. Subrata Sen, Dr. Ralf Krahe, Dr. Kenneth Tsai, and Dr. Sadhan Majumder. They have been with me on my academic journey from the beginning. They have provided me with invaluable project and career recommendations, as well as written countless letters of reference. Thank you for your support and advice over these past few years.

I also want to acknowledge my friends at the Graduate School of Biomedical Sciences (GSBS), the legacy Human and Molecular Genetics (HMG) Program, the Community Outreach Program, and the Department of Translational Molecular Pathology at MD Anderson Cancer Center. I have formed so many life-long friendships through GSBS and the HMG Program while developing useful collaborative and leadership skills. I appreciate how close-knit our HMG community was and will remember all of the fun times we’ve had together. The Outreach Program will always have a special place in my heart. I have enjoyed working with GSBS students, faculty, and staff to organize fun and educational events for the Houston community. Seeing bright young
minds so enthusiastic about learning gives me constant inspiration to become an educator one day. I am grateful for the invaluable leadership experiences and opportunities that I would not have otherwise been given if I had not joined the Outreach Program. I also want to thank GSBS and MD Anderson faculty and staff for taking their time to help graduate students, like myself, navigate their way through school, with special thanks to Karen Weinberg, Cheryl Spitzenberger, Elisabeth Lindheim, Lily D’Agostino, Bert Shaw, and Barbara Sadeghi.

Finally, I want to thank the love of my life, Josh. We have powered through college, medical school, residency, graduate school, and long-distance. To say that our relationship is unique would be an understatement. Thank you for being my shoulder to cry on, my ray of sunshine, and my rock. With you by my side, all my insecurities dissolve and I feel like I can do anything.
Breast cancer is the most commonly diagnosed cancer in women in America. Ductal carcinoma in situ (DCIS), one of the earliest pre-invasive forms of invasive ductal carcinoma (IDC), has a 30-50% risk of progressing to IDC. Understanding the mechanisms regulating progression from DCIS to IDC would help identify biomarkers to stratify patients at higher risk of progression or metastasis. Cumulative literature suggests the earliest phase of dissemination from the primary tumor is driven by the epithelial-mesenchymal transition (EMT) program. 

DEAR1 is a tumor suppressor gene which is mutated, undergoes loss of heterozygosity in breast cancer, and is downregulated in DCIS lesions and IDC. DEAR1 regulates acinar morphogenesis and cell polarity and is a negative regulator of TGF-β-driven EMT through inhibition of SMAD3. Studies have now demonstrated that induction of EMT promotes acquisition of stem/progenitor cell properties, further adding to the repertoire of cellular regulation by EMT mediators.

I demonstrate that loss of DEAR1 in human mammary epithelial cells (HMECs) and DCIS cells results in a mammosphere phenotype independent of the canonical TGF-β pathway, suggesting that DEAR1 regulates stem/progenitor cell properties. DEAR1-knockdown (KD) HMEC mammospheres express high levels of stem/progenitor cell marker aldehyde dehydrogenase (ALDH1) and display a basal-like phenotype through repression of CD24 and EpCAM expression. There is significant upregulation of master EMT and stem cell regulators, including SNAI2, in DEAR1-KD HMECs and I show that DEAR1 binds to and promotes polyubiquitination of SNAI2. I reveal a novel DEAR1-SNAI2 axis that partially regulates stem/progenitor cell properties in HMECs and demonstrate a significant association between loss of DEAR1 in basal-like/triple-negative breast cancers (TNBC), early-age of onset, and risk
of shorter time to metastasis in TNBC. Additionally, I identify a possible mechanism governing DEAR1 regulation in mammary epithelial cells through miRNAs miR-10b and miR-196b.

Results herein demonstrate that DEAR1 promotes stem/progenitor cell properties partially through TGF-β-mediated EMT and also through SNAI2 independently of TGF-β-SMAD3 signaling. I hope to use this understanding of DEAR1 and its regulation of cell polarity, EMT, and stemness to stratify high risk patients who would benefit from more aggressive or targeted therapy.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCEPTANCE PAGE</td>
<td>I</td>
</tr>
<tr>
<td>TITLE PAGE</td>
<td>II</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>III</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>IV</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>VI</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>VIII</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>X</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>XIII</td>
</tr>
</tbody>
</table>

## CHAPTER 1: BACKGROUND AND SIGNIFICANCE  

1.1 *Ductal Carcinoma In Situ and Invasive Ductal Carcinoma*  

1.2 *Epithelial-Mesenchymal Transition (EMT) and its role in carcinogenesis*  

1.3 *Stem cells and cancer*  

1.4 *TRIM proteins and their roles in cancer*  

1.5 *Ductal Epithelium-Associated RING Chromosome 1 (DEAR1)*  

## CHAPTER 2: LOSS OF DEAR1 CONFFERS STEM AND/OR PROGENITOR CELL PROPERTIES  

2.1 *Introduction*  

2.2 *Results*  

2.3 *Discussion*  

## CHAPTER 3: DEFINING THE MECHANISM BY WHICH DEAR1 MEDIATES STEM AND/OR PROGENITOR CELL PROPERTIES  

3.1 *Introduction*  

viii
3.2 Results 62
3.3 Discussion 72

CHAPTER 4: LOSS OF DEAR1 CORRELATES WITH POOR OVERALL SURVIVAL IN INVASIVE BREAST CANCER AND TIME TO METASTASIS IN TRIPLE NEGATIVE BREAST CANCER 75
4.1 Introduction 76
4.2 Results 77
4.3 Discussion 84

CHAPTER 5: POST-TRANSCRIPTIONAL REGULATION OF DEAR1 BY MICRORNAS 89
5.1 Introduction 90
5.2 Results 93
4.3 Discussion 97

CHAPTER 6: MATERIALS AND METHODS 104

CHAPTER 7: DISCUSSION AND FUTURE DIRECTIONS 111

REFERENCES 123

VITA 143
LIST OF FIGURES

Figure 1.1: Cancer cell of origin theory ................................................................. 6

Figure 1.2: Cancer and normal stem cell characteristics ........................................ 17

Figure 1.3: Illustration of DEAR1 gene, protein, and mutation residues ..................... 23

Figure 1.4: Illustration of DEAR1 regulation of cell polarity and EMT ....................... 27

Figure 2.1: Mammosphere assay ............................................................................. 32

Figure 2.2: Mammosphere formation in MCF10A cells ............................................ 35

Figure 2.3: Stable DEAR1-KD in 76N-E6 HMECs results in increased mammosphere number and size in the presence of TGF-β ................................................................. 36

Figure 2.4: Increased primary mammosphere formation in DEAR1-KD HMECs is partially dependent on SMAD3 signaling in the presence of TGF-β ................................................................. 38

Figure 2.5: ALDH1 expression in 2D and 3D mammosphere culture ......................... 41

Figure 2.7: Mammospheres contain bipotent cells ................................................... 45

Figure 2.8: DEAR1-KD HMECs contain a larger population of basal-like cells ............ 46

Figure 2.9: DEAR1 knockdown in DCIS.COM cells results in generation of significantly larger primary mammospheres compared to control vectors ................................. 49

Figure 2.10: DEAR1 knockout in DCIS.COM cells also generates larger primary mammospheres .......................................................................................................................... 51

Figure 2.11: All DCIS.COM mammospheres formed secondary and tertiary mammospheres, but DEAR1-KO clones consistently formed larger mammospheres with each passage ............... 53

Figure 2.12: TGF-β inhibits formation of DCIS mammospheres .................................. 55

Figure 3.1: Loss of DEAR1 in HMECs results in increased basal expression of SNAI2 and ZEB2 mRNA ................................................................................................................. 63
Figure 3.2: *SNAI2*, *ZEB1*, and *ZEB2* basal mRNA levels are significantly higher in *DEAR1*-KD clones and *SNAI2* mRNA expression remains increased in *DEAR-SMAD3* double KD cells...65

Figure 3.3: SNAI2 promoter-activated luciferase expression is inhibited by ectopic DEAR1 expression in HEK293T cells.................................................................68

Figure 3.4: Loss of DEAR1 in HMECs result in increased SNAI2 protein; DEAR1 binds to and polyubiquitinates SNAI2 ...........................................................................69

Figure 3.5: *SNAI2* knockdown in *DEAR1*-KD HMECs did not fully rescue mammosphere phenotype ....................................................................................................71

Figure 4.1: *DEAR1* LOH and *SNAI2* alterations are frequent in multiple cancer types in cBioPortal ........................................................................................................................................78

Figure 4.2: *DEAR1* LOH and mRNA downregulation occurs in 17% of invasive breast carcinoma patients in the METABRIC cohort.................................................................79

Figure 4.3: *SNAI2* amplification and mRNA upregulation occurs in 12% of invasive breast carcinoma patients in the METABRIC cohort........................................................................80

Figure 4.4: Alterations in both *DEAR1* and *SNAI2* serve as a more significant predictor of overall survival in the METABRIC cohort ...........................................................................81

Figure 4.5: *DEAR1* and *SNAI2* mRNA expressions are altered in basal-like and triple negative breast cancer ...........................................................................................................82

Figure 4.7: Loss of DEAR1 protein expression significantly correlates with younger age at diagnosis and shorter time to metastasis in TNBC patients........................................86

Figure 5.1: miR-10b and miR-196b were predicted by at least three prediction algorithms to target *DEAR1* ..............................................................................................................92

Figure 5.2: miR-10b and miR-196b are upregulated in cancer cell lines that lack DEAR1 protein expression..................................................................................................................94
Figure 5.3: Expression of miRNA mimics significantly reduces luciferase activity in HEK293T cells containing putative and standard miRNA binding sites ................................................................. 96

Figure 5.4: Expression of miRNA mimics significantly reduces luciferase activity in HEK293T cells containing the entire DEAR1 3'UTR sequence containing miRNA binding sites ............ 98

Figure 5.5: Transient transfection of miRNA mimics into cell lines that express DEAR1 protein result in reduction of DEAR1 expression ................................................................. 99

Figure 5.6: Mutations in the putative miR-10b binding site in DEAR1 3' UTR results in rescued luciferase activity ........................................................................................................ 100

Figure 7.1: DEAR1 is a regulator of cell polarity, EMT, and stem/progenitor cell properties... 115
LIST OF TABLES

Table 2.1: Loss of DEAR1 in HMECs does not confer self-renewal potential.......................... 43

Table 4.1: Patient demographics and tumor characteristics of TNBC TMA............................. 85
CHAPTER 1: Background and significance
1.1 Ductal Carcinoma In Situ and Invasive Ductal Carcinoma

Breast carcinoma remains the most commonly diagnosed cancer in women in the United States and is the third-leading cause of cancer-related deaths, led by colorectal and lung carcinomas. It is estimated that there will be approximately 266,120 new cases and 40,920 deaths from breast cancer in 2018. The median age of diagnosis is 62 years of age and the average 5-year survival rate is nearly 90%. The majority of patients are diagnosed when the disease is at an early, localized stage in which the 5-year survival rate is quite favorable (nearly 99%); however, if diagnosed after the cancer has metastasized the 5-year survival rate significantly drops to 27% [1]. This highlights the importance of diagnosing and rendering appropriate treatment when the disease is at its earliest stage to reduce morbidity and mortality.

Breast cancer subtypes

Breast carcinomas have been traditionally categorized into subtypes based on expression of hormone receptors for ER (estrogen), PR (progesterone), and growth factor receptor HER2 using immunohistochemistry (IHC) staining of tumor tissues. This standard method of classification has been further elucidated by Perou and colleagues through gene expression profiling to correlate gene expression patterns with hormone receptor status classification and clinical outcome, categorizing tumors into intrinsic subtypes that also corresponds well with the IHC status [2-4]. Tumors that fall into the luminal A subtype demonstrate high expression of ER and/or PR but lack HER2 overexpression. This tumor subtype is typically low-grade, slow-growing, has the most favorable prognosis, and is the most common subtype, accounting for 30-50% of all breast cancer cases. Luminal A tumors are characterized by low expression of proliferation genes, low Ki67, and high expression of CK8/18 luminal cytokeratin, GATA3, and BCL2. [5, 6]. The luminal B subtype is also ER/PR positive, can be either HER2 negative or positive, and has high Ki67 staining. One key difference between luminal A and luminal B tumors is that luminal B tumors tend to exhibit higher expression of proliferation-related genes such as v-MYB, GGH, LAPTMB4, NSEP1, and CCNE1 and lower expression of ER-related genes [6, 7]. Activation of growth-promoting pathways, such as FGFR1
and PI3K, are also implicated in luminal B cancers [8]. HER2 positive tumors are generally ER/PR negative and overexpress HER2, a tyrosine kinase receptor mapped to Chromosome 17q21. HER2 positive tumors often have acquired TP53 mutations, are of higher grade, are highly proliferative, and, in the absence of treatment, have a worse prognosis compared to luminal subtypes [5, 6]. According to the SEER database, approximately 10% of all breast cancer cases diagnosed in 2010 were of the HER2 subtype [9]. The aforementioned subtypes of breast cancer express surface markers that can be utilized for targeted therapies (e.g. trastuzumab/Herceptin against HER2 positive tumors). The triple negative breast cancer (TNBC) subtype is characterized by lack of ER, PR, and HER2 expression, making this disease difficult to apply targeted hormone therapies. TNBC patients have poor prognosis compared to luminal and HER2 positive subtypes. TNBCs are associated with BRCA1 mutations, younger age of diagnosis, high grade tumors, higher risk of metastasis and recurrence, and high proliferative indices [10, 11]. TNBCs can be further categorized into subtypes based on differences in gene expression profiles and biological pathways. One well-known subtypes of TNBC is the basal-like classification that accounts for up to 70% of TNBCs and is characterized by high expression of myoepithelial/basal markers (e.g. cytokeratin 5 & 14), TP53 mutations, RB inactivation, and expression of genes involved in matrix remodeling [5, 6]. It is imperative that we acquire a deep understanding of the biology of this disease in order to provide more effective treatment options.

**Ductal carcinoma in situ**

Various factors can contribute to the risk of breast cancer, including age, family history, genetic alterations (e.g. TP53, HER2, BRCA1/2), and hormone replacement therapy [12]. The two subtypes of breast carcinomas are lobular and ductal carcinomas, the latter constituting approximately 80% of breast carcinomas worldwide [4]. The current linear model of breast cancer progression begins with the evolution of normal mammary epithelial cells to flat epithelial atypia, followed by atypical ductal hyperplasia (ADH) that can progress to ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC). The development of mammography and its utility in clinical screenings since the 1980s has led to an increase in the incidence rate of breast cancer.
annually [13, 14]. Approximately 25% of newly diagnosed cases are DCIS, one of the earliest, pre-invasive forms of IDC that is considered Stage 0 breast cancer [12, 14]. DCIS is defined as a neoplastic proliferation of mammary epithelial cells that remains confined within the ductal-lobular system and the non-obligate precursor of IDC. Although DCIS does not inherently progress to invasive ductal carcinoma, if left untreated up to 60% of low-grade DCIS can progress to invasive disease within 10 years [12, 13, 15]. The current standard treatment for DCIS is surgery followed by radiation therapy. This treatment regimen reduces recurrence risk to approximately 8%, in which about half of these recurrences are invasive and have up to a 20% chance of metastasizing within 8 years [12]. It remains unknown how best to identify which cases are at risk for progression to IDC and which cases will remain indolent, posing the issue of over- or under-treatment of this disease.

There are known risk factors that contribute to recurrence, such as surgical margins, nuclear grade, age of onset, and presence of comedonecrotic lesions [12]. However, the precise mechanisms and pathways responsible for recurrence and progression are still poorly understood. Several studies have demonstrated that while there are few dramatic genetic and/or expression profile differences between synchronous DCIS and IDC [4, 14, 16, 17], there seem to be significant differences between pure DCIS and DCIS with invasive components, as well as between low-grade and high-grade DCIS, the latter of which has a higher risk of local recurrence, metastasis after invasive recurrence, and poor disease-free survival [13, 18]. The fact that pure DCIS and DCIS with invasive components harbor different genetic aberrations and that DCIS associated with invasive components and matched IDC samples demonstrated a significantly greater number of predicted driver mutations compared to pure DCIS, such as \textit{BRCA2}, \textit{FGFR2}, \textit{EPHA1}, \textit{DCLK3}, and \textit{PTPRB}, suggests that progression from DCIS to invasive carcinoma is driven by the selection of cells that have accumulated an advantageous repertoire of genetic alterations [14, 16, 17]. There is also strong evidence that genetic and/or epigenetic alterations in stromal cells that comprise the extracellular matrix (ECM) can have a significant impact on invasion, migration, and metastasis [4, 19-21].
Breast cancer cell of origin theories

It is well understood that DCIS/IDC is a heterogeneous disease, and the complexity of the genetic and/or epigenetic changes that occur in cells that form tumors make it difficult to treat. There are currently two popular hypotheses, not undeniably mutually exclusive, that demonstrate the heterogeneous nature and cell origin of breast carcinogenesis (Figure 1.1). Both hypotheses agree on the concept that tumors are formed from one transformed cell through acquisition of mutations, but they pose different explanations regarding the target cell that undergoes transformation, heterogeneity, and resistance. The first is termed the sporadic clonal evolution model (Figure 1.1A), which postulates that all cells are targets of random mutagenesis. These cells accumulate mutations over time, and advantageous mutations that provide cells with the ability to proliferate and form a tumor are selected for. Heterogeneity, based on this theory, comes from the multitude of various mutations individual cells undergo throughout their lifecycles. Resistance to chemotherapeutic drugs is believed to occur through selection of specific clones that have acquired the necessary mutations for survival [21, 22]. The second hypothesis is termed the cancer stem cell model in which a normal mammary stem cell is transformed into a mammary cancer stem cell that retains its ability to self-renew, creating more cancer stem cells, as well as generates differentiated, cancerous progeny of various subtypes (Figure 1.1B). Based on this second model, heterogeneity is generated through the differentiation of cancer stem and/or progenitor cells into lineage-specific mammary cells and drug resistance is an inherent property of cancer stem cells due to their quiescent nature [21, 23]. Though the two models of breast carcinogenesis differ in respect to the origin of the tumor-initiating cell, they are not necessarily mutually exclusive. It is possible that a non-stem cell is transformed and acquires stem cell-like properties through accumulations of mutations. Conversely, it is possible that a transformed stem cell acquires mutations over time and is selected for during clonal evolution. The fact remains that heterogeneous tumors are difficult to treat due to the varying genetic alterations of each cell in the bulk population. A clear understanding of precise pathways and
Figure 1.1: Cancer cell of origin theory. The cancer cell origin theory postulates that tumors arise from one single cell that has acquired mutations necessary for transformation. A) The clonal evolution model proposes that any normal cell is a target of transforming mutations. Over time, transformed cells acquire selective advantages due to accumulation of various mutations that allow them to proliferate. B) The stem cell model hypothesizes that tumors are analogous to organs with uncontrolled growth, therefore the cell of origin must be similar to stem cells are capable of self-renewal and generating other cell types.
mechanisms that drive DCIS development, as well as progression to IDC, will help determine the most effective forms of therapies for these patients.

1.2 Epithelial-Mesenchymal Transition (EMT) and its role in carcinogenesis

Epithelial cells, which constitute the lining of major organs and generate glandular structures, have the unique ability to form sheets and tubes that are essential to organ function and are characteristically immobile [24]. Maintenance of these structures depends highly on cell-cell adhesion, accomplished through formation of cellular junctions and apical-basal polarity, which is designated through localization of particular polarity protein complexes and is unique to epithelial cells [25-27]. On the other hand, mesenchymal cells are highly mobile and are capable of generating an extracellular matrix to support epithelial structures [24]. In certain circumstances, such as during organ development or wound healing, epithelial cells shed their epithelial phenotype and acquire mesenchymal characteristics. This process is called EMT that begins with loss of cell polarity and adhesion followed by acquisition of migratory and invasive potential.

Apical-basal polarity and cell adhesion

Maintenance of cell polarity is achieved through various protein complexes that dictate apical and basal regions. These polarity protein complexes also associate with protein structures that form junctions between cells to achieve cell-cell adhesion. The PAR complex, consisting of PAR6, PAR3, and aPKC proteins, and Crumbs complex, consisting of CRB, PALS1, and PATJ proteins, segregate to the apical portion of epithelial cells and are associated with tight junctions. The Scribble complex, which is comprised of SCRIB, DLG, and LGL proteins, localize to the basolateral portion and are associated with adherens junctions [25, 27]. Tight junctions and adherens junctions are crucial mediators of cell-cell contact. While tight junctions function to regulate transportation of ions between neighboring cells through association of occludins, claudins, and ZO proteins, adherens junctions initiate and maintain cell-cell adhesion mediated by E-cadherin and catenins [28]. Specific junctions called gap junctions and desmosomes are
localized to the lateral portion of epithelial cells. Desmosomes are responsible for providing strength between adjoining cells through protein complexes that consist of members of the desmocollin and desmoglein family of proteins that link to plakophilins and plakoglobins of the armadillo family. Armadillo proteins are connected to desmoplakin proteins that are directly bound to intermediate filaments within the adjacent epithelial cell [27]. Epithelial cells are also anchored to the basement membrane through hemidesmosomes that are connected to cytokeratins within the cell and bound to transmembrane proteins called integrins that bind to the ECM [27]. Together, these complexes regulate and maintain cell polarity and adhesion that is characteristic of an epithelial phenotype.

*Types of EMT*

The discovery that epithelial cells can convert to a mesenchymal phenotype through a process called EMT was primarily established through observations made in chicken embryos by Elisabeth Hay in the early 1980s [29]. EMT is a highly regulated process that is crucial during embryogenesis and development, and is activated throughout the adult lifespan during wound healing and scar tissue formation. EMT is a mechanism by which immobile epithelial cells lose their apical-basal polarity, detach from neighboring cells and the basement membrane through disruption of the cellular adhesion complexes discussed above, acquire migratory potential, lose epithelial markers such as cytokeratins and E-cadherin, and display a mesenchymal-like phenotype including overexpression of vimentin and N-cadherin. Activation of the EMT process has now been demonstrated as a major driver of invasion and metastasis in epithelial cancers [24-27, 30].

