# THE ROLE OF E2F5
# IN MAMMARY GLAND DEVELOPMENT AND BREAST CANCER PROGRESSION

By

Briana To

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physiology – Doctor of Philosophy

2020

## ABSTRACT

THE ROLE OF E2F5 IN MAMMARY GLAND DEVELOPMENT AND BREAST CANCER PROGRESSION

By

Briana To

Dysregulation of mammary developmental processes have resulted in breast cancer development and progression. Thus, understanding normal mammary gland development is critical to understanding how cancer cells are initiated and maintained. Transcriptomic profiling of various mammary gland developmental phases, including virgin, pregnancy, lactation and involution, revealed distinct gene expression patterns associated at each stage. Pathway analysis predicted a role for transcription factors E2F1-4 during different stages of development. The importance of E2Fs in mammary development was confirmed by the defects observed in the mammary glands of mice deficient for various E2Fs.

To examine if compensation occurs among activator E2Fs in the mammary gland, we analyzed mammary gland development in double E2F knockout mice. Our analysis revealed that compensation does occur among activator E2Fs in the mammary gland. However, this compensation appears to be very specific, as E2F2 can compensate for E2F1 but not E2F3 loss. Although the role of E2F1-4 has been characterized in mammary gland development, little is known about E2F5, an E2F demonstrating repressor activity. Using bioinformatic analysis, we predicted a role for E2F5 in terminal end bud differentiation and developmental stages. To further examine the role of E2F5 in normal mammary gland development, we generated a mammary-specific E2F5 knockout mouse model (E2F5CKO). Analysis of mammary gland development in E2F5CKO mice reveal

only modest mammary gland defects. However, we found that E2F5CKO animals develop mammary tumors after a prolonged latency. Using transcriptomic profiling, we identified oncogenes that are dysregulated in E2F5CKO tumors, potentially contributing to tumorigenesis. Through these studies, we have identified a novel role of E2F5 as a tumor suppressor in breast cancer and further elucidated roles of E2Fs in mammary gland development.

This work is dedicated to my grandpa: Kiet Ma

## ACKNOWLEDGMENTS

I would like to acknowledge those that have inspired me to become a scientist starting with the members of Dr. Philip Berman's lab at the University of California Santa Cruz. My time in that lab helped me realize my passion for research. I would like to acknowledge my mentor at Michigan State, Dr. Eran Andrechek. Thank you for giving me the opportunity to work with you. Thank you for always being supportive especially when I was juggling medical school and research. I admire your tenacity and I hope I will be able to handle adversity as well as you in the future. Thank you for always trying to motivate me to become a better scientist and to step outside my comfort zone to become a better presenter.

Thank you to Lauren Aitch and the Aitch foundation for funding my research. I am honored to have been selected as one of the Aitch fellowship recipients. Thank you to all the members of the Andrechek lab past and present.

A special thank you to my committee: Dr. Kathy Gallo, Dr. Karen Liby, Dr. Daniel Isaac and Dr. Hua Xiao for letters of recommendations and all the valuable feedback over the years. I would also like to acknowledge all those that attended the Cancer Research Network on a regular basis. Thank you for creating a great community for cancer research. I would like to thank the College of Natural Sciences for the continuous fellowships as well as the department of Physiology for all their support. Thank you to Dr. Justin McCormick and Bethany Heinlen for giving me the opportunity to be part of the DO-PhD program.

Thank you to my family for their constant support. Thank you to my friends: Aaron Chow, Yajing Ji, Sean Misek, Laura Harding and Scott Jackson who have been my second family in Michigan. You guys helped keep me sane throughout the years. Finally, thank you to Jonathan Rennhack for always having my back. You motivated me during the first week of my rotation in the Andrechek lab and have continued to motivate and support me since. Thank you for answering a million of my questions and always helping me see the positive side of things. I aspire to become a great scientist like you.

# TABLE OF CONTENTS

## LIST OF FIGURES

## KEY TO SYMBOLS AND ABBREVIATIONS

TEB – Terminal End Bud

ER – Estrogen Receptor

PR – Progesterone Receptor

HER2 – Human Epidermal Growth Factor Receptor 2

TNBC – Triple negative breast cancer

PDL-1 – Programmed Cell Death Ligand 1

FISH – Fluorescence in situ hybridization

PARP – Poly(ADP-ribose)

BRCA1/2 – Breast Cancer 1/2

VEGFR – Vascular Endothelial Growth Factor Receptor

VEGF – Vascular Endothelial Growth Factor

CCR7 – C-C Chemokine Receptor Type 7

CXCR4 – C-X-C Chemokine Receptor Type 4

CXCL12 – C-X-C Motif Chemokine Ligand 12

CCL21 – C-C Motif Chemokine Ligand 21

GEM – Genetically Engineered Mouse models

NSG – NOD Scid Gamma

PDX – Patient Derived Xenograft

HMEC – Human Mammary Epithelial Cells

ssGSEA – Single Sample Gene Set Enrichment Analysis

TCGA – The Cancer Genome Atlas

E2F5CKO – E2F5 Conditional Knockout

PARP – Poly ADP Ribose Polymerase

FBS – Fetal Bovine Serum

DFS – Disease Free Survival

PyMT – Polyoma Middle T

MMTV – LTR- Mouse Mammary Tumor Virus Long Terminal Repeat

EGFR – Epidermal Growth Factor Receptor

EMT – Epithelial Mesenchymal Transition

qRT-PCR – quantitative Reverse Transcription Polymerase Chain Reaction

LT – Large T antigen

st – Small T antigen

ChIP- Seq – Chromatin Immunoprecipitation Sequencing

Rb – Retinoblastoma

CRISPR - Clustered Regularly Interspaced Short Palindromic Repeats

# CHAPTER 1:

# INTRODUCTION

## Mammary Gland Development

The breasts are composed of glandular and fibrous tissues that are embedded in a fatty matrix, along with a network of blood and lymphatic vessels and nerves. The mammary glands are modified sweats glands that consist of a series of 15-20 secretory lobules connected to ducts [1]. The mammary epithelium is composed of two cell types that are arranged into two layers, the inner layer of luminal epithelial cells and an outer layer of myoepithelial cells [2]. Much of what is known about mammary gland development has been derived from studies in mice. At birth, the mammary gland begins as rudimentary ductal tree and is in a relatively quiescent state until puberty [3]. At puberty, which is approximately 4 weeks of age in mice, the production of hormones and growth factors promote cell division and the formation of terminal end buds (TEBs) [3]. These specialized structures drive ductal tree extension into the fat pad. An extensive ductal network is developed from the rudimentary mammary epithelium through bifurcation and branching. Next, extensive remodeling of the mammary gland occurs during pregnancy when differentiation of alveolar buds enable milk secretion [3] . During lactation, alveolar buds differentiate into lobular alveoli that function for milk production. Following lactation, widespread death of alveolar secretory epithelial cells occurs in a process known as involution [3].