The EMT process can be distinguished into three types based on the specific context in which it occurs. Type I EMT, also known as developmental EMT, is strictly controlled and is first witnessed during gastrulation, a process that results in the generation of the three germ layers in a fertilized egg. Type I EMT is most notably controlled through Wnt signaling, which can be modulated through Nodal and Vg1, members of the transforming growth factor beta (TGF-β) superfamily of proteins [24, 30]. A type II EMT is associated with wound healing and fibrosis, a
phenomena characterized by a prolonged inflammatory response due to injury. It has been observed that during fibrosis, cells can undergo what is termed as a “partial EMT” in which cells express mesenchymal markers but still maintain a relatively epithelial phenotype. This partial EMT has been demonstrated to be regulated by the release of TGF-β, platelet-derived growth factor (PDGF), epidermal growth factor (EGF), and fibroblast growth factor 2 (FGF-2) by macrophages and fibroblasts, which are recruited to sites of injury. Inflammatory cells can also secrete metalloproteases, notably MMP-2, MMP-3, and MMP-9, to facilitate the release of epithelial cells from the basement membrane [30]. Finally, type III EMTs occur in the context of carcinogenesis where transformed, carcinoma cells acquire a mesenchymal phenotype and become highly invasive. The ability to break from neighboring cells and the basement membrane allows cancer cells to initiate the multistep metastatic process by which cells are required to intravasate into the bloodstream, survive circulation, leave the bloodstream (extravasation), generate micrometastases at a distant site, and finally colonize the secondary tumor [30, 31]. It was observed that cells from secondary tumors resemble their primary tumors; the fact that they have reverted back to an epithelial phenotype and no longer express mesenchymal markers suggests that to form micrometastases after extravasation requires migratory mesenchymal cells to convert back into epithelial cells through the reverse process termed mesenchymal-epithelial transition (MET) [24-27, 30, 32]. Despite being classified based on biological function, the EMT process in all three types share similar regulatory elements in terms of activation and signaling pathways that mediate this physical phenomena.

**Key steps for initiation of the EMT program**

It is agreed upon that initiation of EMT consists of a set of crucial events that includes: 1) loss of adhesion and cell polarity, 2) changes to the cytoskeleton that mediate an enhanced motility phenotype, and 3) differential gene expression as cells shift from epithelial to mesenchymal phenotype [24-27, 30].

Loss of cell-cell contacts can be achieved through transcriptional regulation, post-transcriptional regulation (e.g. ubiquitination and proteasomal degradation), or through inhibition
of maintenance pathways affecting components of adhesion junctions [25, 27]. For example, loss of E-cadherin gene expression, through copy number alterations, mutations, or promoter methylation, or phosphorylation of β-catenin, which binds directly to E-cadherin at adherens junctions, through activation of Wnt signaling leads to destabilization and decreased cell adhesion [27]. Loss of E-cadherin is observed frequently in breast carcinomas, is associated with high grade tumors and invasive behavior, and confers poor prognosis [33, 34]. Loss of cell adhesion also confers loss of cell polarity as polarity complexes also undergo dissolution [25].

Rearrangement of the cytoskeleton is required for cells to elongate, thereby acquiring a spindle-shaped morphology that is characteristic of mesenchymal cells, and present directional polarity. Cytoskeletal reorganization is regulated by RHO GTPases, which modulate actin filaments, and RAC1 and CDC42, which are involved in lamellipodia and filopodia formation and recruited to the leading edge of the migrating cell through PI3K signaling [25, 27].

EMT is initiated by extracellular signals and processed through activation of membrane receptors that in turn lead to activation of intracellular signaling pathways that eventually cause changes in gene expression to achieve the physical changes discussed above [27]. Changes in gene expression to alter cellular adhesion include repression of junction proteins, such as cytokeratins, cadherins, claudins, occludins, and integrins. Activation of the EMT program also results in upregulation of mesenchymal markers, such as N-cadherin, NCAM, vimentin, and fibronectin, as well as metalloproteases to facilitate ECM breakdown [25-27].

**TGF-β is a potent inhibitor of cell proliferation and inducer of EMT**

There are several known inducers of the EMT program, including growth factors EGF, hepatocyte growth factor (HGF), insulin growth factor 1 (IGF-1), and FGF. The most well characterized and studied is the TGF-β superfamily of proteins, which includes TGF-β1, TGF-β2, and several bone morphogenetic protein (BMP) cytokines [25]. De Larco and Todaro first publicized the discovery of TGF-β in 1978 [35]. Shortly after, its role in inhibiting proliferation in BSC-1 monkey kidney cells was established by Holley and colleagues [36]. In the 1980s, Moses and colleagues demonstrated that TGF-β could inhibit proliferation of several cell types [37] and
Silberstein and Daniel were the first to show that TGF-β can inhibit mammary gland development in mice and that this inhibition was reversible with removal of TGF-β ligand [38]. TGF-β has been shown to induce expression of cell cycle inhibitors, namely p15\(^{NK14B}\), an inhibitor of CDK4 and CDK6, and c-Myc [39-41]. Despite its obvious role in inhibiting cell proliferation \textit{in vivo} and \textit{in vitro}, it was observed that TGF-β, most notably TGF-β1, is upregulated in multiple cancers [39, 42]. It became clear that TGF-β plays dual roles depending on cellular contexts [39, 43-45]. In a study by Basolo et al, TGF-β inhibition affected proliferation of untransformed or spontaneously immortalized mammary epithelial cells but not cell lines that were transformed by the SV40-large-T antigen, suggesting that inactivation of the tumor suppressors Rb and/or p53 can subvert the anti-proliferative effect of TGF-β signaling [46]. Further studies provided evidence that TGF-β signaling inhibits growth in the context of normal cells and at early stages of tumorigenesis, but due to the accumulation of genetic alterations and genomic instability during carcinogenesis, TGF-β signaling becomes tumor promoting at later stages of tumorigenesis and its role in initiating EMT becomes an important mediator of metastasis [39].

TGF-β ligands bind to combinations of type I and type II TGF-β receptors, which are transmembrane serine/threonine kinases [39]. Currently there are seven and five known type I and type II receptors, respectively, and 29 ligands in mammals. Signaling is initiated when a ligand binds and activates a type II receptor kinase of the heteromeric receptor complex, resulting in phosphorylation and activation of a type I receptor. Different ligands have varying degrees of affinity for different receptor combinations, resulting in activation of specific effector proteins in order to mediate several signaling pathways. For example, TGF-β1 and TGF-β3 ligands primarily bind to type II homodimer receptor complexes whereas TGF-β2 binding requires a type I/II heteromeric complex. Type II TβRII activation of type I TβRI/ALK5 activates SMAD2/3 effector proteins whereas activation of type I ALK1 receptor activates SMAD1/5 signaling [45]. The SMAD pathway is discussed in further detail below.
**Canonical TGF-β signaling through SMADs**

The most well-known and understood TGF-β effector pathway occurs through the activation of SMAD proteins, also referred to as the TGF-β canonical pathway. There are eight known SMAD proteins. Receptor-activated SMADs, or R-SMADs, include SMAD1, 2, 3, 5, and 8. SMAD2 and SMAD3 are the main effector proteins of the canonical TGF-β signaling pathway and are activated through phosphorylation of its C-terminal domain by TβRI and activin receptor, ActRIB. SMAD4 can form a heterotrimeric complex with SMAD2 and/or SMAD3 to facilitate translocation into the nucleus to activate the EMT transcriptional program. Translocation of SMAD2 and SMAD3 is mediated through a nuclear localization sequence recognized by importin-β. Upon translocation into the nucleus, SMAD2 and SMAD3 interactions with the CBP/p300 transcriptional coactivator and specific transcription factors, such as c-Jun, STAT3, and LEF/TCF, in addition to their DNA-binding capabilities, allow for specific binding to promoter regions and transcriptional activation of EMT-related genes [39, 45]. The majority of TGF-β-mediated transcriptional activation of EMT is mediated through SMAD3-SMAD4 activity. Upon ligand binding, TGF-β receptors are active for up to 4 hours; the longer the receptors remain active, the longer R-SMADs remain in the nucleus to regulate gene transcription, including activation of key EMT transcription factors such as SNAI, TWIST, and ZEB proteins [45]. SMADs also activate expression of mesenchymal genes, including fibronectin, vimentin, and collagens [25].

Regulation of SMAD proteins can be accomplished in several ways. SMAD6 and SMAD7 are inhibitory SMADs that repress R-SMAD signaling by competitively binding to the intracellular domain of type I receptors to prevent R-SMAD phosphorylation [25, 39]. Inhibitory SMADs are activated by TGF-β signaling, generating a negative feedback loop. SMAD7 can also recruit SMAD ubiquitylation regulatory factor 1/2 (Smurf1/2), a member of the HECT family of E3 ubiquitin ligases, to type I receptors to induce receptor ubiquitination and proteasomal degradation. Phosphorylated SMAD3 has been shown to interact with the SCF ubiquitin ligase complex that facilitates SMAD3 nuclear exportation and ubiquitin-mediated degradation [45].
Non-SMAD signaling and other inducers of EMT

In addition to SMAD-mediated signaling, TGF-β has been shown to activate RHO GTPases, which mediate cytoskeletal changes during EMT. TGF-β also activates the PI3K/AKT pathway to induce mTORC1 and mTORC2 proteins that are involved in motility and acquisition of mesenchymal traits, respectively. Additionally, TGF-β signaling also contributes to activation of MAPK pathways mediated by ERK, JNK, and p38 [25, 26, 39].

Aside from TGF-β signaling, there are several other pathways that can mediate induction of EMT, including the PI3K/AKT and MAPK pathways. For example, RAS signaling has been shown to induce expression of SNAI1/2 transcription factors and RHO GTPases [25]. FGF and HGF signaling has also been demonstrated to initiate EMT through expression of SNAI proteins [47]. Furthermore, other signaling pathways that have been implicated in regulating self-renewal, stem cells, and play crucial roles during embryogenesis and development, such as Wnt, Notch, and Hedgehog signaling, also induce an EMT [25, 26, 39, 48]. Wnt signaling inactivates the destruction complex and inhibits GSK3β-mediated phosphorylation of β-catenin, which frees it from its interaction with E-cadherin and α-catenin resulting in reduction of cell-cell adhesion [27]. Activation of Notch signaling through increased Jagged1 expression has been reported to induce EMT through increased expression of the EMT transcription factor and master regulator, SNAI1, and repression of E-cadherin expression [48].

There is also evidence of crosstalk between canonical TGF-β-SMAD signaling and other pathways that mediate EMT. Activation of JNK by TGF-β can enhance SMAD3 nuclear transportation and can lead to c-Jun-mediated cooperation with SMAD3 and SMAD4 complexes in the nucleus and transcriptional activation [49]. RAS/MAPK signaling, although induces production of TGF-β1, can phosphorylate SMAD3 to inhibit its translocation into the nucleus [44, 50]. As mentioned briefly, SMAD proteins can interact with the LEF/TCF transcription factors that are stimulated in the context of Wnt signaling [51]. Activation of ERK can also phosphorylate and activate SMAD2 for nuclear translocation [45].
Transcription factors driving the EMT program

Several transcription factors are expressed in response to initiation of the EMT program, most notably SNAI1/2, TWIST1/2, ZEB1/2, and FOXC2. These EMT-related transcription factors regulate the expression of target genes that are involved in carrying out the transition process, including repression of epithelial markers, polarity genes, and cell adhesion genes and upregulation of mesenchymal markers and metalloproteases. Some transcription factors are also involved in regulating the expression of each other. For example, ZEB1 gene is a direct target of SNAI1 [25], and both transcription factors repress E-cadherin expression.

SNAI1 and SNAI2 have been shown to be highly active during all three types of EMTs and can be activated through several EMT-inducing pathways, including TGF-β, Wnt, and Notch. Both SNAI proteins are transcriptional repressors and contain a zinc-finger c-terminal domain that binds to E-box sequences in promoter regions of target genes. SNAI1/2 can also modulate gene expression through recruitment of histone modification complexes [25, 27]. SNAI proteins are regulated via post-translational modifications that either affect their localization or degradation. GSK3β, in the absence of Wnt signaling, can phosphorylate SNAI1 to mediate its nuclear export and proteasomal degradation [25]. The tumor suppressor p53 has been demonstrated to interact with SNAI2 and MDM2, resulting in ubiquitin-mediated degradation of SNAI2 [52]. ZEB1 and ZEB2 also bind to E-box DNA sequences to mediate gene repression that often involves a co-repressor called c-terminal binding protein (CTBP), and can be activated through TGF-β, Wnt, and RAS signaling. Both SNAI and ZEB proteins are known for their role in repressing E-cadherin expression. ZEB1 can also function as a transcriptional activator when it is in a complex with the SMAD transcriptional coactivator CBP/p300 [25]. Members of the microRNA family miR-200 regulate ZEB expression post-transcriptionally; miR-200a, b, and c and miR-205 directly target ZEB1/2 mRNA [53]. TWIST1/2 are EMT-regulators but also mediate cell differentiation [47] and are activated under hypoxic conditions through HIF1α. Similar to SNAI proteins, TWIST proteins are regulated through post-translational modifications [25].
**EMT and carcinogenesis**

TGF-β and EMT-related genes have been found to be upregulated in epithelial cancers, as well as downregulation of E-cadherin, and correlated with increased tumor grade, prognosis, and metastasis [33, 34, 54, 55]. Immunohistochemistry revealed high TWIST and SNAI2 expression in breast carcinomas. Increased SNAI1 expression in particular correlated with increased tumor grade, metastasis, recurrence, and poor outcome [54]. A particular subset of triple negative breast cancers is associated with a mesenchymal phenotype and express EMT-related proteins, including vimentin, fibronectin and EMT transcription factors [56]. It was discovered by Mani and colleagues that induction of the EMT program through overexpression of EMT transcription factors or exposure to TGF-β generated mammary epithelial cells that have acquired stem cell properties, including enhanced self-renewal and differentiation capacities, expression of the CD44+/CD24- phenotype, and formation of *in vitro* mammospheres, a mass of mammary cells that consists of stem, progenitor, and differentiated cells through an assay that enriches for stem/progenitor cells in anchorage-independent conditions [57, 58]. This discovery brings to light the question of whether EMT in the context of cancer can generate cancer stem cells, thereby contributing to tumor heterogeneity and chemotherapy resistance.

### 1.3 Stem cells and cancer

The concept of a cancer cell of origin began in the 1930s by Furth and Kahn, who showed that a single murine tumor cell could establish formation of a new tumor in a different mouse [59]. The stochastic evolution model was developed in the early 1970s and proposed that tumors arise from a single cell that has acquired an array of genetic and molecular alterations that allow it to proliferate uncontrollably and ignore growth inhibition signals. The alternative theory of cancer stem cells as the cell of origin has been established for nearly five decades, beginning with studies that demonstrated how teratomas, tumors developed from germ cells, generate
differentiated cells of multiple cell types that follow a similar lineage path as a normal developing germ cell [60].

Embryonic stem cells are undifferentiated cells that are capable of generating the three germ layers and all cell types in the developing organism. Stem cells are capable of both symmetric and asymmetric division; that is, they can self-renew and generate an identical stem cell progeny or they can generate a daughter cell that exhibits a more lineage-restricted and mature phenotype (multipotent progenitor, lineage-restricted progenitor, or differentiated cell) [61]. In adulthood, there exist populations of stem cells that are capable of tissue regeneration in the case of cell turnover or tissue damage. Stem cell populations have been identified and isolated from intestinal crypts, bulge region of hair follicles, and stem cell niches of the skin and mammary gland [62]. Stem cells isolated from the breast epithelium are able to generate a fully functional mammary ductal system when transplanted into a cleared mammary fat-pad in mice. Self-renewing stem cells are quiescent and rare. About 0.01% of bone marrow cells are believed to be from a hematopoietic stem cell lineage [63] and may survive for extensive periods of time compared to their more differentiated progeny. This longevity provides more possibilities of incurring transforming mutations [64].

The theory of cancer stem cells as the target cell for carcinogenic transformation was pioneered in the context of acute myeloid leukemia. Dick and colleagues demonstrated that there is a rare population of CD34+/CD38− cells within the transformed hematopoietic cell hierarchy with the ability to form tumors upon transplantation into immunocompromised NOD/SCID mice. These rare cancer cells were able to generate differentiated progeny and self-renew, similar to hematopoietic stem cells [65]. These carcinogenic stem-like cells are termed cancer stem cells (CSCs) and have been isolated in solid tumors as well, including breast carcinomas [66]. Similar to normal stem cells, CSCs have the ability to self-renew and give rise to differentiated progeny [64, 67-69]. Progenitor cells, though unable to self-renew, could gain self-renewal capacity through mutations and become a target for neoplastic transformation (Figure 1.2) [64].
Figure 1.2: Cancer and normal stem cell characteristics. Normal stem cells can become cancer stem cells through acquisition of mutations that would inhibit of apoptotic signals and allow for uncontrolled growth. CSCs, like their normal counterparts, can self-renew and asymmetrically differentiate to generate progenitor cells and mature differentiated cells that form the bulk of the tumor. Several studies have now shown that mutations can occur that push mature cells to become dedifferentiated, including initiation of the EMT process.
The mammary ductal system is maintained by stem cells that are believed to reside in and originate from the myoepithelial/basal cell population adjacent to the basement membrane and can generate luminal or myoepithelial/basal cells that form branching ducts within the mammary gland and end in terminal ductal-lobular units [70-72]. The interior of the ducts is composed of a single layer of luminal epithelial cells encompassed by myoepithelial cells [71, 73]. Many studies aimed to identify the putative stem and progenitor cell populations within the mammary gland architecture. Several marker combinations have been used to classify distinct populations of cells, including CD49f (α-6-integrin), which is expressed in cells located in the basal region; epithelial cell adhesion molecule (EpCAM) that is expressed by cells in the luminal compartment as well as cells in the underlying myoepithelial/basal region; mucin 1 (MUC1), a luminal cell surface marker; and luminal- and basal-specific cytokeratins such as CK8/18 and CK14, respectively. Putative stem and progenitor cell populations can also be identified based on cells’ ability to exclude Hoechst 33342, a fluorescent dye, through overexpression of ABC drug transporters [71, 74, 75]. Al-Hajj and colleagues demonstrated that breast cancer stem cells could be identified by a CD44+/CD24- surface adhesion marker phenotype sorted from Lineage-negative (Lin-) breast tumor cells; these cells, when injected at low density, formed tumors with as few as 1,000 cells whereas CD44+/CD24+ cells were not able to form any tumors [66]. Furthermore, CD44+/CD24- cells proliferated to produce a heterogeneous population and generated CD44+/CD24-, CD44-/CD24+, and CD44+/CD24+ cell phenotypes [76]. CD44+/CD24- cells have been shown to associate with poor overall survival and disease-free survival [77] and are detected in basal-like and HER2 overexpressing breast tumors [78]. Another widely used stem and progenitor cell marker is aldehyde dehydrogenase 1 (ALDH1), a retinol-oxidizing enzyme that is expressed at high levels in both normal and breast cancer stem cells. Injection of as few as 500 ALDH1+ invasive ductal carcinoma cells into cleared mouse mammary fat-pads generated tumors [79]. Interestingly, there is little overlap between ALDH1+ and CD44+/CD24- cells, but cells that are ALDH1+/CD44+/CD24- are highly metastatic and aggressive compared to ALDH1+ or CD44+/CD24- alone. Although each phenotype associates
with metastasis in vitro and in vivo, CD44+/24- cells correlated with proliferative potential whereas ALDH1+ correlated with migration and invasion [80].

Studies have now identified several signaling pathways deregulated in cancers that are also associated with regulating stem cell self-renewal, such as Wnt, Notch, Hippo, Hedgehog, and TGF-β signaling pathways [15, 64, 81-83]. In 2008, Mani et al. established a link between induction of EMT and stem cell properties. It was demonstrated that immortalized human mammary epithelial cells (HMLEs) forced to undergo EMT either through the overexpression of master regulators of EMT or through exposure to TGF-β was sufficient to induce stem cell phenotypes such as enhanced mammosphere formation, in vitro differentiation, and increased CD44+/CD24- cell population. It was also observed that the existing population of stem cells in HMLEs overexpressed mesenchymal markers, including N-cadherin and vimentin, and EMT transcription factors, including SNAI1, ZEB2, and TWIST1 [57]. Other studies of EMT-related transcription factors have confirmed a relationship between overexpression of master regulators of EMT and a stem cell phenotype. Overexpression of TWIST in normal human mammary epithelial cell line MCF10A and luminal breast cancer cell line MCF-7 generated breast epithelial cells that could exclude Hoechst dye and increased the number of CD44+/CD24- cells through the ability of TWIST to mediate CD24 transcriptional activity [84]. Interestingly, studies related to TWIST and cancer stem cell properties revealed that TWIST-mediated stemness can be independent of EMT [85]. Other EMT regulators, such as SNAI2, were also shown to play a large role in dedifferentiation in vitro and in vivo [86-88]. Thus, transcription factors that mediate EMT through global reprogramming of gene expression could allow for the acquisition stem cell properties.

1.4 TRIM proteins and their roles in cancer

Tripartite motif (TRIM) proteins constitute a large family of proteins that contain a RING finger, B-box, and coiled-coil domains, thus are also called RBCC proteins, and are involved in
mediating protein-protein interactions and formation of large protein complexes. TRIM/RBCC proteins have been demonstrated to play roles in a variety of cellular processes that include cell growth, differentiation, cell death, gene expression, protein turnover, and viral immunity. There are over 100 known TRIM proteins in mammals, which are categorized into subfamilies (C-I through C-XI) based on their structural organization [89]. While there exists low sequence homology between TRIM family members, the organization of the RBCC domains in the N-terminus are conserved: the RING finger domain begins 10-20 amino acid residues from the first methionine and is followed by one to two B-box domains and a coiled-coil domain. The C-terminus is much more variable among TRIM proteins, though a few family members harbor PRY/SPRY domains that are hypothesized to play roles in autoimmunity and viral restriction [90-93].

The RING (Really Interesting New Gene) finger domain is formed by a sequence of cysteine and histidine residues, binds zinc atoms, and is found on many E3 ubiquitin ligases, such as BRCA1 and MDM2. E3 ubiquitin ligases are enzymes involved in the regulation of protein activity and localization through their ability to ubiquitinate target proteins. E3 ligases mediate the transfer of ubiquitin proteins from an E2 ubiquitin conjugating enzyme to target substrates, specifically on the substrates’ lysine (Lys) residue. The exact location of the lysine residue subjected to ubiquitin conjugation, as well as the number of ubiquitin moieties, play a crucial role in determining protein fate: polyubiquitination of Lys-48 signals for 26S proteosomal degradation whereas monoubiquitination or ubiquitination at Lys-63 controls protein localization and activation/deactivation [89, 90, 94]. In addition to ubiquitin, E3 ligases have been demonstrated to conjugate other ubiquitin-like moieties, including SUMO, Nedd8, and ISG15, to target substrates to regulate their functions. This conjugation ability through the RING finger domain allows TRIM proteins to regulate protein turnover, localization, activity, and gene transcription through activation of histone modifiers and transcription factors [90, 95].

Following the RING finger domain is the B-box domain that also consists of a series of cysteine and histidine residues and, similar to the RING finger domain, also binds zinc atoms. All
TRIM proteins contain one to two B-box domains (labeled B-box1 and B-box2), and almost always possess a B-box2 domain which is hypothesized to function in ubiquitination and binding of DNA, RNA, and proteins. Following the B-box domain is always the coiled-coil domain that mediates protein-protein interactions [90]. In addition to binding to target substrates for ubiquitination, TRIM E3 ligases can ubiquitinate themselves or other TRIM proteins as a method of self-regulation [89, 90].

There are several notable TRIM E3 ubiquitin ligases that play roles in carcinogenesis through their involvement in the regulation of gene expression, cell proliferation and growth, and apoptosis. Though the majority of TRIM proteins involved in tumorigenesis display oncogenic properties, it was observed that their specific roles are context dependent. TRIM19, also known as PML, is a tumor suppressor protein that generates PML nuclear bodies, a site at which transcription, DNA damage repair, and replication occurs, and has been demonstrated to bind to and inhibit MDM2 activity to stabilize p53 in the nucleus [89, 96]. However, the PML-RARα (Retinoic Acid Receptor Alpha) fusion protein, observed in the majority of acute promyelocytic leukemia patients, functions as a dominant negative protein that inhibits wild-type PML function and localization [89]. More recently, in addition to its role in carcinogenesis, PML has been identified as a regulator of pluripotency and stemness mediated through TBX3 and TGF-β signaling and binds stem cell regulators including OCT4 and c-Myc [97, 98]. An earlier study also identified PML as a regulator of luminal progenitor cell populations mediated through STAT3 and STAT6 [99].

Members of the TIF1 family, such as TRIM24 and TRIM28, are overexpressed in breast and gastric cancers, respectively, and are associated with poor prognosis. TRIM24 harbors a PHD finger domain and bromodomain that mediate interaction with chromatin remodeling nuclear proteins and, similar to TRIM19, can interact with RARα and often undergoes fusion with other proteins. The TRIM24-FGFR1 fusion protein contains TRIM24’s RBCC domain and FGFR1’s tyrosine kinase domain, allowing for constitutive activation of FGFR1 that promotes transformation. However, TRIM24 seems to function as a tumor suppressor specifically in hepatic
cancer, as homozygous deletion of Trim24 in mice resulted in increased incidence of tumors in the liver [89].

TRIM27, also known as RFP (RET Finger Protein), is another TRIM protein that seems to have both oncogenic and tumor suppressive functions. While studies have shown that TRIM27 potentially interacts with the PML-RARα fusion protein, is overexpressed in multiple cancer cells, and correlates with ERBB2 expression in breast carcinomas, it was also demonstrated to promote apoptosis. TRIM32, a regulator of miRNA biogenesis and neural differentiation, also demonstrates polar roles in carcinogenesis. Its tumor suppressive function is suggested through its mediation of TNF (tumor necrosis factor) –induced apoptosis. However, several reports of its overexpression in squamous cell carcinomas suggest that it could act as an oncogene [89].

Several TRIM proteins have also been shown to mediate the activity and stability of well-known tumor suppressor, p53. TRIM19 and TRIM13 function to stabilize p53 through the inhibition of its main regulator, MDM2 (Mouse Double Minute 2). TRIM24, TRIM27, TRIM28, and TRIM29 negatively regulates p53 through various mechanisms; TRIM24 binds to an ubiquitinates p53, TRIM27 and TRIM28 stabilize MDM2 to promote p53 ubiquitination, and TRIM29 mediates p53 inhibition through its transport out of the nucleus [89].

It is quite clear that TRIM E3 ubiquitin ligases play a major role in carcinogenesis, largely through their abilities of controlling the activity, stability, and expression of other tumor suppressor and tumor-promoting proteins. In the next section, I will report on another TRIM protein, TRIM62 (also known as DEAR1 – Ductal Epithelium-Associated RING Chromosome 1), which has been shown to be a major regulator of acinar-morphogenesis and cell polarity, and negative regulator of TGF-β-mediated EMT.

1.5 **Ductal Epithelium-Associated RING Chromosome 1 (DEAR1)**

*Ductal Epithelium-Associated RING Chromosome 1 (DEAR1)* is a tumor suppressor gene discovered in the Killary Lab and identified through suppression subtractive hybridization that
Figure 1.3: Illustration of DEAR1 gene, protein, and mutation residues. The DEAR1 gene encodes a protein product that contains five domains and has sequence similarities to the family of TRIM E3 ubiquitin ligases. Sequence analysis revealed that DEAR1 undergoes an array of mutations across all five domains, the majority of which are non-synonymous mutations, in various cancer types. Permission was given to use this figure from Nanyue Chen et al, “DEAR1 is a Chromosome 1p35 Tumor Suppressor and Master Regulator of TGF-beta-Driven Epithelial-Mesenchymal Transition”. Cancer Discovery 2013:1172-1189. License number: 4367830052032.
resulted in a partial cDNA sequence that mapped to Chromosome 1p35.1, a region that resides in the 1p interval that frequently undergoes loss of heterozygosity (LOH) in epithelial cancers, including breast carcinoma [4]. Sequencing revealed that DEAR1 shares sequence similarities to the TRIM/RBCC protein family, thus DEAR1 is also annotated as TRIM62. The DEAR1 gene encodes a 475 amino acid protein predicted to be 54kDa in size and harbors 5 domains: RING finger, B-box2, coiled-coil, PRY, and SPRY domains (Figure 1.3) [100]. Studies in our lab and by Huang et al. demonstrates that DEAR1 has E3 ubiquitin ligase activity and localizes in the cytoplasm; it has also been shown that DEAR1 has the ability to self-polyubiquitinate in HEK293T cells and that the ligase activity is dependent on the RING domain [101]. DEAR1 is expressed in normal glandular epithelium of tissues such as the bladder, kidney, breast, and prostate. In a cohort of 14 DCIS samples with associated normal and invasive ductal carcinoma samples, DEAR1 protein was highly expressed in the normal mammary epithelium but its expression is downregulated in 71% of cases; 50% of the 10 cases with associated IDC also showed loss of expression of DEAR1. DEAR1 expression was demonstrated to be downregulated in 75% of breast cancer cell lines assayed, including two lines of the 21T series derived from a 36 year-old female with invasive ductal adenocarcinoma. Sequencing of three cell lines derived from the 21T series revealed that all three harbored a non-conservative missense mutation, R187W, which maps to the third exon that encodes the coiled-coil domain that is predicted to mediate protein-protein interactions [100].

Functional assays were performed using 3D matrigel culture to assess the importance and relevance of the R187W mutation, as well as expression of DEAR1, in acini formation and pathogenesis. The 21MT cell line that harbors the R187W mutation formed large, disorganized, multi-acinar structures when plated in 3D matrigel. By expressing the wildtype (WT) DEAR1 protein, the mutant phenotype was rescued and the 21MT cells were able to form proper acini similar to those formed by normal immortalized human mammary epithelial cells (HMECs). Staining of the acini revealed that the acinar structures formed from 21MT cells have disorganized expression of E-cadherin and α-6-integrin, which is a marker of proper polarity.
signal. Ki67 staining suggests that DEAR1 does not influence proliferation but rather apoptosis as shown by active Caspase 3 staining. These experiments were also carried out in MCF-7 breast cancer cells, which do not express detectable DEAR1 protein, and results were similar to what was demonstrated in 21MT cells. To further demonstrate that DEAR1 expression is important for the regulation of proper acini formation, lentiviral shRNA knockdown (KD) of DEAR1 was performed on immortalized HMECs. Results showed that when DEAR1 expression is lost, cells were unable to form normal acini in matrigel following 16 days of culture in contrast to control HMECs, which formed proper acini. These irregular acini showed loss of apical-basal polarity as demonstrated by improper expression of α-6-integrin [100].

These initial studies that aimed to characterize DEAR1 and ascertain its function revealed that DEAR1 is a crucial regulator of apical-basal polarity and its expression is necessary for proper acinar morphogenesis. Data from the DEAR1 knockout mouse model suggests that DEAR1 is a bona fide tumor suppressor. DEAR1 hetero- and homozygous knockout (KO) mice formed late onset tumors in 12.9% and 17.7% of cases, respectively, whereas only 4% of DEAR1 WT mice developed tumors. DEAR1-KO mice developed adenocarcinomas in multiple organs including mammary, pancreatic, lung, and liver, as well as sarcomas and lymphomas. Interestingly, DEAR1 heterozygous mice developed tumors with a similar frequency as DEAR1 homozygous knockout mice, suggesting that DEAR1 might function as a haploinsufficient tumor suppressor [102].

Since loss of cell polarity is strongly associated with early steps in initiating EMT, it was postulated that DEAR1 may also play a role in mediating the EMT process. Western blot analysis indicated that loss of DEAR1 in the presence of TGF-β resulted in increased mesenchymal phenotypes, as indicated by increased expression of Vimentin and N-cadherin. DEAR1-KD clones in the presence of TGF-β exhibited increased cell motility as assayed through migration assays in matrigel. Scratch wound assays confirmed that loss of DEAR1 after treatment with TGF-β resulted in rapid wound closure compared to untreated WT and KD cells. Through co-IP experiments, it was demonstrated that DEAR1 binds to SMAD3, a crucial co-effector protein in
the canonical TGF-β pathway, and results in polyubiquitination of SMAD3. DEAR1 and SMAD3 expressions are inversely correlated in the DEAR1 knockout mouse model, as well as in human breast tumors. 

In vitro assays revealed that DEAR1 KD resulted in increased phosphorylated SMAD3 in the presence of TGF-β and SMAD3 levels are reduced in DEAR1 overexpressing MCF-7 cells. These experiments show that in addition to its role in regulating cell polarity and acinar morphogenesis, DEAR1 is also a negative regulator of TGF-β-mediated EMT through the inhibition of SMAD3 (Figure 1.4) [102, 103].

In addition to the R187W mutation, DEAR1 was shown to harbor mutations in a variety of carcinomas that are associated with Chromosome 1p LOH, such as lung squamous cell, pancreatic, and renal carcinomas, and these mutations can be found across all five domains (Figure 1.3) [102]. Deep sequencing revealed that DEAR is mutated in 71% of pure DCIS and DCIS with microinvasive lesions. Approximately 8% of the variants in pure DCIS were found within exonic regions, the majority of which were not reported in variant databases [Reuther, in preparation]. The D106V variant was shown to enhance the transcriptional activity of SMAD3 compared to wildtype DEAR. Functional assays utilizing SKRB3 breast cancer cell line, which does not express detectable levels of DEAR1 protein, were performed to study the effects of the variants. Results indicated that while expression of wildtype DEAR1 allowed for proper acini formation, cells expressing the R254Q variant were unable to form acini, similar to that of the R187W mutation previously discussed [Reuther, in preparation].

Since DEAR1 is mapped to a region that is frequently observed to undergo LOH in epithelial cancers and copy number alterations (CNA) in breast, lung, pancreatic, and colorectal cancers, its expression status may be of value for determining prognosis in cancer patients. DEAR1 expression was screened in a cohort of 158 stage I/II breast cancer patients who have had breast conservation surgery followed by radiotherapy between the ages of 25-49 years. Approximately 56% of these samples demonstrated loss of DEAR1 expression. Although DEAR1 loss did not correlate with ER or HER2 status, tumor size, metastasis, or BRCA1/2 mutations, loss of DEAR1 expression correlated with family history of breast cancer, PR-negative tumors,
Figure 1.4: Illustration of DEAR1 regulation of cell polarity and EMT. DEAR1 functions as a master regulator of acinar morphogenesis through regulation of apical-basal cell polarity, an important aspect of maintaining cellular architecture. DEAR1 also negatively regulates EMT through the canonical TGF-β signaling pathway by polyubiquitinating effector signaling protein, SMAD3, and signaling for proteasomal-mediated degradation, thus limiting the availability of SMAD3 to translocation into the nucleus to transcriptionally activate EMT-related genes. In the absence of DEAR1, through loss of function mutations or loss of expression, SMAD3 is phosphorylated and activated in the presence of TGF-β and an EMT occurs.
and the triple negative phenotype. DEAR1 loss of expression also significantly predicted local recurrence. At a 5-year follow-up, DEAR1 expression correlated with a 95% local recurrence-free survival, which remained consistent over the course of a 15-year follow-up. However, tumors that had loss of expression demonstrated a drop in recurrence-free survival to 58% after 15 years of follow-up [100].

Other studies in different cell types confirm what the Killary Lab has reported on DEAR1 and its role in the initiation and progression of cancer. In a study by Quintás-Cardama et al., DEAR1 was shown to be expressed in immortalized normal bronchial epithelial cells but loses expression in 66% of non-small cell lung cancer (NSCLC) cell lines tested and 86% of NSCLC tumors (n=214). A separate cohort of 72 NSCLC patients demonstrated DEAR1 loss of expression in 62% of tumor samples. Patients with low DEAR1 staining relapsed at a significantly faster rate compared to patients with high DEAR1 staining (5.1 v 2.87 years, p=0.049). Using a DEAR1-deficient K-RasLA1 mouse model, this study showed that compound mutant mice experienced reduced lifespans and a higher number of lung tumors compared to that of K-Ras and p53+/−-K-Ras mutant mice and identified DEAR1 as a haploinsufficient gene that synergizes with K-RasG12D mutation to promote invasion and metastasis in the context of lung adenocarcinomas [104]. Another study in acute myeloid leukemia showed that loss of DEAR1 expression correlated with poor overall survival and shorter disease-free survival in acute myeloid leukemia compared to healthy controls [105, 106].

A study on cervical cancer by Liu and colleagues demonstrated that DEAR1 expression is reduced in early stage cervical cancer samples, correlated with tumor size, recurrence, differentiation grade, and is downregulated in cervical cancer cell lines. Experiments using SiHa and HeLa, cervical cancer cell lines that do not express detectable levels of DEAR1 protein, revealed that overexpression of DEAR1 resulted in inhibition of proliferation, colony formation, migration, invasion, tumor growth, and pulmonary metastases. The authors identified a negative correlation between DEAR1 expression and MAPK/JNK signaling and demonstrated that overexpression of DEAR1 results in repression of c-Jun and its downstream transcriptional
targets such as SNAI2 and Cyclin D1. This study identified another pathway involved in tumor initiation and progression potentially regulated by DEAR1 in the context of cervical cancer [107].

In addition to its role in cancer, DEAR1 was shown to mediate immunity through induction of NF-κB signaling that requires the RING, B-box and coiled-coil domains [108, 109]. A study by Cao et al. identified CARD9, a mediator of cytokine production and T-cell response, as a target of DEAR1 ubiquitination that is dependent on DEAR1’s RING domain. This study discovered that ubiquitination of CARD9 at the Lys125 residue is crucial for CARD9 activity and that abrogation of CARD9-DEAR1 interaction through loss of DEAR1 or CARD9 small molecule inhibitor, results in repression of CARD9 signaling, inhibition of cytokine production and signaling, and susceptibility to infection [108, 110].

The data collected thus far indicate that DEAR1 is a bona fide Chromosome 1p tumor suppressor that undergoes mutation, copy number alterations, and loss of expression in a variety of epithelial cancers, including breast cancer. Evidence indicates that DEAR1 plays an important role in maintaining cell polarity, regulating acinar morphogenesis, and negatively regulating the TGF-β-mediated EMT program and may play a critical role in inhibiting tumor initiation and progression. Since loss of DEAR1 releases the inhibitory effect on TGF-β-mediated EMT, which has now been shown to confer stem cell properties in both normal and cancer cells, I hypothesize that DEAR1 may also regulate stem and/or progenitor cell properties. Results herein demonstrate that DEAR1 regulates progenitor cell phenotypes in both the absence and presence of TGF-β signaling and identifies a novel regulatory mechanism of the EMT and stemness regulator, SNAI2, through DEAR1 that is independent of the canonical TGF-β pathway, which may play a role in mediating the stem cell-like phenotype. Additionally, I provide preliminary data that suggests DEAR1 expression may be regulated by microRNAs, specifically miR-10b and miR-196b.
CHAPTER 2: Loss of DEAR1 confers stem and/or progenitor cell properties
2.1 Introduction

Several groups have demonstrated the link between EMT initiation and the acquisition of stem cell properties in human breast cells. This discovery was pioneered by Mani and colleagues when they observed that HMLEs forced to undergo EMT through overexpression of EMT regulators SNAI1 and/or TWIST1 or exposure to TGF-β resulted in enhanced mammosphere formation, in vitro differentiation capabilities, and a greater number of CD44+/CD24- expressing cells. When HMLE cells were sorted based on CD44 and CD24 surface marker expression, the CD44+/CD24- population showed higher expression of mesenchymal markers and EMT markers and transcription factors, including vimentin, fibronectin, SNAI1, and ZEB2 compared to their CD24+ counterparts. Furthermore, the sorted CD44+/CD24- cells were the only cells capable of generating mammospheres. It was also demonstrated that H-Ras transformed HMLEs (HMLER cells) undergoing an EMT were able to generate tumors in cleared mouse mammary fat-pads when injected at low densities, providing evidence that EMT can induce a cancer stem cell phenotype [57, 76]. Morel et al. used HMLER cells to demonstrate that initiation of EMT or activation of MAPK can induce generation of a larger population of CD24- cells and showed that CD24+ cells can be converted to CD24- cells through exposure to TGF-β [76].

The mammosphere assay is an in vitro protocol that was derived from the neurosphere assay first established by Reynolds and Weiss in 1996 [111] to functionally identify stem and/or progenitor cells. Cells grown in 2D culture are dissociated into single cells and plated on ultra-low attachment plates in Mammary Epithelial Cell Growth Medium (MEGM) supplemented with EGF, hFGF, and heparin. Cells able to survive and proliferate in the low adherent condition are expected to have stem/progenitor cell properties and form spheres, termed “mammospheres”, that are comprised of stem, multipotent and/or bipotent progenitor, and differentiated cells (Figure 2.1) [58, 112, 113]. The mammosphere assay can also be used to study self-renewal and differentiation capacity, two noteworthy characteristics of normal stem cells. An in vitro assay to demonstrate self-renewal capacity is through serial passaging of primary mammospheres for
Mammosphere forming efficiency (MFE): (number of mammospheres / number of cells seeded) x 100
Normal mammary epithelial cells: 0.1-0.7%
Breast cancer cell lines: 1-4%

**Figure 2.1: Mammosphere assay.** The mammosphere assay is an *in vitro* method of enriching for cells with stem or progenitor cell properties, which are capable of surviving and proliferating in non-adherent conditions. Mammary epithelial cells grown in 2D culture are dissociated and plated in ultra-low attachment plates in MEGM supplemented with EGF, FGF, and heparin. Mammospheres are spheres of cells that consist of stem or progenitor cells and differentiated progeny. Mammosphere forming efficiency (MFE) is calculated as a percentage by counting the number of mammospheres larger than 50μM divided by the number of cells seeded, multiplied by 100. The MFE range for normal mammary epithelial cells is 0.1-7% whereas the MFE for breast cancer cell lines is 1-3%. Primary mammospheres could be serially passaged to generate secondary and tertiary mammospheres to determine self-renewal capacity and stained to study differentiation marker expression to evaluate differentiation capacity [58].
generations. Over time, cells with the ability to self-renew would generate an increasing number of mammospheres upon re-plating. Progenitor cells, which do not have self-renewal capacities but have higher proliferation potential, would result in the generation of larger mammospheres but serial passaging would result in a decrease in mammosphere quantity over time [58, 112, 114]. In vitro differentiation assays include flow cytometry sorting based on combinations of epithelial surface markers to determine changes in population of different epithelial lineages, immunofluorescence staining of mammospheres using luminal and basal cytokeratin markers, and colony forming assays to identify multipotent progenitor cells [57, 58, 70, 114]. An in vivo method to test for stemness is through a limited dilution assay in which different densities of cells are injected into cleared mammary fat-pads. At low density, cells with self-renewal potential and differentiation capacity can regenerate the mammary ductal system. This method has also been utilized to test the identity of cancer stem cells [57, 58, 79, 115].

Although DEAR1 has been demonstrated to be a bona fide tumor suppressor in breast carcinoma, its specific role in carcinogenesis and biological pathways remains unclear. Given that DEAR1 is a regulator of acinar morphogenesis through its control of apical-basal cell polarity [100] and is a negative regulator of EMT initiation through the canonical TGF-β pathway by binding and polyubiquitinating SMAD3 [102] and that EMT has been directly linked to the acquisition of a stem cell phenotype in immortalized human mammary epithelial cells, I hypothesize that DEAR1 could play a role in mediating the acquisition of stem cell properties. Using the mammosphere assay described above, I present data indicating that loss of DEAR1 expression in human mammary epithelial cells and DCIS cells enhances stem/progenitor cell properties.

### 2.2 Results

Utilizing stable DEAR1 KD MCF10A cells, a commonly used spontaneously immortalized non-tumorigenic human mammary epithelial cell line, I observed that loss of DEAR1 (DshR)
resulted in a significant increase in mammosphere quantity compared to control vector cells (CshR) (DshR 23.11±5.18 vs CshR 6.125±2.75, p<0.0001, Figure 2.2). Stable DEAR1 KD in another immortalized, non-tumorigenic mammary epithelial cell line, 76N-E6, produced significantly greater numbers of primary mammospheres compared to control vector cells (DshR 11.875±7.63 vs CshR 0.96±1.60, p<0.0001). When 76N-E6 cells are treated with 4ng/mL TGF-β for 48 hours prior to plating for mammospheres, all three DEAR1 KD clones formed significantly more mammospheres compared to TGF-β-exposed control vector cells (DshR 28.36±12.21 vs CshR 4.91±3.22, p<0.0001, Figure 2.3A & C).