## Breast Cancer

Breast cancer is the most commonly diagnosed malignancy in women worldwide. Most malignancies in the breast begin in cells which form secretory lobules and ducts and as such, are classified as either ductal or lobular breast cancer. Of these two types of

breast cancer, ductal breast cancer is much more common than lobular breast cancer, with 90% of cases arising from ductal cells. Both these breast cancer types can be classified as in situ or invasive [4]. In situ breast cancer denotes cancerous cells that are limited to its original structure in lobules or ducts, while invasive breast cancer is defined as cancerous cells that have invaded surrounding tissues. Staging of breast cancer helps predict prognosis and guide plan of care. There are various factors that are used to calculate the stage of breast cancer including the size of the tumor, whether the cancer has spread to other regions of the body, and lymph node involvement. Localized breast cancer is generally categorized as stage 0 while stage I-IV are reserved for invasive breast cancer. Breast cancer spread and involvement to other tissues is more extensive with increasing stage number. Not surprisingly, as breast cancer progresses from stages 2, 3 or 4, the survival rate is significantly decreased. The five-year survival rate of stage 0 breast cancer is 100% and drops to 22% in stage 4 breast cancer [5] . In stage 4 breast cancer, the disease has spread beyond the breast to other organs in a process known as metastasis.

**Breast Cancer Heterogeneity**

Breast cancer is a very heterogeneous disease with both intra- and inter-tumor heterogeneity. Various classifications are used to help better understand and define this heterogeneity. Clinically, one of the most widely used classification of breast cancer subtypes is based on the presence or absence of Estrogen Receptor (ER), Progesterone Receptor (PR) and/or Human Epidermal Growth Factor Receptor 2 (HER2) as determined by immunohistochemistry and/or Fluorescence in situ hybridization (FISH).

The combined absence of these three receptors is known as triple negative breast cancer (TNBC). Receptor status is one of the prognostic factors used clinically, with positive ER and PR expression being generally associated with better outcomes. In addition, receptor status is also used as predictive markers in early breast cancer. Thus, it helps identify patients who may benefit from endocrine therapy or HER2- targeted therapy [6]. The absence of receptors limits the treatment of TNBC to surgery, radiation therapy and general chemotherapy. However, in more advanced disease, immunotherapy is available in combination with chemotherapy in TNBC that express programmed cell death ligand (PDL-1) [7]. In the subset of TNBC patients with germline BRCA mutations, inhibitors of poly(ADP-ribose) (PARP) may be a therapeutic option [8].

Alternatively, breast cancer can be further divided into molecular subtypes based on genomic profiling. Unique gene expression profiles stratify breast tumors into molecular subgroups including luminal A, luminal B, HER2-enrich, normal-like, basal-like and claudin low groups. While this classification is more widely utilized in research settings, it has become more widely adapted in the clinical setting as they offer additional prognosis and predictive abilities. Integrating these molecular subtypes with their clinically relevant markers demonstrate that luminal A and luminal B tumors are ER positive, HER2- enriched subtypes is HER2+ and the claudin low and basal subtypes are triple negative.

**Breast Cancer Oncogenes and Tumor Suppressors**

The discovery of oncogenes and tumor suppressors have been critical in driving the understanding of breast cancer biology. Genomic analysis of breast cancer patients

has revealed several frequently deregulated oncogenes including HER-2, MYC and Cyclin D1. The HER-2 gene is amplified and overexpressed in 20-30% of invasive breast cancer and is rarely amplified in benign breast disease [9–12]. The HER2 gene encodes for Human Epidermal Growth Factor Receptor 2, a receptor tyrosine kinase that can activate various downstream signaling cascades that mediate cell proliferation, differentiation, apoptosis, and motility. Another oncogene, Myc, is amplified in 15-25% of breast tumors [13]. Myc is a transcription factor that mediates the expression of genes involved in various biological functions including cell proliferation, differentiation and survival. Cyclin D1 is reported to be overexpressed in 40-50% and amplified in 10-20% of breast cancer [14]. In its canonical role, cyclin D1 mediates cell cycle progression by forming a protein complex with CDK4/6. Together, this complex will phosphorylate pRb, resulting in its release on transcription factor E2F. The activation of E2Fs mediates transcription of genes required for cell cycle progression [15]. In addition to their amplification/overexpression in human breast cancer, overexpression of Her2/neu [16,17], Myc [18] and Cyclin D1 [19] in the mouse mammary epithelium results in multifocal mammary tumors, further affirming their roles as oncogenes in breast cancer.

In contrast to oncogenes, tumor suppressors are defined as genes whose loss of function or inactivation contributes to cancer development.

One of the most well-known tumor suppressors is BRCA1/2. Mutation in either of these genes account for the majority of hereditary breast and ovarian cancer [20]. The BRCA1 and BRCA2 genes encode proteins that have essential roles in DNA damage repair, cell- cycle arrest, apoptosis, genetic instability and transcriptional activation. BRCA1/BRCA2 driven breast cancer is also associated triple negative breast cancer. To

further validate their roles as tumor suppressors, susceptibility in tumor formation is examined in mouse models with a deletion or inactivation of candidate genes. Indeed, inactivation of BRCA1 and BRCA2 through various strategies have promoted mammary tumor formation in mice [21,22].

**Tumor Metastasis**

Metastasis is the main cause of mortality in breast cancer patients. In order for cells to metastasize to distant organs, tumors cells must undergo a multi-step process. This process is known to be highly inefficient given that proper coordination is required in each step in order for the tumor cells to progress to the next stage. In the first step of metastasis, tumors must develop invasive properties that allow them to invade into adjacent tissue. Next, tumor cells migrate into vessels by a process known as intravasation. Survival in the circulatory system may be challenging as the tumor cell needs to overcome the harsh environment. Once the tumor cell has reached its destination, it will extravasate and colonize distant organs. More recently, studies have shown that perimetastatic niches help tumor cells colonize and proliferate at distant sites.

Generally, the first site that breast cancer spreads to are local lymph nodes. The most common sites of distant metastasis are bone, brain, liver and lungs [23]. Previous studies have identified transcriptional profiles of metastasis that colonize the lung, bone and brain [24–26]. These studies suggest that the metastasis to specific sites are driven by distinct transcriptional changes.

## Lymphatic Metastasis

Tumor cells can metastasize to distant sites by traveling through blood vessels or the lymphatics system. In breast cancer, lymphatic metastasis remains the most important prognostic indicator for survival in patients diagnosed with breast cancer. Lymph node status also plays a major role in determining the management of breast cancer patients. While blood vessels deliver oxygen and nutrients to tissues, lymphatic vessels collect protein rich fluid from tissues and serves as conduits for the immune and gastrointestinal system. The different functions of these vessels are reflected in their inherent physiology and structure. Unlike blood vessels, lymphatic vessels are designed to be highly permeable with a discontinuous basement membrane and loose cell-cell junctions. In addition, the flow rate within lymphatic vessel is 100-500x slower than blood vessels and has less shearing stress [27]. Furthermore, the composition of the lymph contains 1000-fold higher concentration of hyaluronic acids, a molecule with potent cell-protecting and pro-survival properties [28,29]. These structural and mechanical characteristics provide an ideal environment for tumors cells to intravasate and survive.