We previously showed that DEAR1 binds specifically to SMAD3 and promotes its polyubiquitination. DEAR1-SMAD3 double KD HMECs were demonstrated to rescue the mesenchymal phenotype presented in DEAR1 KD HMECs. To determine if the mammosphere phenotype is dependent on SMAD3 signaling, we plated non-treated and TGF-β-treated DEAR1-SMAD3 double KD 76N-E6 cells in mammosphere growing conditions. In the absence of TGF-β exposure, DEAR1-SMAD3 double KD 76N-E6 cells generated as many mammospheres as DEAR1-KD cells (DshR-shSM3 19.46±6.95 vs DshR 17.92±10.75, p-value=0.56, Figure 2.4A & C). In the presence of TGF-β signaling, we observed that DEAR1-SMAD3 double KD 76N-E6 cells formed significantly fewer primary mammospheres compared to DEAR1 KD alone (DshR-shSM3 15.25±4.91 vs DshR 30.13±10.01, p<0.0001, Figure 2.4A & C). This data suggests that the mammosphere phenotype mediated by loss of DEAR1 is partially dependent on SMAD3 expression since we observed partial rescue of the mammosphere phenotype in DEAR1-SMAD3 double KD cells in the presence of TGF-β. However, since the average number of primary mammospheres formed by DEAR1-SMAD3 double KD 76N-E6 cells was similar to that of DEAR1 KD alone in the absence of TGF-β signaling, we believe that loss of DEAR1 alone contributes to the mammosphere phenotype independent of the TGF-β-SMAD3 axis. To confirm that mammospheres formed by HMEC clones are enriched for stem-like cells, whole spheres were harvested and fixed onto glass microscope slides for immunofluorescence staining.
Figure 2.2: Mammosphere formation in MCF10A cells. A preliminary mammosphere assay using immortalized human mammary epithelial cell line MCF10A reveals stable DEAR1-KD results in increased number of primary mammospheres. MCF10A control vector and DEAR1-KD pools were grown in 2D culture and plated in mammosphere-growing conditions (1000 cells/well in an ultra-low attachment 96-well plate). After 14 days, mammospheres larger than 50μM were counted. DEAR1-KD MCF10A cells formed significantly more cells compared to control vector cells (****p<0.0001).
**A**

76N-E6 primary mammospheres

<table>
<thead>
<tr>
<th>No. mammospheres (&gt;50μM)</th>
<th>CshR1</th>
<th>CshR2</th>
<th>DshR1</th>
<th>DshR2</th>
<th>DshR3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No TGF-β</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGF-β</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B**

TGF-β

[Image of Western blot showing DEAR1 and β-actin]

**C**

Average number of 76N-E6 primary mammospheres

<table>
<thead>
<tr>
<th></th>
<th>No TGF-β</th>
<th>TGF-β</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial</td>
<td>1 2 3</td>
<td>1 2 3</td>
<td>p-value</td>
</tr>
<tr>
<td>CshR1</td>
<td>0.25±0.5</td>
<td>0.25±0.5</td>
<td>0.25±0.5</td>
</tr>
<tr>
<td>CshR2</td>
<td>3.75±2.06</td>
<td>0±0</td>
<td>0.25±0.5</td>
</tr>
<tr>
<td>DshR1</td>
<td>12±2.71</td>
<td>25.5±2.65</td>
<td>10.75±2.5</td>
</tr>
<tr>
<td>DshR2</td>
<td>0.25±0.5</td>
<td>0.25±0.5</td>
<td>0.25±0.5</td>
</tr>
<tr>
<td>DshR3</td>
<td>14±3.92</td>
<td>3.25±1.26</td>
<td>5.75±2.22</td>
</tr>
</tbody>
</table>

**D**

[Images of mammospheres under No TGF-β and TGF-β conditions]
Figure 2.3: Stable DEAR1-KD in HMECs results in increased mammosphere number and size in the presence of TGF-β.  A) Two control vector clones (CshR) and three DEAR1-KD clones (DshR) were grown in 2D culture in two conditions: non-treated and treated (4ng/mL TGF-β for 48h). Cells grown in 2D were trypsinized and plated in 96-well ultra-low attachment plates at 1000 cells/well density. The bar graph represents the average number of primary mammospheres formed from each cell line and condition across three trials. In the non-treated cohort, two DEAR1-KD clones formed significantly more mammospheres compared to control vector cells. In the presence of TGF-β, all three DEAR1-KD clones formed significantly more spheres compared to treated control vector cells. Statistical analysis was performed using one-way ANOVA multiple comparison (compared to CshR1). B) Western blot analysis shows that DEAR1 expression is reduced in DEAR1-KD HMECs. C) Average number of primary mammospheres generated by control vector and DEAR1-KD HMEC clones in each trial and corresponding p-values. DEAR1-KD HMECs grown under TGF-β exposure form more primary mammospheres compared to treated control vector HMECs. D) HMEC mammospheres larger than 50uM were counted using a Zeiss microscope. Images were taken at 10X magnification. This assay was performed in triplicate, and data presented represents the average number of mammospheres across all three trials. P-value ***<0.001, ****<0.0001.
### A 76N-E6 primary mammospheres

![Graph showing the number of mammospheres](image)

- **Cell lines**: CshR, DshR, DshR/shC, DshR/shSM3
- **Conditions**: No TGF-β, TGF-β
- **Results**: ns, ****

### B

**Western Blot Analysis**

- **Proteins**: SMAD3, DEAR1, β-actin
- **Conditions**: TGF-β

### C

#### Average number of 76N-E6 primary mammospheres

<table>
<thead>
<tr>
<th>Trial</th>
<th>No TGF-β</th>
<th>TGF-β</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>CshR</td>
<td>2±1.51</td>
<td>1.5±1.2</td>
</tr>
<tr>
<td>DshR</td>
<td>13±3.55</td>
<td>19.25±4.33</td>
</tr>
<tr>
<td>DshR-C</td>
<td>9.88±2.95</td>
<td>32±4.28</td>
</tr>
<tr>
<td>DshR-SM3</td>
<td>22.25±1.31</td>
<td>24.25±4.77</td>
</tr>
</tbody>
</table>

### D

![Images showing mammosphere formation](images)
Figure 2.4: Increased primary mammosphere formation in DEAR1-KD HMECs is partially dependent on SMAD3 signaling in the presence of TGF-β. A) Untreated DEAR1-KD cells formed significantly more primary mammospheres compared to untreated control vector cells; untreated DEAR1-SMAD3 double KD did not show a significant difference in mammosphere quantity compared to DEAR1-KD alone. However, in the presence of TGF-β signaling, DEAR1-SMAD3 double KD cells formed significantly fewer mammospheres compared to loss of DEAR1 alone, suggesting that the mammosphere phenotype is dependent on SMAD3 signaling only in the context of TGF-β exposure. B) Western blot analysis confirms DEAR1 KD in HMECs and loss of SMAD3 expression in the DEAR1-SMAD3 double KD clone. C) Average number of primary mammospheres generated by control vector and DEAR1-KD HMEC clones in each trial and corresponding p-values. DEAR1-SMAD3 double KD HMECs grown under TGF-β exposure consistently form fewer primary mammospheres compared to DEAR1-KD HMECs. D) HMEC mammospheres larger than 50uM were counted using a Zeiss microscope. Images were taken at 10X magnification. This assay was performed in triplicate, and data presented represents the average number of mammospheres across all three trials. P-value ****<0.0001.
ALDH1 is an established marker for stem and progenitor cells, and is overexpressed in both normal and cancer stem cells [79]. ALDH1 expression is barely detectable in HMEC clones grown in 2D culture conditions (Figure 2.5A). However, HMECs grown in 3D mammosphere conditions result in high expression of ALDH1 (Figure 2.5B), indicating that in our 3D HMEC mammosphere assay, cells with stem and/or progenitor properties are enriched and importantly, that loss of DEAR1 results in a higher number of cells with stem and/or progenitor properties as demonstrated by the increase in the number of mammospheres formed by DEAR1 KD clones.

Stem cells are distinguished from other cell types through their ability to self-renew (forming more stem cells) and to differentiate into other cell types (multipotent progenitor cells, bipotent progenitor cells, and mature differentiated cells) [58, 113, 114]. In order to determine if loss of DEAR1 contributed to self-renewal in our HMEC model, primary mammospheres were passaged by disseminating spheres into single cell suspensions and re-plated at the same density to form secondary and tertiary mammospheres. According to Dontu et al., the number of spheres should increase after multiple passages if there are stem cells, which have the ability to generate more stem cells, in the population [58]. I observed that loss of DEAR1 in HMECs did not allow for the acquisition of self-renewal capabilities, even in the presence of TGF-β signaling, as indicated by the unaltered or reduced fold-change (DshR 0.59±0.94 vs CshR 1.08±0.55, p=0.55, Figure 2.6).

Another well-known characteristic of stem and progenitor cells is the ability to differentiate into other cell types. There are two types of differentiated mammary epithelial cells that can be generated by mammary bipotent progenitor cells: luminal and myoepithelial/basal cells. To determine if DEAR1 loss could promote bipotential capacity, or the capacity to differentiate into either luminal or myoepithelial/basal lineages, primary mammospheres were fixed onto microscope slides and stained by immunofluorescence staining for cytokeratin (CK) 8/18, a
Figure 2.5: ALDH1 expression in 2D and 3D mammosphere culture. ALDH1, a marker of normal and cancer stem cells, is expressed in HMEC cells grown in mammosphere conditions, but not in 2D culture. A) 76N-E6 control and DEAR1-KD clones were grown in 2D culture in the absence of TGF-β. When grown in these conditions, expression of ALDH1 is minimal (white arrows). B) When 76N-E6 control and DEAR1-KD clones were grown in mammosphere conditions, ALDH1 was highly expressed in mammosphere-forming cells, even in the control mammospheres. A similar trend was observed when cells were grown in either 2D or 3D mammosphere culture in the presence of TGF-β signaling (C&D). These results indicate that the mammospheres formed using this assay are enriched for stem- or progenitor-like cells. Images were taken at 10X magnification.
Table 2.1: Loss of DEAR1 in HMECs does not confer self-renewal potential. Untreated and TGF-β-treated primary mammospheres were dissociated into single-cell suspensions and replated in mammosphere growing conditions to generate secondary mammospheres. One *DEAR1*-KD clone (DshR1) was able to generate more secondary mammospheres in the untreated cohort, however this data could not be replicated in other trials. All other clones did not demonstrate self-renewal capacities. This indicates that loss of DEAR1 does not affect self-renewal capacities in HMECs. Self-renewal assays were performed in three independent trials.

<table>
<thead>
<tr>
<th>HMEC clone</th>
<th>Passage</th>
<th>MFE</th>
<th>Fold Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>CshR1</td>
<td>1</td>
<td>0.325%</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.267%</td>
<td></td>
</tr>
<tr>
<td>CshR2</td>
<td>1</td>
<td>0.225%</td>
<td>1.47</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.33%</td>
<td></td>
</tr>
<tr>
<td>DshR1</td>
<td>1</td>
<td>1.55%</td>
<td>1.68</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.6%</td>
<td></td>
</tr>
<tr>
<td>DshR2</td>
<td>1</td>
<td>0.675%</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.067%</td>
<td></td>
</tr>
<tr>
<td>DshR3</td>
<td>1</td>
<td>1.275%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HMEC clone</th>
<th>Passage</th>
<th>MFE</th>
<th>Fold Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>CshR1</td>
<td>1</td>
<td>0.325%</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.2%</td>
<td></td>
</tr>
<tr>
<td>CshR2</td>
<td>1</td>
<td>0.075%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>DshR1</td>
<td>1</td>
<td>1.85%</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.4%</td>
<td></td>
</tr>
<tr>
<td>DshR2</td>
<td>1</td>
<td>1.5%</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.8%</td>
<td></td>
</tr>
<tr>
<td>DshR3</td>
<td>1</td>
<td>2.7%</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.0%</td>
<td></td>
</tr>
</tbody>
</table>
luminal cell surface marker, and CK5, a myoepithelial/basal cell surface marker [15, 62, 77]. When stained for these surface markers, I observed a population of cells within the mammospheres that co-express both CK8/18 and CK5, suggesting that these cells have bipotential capacity [57], but did not demonstrate a significant difference in the number of bipotent cells between CshR and DshR clones (no treatment p=0.25, Figure 2.7A; TGF-β p=0.083, Figure 2.7B). FACS analysis using EpCAM and CD49f (α-6-integrin) revealed that when DEAR1 expression is lost in HMECs, there is a significant shift in the percentage of cells in the EpCAM+/CD49f+ population, indicative of luminal progenitor/differentiated cells, to a EpCAM−/low/CD49f+ phenotype, indicative of basal progenitor and/or bipotent cells [73, 88, 115], in two DEAR1 KD clones (DshR 3.82%±2.50 vs CshR 0.54%±0.28, p=0.04, Figure 2.8A&B). Furthermore, a FACS analysis by Dr. Nanyue Chen, an Assistant Professor in our laboratory, using CD44 and CD24 surface markers revealed that DEAR1-KD clones have a higher CD44:CD24 ratio compared to control vector clones due to loss of CD24 expression (DshR 45.22±22.59 vs CshR 95.88±19.12, p=0.01, Figure 2.8C). Together, this data suggests that although loss of DEAR1 may not result in a stem cell phenotype, loss of DEAR1 can affect progenitor cell properties and regulate basal cell fate.

To study if loss of DEAR1 in a different mammary epithelial cell line results in acquisition of a stem/progenitor cell phenotype, we utilized MCF10DCIS.COM cells (DCIS.COM), a cell line derived from MCF10A cell line series that generates necrotic DCIS lesions when injected into immunocompromised mice [116-119]. DCIS.COM cells express high levels of DEAR1, thus we knocked down DEAR1 expression using the GIPZ-GFP shTRIM62 lentiviral vectors from Dhharmacon to ascertain any effects on the mammosphere phenotype. Two control vector clones (CshR-7 and CshR-20) and one DEAR1-KD clone (DshR 9-7) were selected based on protein and mRNA expression (Figure 2.9A). Results show that the number of primary mammosphere formed by control vector and DEAR1-KD clones were not consistently significantly different when
Figure 2.7: Mammospheres contain bipotent cells. 76N-E6 control and DEAR1-KD clones grown in 3D mammosphere culture in the absence (A) or presence (B) of TGFβ were stained using luminal and myoepithelial/basal cytokeratin surface markers, CK 8/18 and CK 5, respectively. All four categories of mammospheres contained a small population of cells that co-express both CK8/18 and CK5 (white arrows). Images were taken at 20X magnification.
Figure 2.8: DEAR1-KD HMECs contain a larger population of basal-like cells. A&B) WT and DEAR1-KD HMECs were sorted based on expression of EpCAM and CD49f surface markers. Previous studies have categorized epithelial cells into distinct lineages based on surface marker expression profiles. EpCAM$^{hi}$/CD49f$^+$ signifies mature luminal cells while EpCAM$^{hi}$/CD49$^+$ are luminal progenitor cells; EpCAM$^{lo}$/CD49f$^+$ cells are categorized as mature basal cells and EpCAM$^+$/CD49f$^+$ as basal progenitor cells. FACS data revealed that loss of DEAR1 expression HMECs generates a significant increase in the population of basal mature and progenitor cells that are EpCAM$^{lo}$/CD49f$^+$. C) FACS data indicates that there is significant repression of CD24 surface marker expression in both DEAR1-KD clones compared to control vectors. P-value **<0.01, ****<0.0001.
plated at a density of 500 cells/well (DshR 43±10.64 vs CshR 40.56±10.3, p=0.60), however we observed a consistent significant increase in the average size of mammosphere formed by the DEAR1-KD DCIS clone (DshR 112.09uM±31.60 vs CshR 77.27uM±17, p<0.0001, Figure 2.9B) when plated at a density of 1000 cells/well. Dr. Nanyue Chen also generated DEAR1 knockout (KO) DCIS clones. With his permission, I used these clones and respective wildtype control clones to test mammosphere formation. Similar to the KD experiments, knockout clones did not consistently form more primary mammospheres compared to controls (KO 63.67±14.8 vs WT 56.96±12.16, p=0.59, Figure 2.10A). Results indicated that, on average, DEAR1-KO DCIS clones consistently formed larger mammospheres compared to DEAR1-expressing DCIS clones (KO 113.02uM±52.93 vs WT 82.96uM±20.66, p<0.0001) and that all clones formed at least one or more mammosphere(s) larger than 150uM (Figure 2.10B). Mammosphere size has been previously correlated with progenitor cell population [58, 112]. DCIS primary mammospheres were serially passaged to form secondary and tertiary mammospheres. All cell lines could be propagated as secondary and tertiary spheres, but DEAR1-KO mammospheres continued to form larger mammospheres compared to DCIS WT mammospheres (KO 134.8uM±58.19 vs WT 98.99uM±28.61, p<0.0001, Figure 2.11A&B). DEAR1 knockout in DCIS cells, however, demonstrated a different phenotype when exposed to TGF-β compared to the phenotype caused by loss of DEAR1 in HMECs. Using two control and two DEAR1-KO DCIS clones, cells were plated and treated with DMSO, TGF-β inhibitor, TGF-β, or both TGF-β and inhibitor. As shown in Figure 2.12A, treatment with TGF-β inhibitor resulted in an increase in mammospheres in all clones while treatment with TGF-β resulted in significant inhibition of the number of primary mammospheres in control clones (treated 17.17±5.6 vs untreated 63.83±14.99, p<0.0001) and reduction of primary mammospheres formed by DEAR1-KO clones, though not at a significant level (treated 59±4.73 vs untreated 71.67±7.58, p=0.2). However, TGF-β-treated DEAR1-KO clones formed significantly more mammospheres compared to TGF-β-treated control clones (KO 59±4.73 vs WT 17.17±5.6, p<0.0001). There was not a significant difference in size between TGF-β-treated DEAR1-KO and control DCIS clones (KO 104.86uM±42.60 vs WT
**Figure 2.9: DEAR1 knockdown in DCIS.COM cells results in generation of significantly larger primary mammospheres compared to control vectors.** A) Stable DEAR1-KD and control vector DCIS clones were generated using Dhharmacon lentiviral shRNA vectors. DEAR1 loss was confirmed by western blot analysis and qPCR. B) DCIS clones were plated in mammosphere growing conditions at either 500 or 1000 cells/well densities. DEAR1 KD did not increase primary mammosphere quantity, however at 1000 cells/well, DEAR1-KD clone DshR 9-7 formed significantly larger mammospheres compared to the two control vector clones. P-value *=0.01-0.05, ** <0.01, **** <0.0001.
A

DCIS mms (#)

No. mammospheres (>50μM)

DCIS WT

DCIS DR1 KO

DCIS clones

<table>
<thead>
<tr>
<th>MFE (%)</th>
<th>500 cells/w</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC12 (WT)</td>
<td>10.4%</td>
</tr>
<tr>
<td>DC57 (WT)</td>
<td>14.55%</td>
</tr>
<tr>
<td>DC75 (WT)</td>
<td>13.35%</td>
</tr>
<tr>
<td>DC17 (KO)</td>
<td>11.05%</td>
</tr>
<tr>
<td>DC21 (KO)</td>
<td>13.1%</td>
</tr>
<tr>
<td>DC23 (KO)</td>
<td>9.35%</td>
</tr>
<tr>
<td>DC26 (KO)</td>
<td>16.85%</td>
</tr>
<tr>
<td>DC35 (KO)</td>
<td>11%</td>
</tr>
<tr>
<td>DC44 (KO)</td>
<td>11.55%</td>
</tr>
</tbody>
</table>

B

DCIS mms (size)

Average size (μM)

DCIS WT

DCIS DR1 KO

DCIS clones

DCIS mms (500 cells/w)

Percentage

DCIS KO clones

>151μM

126 - 150μM

101 - 125μM

76 - 100μM

50 - 75μM
Figure 2.10: *DEAR1* knockout in DCIS.COM cells also generates larger primary mammospheres. A) DCIS cells were plated at a density of 500 cells/well in low-attachment mammosphere conditions. Similar to the *DEAR1*-KD DCIS clone, knockout of *DEAR1* by CRISPR technology does not consistently result in an increase in primary mammosphere number. B) Three of six *DEAR1*-KO clones consistently generate significantly larger mammospheres compared to *DEAR1*-expressing DCIS clones and all *DEAR1*-KO clones form multiple mammospheres larger than 150μM. P-value ***<0.001, ****<0.0001.
Figure 2.11: All DCIS.COM mammospheres formed secondary and tertiary mammospheres, but DEAR1-KO clones consistently formed larger mammospheres with each passage. A) DCIS clones demonstrate differences in self-renewal capacities. Both control and DEAR1-KO clones are capable of generating secondary and tertiary mammospheres. B) DEAR1-KO clones continue to generate larger mammospheres after the second and third passages. The consistent increase in mammosphere size in DCIS cells with reduced or no DEAR1 expression indicates that loss of DEAR1 could regulate progenitor cell proliferation.
93.53uM±32.37, p=0.23) although DEAR1-KO clones continued to generate larger mammospheres in the DMSO and TGF-β inhibitor cohorts (KO 167.36uM±70.44 vs WT 95.29uM±30.83, p<0.0001 and KO 167.3uM±67.39 vs WT 125.19uM±56.05, p<0.01, respectively; Figure 2.12B). Overall this data suggests that loss of DEAR1 expression regulates progenitor cell properties as demonstrated by generation of larger mammospheres by DEAR1-KO clones that is not regulated through the TGF-β pathway.

2.3 Discussion

Since loss of DEAR1 resulted in an EMT signature and increased migration in the presence of TGF-β in HMECs, I hypothesized that loss of DEAR1 would also allow cells to acquire stem/progenitor cell properties. Stable DEAR1 KD in HMECs resulted in significantly greater number of mammospheres and exposing DEAR1-KD HMECs to TGF-β1 enhanced the mammosphere phenotype. These results confirm results of experiments performed in HMECs that DEAR1 prevents mammosphere formation independently of its function in regulating the canonical TGF-β pathway. Further evidence of this theory was provided when I tested whether the mammosphere phenotype is dependent on SMAD3. Utilizing DEAR1-SMAD3 double KD HMECs, I observed that non-treated double KD cells did not form significantly fewer mammospheres compared to DEAR1 KD alone, but that TGF-β1 treated DEAR1-SMAD3 double KD cells formed approximately half the number of primary mammospheres compared to DEAR1 KD alone. This data suggest that the mammosphere phenotype is only partially dependent on SMAD3 in the context of TGF-β signaling, and provide further evidence that mammosphere formation may be dependent on DEAR1 independently of TGF-β-mediated EMT. There are several known pathways that have been demonstrated to play roles in the regulation of stemness, such as the Hippo, Wnt, Notch, and MAP kinase pathways [15, 64, 69, 81]. It is possible that DEAR1 may regulate one of these alternative pathways to regulate stem/progenitor cell properties. It was also shown that overexpression of TWIST, a known EMT regulator, in
A. **DCIS mms (#)**

- **DCIS clones:** DC12, DC57, DC23, DC26

- **Comparison:** DCIS WT vs. DCIS DR1 KO

- **Statistics:**
  - DC12: Significant difference (P<0.0001)
  - DC57: Significant difference (P<0.0001)
  - DC23, DC26: Not significant

B. **DCIS mms (size)**

- **Average size (μM):**
- **Statistics:**
  - DC12, DC57, DC23, DC26: Not significant (N.S.)

*Compared to DC12 TGFb+*
Figure 2.12: TGF-β inhibits formation of DCIS mammospheres. TGF-β treatment results in reduced mammosphere quantity in DCIS.COM cells, however DEAR1-KO DCIS.COM clones treated with TGF-β formed significantly more primary mammospheres compared to treated control clones. A) DCIS.COM clones pre-treated with TGF-β and maintained in mammosphere growing conditions supplemented with TGF-β formed significantly fewer mammospheres compared to untreated cells (p<0.0001). However, DCIS-KO DCIS clones formed significantly more primary mammospheres compared to treated DC12 control clone (p<0.0001). Treatment with TGF-β and TGF-β inhibitor rescued the phenotype. B) Although DMSO-treated DEAR1-KO DCIS clones still formed larger mammospheres compared to DMSO-treated control clones, TGF-β-treated DEAR1-KO DCIS clones failed to generate larger mammospheres compared to control clones.
human mammary epithelial cells can confer stemness independently of EMT [84, 85]. Thus, it is possible that DEAR1, a regulator of EMT, can affect stem/progenitor cell properties without initiation of an EMT program.