## Mechanism of Lymphatic Metastasis

Tumor metastasis via the lymphatic system was once thought to be a passive process where detached tumor cells were passively collected in the lymphatic vessels and subsequently drained into sentinel lymph nodes. However, increasing evidence has demonstrated that the process of tumor cell intravasation into lymphatic vessels can be facilitated by chemokines ligands and receptors expressed by tumor and stromal cells. In breast cancer, two chemokine receptors, CXCR4 and CCR7 have been implicated in

regulating this process [30]. The expression of CXCL12 and CCL21, chemokine ligands of CXCR4 and CCR7 receptors, by lymphatic endothelial cells can attract tumors cells into lymphatic vessels. Furthermore, overexpression of these ligands in breast cancer cells increase migration and invasiveness *in vitro*. [31]. Conversely, inhibition of CXCR4 results in decreased lymph node metastasis in an orthotopic mouse model [32,33]. Clinically, elevated expression of the chemokine receptors CCR7 and CXCR4 is associated with lymph node metastasis [34,35].

Upon stimulation by secreted factors, tumors cells can intravasate into pre-existing or newly formed lymphatic vessels. The process in which new lymphatic vessels are formed, also known as lymphangiogenesis, can be induced by tumor cells and the microenvironment. Mechanistically, lymphangiogenesis is largely driven by the VEGFR-3 pathway, with VEGF-C and VEGF-D being the main regulators of this pathway [36–39]. VEGF-C and VEGF-D are two secreted glycoproteins that are expressed in tumor and stromal cells, while VEGFR-3, a tyrosine kinase receptor, is expressed primarily on the lymphatic endothelium. VEGF-C and VEGF-D can promote the formation of intratumoral lymphatic vessels by binding to VEGFR-3. In a transplantation mouse model, VEGF-C increases intratumoral lymphangiogenesis and the incidence of lymph node metastasis [37]. However, it remains ambiguous whether intratumoral lymphatic vessels are functional as they commonly appear to be collapsed[40,41]. VEGF-C and VEGF-D can also induce morphological changes in peritumoral lymphatics that support the entry of tumors cells into lymphatic vessels. Furthermore, a number of studies show that elevated VEGF-C is reported in 30-40% of breast cancers and is correlated with high incidence of lymphovascular invasion, lymph node metastasis and lower disease-free survival (DFS)

[42–44]. Despite increasing efforts to study the process of lymph node metastasis and lymphangiogenesis in breast cancer, there remains a multitude of questions that need to be addressed. For example, the genetic events driving lymph node metastasis in breast cancer is still ambiguous. In addition, the importance and requirement of lymph node metastasis in promoting metastatic spread to distant organs is unclear. To begin answering these questions, it is important to utilize the appropriate model system and experimental tools. There are several model system choices available, including transplantation and spontaneous models.

**Genetically Engineered Mouse Models**

Genetically engineered mouse models are a widely used system to study breast cancer development, progression and metastasis. Conventional transgenic mouse models use mammary-specific promoters to drive the expression of oncogenes. One of the commonly used mammary-specific promoters include mouse mammary tumor virus long terminal repeat (MMTV-LTR). The MMTV-LTR is a regulatory element derived from the mouse mammary tumor virus which is present in the milk of infected mice [45]. Although MMTV is active throughout the ductal epithelial cells in nulliparous female, their expression is elevated by steroid hormones and prolactin during pregnancy and lactation [46]. In addition, these hormones and prolactin promotes the proliferation of luminal epithelial cells, thus the expression of MMTV is higher these epithelial cell types. However, MMTV is still active in basal epithelial cells [47]. It is important to note that MMTV expression is not exclusive to mammary epithelium as it also occurs in other secretory tissues including salivary glands. Using different lengths of MMTV-LTR in

transgenic mice can lead to different expression patterns of the transgene. For example, shorter variant of MMTV promoter have been shown to activate a widespread expression of the transgene while a longer variant demonstrates a more confined expression of the transgene to the mammary gland and salivary glands [17,48,49]. Numerous transgenic mouse models of breast cancer have been developed using MMTV-LTR to drive the expression of oncogenes including Myc [18], Ha-Ras [50], Cyclin D1 [19], HER2/ErbB2 [16,17] and polyomavirus middle T antigen (PyMT) [51] in the mammary epithelial cells. In addition to models driven by oncogene activation and expression, numerous knockout strains (BRCA, p53) rely on disabling tumor suppressor gene function by deleting the gene or by introducing a point mutation that renders the gene defective. To generate a conditional knockout targeting the mammary epithelium, Cre/lox recombination system is commonly utilized. This system relies on transgenic mice that expresses Cre recombinase under a mammary specific promotor (eg MMTV) and a transgenic mice in which the targeted gene is floxed by two loxP sequences. Crossing these two mouse strains leads a transgenic mice that expresses Cre in the mammary gland which in turn will cut at the loxP site and cause an excision in the targeted gene [52]. In contrast to the transgenic models, knockout mouse models generally develop tumor after a longer latency, lower multifocality and lower penetrance. Surprisingly, although numerous GEMs have been generated and characterized, many of those models lack the ability to model metastasis. Of the mouse models that do metastasize, pulmonary metastases are most commonly observed and very few mouse models have reported lymph node involvement. One of the most utilized models of metastasis that also has lymph node involvement is MMTV-PyMT FVB. The MMTV-PyMT mouse model uses the MMTV promoter to

overexpress Polyoma middle T-antigen in the mammary gland, leading to multifocal adenocarcinoma with a median latency of 53 days. These mice are highly metastatic with 100% of the mice developing lung metastasis and 30% of the mice with lymph node metastasis by 13.5 weeks [53]. The metastatic burden of the transgenic mice may be influenced by the background strain of the transgenic mice. For instance, Lifsted *et al.* interbred MMTV-PyMT mice into various strains of mice. Their results demonstrates a difference in tumor latency and metastatic index in the various strains relative to MMTV-PyMT on the FVB background [54].

Another commonly used GEM to study metastasis is MMTV-Neu. This model is based on the overexpression of the rat form of human HER2 (Neu, ErbB2). To model HER2+ve breast cancer, numerous variations have been generated, including various mutant alleles and inducible systems. One of these variations include the original MMTV-NeuNT model which consist of a point mutation, resulting in constitutive activation and rapid tumor formation and metastasis [17]. Examining all the various Neu models, the tumor latency and rate of metastasis varies widely. Similar to MMTV-PyMT, the genetic background of the MMTV-Neu model also impacts the metastatic potential, either raising or lowering the rate of metastasis depending upon the background. One of the MMTV Neu models that has been observed to develop lymph node involvement is BALB/c MMTV-NeuNT [55,56]. In this model, 100% of the mice develop multifocal tumors with a median tumor latency of 21 weeks with lymph metastasis in the neck, lateral thoracic and axillary lymph nodes.

Lymph node metastasis have also been reported in the MMTV-Wnt-1 transgenic mouse model. Similar to MMTV-PyMT and MMTV-Neu GEMs, this model also forces

overexpression of the oncogene gene Wnt-1 in the mammary epithelium [57]. The MMTV-Wnt-1 mouse model have a 60% tumor incidence with a median tumor latency of 8 months. In their manuscript, they noted that when the primary tumor was resected, the majority of MMTV-Wnt-1 mice develop lymph node and lung metastasis.

Genomic profiles revealed that key attributes in human breast cancer subtypes were conserved among mouse models [58]. Not surprisingly, not one mouse model was able to recapitulate the heterogeneity across the human breast cancer subtypes. Thus, there is no one mouse model that is superior in modeling human breast cancer.

**Cell Line Models**

Implantation of breast cancer cell lines into mice is a commonly used model of tumor progression and bmetastasis. Many cell lines have been well characterized and have publicly available multi-omic datasets. There are numerous human breast cancer cell lines commercially available and the choice of cell line is critical when designing an experimental study. This is especially important when choosing a cell line to study lymph node metastasis, given that human breast cancer cell lines have varying degrees of metastatic potential and different sites of metastasis. One commonly used breast cancer cell line for metastatic studies is MDA-MB-231, despite the numerous caveats associated with this cell line which include being genetically distinct from basal-like metastatic breast cancer patients [59]. Orthotopic injection of MDA-MB-231 into the mammary fat pad frequently metastasizes to the lung and lymph nodes [60]. Differences in metastatic potential have been reported in various immunodeficient mouse strains including nude athymic and NSG mice [61,62].

An additional cell line that has a high propensity for lymph node metastasis is MDA-MB-468LN. This cell line is a variant isolated from a lung metastasis derived by orthotopic injection of human breast cancer cell line MDA-MB-468. It produces wide-spread lymph node involvement in 100% of the orthotopic injected nude athymic mice [63]. To accelerate metastasis to the lymph nodes and lungs, Lee *et al*. demonstrated that pre-treatment of animals with tumor- conditioned media results in accelerated metastasis to the lymph nodes and lungs in two different triple negative breast cancer cell lines, MDA-MB-231 and SUM-149 [64]. While the specific tumor-derived factors mediating the accelerated metastasis were not identified, the investigators did observe increased phosphorylation of VEGFR and EGFR, resulting in enhanced angiogenesis and lymphangiogenesis.

Another option for cell lines to model metastasis is the use of syngeneic cell lines. One of the most commonly used mouse mammary carcinoma cell lines for the development \metastasis is 4T1. This highly invasive and metastatic cell line was derived from the mammary tumor tissue of a BALB/cfC3H mouse [65]. Lung and lymph node metastases was observed in >90% of the lymph nodes that were harvested as early as 14-18 days after injection of 4T1 into the abdominal mammary gland [66]. Due to the aggressive nature of these tumors, resection of the primary tumor may extend the life span of the tumor- bearing mice, potentially allowing for increased metastasis to develop [67].

There are numerous routes to introduce tumor cells into mice depending on what stage of metastasis is being examined and/or the site of metastasis targeted. For example. intravenous injection of tumors via the tail is demonstrating the ability of the

tumor cells to extravasate from the vessel and colonize metastatic sites. In contrast, injection of tumor cells directly into the mammary fat pad is able to mimic the natural progression of the metastatic stages. In addition, orthotopic injection is also the commonly used injection site for the development of lymph node metastasis.

To model only the later stages of lymph node metastasis, tumor cells may be transplanted directly into the lymph node (eg intra-axillary) and monitored for dissemination to distant organs. For example, Leslie *et al*., introduced 4T1 cells directly into the draining lymph node and subsequently characterized the tumor progression in the lymph node. Furthermore, they demonstrated that 4T1 cells were able to disseminate and colonize the lung.

Cell line models are well characterized and reliable models of metastasis. However, one of the major pitfalls is the need for an immunodeficient host. Previous studies have demonstrated that the immune system appears to play an active role in mediating metastasis. In addition, many of the immortalized cell lines have been grown for decades, and concerns have arisen over the loss of heterogeneity and clonal selection while cultured on plastic. Indeed, it has been demonstrated that gene expression from cell lines grown in two-dimensional culture differs from cell lines grown in three-dimensional culture or grown in xenograft models.

**Patient-Derived Xenograft Models**

Although cell lines are powerful experimental models, many have questioned their ability to recapitulate the biology of human breast cancer and the effects of two-dimensional culture. Some of these concerns have been alleviated by the development

of Patient-Derived Xenografts (PDX). PDXs are collected from patients and propagated in mice, thus they avoid the changes that may occur with *in vitro* cultures. Unlike many of the cell lines that were established from metastatic sites such as pleural effusions, a majority of the PDXs are derived directly from the primary tumor. In addition, PDXs can generally recapitulate the tumor biology of the original tumor including genomic, transcriptomic and proteomic expression patterns. Clinically, PDXs have retained histopathological patterns as well as the status of clinical molecular markers such as ER, PR and HER2 [68]. Therefore, PDXs overcome some of the clinical limitations that are associated with tumors derived from cell line transplantations.

PDX models most commonly metastasize to the lung and lymph nodes. Interestingly, PDXs demonstrate similar metastatic site specificity when compared to tumor of origin [68]. PDX models have been used in several studies to model metastasis [69–71]. For example, Bockhorn *et al*. implanted a pulmonary metastatic lesion into the mammary fat pad of mice to examine the characteristic of the resulting primary mammary tumor [72]. Interestingly, relative to the parental primary tumor, when parental lung metastatic tumors were injected into the mice, they grew more slowly and had a reduced metastatic burden. Furthermore, this study identifies differentially expressed mRNA and miRNA between the parental primary tumor and pulmonary metastasis derived tumor. One of the notable findings is the upregulation of miR-138 in the pulmonary metastasis-derived tumor and its subsequent role in mediating invasion and EMT in breast cancer cell lines. Their results suggest that mir-138 may exert its control over metastasis by targeting EZH2, a gene that is critical in regulating self-renewal and tumor progression. Similar to human breast cancer cell lines, one of the main pitfalls to PDX models is the

lack of immune cell diversity. Thus, this model system also has limitations in reproducing the interaction between tumor cells and its immune environment. Other limitations to PDX models include low take rate and long incubation period.

**-OMIC Studies in Lymphatic Metastasis**

Genomic, transcriptomic, proteomic and epigenomic studies have been valuable tool in identifying potential drivers of lymphatic metastasis [84]. At a gene expression level, several studies have revealed differentially regulated genes when comparing primary mammary tumors and matched lymph node metastases [85–91]. Many of the genes identified in these studies appear to be differentially regulated in a way that favors the multi-stage process of metastasis. These genes are associated with cell basement membrane function, ECM remodeling, tumor recruitment, cell signaling and EMT transition [92]. Furthermore, in a study comparing primary tumors and matched lymph node metastases from 26 patients, a set of 79 differentially expressed genes was able to predict clinical outcome in node-positive breast cancer patients [88].