Although self-renewal studies revealed that loss of DEAR1 does not confer self-renewal capacity, this does not eliminate the possibility that DEAR1 may regulate progenitor cell properties. I identified a small population of bipotent cells within primary mammospheres that co-express CK8/18, a luminal specific surface marker, and CK5, a basal/myoepithelial specific surface marker. This was also demonstrated by Mani and colleagues using CK14 and CK8/18 antibodies on HMLE mammospheres [57]. Wang et al. demonstrated that BT474, MDA-MB-361, and MCF7 breast cancer cell lines expressed CK5 when grown as mammospheres but not in 2D culture [120]. Primary mammospheres also stained highly for ALDH1, indicating that cells grown in mammosphere conditions are enriched for stem and/or progenitor cell properties, as ALDH1 has been demonstrated as a marker of both stem and progenitor cells in a variety of cell types and is associated with migration and invasion [77, 79, 80]. Loss of DEAR1 also resulted in a slight increase in the population of basal progenitor cells (EpCAM−/CD49f+) in 2 of 3 DEAR1-KD HMEC clones. In a study by Proia and colleagues, mutations in the tumor suppressor and breast cancer susceptibility gene, BRCA1, resulted in a significant increase in the proportion of EpCAM+/CD49f+ cells, suggesting that mutation of BRCA1 leads to altered progenitor cell fate, mainly through increased population of basal progenitor cells [88]. Other studies have shown that BRCA1 mutations or deficiency results in increased incidence of basal-like tumors and induction of the EMT program through activation of TWIST [112, 121, 122]. Furthermore, in a FACS analysis using CD44 and CD24 surface markers, loss of DEAR1 in two DEAR1-KD clones resulted in decreased CD24 expression leading to an increase in the CD44:CD24 ratio. CD24 expression is associated with the luminal epithelial cell phenotype and its loss of expression has been observed in metastatic breast cancer cells [77, 80, 123]. Although we did not see an increase in CD44 expression in DEAR1-KD HMECs, it is important to note that utilizing surface
markers to identify stem and progenitor cell populations is context-dependent and marker combinations vary between different cell types and cell lines. In a study by Smart and colleagues, using CD49f and CD24 provided better stratification than CD44 and CD24 [124]. As a whole, these results indicate that although DEAR1 may not affect self-renewal, it may still function as a regulator of progenitor cell properties and that loss of DEAR1 increases the number of progenitor cells capable of forming mammospheres. It is also interesting that loss of DEAR1 in HMECs causes a shift in the population of differentiated basal/basal progenitor cells since our previous study in a cohort of young female breast cancer patients revealed that loss of DEAR1 significantly correlated with triple negative breast cancer, which contains a subset of basal-like cancers that is associated with an EMT signature [6, 100].

To study the effects of DEAR1 on stem/progenitor cell properties, I utilized a DCIS cell line, MCF10DCIS.COM, that demonstrates mammosphere forming capabilities [15, 81] and expresses wildtype DEAR1. Loss of DEAR1 expression, through either lentiviral shRNA KD or CRISPR-mediated knockout, in DCIS.COM cells did not result in consistent changes in primary mammosphere quantity, but rather a significant increase in mammosphere size compared to respective control cells, even after several passages. The difference in mammosphere phenotype between DCIS cells and HMECs could be due to difference in cell lines and inherent differences between genomes (one being a stage 0 breast cancer cell line and the other an immortalized human mammary epithelial cell line). This data is consistent with what has been established regarding mammosphere quantity and size. While the number of mammospheres is an indication of self-renewal capacity, the size of mammospheres infers a larger population of proliferative progenitor cells [112]. This data provides further evidence that DEAR1 regulates progenitor cell properties in HMECs and DCIS cells.

Another phenotype that differed between DEAR1 loss in DCIS cells and HMECs was the effect of TGF-β on mammosphere formation. Interestingly, exposure to TGF-β caused an opposite effect on mammosphere formation in DCIS cells compared to HMECs in that the number of mammospheres was significantly reduced in both DEAR1-WT and DEAR1-KO DCIS clones,
although the number of mammospheres formed in DEAR1-KO clones was higher than those of DEAR1-WT clones. A possible explanation for this phenomenon could be due to the tumor suppressive role of TGF-β that is highly context-dependent. A study by Wang and colleagues demonstrated that TGF-β pre-treatment in BT474 and MDA-MB-361 cell lines resulted in higher MFEs, but pre-treatment in MCF7 cells resulted in decreased MFE compared to untreated controls [120]. Several studies have demonstrated that TGF-β inhibits proliferation in multiple cell types, including cancer cell lines, in part through its regulation of p15INK4B, p21, and c-Myc [39]. It is now clear that TGF-β acts as a tumor suppressor in normal development and in the earlier stages of cancer, but switches to an oncogenic role in late stage cancer and facilitates metastasis through its involvement in inducing EMT [30, 39, 43]. Since 76N-E6 cells were immortalized by HPV E6, the inactivation of p53 may have also protected these HMECs from TGF-β’s growth inhibition [46].

Together, this data provides evidence that loss of DEAR1 in HMECs and DCIS confers progenitor cell properties that is independent of TGF-β-mediated EMT. Further experiments could be designed to understand the specific impacts of DEAR1 loss on progenitor cell fate and elucidate the effects of DEAR1 on differentiation capabilities \textit{in vivo}. 
CHAPTER 3: Defining the mechanism by which DEAR1 mediates stem and/or progenitor cell properties
3.1 Introduction

Several known oncogenic pathways have been implicated in mediating stemness in mammary epithelial cells and the formation of mammospheres in vitro, including Notch, Hippo, MAPK/Akt, and TGF-β signaling pathways [15, 57, 81, 114, 125]. Transcription factors that mediate the EMT program have been demonstrated to also play roles in self-renewal, differentiation, and determining cell fate [57]. It is unknown whether expression of these factors is the cause or a consequence of acquiring stem cell properties, though some studies have demonstrated that overexpression of specific transcription factors results in an increased population of stem and/or progenitor cells.

In particular, overexpression of TWIST1, ZEB1, FOXC2, and SNAI1/2 proteins in human mammary epithelial cells was shown to generate cells with stem-like phenotypes, including CD44+/CD24- marker profile and mammosphere formation [126, 127]. TWIST1 overexpression resulted in reduction of CD24 epithelial differentiation marker and expression of BMI1, a stem cell factor involved in self-renewal [84, 128]. ZEB1 expression has been found to correlate with poorly differentiated pancreatic, breast, and colorectal tumors [56, 129, 130]. Repression of ZEB1 in pancreatic cancer cell lines resulted in reduced CD24+/CD44+ pancreatic stem cell population, failure to form tumorspheres, tumor-initiating capacity assayed by in vivo limited dilution assay, and inhibition of stem cell factors including BMI1 and SOX2 [129]. Furthermore, KD of ZEB1 expression in metastatic breast cancer cells resulted in repression of epithelial markers MUC1 and CD24 [130]. FOXC2, another prominent EMT regulator, was highly expressed in the CD44+/CD24- stem cell population and ectopic expression resulted in HMLER cells gaining stem cell properties including increased mammosphere formation and a significant shift from CD44-/CD24+ surface marker expression to CD44+/CD24- phenotype. Additionally, FOXC2 induces expression of genes that are enriched in basal tumors [127]. SNAI1 negatively regulates miR-34, which inhibits stemness and promotes cell differentiation [126]. SNAI2 was shown to cooperate with SOX9 in murine mammary luminal epithelial cells to drive cells into a more de-differentiated state [87]. SNAI2 KD in HMECs led to high expression of luminal markers, including
EpCAM, E-cadherin, MUC1, and CD24, implying that SNAI2 is a repressor of luminal differentiation [72]. Overexpression of SNAI2 in MCF10A cells resulted in an increase in the CD44+/CD24- population and acquired mammosphere-forming capabilities; this observation was not replicated in the MCF-7 luminal breast carcinoma cell line, suggesting that SNAI2-mediated acquisition of stem-like properties is context-dependent [72, 86]. Furthermore, SNAI2 overexpression was observed in breast carcinomas of the basal subtype, often associated with BRCA1 mutation, and demonstrated high expression of stem cell factors CD133 and BMI1 [88, 131].

We have preliminary data showing that TGF-β exposure and loss of DEAR1 in 76N-E6 HMECs resulted in a significant increase in SNAI2 and ZEB2 mRNA levels. Since we observed a mammosphere phenotype with loss of DEAR1 in the absence of TGF-β signaling in HMECs, we hypothesize that these EMT-related transcription factors are upregulated in the context of DEAR1 downregulation and are, in part, responsible for the mammosphere phenotype.

3.2 Results

Utilizing 76N-E6 HMECs, control and DEAR1-KD clones were grown in 2D culture in the absence and presence of TGF-β for qRT-PCR analysis. Results indicate that loss of DEAR1 led to a significant increase in basal SNAI2 and ZEB2 mRNA compared to untreated control cells (1.44 fold change, p=0.0002 and 13.28 fold change, p<0.0001, respectively) and that this is enhanced when cells were grown in the presence of TGF-β (3.11 and 76.56 average fold change, respectively, p<0.0001, Figure 3.1A&B), confirming our previous observations. Since miR-200 family members negatively regulate ZEB proteins, I evaluated expression levels of miR-200b and miR-200c to determine if loss of DEAR1 affected these miRNAs known to inhibit ZEB2. qRT-PCR results show that miR-200b levels decrease in DEAR1-KD cells after TGF-β exposure, but not significantly different compared to the basal level of expression in the untreated control vector cells (1.2 fold change, p=0.89, Figure 3.1C) while miR-200c levels steadily increased after
Figure 3.1: Loss of DEAR1 in HMECs results in increased basal expression of *SNAI2* and *ZEB2* mRNA. A) Basal *SNAI1* mRNA expression in a DEAR1-KD clone is significantly higher compared to a control vector clone. *SNAI2* levels increase after TGF-β treatment in both CshR and DshR clones, however the increase in DshR clone remains significantly greater than the treated CshR clone. B) Basal *ZEB2* mRNA expression in the DEAR1-KD clone is significantly higher compared to control and, similar to *SNAI2* levels, also increase after treatment with TGF-β that is significantly greater compared to treated control. C-D) miR-200b and miR-200c, negative regulators of ZEB1, expressions are significantly higher at basal level in DEAR1-KD clone compared to control vector, indicating that loss of DEAR1 affects ZEB1 mRNA expression independently of miRNA regulation. P-value **<0.01, ***<0.001, ****<0.0001.
TGF-β exposure (1.89 fold change, $p<0.0001$, **Figure 3.1D**). At basal level, in which cells are not exposed to TGF-β, DEAR1-KD cells express higher levels of miR-200b (1.45 fold change, $p=0.01$) and miR-200c (1.67 fold change, $p=0.003$) compared to untreated control cells, indicating that loss of DEAR1 mediates overexpression of ZEB2 independently of miR-200 family expression.

Next I performed qRT-PCR using two control vector clones and three DEAR-KD clones to evaluate basal expression of SNAI2, ZEB1, and ZEB2 mRNA. As shown in **Figure 3.2A**, all 3 DEAR1-KD clones exhibited significantly higher expression of SNAI2 and ZEB1 mRNA (for $p$-values please refer to table listed in **Figure 3.2A**). Further assays using untreated and TGF-β-treated HMECs confirmed that loss of DEAR1 resulted in basal upregulation of SNAI2, ZEB1, and ZEB2 baseline mRNA and that exposure to TGF-β increased these mRNA levels in different clones at various times after exposure (**Figure 3.2B–D**). Since I observed increased primary mammospheres with DEAR1-KD alone even in the context of SMAD3 downregulation, I asked whether these same transcription factors are affected in DEAR1-SMAD3 double KD cells. qRT-PCR results indicate that SNAI2 baseline expression in DEAR1-KD and DEAR1-SMAD3 double KD cells remained significantly upregulated (1.93 fold change, $p=0.0001$; 1.42 fold change, $p=0.002$, respectively) compared to the control vector cell line (**Figure 3.2E**). This demonstrates that SNAI2 expression is not dependent on TGF-β-SMAD3 signaling and that loss of DEAR1 alone can confer upregulation of SNAI2 mRNA level. Since SNAI2 mRNA was not affected by loss of SMAD3 and has been shown to be involved in driving basal cell lineage and mammosphere formation [86, 88], it became a potential candidate gene mediating stemness in the context of DEAR1 downregulation.

To test if DEAR1 regulates SNAI2 expression at the transcriptional level, HEK293T cells were co-transfected with pGL3-SNAI2 luciferase vector that contains the SNAI2 promoter region upstream of the luciferase gene and pcDNA-DEAR1 or pcDNA-CNTL. Results consistently show
Figure 3.2: SNAI2, ZEB1, and ZEB2 basal mRNA levels are significantly higher in DEAR1-KD clones and SNAI2 mRNA expression remains increased in DEAR-SMAD3 double KD cells. A) Baseline expressions of SNAI2 and ZEB1, and ZEB2 mRNA are significantly higher in all three DEAR1-KD clones compared to control vectors. Table below panel A summarizes fold changes across all three clones and corresponding p-values. B-D) ZEB1, ZEB2, and SNAI2 mRNA levels across two control vector clones and three DEAR1-KD clones. In this analysis, all three DEAR1-KD clones exhibit significantly higher basal levels of ZEB1, ZEB2, and SNAI2 mRNA compared to control vectors and treatment with TGF-β results in an increase of mRNA levels at various time points in different clones. E) Analysis of DEAR1-SMAD3 double KD cells reveals that SNAI2 mRNA expression remains significantly higher compared to control cells, suggesting that SNAI2 expression could be independent from SMAD3 expression in the absence of TGF-β signaling. P-value *=0.1-0.5, **<0.01, ***<0.001, ****<0.0001.
that increasing concentrations of DEAR1 (200, 300, and 400ng) repressed luciferase activity compared to control cells at both 24 hours post-transfection (0.64, 0.58, and 0.40 fold change, respectively; p<0.001 vs pcDNA, Figure 3.3A) and 48 hours post-transfection (0.48, 0.34, and 0.24 fold change, respectively; p<0.001 vs pcDNA, Figure 3.3B). To test if DEAR1 may also regulate SNAI2 at the protein level, control and DEAR1-KD HMECs were used for western blot analysis. Results indicate that SNAI2 basal protein level was increased in all three DEAR1-KD clones and that this upregulation was enhanced in the presence of TGF-β (Figure 3.4A). To test if DEAR1 associates with SNAI2, HEK293T cells were co-transfected with HA-DEAR and Myc-SNAI2 vectors for co-immunoprecipitation assays. Results reveal that DEAR1 associates with SNAI2 (Figure 3.4B). Furthermore, a ubiquitination assay in HEK293T cells indicate that ectopic expression of DEAR1 promotes polyubiquitination of SNAI2 (Figure 3.4C). To test if a point mutation, C11A, in the RING domain that has been shown by others to be a loss-of-function mutation and utilized in other studies investigating DEAR1’s ligase function [108, 109] has an effect on DEAR1’s ability to polyubiquitinate SNAI2, I utilized a DEAR1-C11A mutant plasmid generated by Dr. Balasenthil. The C11A mutation did not affect ubiquitination of SNAI2 (Figure 3.4D) and further studies are required to determine which domains of DEAR1 are crucial for target binding and ubiquitination. Cumulatively, these results indicate that DEAR1 is a negative regulator of SNAI2 expression independent of the canonical TGF-β pathway and that SNAI2 is a good candidate for mediating the stem/progenitor cell properties in HMECs.

To determine if DEAR1-mediated SNAI2 expression affects mammosphere formation, I used DshR clones to perform lentiviral shRNA KD of SNAI2, generating DEAR1-SNAI2 double KD pooled HMECs that were used for mammosphere assays. As shown in Figure 3.5, loss of SNAI2 in either of the DEAR1-KD clones did not significantly reduce the quantity of mammospheres formed compared to the pooled DEAR1-KD control vector (DshR-cntl), although there was a significant decrease in the number of mammospheres formed by DshR1-shSNAI2 compared to the parental DshR1 clone (4.5 vs 29.25, p<0.0001). Although the DshR3-shSNAI2
Figure 3.3: SNAI2 promoter-activated luciferase expression is inhibited by ectopic DEAR1 expression in HEK293T cells. A) Cells co-transfected with pGL3-SNAI2 luciferase vector and increasing concentrations of pcDNA-DEAR1 demonstrate significant reduction of luciferase activity. Cells were harvested 24 hours post-transfection. B) The same inhibitory trend is observed in cells harvested 48 hours post-transfection. P-values ***<0.001, ****<0.0001.
Figure 3.4: Loss of DEAR1 in HMECs result in increased SNAI2 protein; DEAR1 binds to and polyubiquitinates SNAI2. Loss of DEAR1 in HMECs results in an increase in baseline SNAI2 protein expression and is further enhanced in the presence of TGF-β. Co-IP and ubiquitination assays in HEK293T cells reveal that DEAR1 and SNAI2 interact to promote SNAI2 polyubiquitination. A) Western blot analysis of SNAI2 expression show that SNAI2 protein increases in all untreated DEAR1-KD HMEC clones compared to untreated control HMECs. Treatment with TGF-β enhances SNAI2 expression in all clones. B) Co-IP using HEK293T cells co-transfected with HA-DEAR1 and Myc-SNAI2 vectors reveal that DEAR1 and SNAI2 interact. Cell lysates were pulled down using anti-HA antibody and western blot analysis using anti-Myc antibody shows a band corresponding to SNAI2 only in cells transfected with DEAR1 and SNAI2. C) HEK293T cells were transfected with pcDNA-DEAR1, Myc-SNAI2, and 8x-HA-ubiquitin. Cell lysates were harvest 48 hours post-transfection and pulled down with anti-Myc antibody. Western blot analysis using anti-HA antibody shows that SNAI2 is polyubiquitinated only in the presence of DEAR1. D) The ubiquitination assay was repeated using a RING domain mutant, DR1-C11A, to determine if the interaction and polyubiquitination is dependent on DEAR1 RING-finger domain function. Western blot analysis shows that DR1-C11A still interacts and polyubiquititates SNAI2, suggesting that the interaction between DEAR1 and SNAI2 may be indirect.
Figure 3.5: *SNAI2* knockdown in *DEAR1*-KD HMECs did not fully rescue mammosphere phenotype. Using two *DEAR1*-KD clones, stable *SNAI2* KD was achieved by lentiviral GFP-shRNA vectors. Western blot analysis confirmed *SNAI2* downregulation after two rounds of GFP sorting. Control vectors, parental DshR clones, DshR control pool, and DshR shSNAI2 pooled cells were grown in mammosphere conditions. Overall, there were no significant differences between DshR-cntl and DshR shSNAI2 mammospheres, although there is a decreasing trend in DshR shSNAI2 mammospheres. P-values ****<0.0001.
pooled cells formed fewer primary mammospheres compared to DshR-ctl pooled cells, this difference was not significant (4.75 vs 9.5, p=0.17). This data suggests that loss of SNAI2 is not sufficient to completely rescue the mammosphere phenotype and that other factors may be involved in regulating the stem cell phenotype in HMECs.

### 3.3 Discussion

Using qRT-PCR, I confirmed initial experiments demonstrating that exposure to TGF-β signaling in the context of DEAR downregulation led to an increase in SNAI2 and ZEB2 mRNA. Results also showed that basal levels of SNAI2 and ZEB2 were significantly higher in DEAR1-KD clones, compared to control vector cells, indicating that loss of DEAR1 alone was sufficient to cause an upregulation of mRNA levels of EMT regulators. In regards to ZEB2 mRNA expression, I eliminated the possibility that its negative regulators miR-200b and miR-200c affected ZEB2 mRNA at baseline levels. Analysis of EMT regulators in two control vector and three DEAR1-KD HMEC clones using qRT-PCR revealed that loss of DEAR1 demonstrated upregulation of SNAI2, ZEB1 and ZEB2 mRNA in the absence TGF-β signaling. Importantly, SNAI2 mRNA expression was not significantly affected in DEAR1-SMAD3 double KD cells, indicating that SNAI2 may be an EMT-related/stemness factor that mediates the mammosphere phenotype in DEAR1-KD clones independently of the TGF-β-SMAD3 axis [86, 87, 121].

Luciferase assays indicated that ectopic expression of DEAR1 inhibits expression of luciferase driven by the SNAI2 promoter, demonstrating that DEAR1 transcriptionally inhibits SNAI2 expression, though whether this is a direct or indirect mechanism requires further investigation. This data corroborates with the discovery from Liu et al. that DEAR1 indirectly negatively regulates SNAI2 by regulating c-Jun in cervical cancer cells, although the precise mechanism of c-Jun inhibition by DEAR1 was not described [107]. Other studies have identified regulatory pathways that affect SNAI2 at the transcriptional level, including SPARC/Akt in melanoma cells, NF-κB pathway activation in breast cancer cells, and ER-mediated MTA3.
activation which complexes with the transcriptional corepressor unit, Mi-2/NuRD, in mammary epithelial cells [132-135]. Unpublished data by Dr. Balasenthil in our lab suggests that DEAR1 plays a role in mediating ER stabilization, thus serving as a potential pathway by which DEAR1 affects SNAI2 transcriptional activation. A study by Ye et al. identified another mechanism by which ERα mediates SNAI2 transcriptional repression through recruitment of histone deacetylase 1 (HDAC1) through nuclear receptor co-repressor 1 (NCOR1) binding to the estrogen-response element sequences found in the SNAI2 promoter [136]. It would be interesting to evaluate if DEAR1 participates in any of these pathways that transcriptionally mediate SNAI2 expression. In addition to increased SNAI2 mRNA, I also observed an increase in SNAI2 protein in DEAR1-KD HMEC clones in the absence of TGF-β signaling. Co-IP and ubiquitination assays indicate that DEAR1 binds to SNAI2 and promotes its polyubiquitination, although data suggests that this mechanism of SNAI2 regulation by DEAR1 may be indirect since there is still polyubiquitination of SNAI2 by the DEAR1 C11A RING domain mutant. Alternatively, a single point mutation may not be sufficient to abrogate DEAR1’s ability to polyubiquitinate its target proteins; more studies using deletion constructs will provide insight in regards to which domain(s) is/are crucial for target recognition and ubiquitination. DEAR1 may act through other post-translational regulators of SNAI2, such as BRCA1, p53/MDM2, or Wnt/GSK3β and β-Trcp1 E3 ubiquitin ligase [52, 88, 137]. Unpublished data by Dr. Balasenthil suggests that DEAR1 binds to and stabilizes p53, which regulates SNAI2 polyubiquitination by MDM2 and consequential proteasomal degradation [52].