Epigenetic studies on lymph node metastasis have revealed a potential role for epigenetic alterations in regulating lymph node metastasis [93–95]. Several studies examined the methylation pattern differences between primary breast tumors and matched lymph node metastases in patients. The methylation status of several genes including DFNA5 [96], ID4 [97] and CDH1[98] have been found to be correlated with lymph node metastasis in breast cancer patients. One caveat is that many of these studies have focused on subsets of cancer genes instead of examining genome-wide DNA methylation patterns [93,99–101]. This in part may explain the discrepancy in which

studies have found higher heterogeneity between primary breast tumor and lymph node metastasis while other studies have found more conserved methylation patterns between the two sites. More recently, whole genome methylation analysis was performed on triple negative breast cancer primary samples, matched normal adjacent tissue and matched lymph node metastases [102]. This study was unique in their integration of gene expression data from their previous studies with methylation data on matched samples. Gene expression analysis revealed 83 genes with altered expression in primary breast tumors and their matched lymph node metastases. Of those 83 lymph node metastasis associated genes, 18 were validated to have alterations in methylation [102].

Recently, a genome-wide methylation study in human breast cancer cell lines revealed widespread hyper and hypomethylation events that were unique to a highly metastatic cell line [94]. Further work identified the chromatin modifier HDAC11 as a mediator of lymph node metastasis [103]. Inhibition of HDAC1 resulted in decrease lymph node tumor formation. Moreover, inhibition of HDAC11 in a mouse model system also resulted in increased pulmonary metastasis.

Although these studies have identified differential gene expression and methylation patterns, many of them did not demonstrate the functional effects of these alterations on lymph node metastasis. Follow-up functional studies are essential to determine if the identified genetic events contribute to the regulation of mediating lymphatic metastasis. In addition, integration of various data platforms (gene expression, proteomics, methylation, etc) will offer a more robust analysis. However, many of these studies have been conducted in bulk tumor tissues and the complex cellular heterogeneity of the tumor may mask the signal from the tumor cells that are metastasizing to the lymph

node. The development of single cell sequencing technology allows for a better understanding of cell population heterogeneity and tumor evolutionary changes. Moving forward, the use of single cell sequencing may help provide additional insight into biological changes associated with lymph node metastasis. Recently, Bao *et al.* performed single cell sequencing on morphologically distinct areas of primary breast cancer and matched lymph node metastasis [104]. Their data suggests that the lymph node metastasis was derived from the clonal expansion of a single tumor cell found at the invasive front of the primary tumor. Furthermore, they identified high level-gain of MCL1, a pro-survival gene, and ch8q amplification as being potential contributors to tumor dissemination to lymph nodes.

**E2F Transcription Factor Family**

E2Fs are a family of transcription factors that regulate cell cycle by controlling the genes required for DNA synthesis. There are nine E2F family members encoded by eight genes. E2F1-E2F3a are classified as transcriptional activators while E2F3b-E2F8 are transcriptional repressors. The transcriptional activity of E2F1-E2F5 are mainly regulated by their interaction with the pocket protein family pRb, p107 and p130 [73]. During the G0/G1 phase of the cell cycle, pRb is bound to activating E2Fs, inhibiting their ability to activate transcription [73]. At the same time, the repressive E2F4-5 binds to p107 and p130 and translocate to the nucleus to repress transcription of target genes [73]. During the G1/S phase, activating E2Fs dissociate from phosphorylated pRb and activate transcription of genes required for the entry into S-phase [73]. Conversely, when p107 and p130 are phosphorylated, it releases E2F4-E2F5. Unlike E2F1-E2F3, E2F4-E2F5 do

not have a nuclear localization signal, so without interaction with the pocket proteins, they are unable to translocate into the nucleus to repress transcription [74]. E2Fs share highly similar DNA binding domains that enables them to directly bind to the E2F consensus sequence (TTCCCGCC) [73,75,76] . In general, E2F1-6 bind to DNA and regulate transcription as a heterodimer with DP proteins DP1 and DP2 [75]. In contrast, E2F7 and E2F8, also known as atypical repressors, bind to DNA without interaction with DP proteins [75]. The expression and localization pattern of E2Fs are cell cycle dependent. Activator E2Fs levels peak at the G1-S phase transition while atypical repressors peak in late S-phase [77]. In contrast, levels of canonical repressors E2F3a-E2F6 remain similar as it is constitutively expressed throughout the cell cycle phases [77]. While most E2Fs are predominantly nuclear, E2F4 and E2F5 show distinct localization dependent on cell cycle phase. In concordance with their role, E2F4-E2F5 remains nuclear in quiescent cells and cytoplasmic in cycling cells [77].

**Role of E2F in Mammary Development and Breast Cancer**

E2Fs play a role in development as demonstrated by the developmental defects arising in E2F knockout mouse models. Depending on which E2F member is dysfunctional, these defects vary in location and severity. In the mammary gland, developmental defects are observed with the loss of E2Fs. E2F1 KO, E2F3 heterozygote and E2F4 KO mice have a decrease in mammary epithelial outgrowth and significant reduction in branching [78]. Interestingly, loss of E2F2 does not result in significant mammary gland changes. Development and cancer are two sides of the same coin, as dysregulation of normal developmental processes can result in breast cancer initiation

and progression. Thus, it is not surprising that E2Fs have also been implicated as having a role in breast cancer development and progression. To study the effects due to the loss of E2Fs in breast cancer, researchers generated E2F knockouts in various murine breast cancer models including MMTV Neu [79], MMTV PyMT [80] and MMTV Myc [81]. The effects of E2F loss on tumor latency, tumor growth and metastasis were observed. Results from these studies demonstrate distinct roles of E2Fs in breast cancer development and metastasis. In the MMTV Neu model, loss of E2F1, E2F2 and E2F3 resulted in delayed tumor onset [79]. However, only loss of E2F1 significantly accelerated tumor growth. In addition, decreased metastatic potential and metastatic burden was observed in E2F1 and E2F2 knockout mice. In the MMTV PyMT mouse model, loss of E2F1 resulted in significant acceleration of tumor onset, while loss of E2F3 resulted in significant delay in tumor onset. In addition, loss of E2F1 and E2F2 in the MMTV PyMT resulted in decreased number of pulmonary metastases. Finally, in the MMTV Myc mouse model, loss of E2F1 resulted in decreased tumor latency and accelerated tumor growth [81]. Both MMTV Myc E2F2 KO and MMTV Myc E2F3 heterozygote mice demonstrated increased tumor latency and decreased tumor incidence [81]. Furthermore, loss of E2F2 in MMTV Myc model resulted in increased pulmonary metastases [82]. Taken together, these studies suggest that activator E2Fs have distinct roles that affect tumor initiation and metastasis. Heterozygous E2F3 increased tumor latency in all three transgenic mouse models. This suggests that E2F3 may play an oncogenic role in tumor development. Interestingly, loss of E2F1 increased tumor latency in MMTV Neu, but decreased tumor latency in both MMTV PyMT and MMTV Myc models. Thus, E2F1 appears to act as an oncogene or tumor suppressor depending on the oncogenic driver.

In addition, these studies highlight the role E2F1 and E2F2 in promoting metastasis. Studies in other cancer types have also demonstrated that E2Fs may behave as an oncogene or tumor suppressor in a tissue-specific manner [83]. Furthermore, these studies demonstrate that alterations of different E2F members results in distinct effects on tumor development. This, in part, may be explained by the fact that E2F members uniquely regulate a subset of targets. Chip-based technologies have identified direct targets of E2Fs, revealing shared and unique targets among E2F members [84–89]. Strikingly, these studies have revealed a broader role of E2F beyond regulating cell proliferation and apoptosis. For example, direct E2F-regulated targets included genes involved in cell differentiation, metabolism, animal development and angiogenesis.

## WORKS CITED

## WORKS CITED


1. Radisky, D.C. and Hartmann, L.C., 2009. Mammary involution and breast cancer risk: transgenic models and clinical studies. *Journal of mammary gland biology and neoplasia*, *14*(2), pp.181-191

2. Cristea S, Polyak K. Dissecting the mammary gland one cell at a time. Nat. Commun. 2018. p. 1–3.

3. McNally S, Stein T. Overview of mammary gland development: A comparison of mouse and human. Methods Mol Biol. 2017. p. 1–17.

4. Li CI, Anderson BO, Daling JR, Moe RE. Trends in Incidence Rates of Invasive Lobular and Ductal Breast Carcinoma. J Am Med Asso. 2003;289:1421–4.

5. Iqbal J, Ginsburg O, Rochon PA, Sun P, Narod SA. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. JAMA - J Am Med Assoc. 2015;313:165–73.

6. Abe O, Abe R, Enomoto K, Kikuchi K, Koyama H, Masuda H, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: Patient-level meta-analysis of randomised trials. Lancet. 2011;378:771–84.

7. Cyprian FS, Akhtar S, Gatalica Z, Vranic S. Targeted immunotherapy with a checkpoint inhibitor in combination with chemotherapy: A new clinical paradigm in the treatment of triple-negative breast cancer. Bosn. J. Basic Med. Sci. Association of Basic Medical Sciences of FBIH; 2019. p. 227–33.

8. Faraoni I, Graziani G. Role of BRCA mutations in cancer treatment with poly(ADP-ribose) polymerase (PARP) inhibitors. Cancers. 2018.

9. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science. 1989;244:707–12.

10. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: Correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987;235:182–91.

11. Allred DC, Clark GM, Molina R, Tandon AK, Schnitt SJ, Gilchrist KW, et al. Overexpression of HER-2/neu and its relationship with other prognostic factors change during the progression of in situ to invasive breast cancer. Hum Pathol. 1992;23:974–9.

12. Mansour EG, Ravdin PM, Dressier L. Prognostic factors in early breast carcinoma. Cancer. 1994;74:381–400.

13. Nass SJ, Dickson RB. Defining a role for c-Myc in breast tumorigenesis . Breast Cancer Res. Treat. 1997. p. 1–22.

14. Steeg PS, Zhou Q. Cyclins and breast cancer . Breast Cancer Res. Treat. 1998. p. 17–28.

15. Lodén M, Stighall M, Nielsen NH, Roos G, Emdin SO, Östlund H, et al. The cyclin D1 high and cyclin E high subgroups of breast cancer: Separate pathways in tumorogenesis based on pattern of genetic aberrations and inactivation of the pRb node. Oncogene. 2002;21:4680–90.

16. Guy CT, Webster MA, Schallert M, Parsonst TJ, Cardifft RD, Muller WJ. Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. Genetics. 1992.

17. Muller WJ, Sinn E, Pattengale PK, Wallace R, Leder P. Single-Step Induction of Mammary Adenocarcinoma in Transgenic Mice Bearing the Activated c-neu Oncogene. Cell. 1988.

18. Stewart TA, Pattengale PK, Leder P. Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell. 1984;38:627–37.

19. Wang TC, Cardiff RD, Zukerberg L, Lees E, Arnold A, Schmidt E V. Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. Nature. 1994;369:669–71.

20. Osborne C, Wilson P, Tripathy D. Oncogenes and Tumor Suppressor Genes in Breast Cancer: Potential Diagnostic and Therapeutic Applications. Oncologist. 2004;9:361–77.

21. Xu X, Wagner KU, Larson D, Weaver Z, Li C, Ried T, et al. Conditional mutation of Brca 1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. Nat Genet. 1999;22:37–43.

22. Cheung AMY, Elia A, Tsao MS, Done S, Wagner KU, Hennighausen L, et al. Brca2 Deficiency Does Not Impair Mammary Epithelium Development but Promotes Mammary Adenocarcinoma Formation in p53+/- Mutant Mice. Cancer Res. 2004;64:1959–65.

23. Whittle JR, Lewis MT, Lindeman GJ, Visvader JE. Patient-derived xenograft models of breast cancer and their predictive power. Breast Cancer Res.

24. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, et al. Genes that mediate breast cancer metastasis to lung. Nature. 2005;436:518–24.

25. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordón-Cardo C, et al. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell. 2003;3:537–49.

26. Paula D. Bos, Xiang H.-F. Zhang, Cristina Nadal, Weiping Shu RRG and DXN. Genes that mediate breast cancer metastasis to the brain. Nature. 2009;459.

27. Rahman M, Mohammed S. Breast cancer metastasis and the lymphatic system (Review). Oncol Lett. 2015;10:1233–9.

28. Ran S, Volk L, Hall K, Flister MJ. Lymphangiogenesis and lymphatic metastasis in breast cancer. Pathophysiology. 2010. p. 229–51.

29. Laurent TC, Fraser JRE. Hyaluronan 1 . FASEB J. Wiley; 1992;6:2397–404.

30. Shields JD, Fleury ME, Yong C, Tomei AA, Randolph GJ, Swartz MA. Autologous Chemotaxis as a Mechanism of Tumor Cell Homing to Lymphatics via Interstitial Flow and Autocrine CCR7 Signaling. Cancer Cell. 2007;11:526–38.

31. Kang H, Watkins G, Parr C, Douglas-Jones A, Mansel RE, Jiang WG. Stromal cell derived factor-1: its influence on invasiveness and migration of breast cancer cells in vitro, and its association with prognosis and survival in human breast cancer. Breast Cancer Res. 2005;7:402–10.

32. Müller A, Homey B, Soto H, Ge N, Catron D, Buchanan ME, et al. Involvement of chemokine receptors in breast cancer metastasis. Nature. 2001;410:50–6.

33. Liang Z, Wu T, Lou H, Yu X, Taichman RS, Lau SK, et al. Inhibition of breast cancer metastasis by selective synthetic polypeptide against CXCR4. Cancer Res. 2004;64:4302–8.

34. Cabioglu N, Yazici MS, Arun B, Broglio KR, Hortobagyi GN, Price JE, et al. CCR7 and CXCR4 as novel biomarkers predicting axillary lymph node metastasis in T1 breast cancer. Clin Cancer Res. 2005;11:5686–93.

35. Kato M, Kitayama J, Kazama S, Nagawa H. Expression pattern of CXC chemokine receptor-4 is correlated with lymph node metastasis in human invasive ductal carcinoma. Breast Cancer Res. 2003;5.