To test if DEAR1-mediated regulation of SNAI2 affects mammosphere formation, SNAI2 expression was knocked down in DEAR1-KD HMECs. Using the DshR1 cells, I showed that although there is a slight reduction, there is no significant difference in mammosphere quantity between DshR1 control and DshR1 shSNAI2 pooled cells. However, there a significant reduction of mammospheres between parental DshR1 and DshR1 shSNAI2 cells. It is possible that the shRNA control vector is causing an unknown off-target effect that may affect mammosphere formation in this particular clone. DshR3 clone did not generate significantly more
mammospheres compared to the control vector clone, which was demonstrated multiple times in other experiments (refer to Chapter 2, Figures 2.3 and 2.4). The DshR3 control pooled cells generated more mammospheres compared to CshR1 and CshR2 control vector cells. Although I observed a reduction of mammospheres formed by DshR3 shSNAI2 pooled cells, this is not a significant difference when compared to DshR3 control pooled cells, potentially due to high error bars and a larger sample size is needed for a more accurate statistical analysis. Alternatively, using pooled cells may not be an accurate representation of effects of loss of SNAI2 in the context of DEAR1 downregulation. Another possibility is that loss of SNAI2 in both DEAR1-KD clones is not sufficient to reverse the mammosphere phenotype and that SNAI2 is an indirect mediator, contrary to what has been previously published.

For the first time, I demonstrate a novel mechanism of SNAI2 regulation at the transcriptional and post-translational level through the tumor suppressor, DEAR1, which is independent of the TGF-β-SMAD3 signaling axis. More work is required to better understand the relationship between DEAR1, SNAI2, and the effects of this regulatory mechanism on stem/progenitor cell properties, in particular the role of SNAI2 on progenitor cell fate and basal cell lineage in the context of DEAR1 regulation and breast carcinogenesis.
CHAPTER 4: Loss of DEAR1 correlates with poor overall survival in invasive breast cancer and time to metastasis in triple negative breast cancer
4.1 Introduction

It was previously demonstrated that DEAR1 protein expression in a cohort of young females between the ages of 25 and 49 years of age diagnosed with invasive breast carcinoma correlates significantly with local recurrence-free survival after a 15-year follow-up (p=0.0334) and associates with PR negative status (p=0.0321), triple negative breast cancer subtype (p=0.0362), and strong family history (p=0.0139) [100]. In a Korean cohort of 70 TNBC patients, DEAR1 was listed as one of the most frequently mutated genes with a frequency of 9%. One somatic variant, c.1094T>G, occurred in 6% of cases and was predicted to be deleterious based on SIFT and PolyPhen2 scores [138]. Studies by Quintás-Cardama et al. observed that loss of DEAR1 expression correlates with poor overall survival (p=0.00038) and shorter survival (p=0.00004) in AML patients. They also observed significant alterations in stem cell regulatory pathways, including Wnt/β-catenin and Notch signaling [106]. These studies indicate the importance of utilizing DEAR1 expression to evaluate prognosis and risk of early-onset TNBC, which has the worst prognosis compared to other breast cancer subtypes [10, 139, 140].

Expression of EMT regulators, including SNAI1/2, TWIST1, and ZEB1, was shown to be deregulated in breast cancer and correlated with poor prognosis, tumor recurrence, and increased aggressiveness [34, 47, 141]. Importantly, the EMT-gene signature is frequent in basal-like tumors, a subgroup of the TNBC subtype that is associated with a poorly differentiated phenotype [141]. High expression and/or amplification of SNAI2 were observed in multiple epithelial cancers, including prostate, lung, and breast carcinomas [1, 142]. SNAI2 is an established transcriptional repressor and mediator of EMT, notably through inhibiting expression of E-cadherin, occludin, claudin, and integrins, and its expression in breast carcinoma correlates with a de-differentiated phenotype and poor prognosis [47]. Furthermore, SNAI2 overexpression was significantly associated with high tumor grade, recurrence, and metastasis in primary breast cancer tissues [54]. The association between SNAI2 and a partially differentiated phenotype in breast carcinoma further provides evidence that SNAI2 could function in regulating epithelial
differentiation in mammary tissue. Additionally, SNAI2 is upregulated in basal-like breast carcinomas associated with BRCA1 mutations [88].

Given that loss of DEAR1 results in an enhanced progenitor cell phenotype and correlates with TNBC and strong family history, upregulation of SNAI2 was shown to associate with the basal-like subtype of TNBC and regulates differentiation, and the discovery of an inverse correlation between DEAR1 and SNAI2 in vitro, I asked whether this correlation exists in breast cancer patients and if I can utilize DEAR1 and SNAI2 expression to determine clinical outcome.

4.2 Results

Utilizing cBioPortal to analyze gene expression profiles across multiple cancer datasets, I observed DEAR1 loss of heterozygosity and SNAI2 amplification and/or overexpression in multiple cancer types (Figure 4.1) [1]. Analysis of the METABRIC cohort of invasive breast carcinomas (n=2509) [1, 140] revealed that DEAR1 LOH and/or loss of expression is observed in 17% of patients and significantly associates with poor overall survival (OS) (p=0.00002939, Figure 4.2). In the same cohort, SNAI2 amplification and/or overexpression occurred in 12% of patients with near significant association with poor OS (p=0.0562, Figure 4.3). DEAR1 LOH/downregulation and SNAI2 amplification/overexpression has a significant tendency to co-occur in this dataset (p<0.001) and alterations in both genes associate with poor OS at higher significance compared to DEAR1 alterations alone (p=0.00002067, Figure 4.4).

Using the breast cancer Gene-Expression Miner (bc-GenExMiner) v4.1 online genomic data mining tool [143, 144], I examined DEAR1 and SNAI2 mRNA in TNBC and non-TNBC samples across multiple published annotated and genomic data since DEAR1 and SNAI2 expressions have demonstrated associations with the TNBC subtype. Results indicate that DEAR1 expression is significantly reduced in basal-like and TNBC samples compared to other breast cancer subtypes (p<0.0001, Figure 4.5A – C). Furthermore, SNAI2 expression is
Figure 4.1: *DEAR1* LOH and *SNAI2* alterations are frequent in multiple cancer types in cBioPortal. A) *DEAR1* LOH is observed in a variety of epithelial cancers, including breast, ovarian, bladder, and pancreatic carcinomas. B) Amplification is the most commonly observed alteration of *SNAI2* in multiple cancer types, including breast, bladder, skin, and pancreatic cancers. Other cancers, such as non-small cell lung and esophagogastric cancers, also harbor *SNAI2* mutations.
Figure 4.2: *DEAR1* LOH and mRNA downregulation occurs in 17% of invasive breast carcinoma patients in the METABRIC cohort (n=2,509). A) Of the 2,491 sequenced samples, *DEAR1* undergoes LOH and/or mRNA downregulation in 415 cases. B) Overall survival for cases with and without *DEAR1* alterations. *DEAR1* LOH/downregulation correlates significantly with poor overall survival with median survival of 122.8 months compared to 168.3 months in the unaltered cohort (p=0.00002939).
Figure 4.3: *SNAI2* amplification and mRNA upregulation occurs in 12% of invasive breast carcinoma patients in the METABRIC cohort (n=2,509). A) Of the 2,491 sequenced samples, *SNAI2* undergoes amplification and/or mRNA overexpression in 303 cases. B) Overall survival for cases with and without *SNAI2* alterations. *SNAI2* amplification/upregulation shows a trend but does not correlate significantly with poor overall survival (p=0.0562).
Figure 4.4: Alterations in both *DEAR1* and *SNAI2* serve as a more significant predictor of overall survival in the METABRIC cohort (n=2,509). A) *DEAR1* LOH/mRNA downregulation and *SNAI2* amplification/overexpression alterations have a significant tendency to co-occur (p<0.001). B) Alterations in both *DEAR1* and *SNAI2* associate with poor OS even more significantly compared to *DEAR1* alterations alone (p=0.00002067).
Figure 4.5: *DEAR1* and *SNAI2* mRNA expressions are altered in basal-like and triple negative breast cancer. A-C) *DEAR1* mRNA level is significantly downregulated in the basal-like and TNBC subtypes compared to other breast cancer subtypes (p<0.0001). D) *SNAI2* mRNA expression is significantly upregulated in TNBC compared to non-TNBC subtype (p=0.0001). E) There is a significant inverse correlation between *DEAR1* and *SNAI2* mRNA expression; on average, *SNAI2* expression is higher in samples where *DEAR1* expression is below the median (p<0.0001).
significantly upregulated in cancers of the TNBC subtype compared to the non-TNBC group (p=0.0001, Figure 4.5D). I also observed an inverse correlation between DEAR1 and SNAI2 expression in breast cancer samples (p<0.0001, Figure 4.5E).

Since there is an evident correlation between DEAR1 expression and TNBC/basal-like phenotype, I analyzed DEAR1 expression in a TNBC tissue microarray (TMA) that consisted of 153 breast cancer samples, 103 of which are TNBC. The majority of patients in this cohort were Caucasian (65%), HER2 negative (98%) and diagnosed at an early stage. Details are summarized in Figure 4.6. Statistical analysis to correlate DEAR1 expression with clinical status and outcomes was performed using all TMA samples (n=153) and in TNBC-only cases (n=103). Results indicate that DEAR1 expression, when dichotomized by the median or distributed into tertiles, is significantly associated with age of diagnosis only in TNBC in which loss of DEAR1 correlated with a younger age of diagnosis (p=0.04 and p=0.01, respectively, Figure 4.7A). When age was dichotomized into ‘younger than 50 years of age’ and ’50 years or older’, loss of DEAR1 significantly associated with patients younger than 50 years of age (p=0.03, Figure 4.7B). There was also a statistically significant correlation between DEAR1 protein expression and time to metastasis in TNBC cases when DEAR1 expression was dichotomized by the median (HR=0.41, 95% CI: 0.17-1.03, p=0.05, Figure 4.7C). Collectively, this data indicates the importance of DEAR1 expression in IDC, specifically in TNBC. With further investigations, we hope to utilize DEAR1 and SNAI2 expression to predict patient outcomes and stratify patients for stringent follow-up and treatment.

4.3 Discussion

At the genomic level, we observed frequent DEAR1 LOH, a common genetic event in cancer that could potentially unmask somatic mutations in tumor suppressor genes when the wildtype allele is lost, and SNAI2 upregulation in the METABRIC invasive breast carcinoma cohort (n=2509). While DEAR1 LOH and/or mRNA downregulation, but not SNAI2 alterations,
<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>N=153</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final DEAR1 (cytoplasm), median (min, max)</td>
<td>N = 153 40 [0, 207]</td>
</tr>
<tr>
<td>Final DEAR1 (membrane), median (min, max)</td>
<td>N = 153 22 [0, 195]</td>
</tr>
<tr>
<td>Tumor Size, median (min, max)</td>
<td>N = 153 2.2 [0.4, 15]</td>
</tr>
<tr>
<td>ER, No. (%)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>106 (69%)</td>
</tr>
<tr>
<td>Positive</td>
<td>46 (30%)</td>
</tr>
<tr>
<td>unknown</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>PR, No. (%)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>139 (91%)</td>
</tr>
<tr>
<td>Positive</td>
<td>13 (8%)</td>
</tr>
<tr>
<td>unknown</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>HER2, No. (%)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>150 (98%)</td>
</tr>
<tr>
<td>Positive</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>unknown</td>
<td>2 (1%)</td>
</tr>
<tr>
<td>TNBC, No. (%)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>49 (32%)</td>
</tr>
<tr>
<td>Yes</td>
<td>103 (68%)</td>
</tr>
<tr>
<td>Age at Diagnosis, median (min, max)</td>
<td>N = 153 54 [26, 83]</td>
</tr>
<tr>
<td>Race, No. (%)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>8 (5%)</td>
</tr>
<tr>
<td>Black</td>
<td>25 (16%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>17 (11%)</td>
</tr>
<tr>
<td>unknown</td>
<td>3 (2%)</td>
</tr>
<tr>
<td>White</td>
<td>100 (65%)</td>
</tr>
<tr>
<td>Pathological Stage, No. (%)</td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>60 (39%)</td>
</tr>
<tr>
<td>IB</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>IIA</td>
<td>43 (28%)</td>
</tr>
<tr>
<td>IIB</td>
<td>19 (12%)</td>
</tr>
<tr>
<td>IIA</td>
<td>16 (10%)</td>
</tr>
<tr>
<td>IIB</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>IIC</td>
<td>4 (3%)</td>
</tr>
<tr>
<td>IV</td>
<td>2 (1%)</td>
</tr>
<tr>
<td>unknown</td>
<td>7 (5%)</td>
</tr>
</tbody>
</table>

**Table 4.1: Patient demographics and tumor characteristics of TNBC TMA.** DEAR1 protein expression in the membrane and cytoplasm was analyzed on a TMA TNBC cohort with Aperio Technologies. Statistical analysis was performed to determine if there was an association between DEAR1 expression and clinical status and outcomes, including tumor size, age, race, and stage.
Figure 4.7: Loss of DEAR1 protein expression significantly correlates with younger age at diagnosis and shorter time to metastasis in TNBC patients. A) Distribution of age at diagnosis in TNBC patients by DEAR1 expression dichotomized by the median or categorized by tertiles of distribution. B) Distribution of DEAR1 expression by age at diagnosis dichotomized at 50 years of age in TNBC patients. C) Kaplan Meier curve for time to metastasis by DEAR1 expression dichotomized by the median in TNBC patients.
significantly correlates with poor OS, *DEAR1* LOH/downregulation in addition to *SNAI2* amplification/overexpression associates with poor OS at greater significance than *DEAR1* alterations alone. These results indicate that *DEAR1* and *SNAI2* expression levels are good predictors of overall survival in invasive breast cancer patients. Since I observed a correlation between *DEAR1* and *SNAI2* expression with respect to a stem/progenitor cell phenotype, it would be meaningful to study if we can utilize their expression to predict chemotherapy response, risk of progression from DCIS to IDC, or stratify patients for therapies that target *DEAR1*-mediated pathways, such as EMT and *SNAI2*.

The *in vitro* data collected thus far indicates that *DEAR1* plays a role in regulating stem/progenitor cell properties, potentially through mediating *SNAI2* expression. The EMT and stem cell phenotype and *SNAI2* overexpression have been related to the basal-like subgroup of the TNBC subtype of breast carcinomas, which exhibit aggressive behavior, poor prognosis, and high risk of metastasis compared to other breast cancer subtypes. Given that *DEAR1* protein expression in a cohort of young female patients (under 50 years of age) with invasive breast carcinoma associated significantly with TNBC and strong family history, I evaluated *DEAR1* and *SNAI2* mRNA expression in TNBC samples using published annotated genomic data collected online from the bc-GenExMiner v4.1 tool. *DEAR1* mRNA expression demonstrated an inverse correlation with *SNAI2* mRNA expression in basal-like cancer and TNBC; *DEAR1* was found to be downregulated in basal-like and TNBC compared to other luminal A, luminal B, HER2-positive, and normal-like breast carcinomas. This provides further evidence that supports the *in vitro* observations demonstrating *DEAR1*'s role in mediating progenitor cell properties and overexpression of *SNAI2* with loss of *DEAR1* in HMECs. I further demonstrate that downregulation of *DEAR1* protein expression significantly correlated with early age of onset in the TNBC TMA, which supports previous data that showed loss of *DEAR1* expression associates with a strong family history of invasive breast carcinoma. Furthermore, loss of *DEAR1* protein expression, when dichotomized by the median, revealed a significant association with shorter time to metastasis in TNBC patients; patients with *DEAR1* expression higher than the median
showed a 59% less risk of metastasis compared to patients with DEAR1 expression lower than the median.

Cumulatively, there is strong evidence demonstrating the importance of DEAR1 copy number loss and expression, in particular with cases of TNBC, in predicting clinical outcome. This data indicates that DEAR1, in addition to SNAI2, can be a potential biomarker to determine prognosis as well as risk of developing early-onset breast cancer and/or metastasis.
CHAPTER 5: Post-transcriptional regulation of DEAR1 by microRNAs
5.1 Introduction

MicroRNAs (miRNAs) are short interfering RNA molecules, approximately 21-25 bases long, which regulate gene expression by binding to target messenger RNA (mRNA) transcripts at the 3’ untranslated region (UTR) to inhibit translation or promote mRNA degradation by nucleases. The first miRNA, lin-4, was discovered and characterized in *C. elegans* in 1993. A few years later, the second miRNA, let-7, was discovered. Since then, thousands of miRNAs were identified in several model organisms, including plants, flies, rodents, and primates [145, 146]. In mammals, primary miRNA transcripts are processed in the nucleus by the Drosha/DCGR8 complex and are exported to the cytoplasm as pre-miRNAs to be further processed by Dicer RNA III endonuclease. The mature miRNA forms a complex with Argonaute, a component of the RNA-induced silencing complex (RISC), to mediate target mRNA binding and regulation [145-147]. While complete complementarity to the target mRNA sequence results in mRNA degradation, partial complementarity results in translational inhibition. The 5’ end of the mature miRNA contains the “seed region” that refers to a sequence of 6-8 bases that exhibits complementary base pairing with the target mRNA sequence located in the 3’UTR. The seed region is crucial for recognition of target mRNA and is used by target prediction programs to identify potential targets [146].

MiRNAs function in many cellular pathways including those involved in cell cycle progression, growth, proliferation, differentiation, and cell death. Approximately 50% of genes encoding miRNAs are located in fragile regions within chromosomes that are vulnerable to genomic alterations. De-regulated miRNAs are often observed in many cancer types, including pancreatic, gastric, lung, and breast cancers [145, 148-150]. Downregulation of tumor suppressive miRNAs that target oncogenes or upregulation of oncomirs that target tumor suppressor genes involved in regulating proliferation and invasion can result in tumor formation and metastasis [145, 146, 150, 151]. For example, the miR-200 family is involved in modulating epithelial-mesenchymal transition (EMT), which is implicated in promoting invasive properties of cancer cells, by downregulating expression of the E-cadherin repressors ZEB1/2. In many
invasive cancers, miR-200 family members are downregulated, allowing for ZEB1/2 activity to repress E-cadherin expression and promote EMT progression [53, 150]. MiR-21 targets multiple tumor suppressor genes, including PTEN and Bcl2, and is upregulated in many cancers including breast cancer, gastric cancer, glioblastoma, pancreatic cancer, and cholangiocarcinoma. MiR-221 is another EMT promoter and is upregulated in a number of cancers including glioblastoma, thyroid cancer, breast cancer, and pancreatic cancer [146, 150-153].

The mechanism by which DEAR1 is downregulated in DCIS and IDC remain unclear; methylation assays reveal that the DEAR1 promoter is not hypermethylated and in some cell lines, DEAR1 mRNA is generated although protein expression is not detected. Currently no miRNAs have been validated to regulate DEAR1 post-transcriptionally. Based on data from literature and miRNA target prediction programs, miR-10b and miR-196b are of particular interest. MiR-10b is upregulated in metastatic breast cancer and targets TIAM1, HOXD10, and RHOC. High levels of miR-10b, induced by EMT initiation and transcription factor Twist, promotes cell invasion by the suppression of HOXD10 and RHOC [148, 149, 154]. MiR-196b became a candidate after a screening of miRNAs in pancreatic cancer cell lines in which it was noted that pancreatic cancer cell lines lacking DEAR1 expression showed high levels of miR-196b. MiR-196b overexpression significantly correlates with poor prognosis in gastric cancer [155] and its overexpression contributes to hematopoietic stem cell maintenance and immortalization of bone marrow progenitor cells in mixed lineage leukemia [156]. Furthermore, overexpression of both miR-10b and miR-196b was observed in tumor-associated endothelial progenitor cells that generate vasculature in high-grade invasive ductal carcinoma [157]. In addition, both candidate miRNAs are predicted to bind the 3'UTR of DEAR1 based on publically available online miRNA target prediction tools (Figure 5.1) [158, 159]. Discovery of a miRNA or group of miRNAs that control DEAR1 levels may lead to a better understanding of DEAR1’s function, explain how DEAR1 expression is lost in breast cancer, and serve as an avenue for therapeutic interventions.
Figure 5.1: miR-10b and miR-196b were predicted by at least three prediction algorithms to target *DEAR1*. A) An online search tool compares results across five target prediction algorithms programs. MiR-10b is predicted to target *DEAR1* in 4 of 5 programs; miR-196b is predicted to target *DEAR1* in 3 of 5 programs. B) Sequence alignment of *DEAR1* 3’UTR with miR-10b and miR-196b. Both miRNAs harbor 7 base complementation with *DEAR1* in the critical 5’ miRNA seed region.
5.2 Results

To determine if miR-10b and miR-196b are valid miRNA candidates that regulate DEAR1 expression, I first analyzed miR-10b and miR-196b expression via qRT-PCR analysis in cell lines that express or lack DEAR1. 76N-F2V, 76N-E6, and MCF10A immortalized, non-tumorigenic human mammary epithelial cell lines demonstrate high DEAR1 protein expression whereas SKBR3, TD474, MCF7, BT20T, H1299, and HCT116 cancer cell lines do not express detectable DEAR1 protein by western blot analysis (Figure 5.2A). q-PCR results reveal that miR-10b is overexpressed in 5 of 6 cell lines that lack DEAR1 protein expression compared to DEAR1-expressing control cell lines (SKBR3: 4994.43 fold change, p<0.0001; T47D: 576.87 fold change, p<0.01; BT20T: 13.65 fold change, p<0.01; H1299: 1169.03 fold change, p<0.0001; Hct116: 832.03 fold change, p<0.0001 vs 76N-F2V; Figure 5.2B). Similarly, miR-196b is overexpressed in 4 of 6 cell lines that lack DEAR1 expression (SKBR3: 13.11 fold change, p<0.0001; T47D: 3.07 fold change, p<0.01; MCF7: 1.95 fold change, p<0.05; Hct116: 4.13 fold change, p<0.01; Figure 5.2C).