36. Mandriota SJ, Jussila L, Jeltsch M, Compagni A, Baetens D, Prevo R, et al. Vascular endothelial growth factor-C-mediated lymphangiogenesis promotes tumour metastasis. EMBO J. 2001;20:672–82.

37. Skobe M, Hawighorst T, Jackson DG, Prevo R, Janes L, Velasco P, et al. Induction of tumor lymphangiogenesis by VEGF-C promotes breast cancer metastasis. Nat Med. 2001;7:192–8.

38. Stacker SA, Caesar C, Baldwin ME, Thornton GE, Williams RA, Prevo R, et al. VEGF-D promotes the metastatic spread of tumor cells via the lymphatics. Nat Med. 2001;7:186–91.

39. Karpanen T, Egeblad M, Karkkainen MJ, Kubo H, Ylä-Herttuala S, Alitalo K. Vascular endothelial growth factor C promotes tumor lymphangiogenesis and intralymphatic tumor growth. Cancer Res. 2001;61:1786–90.

40. Zhang S, Yi S, Zhang D, Gong M, Cai Y, Zou L. Intratumoral and peritumoral lymphatic vessel density both correlate with lymph node metastasis in breast cancer. Sci Rep. Nature Publishing Group; 2017;7:40364.

41. Jain RK. Intratumoral Lymphatic Vessels: A Case of Mistaken Identity or Malfunction? CancerSpectrum Knowl Environ. 2002;94:417–21.

42. Kakeji Y, Koga T, Sumiyoshi Y, Shibahara K, Oda S, Maehara Y, et al. Clinical significance of vascular endothelial growth factor expression in gastric cancer. J Exp Clin Cancer Res. 2002;21:125–9.

43. Mohammed RAA, Green A, El-Shikh S, Paish EC, Ellis IO, Martin SG. Prognostic significance of vascular endothelial cell growth factors -A, -C and -D in breast cancer and their relationship with angio- and lymphangiogenesis. Br J Cancer. 2007;96:1092–100.

44. Gu Y, Qi X, Guo S. Lymphangiogenesis induced by VEGF-C and VEGF-D promotes metastasis and a poor outcome in breast carcinoma: A retrospective study of 61 cases. Clin Exp Metastasis. 2008;25:717–25.

45. Bittner JJ. Some Possible Effects of Nursing on the Mammary Gland Tumor Incidence in Mice. Am J Clin Pathol. 1937;7:430–5.

46. Ross SR. Mouse mammary tumor virus molecular biology and oncogenesis. Viruses. 2010. p. 2000–12.

47. Wagner KU, Ward T, Davis B, Wiseman R, Hennighausen L. Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice. 2001;10:545–53.

48. Hennighausen L, Wall RJ, Tillmann U, Li M, Furth PA. Conditional gene expression in secretory tissues and skin of transgenic mice using the MMTV-LTR and the tetracycline responsive system. J Cell Biochem. 1995;59:463–72.

49. Wagner KU, Wall RJ, St-Onge L, Gruss P, Wynshaw-Boris A, Garrett L, et al. Cre-

mediated gene deletion in the mammary gland. Nucleic Acids Res. 1997;25:4323–30.

50. Sinn E, Muller W, Pattengale P, Tepler I, Wallace R, Leder P. Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: Synergistic action of oncogenes in vivo. Cell. 1987;49:465–75.

51. Guy CT, Cardiff RD, Muller WJ. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Mol Cell Biol. 1992;12:954–61.

52. Kühn R, Torres RM. Cre/loxP recombination system and gene targeting. . Methods Mol. Biol. 2002. p. 175–204.

53. Almholt K, Juncker-Jensen A, Lærum OD, Danø K, Johnsen M, Lund LR, et al. Metastasis is strongly reduced by the matrix metalloproteinase inhibitor Galardin in the MMTV-PymT transgenic breast cancer model. Mol Cancer Ther. 2008;7:2758–67.

54. Lifsted T, Le Voyer T, Williams M, Muller W, Klein-Szanto A, Buetow KH, et al. Identification of inbred mouse strains harboring genetic modifiers of mammary tumor age of onset and metastatic progression. Int J Cancer. 1998;77:640–4.

55. Kobayashi H, Kawamoto S, Sakai Y, Choyke PL, Star RA, Brechbiel MW, et al. Lymphatic drainage imaging of breast cancer in mice by micro-magnetic resonance lymphangiography using a nano-size paramagnetic contrast agent. J Natl Cancer Inst. 2004;96:703–8.

56. Lucchini F, Sacco MG, Hu N, Villa A, Brown J, Cesano L, et al. Early and multifocal tumors in breast, salivary, Harderian and epididymal tissues developed in MMTY-Neu transgenic mice. Cancer Lett. 1992;64:203–9.

57. Li Y, Hively WP, Varmus HE. Use of MMTV-Wnt-1 transgenic mice for studying the genetic basis of breast cancer. Oncogene. 2000. p. 1002–9.

58. Hollern DP, Andrechek ER. A genomic analysis of mouse models of breast cancer reveals molecular features of mouse models and relationships to human breast cancer. Breast Cancer Res . 2014;16:1–16.

59. Liu K, Newbury PA, Glicksberg BS, Zeng WZD, Paithankar S, Andrechek ER, et al. Evaluating cell lines as models for metastatic breast cancer through integrative analysis of genomic data. Nat Commun. 2019;10:1–12.

60. Rose DP, Connolly JM, Liu XH. Effects of Linoleic Acid and γ-Linolenic Acid on the Growth and Metastasis of a Human Breast Cancer Cell Line in Nude Mice and on its Growth and Invasive Capacity In Vitro. Nutr Cancer. 1995;24:33–45.

61. Puchalapalli M, Zeng X, Mu L, Anderson A, Glickman LH, Zhang M, et al. NSG mice provide a better spontaneous model of breast cancer metastasis than athymic (nude) mice. PLoS One. Public Library of Science; 2016;11.

62. Iorns E, Drews-Elger K, Ward TM, Dean S, Clarke J, Berry D, et al. A New Mouse Model for the Study of Human Breast Cancer Metastasis. PLoS One. Public Library of Science; 2012;7:e47995.

63. Vantyghem SA, Allan AL, Postenka CO, Al-Katib W, Keeney M, Tuck AB, et al. A new model for lymphatic metastasis: Development of a variant of the MDA-MB-468 human breast cancer cell line that aggressively metastasizes to lymph nodes. Clin Exp Metastasis. 2005;22:351–61.

64. Lee E, Pandey NB, Popel AS. Pre-treatment of mice with tumor-conditioned media accelerates metastasis to lymph nodes and lungs: A new spontaneous breast cancer metastasis model. Clin Exp Metastasis. 2014;31:67–79.

65. Aslakson CJ, Miller FR. Selective Events in the Metastatic Process Defined by Analysis of the Sequential Dissemination of Subpopulations of a Mouse Mammary Tumor. Cancer Res. 1992;52:1399–405.