To evaluate if each candidate miRNA has affinity to the predicted binding site in the 3'UTR of DEAR1 and if binding to this region leads to translational repression, I utilized a pMIR-REPORT luciferase expression vector. 60bp oligo strands containing the candidate miRNA binding region with restriction enzyme sites flanking the 5' and 3' ends were designed for ligation into the multiple cloning site (MCS) at the 3' end of the luciferase reporter. Oligos contained either the putative binding site found in the 3'UTR of DEAR1 or the standard binding site in which there is perfect complementation to a specific miRNA candidate as positive controls. The luciferase-fusion plasmid and miRNA mimics were co-transfected into HEK293T cells to determine effects on luciferase activity. Results demonstrate that both miR-10b and miR-196b significantly inhibit luciferase activity (p<0.0001, Figure 5.3A&B). Similarly, when the entire DEAR1 3'UTR region was cloned into the pMIR-REPORT vector, addition of miR-10b and miR-196b mimics resulted in significant reduction of luciferase activity compared to respective negative controls (p<0.0001,
Figure 5.2: miR-10b and miR-196b are upregulated in cancer cell lines that lack DEAR1 protein expression. A) Western blot analysis of immortalized, non-tumorigenic human mammary epithelial control cell lines that express DEAR1 and multiple cancer cell lines that do not express detectable DEAR1 protein. B-C) Q-PCR analysis of miR-10b and miR-196b expression in cell lines. Both miRNAs are significantly overexpressed in the majority of cancer cell lines that lack DEAR1 protein expression. U6 snRNA was used as the internal control for miRNA qPCR assays. P-value *=0.05-0.01, **<0.01, ****<0.0001.
Figure 5.3: Expression of miRNA mimics significantly reduces luciferase activity in HEK293T cells containing putative and standard miRNA binding sites. A) HEK293Ts co-transfected with 25ng pMIR-REPORT luciferase vector containing either putative or standard DEAR1 3’UTR sequence and increasing concentration of miR-10b mimic exhibited significant inhibition of luciferase activity compared to cells transfected with a negative control mimic. B) HEK293Ts co-transfected with pMIR-REPORT vector and miR-196b mimic also demonstrated significant reduction of luciferase activity compared to respective controls. Results indicate that miR-10b and miR-196b are capable of binding to and inhibiting luciferase mRNA to reduce protein expression and activity.
To test if overexpression of miR-10b or miR-196b in cell lines that express DEAR1 will affect protein expression, miR-10b and miR-196b mimics were transfected into CCH-1 and MCF10A cell lines. Results indicate that both miRNAs reduce DEAR1 protein expression in both cell lines at various time points (Figure 5.5). This data demonstrates that overexpression of either miR-10b or miR-196b inhibits luciferase activity, results in inhibition of DEAR1 protein expression, and thus are potential candidates for regulating DEAR1 expression.

To show that miRNA binding and subsequent translational repression is dependent on the seed region identified in the DEAR1 3’UTR, a mutant luciferase reporter construct was generated that harbored a 4-base mutation within the putative miR-10b seed region. Luciferase assays were performed in HEK293T cells using the standard, putative, and mutant pMIR-REPORT constructs. In two separate trials, cells co-transfected with mutant pMIR-REPORT vector and miR-10b mimic resulted in the rescue of luciferase activity, which was not significantly different compared to the negative control (Figure 5.6). This demonstrates that inhibition of luciferase activity is dependent on miR-10b and its ability to bind to the seed region within the DEAR1 3’UTR and provides further evidence that miR-10b is a candidate regulator of DEAR1 expression.

**4.3 Discussion**

MiRNAs are short non-coding RNA strands that were first discovered and characterized in worms in the early 1990s. MiRNA sequences and functions are conserved across multiple species and there are currently over 2000 known miRNAs in the human genome [160]. The role that miRNAs play in regulating gene expression has profound and complex effects on major pathways governing development, cell cycle, and carcinogenesis [161].

Studies of differential miRNA expression in breast cancer revealed a panel of deregulated miRNAs that can be utilized to predict subtypes and clinical outcomes.
Figure 5.4: Expression of miRNA mimics significantly reduces luciferase activity in HEK293T cells containing the entire DEAR1 3’UTR sequence containing miRNA binding sites. A) HEK293Ts co-transfected with 25ng pMIR-REPORT luciferase vector containing DEAR1 3’UTR sequence and increasing concentration of miR-10b mimic demonstrated significant repression of luciferase activity compared to cells transfected with a negative control mimic. B) Similar statistically significant results were obtained with HEK293T cells co-transfected with pMIR-REPORT and increasing concentration of miR-196b mimic. This indicates that both miRNAs have affinity to the 3’ UTR of DEAR1 to cause downregulation of luciferase expression.
Figure 5.5: Transient transfection of miRNA mimics into cell lines that express DEAR1 protein result in reduction of DEAR1 expression. A) CCH-1 DCIS cell line transfected with 200nM miR-10b and miR-196b mimics demonstrated reduction of DEAR1 protein expression 48 hours post-transfection. MiR-10b resulted in 30% reduction while miR-196b resulted in 60% reduction, as calculated using ImageJ software. B) MCF10A cell line transfected with 200nM miR-10b mimic also demonstrated 40% reduction of DEAR1 protein expression 48 hours post-transfection whereas transfection with 200nM miR-196b mimic resulted in 40% reduction of DEAR1 protein 72 hours post-transfection. These results show that ectopic expression of miR-10b and miR-196b can inhibit DEAR1 protein expression, making them potential miRNA regulators of DEAR1.
Figure 5.6: Mutations in the putative miR-10b binding site in DEAR1 3’ UTR results in rescued luciferase activity. Two-step PCR mutagenesis was performed using the pMIR-REPORT putative vector to generate a 4-base alteration within the predicted miR-10b seed region. This mutation eliminated repression of luciferase translation and activity. This result indicates that miR-10b binds to this particular region in the DEAR1 3’UTR to repress translation and provides evidence that miR-10b can regulate DEAR1 expression.
Overexpression of miR-155, miR-21, and miR-10b oncomirs is often observed in breast carcinomas. Several groups have demonstrated downregulation of miR-200 family, miR-205, and miR-34a that affect EMT-regulators such as ZEB1/2 and SNAIL1 [146, 147, 151, 152, 154, 156, 160, 161]. A review by Kurozumi and colleagues summarizes a list of miRNAs associated with breast carcinogenesis and miRNA expressions that correlate with breast cancer subtypes. For example, luminal A subtypes often present overexpression of miR-10a, let-7c, let-7f, and miR-191 miR-15b and downregulation of miR-206, miR-15b, and miR-107 while basal-like tumors present overexpression of miR-155, miR17-92, and miR-93 and downregulation of miR-29 and miR-190b [161]. These studies demonstrate an interesting case in favor of utilizing miRNA expression to predict patient prognosis and options for personalized therapy.

Mechanisms behind DEAR1 regulation remain elusive. Since miRNAs are often deregulated in breast carcinomas and DEAR1 is downregulated in DCIS and IDC, we sought to identify candidate miRNAs that might regulate DEAR1 expression post-transcriptionally. Online prediction tools generated several candidate miRNAs based on seed region alignment to the DEAR1 3' UTR. Among the list, miR-10b and miR-196b were of interest based on published literature. Both miR-10b and miR-196b were overexpressed in progenitor endothelial cells associated with tumor vasculature [157]. MiR-10b and miR-196b are overexpressed in most cancer cell lines that lack DEAR1 expression compared to normal human mammary epithelial cell lines that express DEAR1 protein. To test each miRNA’s ability to inhibit gene expression, HEK293T cells were co-transfected with miRNA mimics and a luciferase reporter containing the DEAR1 3' UTR. Results consistently demonstrate that ectopic expression of either miR-10b or miR-196b led to significant repression of luciferase activity compared to cells transfected with a negative control miRNA. Furthermore, transient transfection of either miRNAs in CCH-1 and MCF10A cell lines, which express DEAR1 protein, resulted in reduction of DEAR1 expression.

The gene encoding miR-10b on chromosome 2 is frequently amplified in multiple cancer types, including ovarian, head and neck, pancreatic, prostate, cervical, and breast carcinomas.
Li Ma and colleagues identified miR-10b as an important regulator of breast cancer metastasis both in vitro and in vivo. Interestingly, miR-10b is downregulated in early stage breast carcinomas compared to normal mammary tissue but is overexpressed in some metastatic lesions and cell lines [149]. MiR-10b is also upregulated in other metastatic and invasive tissues including pancreatic and hepatocellular carcinomas, glioblastomas, and neurofibromas [148]. Studies have demonstrated that miR-10b overexpression through TGF-β signaling or TWIST upregulation occurs in late stages of carcinogenesis. Although inhibition of miR-10b with antagonirs did not result in significant changes in primary tumor formation, it led to a significant reduction in lung metastases in mice. Expression of miR-10b also associated with increased stage, lymph node involvement, and higher Ki-67 staining [162]. Since loss of DEAR1 protein expression associates with shorter time to metastasis and with the TNBC subtype, which is at higher risk of aggressive behavior and invasion, miR-10b seems to be a logical negative regulator of DEAR1.

MiR-196b has not been as well studied as miR-10b on its role in breast cancer, however it has been shown to be overexpressed in hematopoietic cancers and associates with poor OS in patients with AML [163] and loss of DEAR1 correlates with poor OS and shorter survival in the same cancer type [105, 106]. Furthermore, overexpression of miR-196b is frequent in colorectal carcinomas and associates with poor prognosis. Additionally, upregulation of miR-196b in colorectal carcinoma cells resulted in increased cancer stem cell properties by activating STAT3 signaling through targeting negative regulators of the JAK/STAT pathway, SOCS1 and SOCS3 [164]. This correlation between miR-196b and stemness is of interest given that we have evidence that DEAR1 also plays a role in mediating stemness in vitro in mammary epithelial cells.

Results herein indicate that miR-10b and miR-196b can target DEAR1 3’ UTR to inhibit protein expression in multiple cell lines. Although more experiments are required to determine if DEAR1 is a bonafide target of miR-10b and/or miR-196b, these findings provide a foundation by
which to continue our investigations into the regulatory mechanisms governing DEAR1 expression.
CHAPTER 6: Materials and methods
76N-E6 human mammary epithelial cells were grown in D-medium which consists of 500mL Minimum Essential Medium Alpha (Corning, MT10022CV), 500mL Ham’s F12K medium (Corning, MT10025CV), 1% FBS (Gibco, 10437028), 2mM L-glutamine (Gibco, 25030081), 10mM HEPES (Gibco, 15630080), 10mg L-ascorbic acid (Sigma, A4403-100MG), 35mg bovine pituitary extract (Life Technologies, 13028), 12.5ng/mL epidermal growth factor (Sigma, E9644), 1ng/mL β-estradiol (Sigma, E2257), 0.0006% ethanolamine (Sigma, E0135), 1ug/mL hydrocortisone (Sigma, H0396), 1ug/mL insulin (Sigma, I-1882), 2.5ng/mL Na selenite (Sigma, S9133), 10mg O-phosphoethanolamine (Sigma, P0503-10MG), 10mg apo-transferin (Sigma, T5391-10MG), and 1ng/mL triiodothyronine (III) (Sigma, T6397). DEAR1-KD and DEAR1-SMAD3 double KD clones were generated previously in the lab by Dr. Nanyue Chen [102]. HEK 293T cells purchased from the MD Anderson characterized cell line core (CCLC) were grown in DMEM (Corning, MT10017CV) and 5% FBS (Sigma, F8192). MCF10A cells were grown in DMEM/F12 50/50 (Corning, MT10090CV) supplemented with 5% horse serum (Gibco, 16050114), 10ug EGF, 0.25mg hydrocortisone, 5mg insulin, and 50ug cholera toxin (CalBioChem, 227036). DCIS.COM cells were propagated in DMEM/F12 medium supplemented with 5ug/mL horse serum, 10ug/mL insulin, and 0.5ug/mL hydrocortisone. Cells were trypsinized with 1X TrypLE Express (Gibco, 12605028) for dissociation and passaging.

Western blot analysis

To obtain whole cell lysates, cells were grown in 2D culture as described above, harvested using 1X SDS sample buffer with 50mM DTT, and sonicated before protein measurements were recorded. Proteins were loaded into a 4-12% SDS-PAGE gradient gel (ThermoFisher Scientific, NW04120BOX), run on the Bolt Mini Gel Tank, and transferred onto a nitrocellulose membrane (BioRad, 1620094). DEAR1 antibody was produced by Bethyl Laboratories (Montgomery, TX, United States). SNAI2/SLUG (C19G7, #9585) and SMAD3 (C67H9, #9523) antibodies were
purchased from Cell Signaling (Danvers, MA, United States). Westerns were normalized using β-actin (A5441) purchased from Sigma (Saint Louis, MO, United States).

**Plasmids**

The pcDNA-DEAR1 and pCMV-HA-DEAR1 constructs were generated previously in the lab [100]. The pCMV-HA-DEAR1-C11A RING mutant vector was generated by Dr. Balasenthil using the Stratagene QuikChange Site-Directed Mutagenesis kit (210518). The SNAI2 cDNA was amplified from the pEGFP-C2-SLUG vector from Dr. Togo Ikuta at the Research Institute for Clinical Oncology using PCR and ligated into a pcDNA-6Myc vector digested with EcoRI and XhoI. pMT123-8x-HA-Ubiquitin was a generous gift from Dr. Dirk Bohmann from Rochester Medical Center. 3x-Flag-Ubiquitin was provided by Dr. Garrison Fathman at Stanford University. pGL3-SNAI2 promoter luciferase plasmid was provided by Dr. Togo Ikuta. For miRNA studies, pMIR-REPORT luciferase plasmid was purchased from Addgene. Luciferase readings were normalized to β-galactosidase (Promega).

**Primers**

To amplify SNAI2 from the pEGFP-C2 backbone via PCR, forward and reverse primers were designed to include EcoRI and XhoI restriction enzyme cut sites, respectively, for ligation into the pcDNA-6Myc backbone (FW primer, 5'-TAACGAAGAATTCATGCCGCGCTCCTTC-3'; RV primer, 5'-TCGTTACTCGAGTCGTGCTACACAGCA-3'). To clone the DEAR1 3'UTR region, forward and reverse primers were designed to include SpeI and HindIII restriction enzyme cut sites, respectively, for ligation into the pMIR-REPORT luciferase plasmid (FW primer, 5'-TAGTCAACTAGTAACACCGGTCCGCATCTAGTC-3'; RV primer, 5'-GCCGGACAAAGCTTTGATCATAACTCATAAAATGGGCAA-3'). Short DNA oligos containing the specific DEAR1 3'UTR sequences where the miRNAs of interest were predicted to bind (termed “putative”) as well as the exact sequences where the miRNAs were predicted to bind (termed
“standard”) were ligated into pMIR-REPORT. To generate mutations in miRNA binding region in the 3’UTR, primers were designed to include a four-base changes.

Co-immunoprecipitation and ubiquitination assays

For co-immunoprecipitation (Co-IP) assays, HEK293T cells were transfected with Myc-SNAI2 and/or HA-DEAR1 plasmids in a 6cm plate, treated with MG-132 proteosome inhibitor (Calbiochem, 133407-82-6) for 2 hours, and harvested at 24 hours post transfection. Mirus TransIT-LT1 transfection reagent (Mirus, MIR 2300) was used at a ratio of 1:3 (DNA:Mirus). Cells were lysed with M-PER lysis buffer (ThermoFisher Scientific, 78501) and incubated with pull-down antibody overnight at 4°C. The following morning, lysates were incubated with protein A/G agarose beads (Santa Cruz, sc-2003) for 2 hours at 4°C followed by three 5-minute washes with RIPA lysis buffer. Proteins were eluted from beads in 30uL of 2X SDS sample buffer for western blot analysis. Anti-HA (H6908) or anti-Myc (M4439) from Sigma were used for pull-down and anti-HA (H3663) from Sigma or anti-Myc (#562) from MBL International Corporation were used for immunoblotting. Ubiquitination assays were set up in a similar manner using HEK293T transfected with Myc-SNAI2, 8x-HA-Ub, and/or DEAR1 plasmids in a 10cm plate. When using the HA-DEAR1-C11A RING mutant plasmid, 3x-Flag-Ub was substituted for 8x-HA-Ub. Cells were harvested at 48 hours post transfection, lysed with 1X SDS, and denatured at 95°C for 10 minutes. Cell lysates were diluted with NP-40 buffer prior to western blot analysis.

Real time quantitative PCR

Cells were treated with 4ng/mL TGF-β1 (EMD Millipore, 616450-1UG) for 3 and 40 hours; cells were harvested for RNA extraction using High Pure RNA Isolation Kit (Roche, 11828665001), followed by a quality check using MD Anderson’s Microarray Core. RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription kit (Applied Biosystems, 4368814). Real time quantitative PCR was performed using TaqMan® Universal Master Mix II, no UNG (Applied Biosystems, 4440047) and TaqMan® probes for TRIM62 (Hs00217089_m1), SNAI2
(Hs00950344_m1), ZEB1 (Hs00232783_m1), ZEB2 (Hs00207691_m1), and GAPDH (Hs02786624_g1) in triplicate and in MicroAmp optical 384-well plates (Applied Biosystems, 4309849). ΔΔCt values were normalized to GAPDH. For miRNA expression, RNA was collected using the miRNeasy Mini Kit (Qiagen, 217004) and reverse transcribed using TaqMan™ MicroRNA Reverse Transcription Kit (Applied Biosystems, 4366596). ΔΔCt values were normalized to U6 snRNP.

**Luciferase Assay**

293T cells were plated at a density of 100,000 cells in a 12-well flat bottom cell culture plate (Corning, 3513) and allowed to adhere overnight. Culture medium was replaced prior to transfection. Cells were transfected using Mirus TransIT-LT1 reagent with 50ng pGL3-SNAI2 luciferase vector, 10ng β-galactosidase, and either pcDNA empty vector or pcDNA-DEAR1 (200, 300, and 400ng). Cells were lysed with 1X Passive Lysis Buffer (Promega, E1941). The Luciferase Assay System (Promega, E1501) was used to measure luciferase activity. β-galactosidase activity, for normalization, was measured using the Galacto-Light Plus Reporter Gene Assay System (Applied Biosystems, T1007). Activities were measured using a 20/20\textsuperscript{n} Luminometer (Turner Biosystems, 998-2036). For miRNA studies, 293T cells were transfected using DharmaFECT transfection reagent (Dharmacon, T-2001-01) with 25ng of pMIR-REPORT (control, DEAR1 3'UTR, putative, or standard), 10ng β-galactosidase, and 100 – 250ng miRNA mimics or 250ng control mimic (Invitrogen, 4464066).

**Mammosphere assay**

The mammosphere protocol was given to us by Dr. Sendurai Mani’s Lab and modified from Dontu et al. [57, 58]. Human mammary epithelial cells were grown in 2D culture in respective normal growing conditions. Cells pre-treated with TGF-β were treated with 4ng/mL TGF-β1 for 48 hours prior to seeding cells in mammosphere growing conditions. Mammosphere medium was made using Mammary Epithelial Growth Medium (LONZA, CC-3051), 1% methylcellulose (400cp,
Sigma, M0430-100G), 20ng/mL bFGF (Corning, 47743-574), 10ng/mL EGF (Sigma, E9644-2MG), and 4ug/mL heparin (Sigma, H3149-50KU), and filtered through a 0.45uM filter (VWR, 28145-481) before use. Cells grown and treated in 2D culture were trypsinized with 1X TrypLE Express and counted using a hemocytometer. Cells were seeded at 1x10^3 cells per well in flat bottom ultra-low attachment 96-well plates (Corning, 3474). Mammospheres were replenished with fresh medium every two days; mammospheres were counted and imaged using a Zeiss light microscope at 10X magnification on Day 14. Secondary and tertiary mammospheres were generated by harvesting primary mammospheres, chemical dissociation with 1X TrypLE Express, and mechanical dissociation with a pipette to achieve single cell suspension. Cells were re-counted and seeded at a density of 500 cells per well.

Statistical analysis

Graphs and statistical analyses were performed using GraphPad Prism 6. Statistical analysis was performed using one-way ANOVA multiple comparison or student’s t-test. Statistical significance refers to p-values less than or equal to 0.05 (*=0.01-0.05, **<0.01, ***<0.001, ****<0.0001).

Immunofluorescence

Whole mammospheres were collected and resuspended in 100uL of 1X PBS (Corning, 21040CV) for cyto spin using Shandon Cytospin 4 (Thermo Scientific, 1000rpm, 8min) onto Shandon Cytospin slides (Thermo Scientific, 5991056) using filter cards (Thermo Scientific, 5991022), then fixed with Safetex cytology spray fixative (Andwin Scientific, PK6000). Slides were stained using standard immunofluorescence staining protocol using 0.3% Triton X-100 in PBS for permeabilization and blocked with anti-Goat serum for one hour at 4C. Anti-CK8/18 (rabbit, Abcam, cat#: ab53280-uL, 1:100), anti-CK5 (mouse, Pierce, cat#: MA5-17057, 1:100), and ALDH1A1 (rabbit, Abcam, cat#: ab52492-10ul, 1:100) were used to detect co-expression of luminal and basal cytokeratins and a marker of stem/progenitor cells. Secondary antibodies goat
anti-mouse (Alexa 488) and goat anti-rabbit (Alexa 555) were used. Nuclei were stained with DAPI (Sigma, 236276). Slides were treated with ProLong Gold anti-fade fluorescent mounting agent (ThermoFisher Scientific, P10144). Fluorescent images were taken using OLYMPUS cellSens Software.

Tumor tissue microarray

In collaboration with Dr. Ignacio Wistuba, Dr. Aysegul Sahin, and Dr. Fei Yang from the Department of Translational Molecular Pathology at MDACC, DEAR1 staining was analyzed on a tissue microarray of 180 breast cancer samples, 103 of which are triple negative breast cancers, using Aperio Technologies. We obtained archival, formalin-fixed and paraffin-embedded (FFPE) material from surgically resected breast cancer specimens from the Breast Tumor Bank at M. D. Anderson Cancer Center from 2001 to 2013 (Houston, TX). Tumor tissue specimens obtained from 180 breast cancers were histologically examined, classified using the World Health Organization (WHO) classification of Breast Tumors and selected for TMA construction. After histologic examination, tumor TMAs were prepared using triplicate 1-mm-diameter cores per tumor. Detailed clinical and pathologic information, including demographic, pathologic TNM staging, overall survival, and time of recurrence was collected. Dr. Suyu Liu and Dr. Laura Rubin from the Department of Biostatistics at MDACC performed statistical analyses to determine if DEAR1 expression correlates with clinical status. Wilcoxon or Kruskal-Wallis rank sum tests were used to study the association between categorical and continuous variables. Survival curves were estimated using the Kaplan-Meier method and differences in survival between groups were assessed using two-sided log-rank tests. Hazard ratios were obtained from univariate Cox proportional hazards regression models. Overall survival and time to metastasis were defined as the time between date of surgery and date of death or date of metastasis, respectively. R version 3.4.2 was used to perform statistical analyses.
CHAPTER 7: Discussion and future directions
Discussion

Breast cancer affects hundreds of thousands of women worldwide. In the United States, breast cancer remains the most commonly diagnosed cancer in women annually [1]. Triple negative breast cancer is the most aggressive subtype of breast cancer, characterized by invasive potential, high risk of metastasis, shorter survival time, and overall poor prognosis [6, 10, 11]. Subgroups of the TNBC subtype include basal-like, immunomodulatory, mesenchymal, mesenchymal stem cell-like, and luminal androgen receptor [165]. Each subgroup is distinguished by differential gene expression profiles. For example, basal-like tumors associate with overexpression of cell cycle genes, DNA damage regulators, and BRCA1 mutations whereas the mesenchymal subgroup is associated with overexpression of EMT regulators [161, 165]. Prognosis of breast cancer when detected early is favorable with a 98.7% 5-year survival rate. However, once the disease has metastasized, the 5-year survival rate drops to 27% [1]. It is crucial that we identify and understand molecular pathways that are involved in the invasion and metastatic process in order better treat breast cancer patients.

One of the earliest events leading up to dissemination of cancer cells from the primary tumor is initiation of the EMT program and loss of apical-basal polarity that mediates invasive potential [32, 126, 141, 166]. DEAR1 is a tumor suppressor mapped to Chromosome 1p35.1, a region within the 1p interval that frequently undergoes loss of heterozygosity in breast cancer [4]. At the molecular level, DEAR1 was demonstrated as a crucial regulator of apical-basal cell polarity and required for proper acinar formation in vitro [100]. Shortly after, DEAR1 was discovered as a negative regulator of TGF-β-mediated EMT through binding and polyubiquitinating the canonical effector protein, SMAD3, leading to repression of transcriptional activation of EMT regulators [102]. In a cohort of 158 early-onset breast cancer patients under the age of 49 years, loss of DEAR1 protein expression significantly correlated with PR negative and TNBC status and associated with strong family history [100]. Presence of DEAR1 protein
expression was a positive indicator of recurrence-free survival after a 15-year follow-up in this cohort, further implicating the importance of DEAR1 in breast carcinogenesis.

The concept of cancer stem cells as tumor-initiating cells has been established for nearly five decades [67]. Within the last ten years, an overwhelming number of studies have connected initiation of EMT or overexpression of EMT transcription factors, including SNAI1/2, ZEB1, and FOXC2, to the generation of cancer stem cells [32, 57, 69, 76, 85, 127, 129, 131, 167, 168]. Since it was previously demonstrated in that DEAR1 mediates invasion and migration through regulation of TGF-β-induced EMT in mammary epithelial cells and that DEAR1 functions as a tumor suppressor by repressing invasive potential in lung adenocarcinoma and ovarian cancer cell lines [102, 104, 107], I hypothesized that DEAR1 may also regulate stem cell properties and that its expression can be utilized to distinguish breast cancer patients who are at risk for progression or metastasis.

To study DEAR1’s role in modulating the stem cell phenotype, I utilized previously established stable DEAR1 KD HMECs for mammosphere assays to assess self-renewal and differentiation properties. While confirming that TGF-β enhances the mammosphere phenotype in the context of DEAR1 downregulation in HMECs and that this is in part regulated through SMAD3 signaling, I also demonstrate that loss of DEAR1 expression enhances mammosphere formation independently of its function in regulating the canonical TGF-β and SMAD3 pathway. This is an important revelation given that not all cancer cells respond in the same manner to TGF-β stimulation. Since its discovery in the late 1970s, several published studies demonstrated the paradoxical influence of TGF-β signaling. While TGF-β is a potent inhibitor of cellular growth and proliferation through its inhibition of C-MYC and CDK4 expression, induction of cell cycle inhibitor p15INK14B, and promotion of apoptosis in several cell lines [37-39], others have shown that TGF-β can also promote cell proliferation in certain cancer cell lines and immortalized cell lines that harbor inactivated RB and TP53 tumor suppressors [44, 46, 120]. I observed this difference in response to TGF-β in the MCF10DCIS.COM cells grown in mammosphere conditions. All DCIS.COM cells treated with TGF-β and grown in mammosphere medium resulted
in a significant reduction of mammospheres compared to control cells treated with DMSO, possibly because DCIS is an early stage of breast cancer and the HMECs utilized in mammosphere assays were immortalized by HPV-E6, thus the growth inhibitory effects of TGF-\(\beta\) does not affect HMECs, which harbor inactivated p53. The mammosphere phenotype was also different in \(DEAR1\)-KD and -KO DCIS cells compared to \(DEAR1\)-KD HMECs in that loss of \(DEAR1\) in DCIS cells resulted in significantly larger mammospheres instead of a difference in mammosphere quantity. Despite the phenotypic difference between DCIS and HMEC mammospheres, larger mammospheres is indicative of a proliferative progenitor cell population [58, 112]. The data suggesting that \(DEAR1\) regulates progenitor cell properties in DCIS corresponds with the data that loss of \(DEAR1\) does not confer the ability to self-renew in HMECs although I identified bipotent and ALDH1-positive cells within HMEC mammospheres. Furthermore, loss of \(DEAR1\) in HMECs resulted in a larger population of EpCAM-/CD49f+ basal progenitor cells and reduction of CD24 luminal marker expression, indicating that \(DEAR1\) might regulate differentiation and progenitor cell properties. The data presented demonstrates that \(DEAR1\), a regulator of EMT, can affect progenitor cell properties without initiation of an EMT program through TGF-\(\beta\) signaling (Figure 7.1). Perhaps cells require additional alterations, such as upregulation of oncogenes, in addition to loss of \(DEAR1\) to generate stem-like cells.

Using qRT-PCR and western blot analysis, I identified SNAI2 as a candidate target of \(DEAR1\) regulation. SNAI2 mRNA and protein levels were elevated in \(DEAR1\)-KD HMECs at basal level in the absence of TGF-\(\beta\) treatment. I demonstrate that ectopic expression of \(DEAR1\) inhibits expression of a luciferase construct containing the SNAI2 promoter, providing further evidence that \(DEAR1\) negatively regulates SNAI2 expression. SNAI2 basal mRNA level was also unaffected by loss of SMAD3 expression in \(DEAR1\)-KD HMECs, indicating that SNAI2 expression can be modulated by other mechanisms independent from the canonical TGF-\(\beta\) pathway. Several studies have demonstrated that overexpression of EMT transcription factors regulate stem and/or progenitor cell properties independent of TGF-\(\beta\) signaling, specifically factors that are expressed and tightly regulated during embryogenesis including SNAI proteins,
Figure 7.1: **DEAR1 is a regulator of cell polarity, EMT, and stem/progenitor cell properties.**

DEAR1 functions as a critical regulator of acinar morphogenesis through regulation of apical-basal cell polarity and inhibitor of TGF-β-mediated EMT in HMECs. Initiation of EMT was shown to induce acquisition of stem/progenitor cell properties and generate CSCs that are resistant to standard chemotherapies and contribute to tumor recurrence. Our data reveals that DEAR1 can regulate stem/progenitor cell properties in HMECs partly through the TGF-β-SMAD3 signaling axis and independently through regulation of master EMT regulator and stemness factor, SNAI2. Furthermore, we identify potential miRNA regulators of DEAR1 protein expression: miR-10b and miR-196b.
which are overexpressed during gastrulation, in neural crest cells, and in undifferentiated mesenchymal cells [27, 72, 134]. SNAI2 expression is also associated with poorly differentiated, metastatic, basal-like and BRCA1 mutant tumors [72, 88]. Studies suggest that SNAI2 is an important regulator of a stem-like, undifferentiated state in cooperation with other transcription factors such as SOX9 [27, 87], and inhibits luminal differentiation through repression of luminal-specific markers such as CD24 and MUC1 [72]. As previously mentioned, Liu et al. observed that overexpression of DEAR1 in human ovarian cancer cell lines resulted in repression of SNAI2 expression mediated through c-Jun [107], providing a mechanism by which DEAR1 may regulate SNAI2 at the transcriptional level to inhibit stem cell properties. Alternatively, DEAR1 may regulate SNAI2 indirectly through stabilization of ERα, which inhibits SNAI2 transcription through recruitment of HDAC1 [136]. Co-IP and ubiquitin assays revealed that in addition to affecting SNAI2 transcription, DEAR1 also binds to and polyubiquitinates SNAI2. Preliminary data suggests that this interaction might not be directly through DEAR1’s RING domain, although it is quite clear that there exists an interaction and that presence of DEAR1 stimulates SNAI2 ubiquitination. Known post-transcriptional regulators of SNAI2 include p53/MDM2, Wnt/GSK3β, β-Trcp1, and BRCA1, although the precise mechanism of this interaction remains unclear [52, 88, 137]. Currently, experiments are underway to determine if the mammosphere phenotype in DEAR1-KD HMECs is regulated through SNAI2 by generating DEAR1-SNAI2 double KD clones. Preliminary data using pooled cells demonstrated a reduction of primary mammospheres compared to DEAR1-KD alone, however not at a significant level.

Another piece of evidence that supports the hypothesis that DEAR1 negatively regulates SNAI2 is the inverse relationship between DEAR1 and SNAI2 mRNA expression in breast cancer samples and that both loss of expression of DEAR1 and overexpression of SNAI2 are significantly associated with the TNBC subtype. Furthermore, according to patient data collected from the METABRIC cohort of 2,509 invasive breast carcinoma patients, DEAR1 undergoes loss of heterozygosity and/or mRNA downregulation in 17% of patients and significantly correlated with poor overall survival. SNAI2 is amplified and/or overexpressed in 12% of patients and has
a significant tendency to co-occur with DEAR1 alterations. Although alterations in SNAI2 alone do not confer a significant correlation with poor overall survival, SNAI2 amplification and/or overexpression in addition to DEAR1 LOH and/or downregulation correlated with poor survival (130 months in the unaltered cohort vs 168 months in the altered cohort) at greater significance when compared to DEAR1 alterations alone, indicating that expression of both DEAR1 and SNAI2 can be reliable markers of prognosis.

Another impactful observation was that loss of DEAR1 protein expression occurred frequently in women who developed TNBC before the age of 50. Furthermore, patients with low DEAR1 protein expression experienced a significantly shorter time to metastasis compared to patients with high DEAR1 expression. I find this data of significant value because TNBC has a tendency to affect women younger than 50 years of age and this subtype, given its aggressive behavior, is at higher risk of metastasis compared to other breast cancer subtypes. This corresponds with our previous data in which loss of DEAR1 associated with the TNBC subtype and strong family history in a cohort of young women diagnosed with IDC. This data indicates that DEAR1 may serve as a predictor of early onset TNBC and an indicator of metastasis risk.

**Future directions**

The data presented herein demonstrate that DEAR1 plays a role in the acquisition of progenitor cell properties and that its expression in breast cancer patients can be potentially useful in determining risk and prognosis. Further work is required to better comprehend the precise mechanism by which DEAR1 governs progenitor cell fate in mammary epithelial cells.

Additional in vitro assays, such as the colony forming cell (CFC) assay to visualize the ability of DEAR1-KD clones to generate luminal and basal-mixed colonies, thus confirming that loss of DEAR1 promotes the generation of bipotent progenitor cells [169], or assaying for co-expression of other markers indicating bipotency, such as surface markers CD10, MUC1, and α-6-integrin [58, 73, 169] and stem/progenitor markers such as Oct4 and Nanog by western blot
analysis or immunofluorescence, can be performed to confirm that loss of DEAR1 confers progenitor cell properties. Another gold-standard assay to study progenitor cell properties is to grow control and DEAR1-KD clones in differentiation medium to determine the potential of each clone to generate lineage-specific cell types, such as ductal or alveolar cells [114]. Other functional assays are required to demonstrate that DEAR1 is an important regulator of progenitor cell properties. One such assay includes a rescue experiment in which ectopic expression of an RNAi-resistant DEAR1 plasmid is performed in DEAR1-KD cells to test mammosphere formation efficiencies.

Since I observed a phenotypic difference in mammosphere formation between HMECs and DCIS.COM cells, it would be of interest to determine if DEAR1 regulates progenitor cell properties in DCIS cells through the same mechanism by which it governs progenitor cell properties in HMECs. One way to do so on a large-scale analysis would be to perform a reverse phase protein array (RPPA) to compare the differential protein expressions between control vector and DEAR1-KD clones from both HMECs and DCIS.COM clones to identify potential pathways that mediate the mammosphere phenotype in DCIS cells. Although data suggests that DEAR1 does not confer stem cell self-renewal capabilities, a labeling assay using PKH26 fluorescent dye can be used to study if loss of DEAR1 affects asymmetric division, a property of stem cells, in HMECs and DCIS. Cells are labeled with PKH26, plated in mammosphere growth conditions, and monitored over time to visualize dye distribution upon division. Asymmetric divisions are determined by fluorescence intensity: if one first-generation daughter cell remains quiescent and the other cell continues to divide, their progeny will display various PKH intensities within the mammosphere; alternatively in symmetric divisions, both cells divide at the same rate and all progeny will have the same relative PKH intensity. This assay was performed in a study by Cicalese and colleagues using an ERBB2 transgenic model to demonstrate that p53 was necessary to mediate asymmetric division; absence or mutation of p53 resulted in increased symmetric division, generating a larger population of stem-like cells [170]. We have unpublished
data by Dr. Balasenthil that shows DEAR1 plays a role in binding to and stabilizing p53. Since the DCIS.COM cell line expresses p53 and DEAR1, it would be of interest to evaluate stability of p53 in the context of DEAR1 loss of expression and determine if this affects symmetric divisions in DCIS cells.

We established that DEAR1 regulates progenitor cell properties independent of TGF-β-SMAD3 signaling potentially through regulating SNAI2 at the transcriptional and post-translational level. At the transcriptional level, it would be of interest to determine if DEAR1 modulates SNAI2 expression through interaction with transcription factors or transcriptional repressor complexes that are known to regulate SNAI2 gene expression, such as c-Jun, Mi-2/NuRD, or NCOR1. Although we have not demonstrated that DEAR1 localizes to the nucleus in HMECs, there are known TRIM proteins that mediate gene transcription including TRIM24 (TIF1α), TRIM19 (PML), and TRIM27 (RFP). TRIM24 was demonstrated to bind to chromatin and histone modifiers, including NCOR1, as well as mediate gene-specific transcriptional control through interaction with the AF-2 domain of nuclear receptors such as ER and RAR [171, 172]. TRIM19 was also demonstrated to activate the nuclear receptor PR, interact with histone acetyltransferase CBP to activate transcription of target genes, as well as negatively regulate gene transcription through association with histone deacetylases through formation of PML-nuclear bodies. The function of TRIM27 is less understood but it was suggested that TRIM27 associates with PML-nuclear bodies and also negatively regulates gene transcription through interaction with RB and histone deacetylase complexes [171, 173]. At the post-translational level, we observed an interaction between DEAR1 and SNAI2 that resulted in SNAI2 polyubiquitination. Data suggests that this interaction was not affected by the C11A point mutation in DEAR1’s RING domain, but more experiments can be done to determine which domain is required for SNAI2 binding. As mentioned previously, there are several known regulators of SNAI2 and it would be interesting to test if DEAR1 complexes with any of these mediators of SNAI2 activity. We have unpublished data that suggests DEAR1 interacts with p53, which has been shown to directly bind
to and promote SNAI2 ubiquitination through recruitment of MDM2. Functional assays utilizing mammospheres are required to confirm our theory that DEAR1 is a regulator of progenitor cell properties through SNAI2 activity. Alternatively, SNAI2 expression may not be the only factor in mediating this phenotype in HMECs. Preliminary data from RPPA using HMEC mammospheres identified potential candidate proteins and pathways that are deregulated in DEAR1-KD HMECs, including Annexin A1 (AnxA1), and Collagen Type VI Alpha 1 Chain (COL6A1). ANXA1 and COL6A1 both have roles in regulating EMT and are overexpressed in triple negative/basal-like breast cancer and cell lines, respectively [139, 174, 175], which correlates well with the discovery that DEAR1 expression is reduced in this subtype of breast cancer. TAZ, a well-known regulator of cell proliferation through the Hippo signaling pathway, has been shown to confer stem cell properties in breast cancer cell lines (e.g. mammosphere formation, self-renewal, and CD44<sup>high</sup>/CD24<sup>low</sup> signature [82, 83, 125, 168, 176]) and was slightly abundant in DEAR1-KD mammospheres compared to controls.

I would also like to further investigate if DEAR1 and SNAI2 expression levels are good predictors of overall survival or risk of early-onset breast cancer or metastasis. To do so, more patient samples are required to test these associations. In particular, we would like to evaluate if DEAR1 and SNAI2 protein expressions demonstrate the same inverse relationship in DCIS and IDC cases as was observed at the mRNA level in basal-like cancers and TNBC. Furthermore, I would like to utilize the same TNBC TMA to stain for SNAI2 protein expression and assess if it strengthens the association between DEAR1 staining with age of diagnosis and time to metastasis. Additionally, I would like to include normal controls and DEAR1 expression at time of diagnosis rather than at time of surgery to better analyze the association between DEAR1 and clinical outcomes and potentially use this information to predict prognosis.

There are currently ongoing clinical trials to study the efficacy of anti-EMT drugs that target specific regulators of the EMT process and also show effectiveness against CSCs that are dependent on EMT activation. Silmitasertib (CX-4945) is a CK2 serine-threonine kinase small
molecule inhibitor undergoing Phase I/II clinical trials that was demonstrated to inhibit SMAD-dependent and -independent TGF-β signaling and reduced expression of EMT-related markers in late stage solid tumors [177]. LY2157299 is a small molecule inhibitor of TGF-β undergoing Phase II/III clinical trials that was shown to be effective in treating TNBCs that have developed resistance to chemotherapy and hypothesized to target residual CSCs that were unaffected by conventional chemotherapy agents [177]. Salinomycin is another small molecule inhibitor that showed efficacy of reducing CSC populations through induction of apoptosis, though the precise mechanism is still being investigated. Natural compounds such as sulforaphane found in broccoli and resveratrol found in grapes were shown to be effective against CSCs through inhibition of the Wnt/β-catenin pathway and metastasis through inhibition of SNAI1/2, respectively [178, 179].

Since I observed a stem/progenitor cell phenotype in cells that have undergone loss of DEAR1 expression in both the absence and presence of TGF-β signaling, using these compounds to test their efficacies on mammosphere formation will provide insight into how we can potentially use DEAR1 expression to stratify patients for personalized targeted therapies in addition to standard chemotherapeutic agents.

Additionally, I would like to expand our investigations regarding transcriptional and post-transcriptional regulation of DEAR1. Several oncogenes and tumor suppressors are upregulated or downregulated, respectively, in cancers due to deregulation of miRNAs. The most notable ones involved in EMT and metastasis includes miR-34a, which negatively regulates SNAI1 expression and is downregulated in breast cancer [147]; miR-200 family members, which target ZEB1 and ZEB2 mRNA and is repressed during invasion and metastasis but re-expressed upon colonization at secondary sites [53, 129]; miR-10b is involved in mediating metastasis, was demonstrated to be upregulated in late-stage breast cancers, and associated with the TNBC subtype [148, 149, 151, 154, 162]; miR-155 expression was shown to be regulated by BRCA1, is overexpressed in breast cancer, and demonstrated to target genes involved in cancer cell migration, invasion, survival, and proliferation [180]. Preliminary data suggests that DEAR1
mRNA could be targeted by miR-10b and miR-196b. Further functional assays can be performed to confirm that these miRNAs target DEAR1, including additional mutation assays, *in vivo* correlational studies, and associative studies using patient samples. Therapeutically, I can test if treatment with antagomir-miR-10b can increase DEAR1 expression in cancer cell lines and the effects on CSC properties.

DEAR1 was demonstrated as a tumor suppressor and critical regulator of apical-basal polarity and EMT initiation. Altogether, results from this study demonstrates a novel role for DEAR1 in the regulation of stem/progenitor cell properties in HMECs and DCIS cells, through modulation of SNAI2, a potent regulator of stemness in mammary epithelial cells. I further identified an inverse correlation between DEAR1 and SNAI2 expression in patient samples, indicating their prognostic significance, and revealed significant associations between DEAR1 expression with age of diagnosis and time to metastasis in TNBC, the most aggressive and difficult to treat breast cancer subtypes. With this information, I hope to utilize our understanding of DEAR1 to stratify patients who are at risk for progression and would benefit from therapies that target DEAR1-regulated pathways.
REFERENCES


16. Shinn Young Kim S-HJ, Min Sung Kim, In-Pyo Baek, Sung Hak Lee, Tae-Min Kim, Yeun-Jun Chung, Sug Hyung Lee: Genomic differences between pure ductal carcinoma in


34. Sivan Elloul MBE, Jahn M. Nesland, Claes G. Tropé, Gunnar Kvalheim, Iris Goldberg, Reuven Reich, Ben Davidson: Snail, Slug, and Smad-Interacting Protein 1 as Novel Parameters of Disease Aggressiveness in Metastatic Ovarian and Breast Carcinoma. American Cancer Society 2004.


Sendurai A. Mani WG, Mai-Jing Liao, Elinor Ng. Eaton, Ayyakkannu Ayyanan, Alicia Y. Zhou, Mary Brooks, Ferenc Reinhard, Cheng Cheng Zhang, Michael Shipitsin, Lauren


68. Dick JE: **Stem cell concepts renew cancer research.** *Blood* 2008, **112**:4793-4807.


78. So Yeon Park HEL, Hailun Li, Michael Shipitsin, Rebecca Gelman, Kornelia Polyak: **Heterogeneity for Stem Cell-Related Markers According to Tumor Subtype and Histologic Stage in Breast Cancer.** *Clin Cancer Res* 2010, **16**:876-878.


81. Gillian Farnie PMW, Robert B. Clarke, Nigel J. Bundred: **Combined Inhibition of ErbB1/2 and Notch Receptors Effectively Targets Breast Ductal Carcinoma In Situ (DCIS) Stem/Progenitor Cell Activity Regardless of ErbB2 Status.** *PlosOne* 2013, 8(2).


87. Wenjun Guo ZK, Joana Liu Donaher, Tsukasa Shibue, Verena Tischler, Ferenc Reinhardt, Shalev Itzkovitz, Aurelia Noske, Ursina Zurrer-Hardi, George Bell, Wai Leong Tam, Sendurai A. Mani, Alexander van Oudenaarden, Robert A. Weinberg: ** Slug and Sox9 Cooperatively Determine the Mammary Stem Cell State.** *Cell* 2012, **178**:1015-1028.


89. Hatakeyama S: **TRIM proteins and cancer.** *Nature Reviews* 2011, **11**:792-804.


97. Christiana Hadjimichael KC, Christoforos Nikolaou, Antonios Klonizakis, Gesthimani-Ioanna Theodossi, Takis Makatounakis, Joseph Papamatheakis, Androniki Kretsovali:


VITA

Uyen Quynh Nguyen Le was born in Saigon, daughter of Dr. Dung Le and Thu Nguyen and sibling of Andrew Le. She graduated from Health Careers High School in San Antonio, TX in 2008. She received her Bachelor of Science degree in Cell and Molecular Biology from The University of Texas at Austin in 2012. After graduating from college, she entered the Ph.D. program at The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences in August 2012.

Permanent address: 135 Rock Springs Ct NE, Atlanta, GA 30306