66. Pulaski BA, Ostrand-Rosenberg S. Reduction of established spontaneous mammary carcinoma metastases following immunotherapy with major histocompatibility complex class II and B7.1 cell-based tumor vaccines. Cancer Res. 1998.

67. Zhang Y, Zhang N, Hoffman RM, Zhao M. Surgically-induced multi-organ metastasis in an orthotopic syngeneic imageable model of 4T1 murine breast cancer. Anticancer Res. 2015;35:4641–6.

68. Derose YS, Wang G, Lin YC, Bernard PS, Buys SS, Ebbert MTW, et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nat Med. 2011;17:1514–20.

69. Powell E, Shao J, Yuan Y, Chen HC, Cai S, Echeverria G V, et al. p53 deficiency linked to B cell translocation gene 2 (BTG2) loss enhances metastatic potential by promoting tumor growth in primary and metastatic sites in patient-derived xenograft (PDX) models of triple-negative breast cancer. Breast Cancer Res. 2016;18.

70. Lawson DA, Bhakta NR, Kessenbrock K, Prummel KD, Yu Y, Takai K, et al. Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature. 2015;526:131–5.

71. Liu X, Taftaf R, Kawaguchi M, Chang YF, Chen W, Entenberg D, et al. Homophilic CD44 interactions mediate tumor cell aggregation and polyclonal metastasis in

patient-derived breast cancer models. Cancer Discov. 2019;9:96–113.

72. Bockhorn J, Prat A, Chang YF, Liu X, Huang S, Shang M, et al. Differentiation and loss of malignant character of spontaneous pulmonary metastases in patient-derived breast cancer models. Cancer Res. 2014;74:7406–17.

73. Iaquinta PJ, Lees JA. Life and death decisions by the E2F transcription factors . Curr. Opin. Cell Biol. 2007. p. 649–57.

74. Lindeman GJ, Gaubatz S, Livingston DM, Ginsberg D. The subcellular localization of E2F-4 is cell-cycle dependent. Proc Natl Acad Sci U S A. 1997;94:5095–100.

75. Dimova DK, Dyson NJ. The E2F transcriptional network: old acquaintances with new faces. Oncogene . 2005;24:2810–26.

76. DeGregori J, Johnson D. Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. Curr Mol Med. 2012;6:739–48.

77. Kent LN, Leone G. The broken cycle: E2F dysfunction in cancer. Nat Rev Cancer. 2019;1.

78. Andrechek ER, Mori S, Rempel RE, Chang JT, Nevins JR. Patterns of cell signaling pathway activation that characterize mammary development. Development . 2008;135:2403–13.

79. Andrechek ER. HER2/Neu tumorigenesis and metastasis is regulated by E2F activator transcription factors. Oncogene . 2015;34:217–25.

80. Hollern DP, Honeysett J, Cardiff RD, Andrechek ER. The E2F Transcription Factors Regulate Tumor Development and Metastasis in a Mouse Model of Metastatic Breast Cancer. Mol Cell Biol . 2014;34:2020.

81. Fujiwara K, Yuwanita I, Hollern DP, Andrechek ER. Prediction and Genetic Demonstration of a Role for Activator E2Fs in Myc-Induced Tumors. 2011

82. Yuwanita I, Barnes D, Monterey MD, O ’reilly S, Andrechek ER. Increased metastasis with loss of E2F2 in Myc-driven tumors. Oncotarget . 2015;6.

83. Chen H-Z, Tsai S-Y, Leone G. RB and E2f: Emerging roles of E2Fs in cancer: an exit from cell cycle control. Nat Rev Cancer . 2009;9.

84. Cam H, Balciunaite E, Blais A, Spektor A, Scarpulla RC, Young R, et al. A common set of gene regulatory networks links metabolism and growth inhibition. Mol Cell. Cell Press; 2004;16:399–411.

85. Wells J, Graveel CR, Bartley SM, Madore SJ, Farnham PJ. The identification of

E2F1-specific target genes. Proc Natl Acad Sci U S A. 2002;99:3890–5.

86. Takahashi Y, Rayman JB, Dynlacht BD. Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. 2000.

87. Jin VX, Rabinovich A, Squazzo SL, Green R, Farnham PJ. A computational genomics approach to identify cis-regulatory modules from chromatin immunoprecipitation microarray data - A case study using E2F1. Genome Res. 2006;16:1585–95.

88. Bieda M, Xu X, Singer MA, Green R, Farnham PJ. Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. Genome Res. 2006;16:595–605.

89. Xu X, Bieda M, Jin VX, Rabinovich A, Oberley MJ, Green R, et al. A comprehensive ChIP-chip analysis of E2F1, E2F4, and E2F6 in normal and tumor cells reveals interchangeable roles of E2F family members. Genome Res. 2007;17:1550–61.

# CHAPTER 2:

# TRANSCRIPTION FACTOR COMPENSATION DURING MAMMARY GLAND DEVELOPMENT IN E2F KNOCKOUT MICE

## Preface

This chapter has previously been published as:

To, Briana, and Eran R. Andrechek. "Transcription factor compensation during mammary gland development in E2F knockout mice." *PloS one* 13.4 (2018): e0194937.

**Abstract**

The E2F transcription factors control key elements of development, including mammary gland branching morphogenesis, with several E2Fs playing essential roles. Additional prior data has demonstrated that loss of individual E2Fs can be compensated by other E2F family members, but this has not been tested in a mammary gland developmental context. Here we have explored the role of the E2Fs and their ability to functionally compensate for each other during mammary gland development. Using gene expression from terminal end buds and chromatin immunoprecipitation data for E2F1, E2F2 and E2F3, we noted both overlapping and unique mammary development genes regulated by each of the E2Fs. Based on our computational findings and the fact that E2Fs share a common binding motif, we hypothesized that E2F transcription factors would compensate for each other during mammary development and function. To test this hypothesis, we generated RNA from E2F1$^{-/-}$, E2F2$^{-/-}$ and E2F3$^{+/-}$ mouse mammary glands. QRT-PCR on mammary glands during pregnancy demonstrated increases in E2F2 and E2F3a in the E2F1$^{-/-}$ mice and an increase in E2F2 levels in E2F3$^{+/-}$ mice. During lactation we noted that E2F3b transcript levels were increased in the E2F2$^{-/-}$ mice. Given that E2Fs have previously been noted to have the most striking effects on development during puberty, we hypothesized that loss of individual E2Fs would be compensated for at that time. Double mutant mice were generated and compared with the single knockouts. Loss of both E2F1 and E2F2 revealed a more striking phenotype than either knockout alone, indicating that E2F2 was compensating for E2F1 loss. Interestingly, while E2F2 was not able to functionally compensate for E2F3$^{+/-}$ during mammary outgrowth, increased E2F2 expression was observed in E2F3$^{+/-}$ mammary glands during

pregnancy day 14.5 and lactation day 5. Together, these findings illustrate the specificity of E2F family members to compensate during development of the mammary gland.

## Introduction

The mouse mammary gland is composed of an arborized epithelial network embedded within a fat pad. At puberty, the mammary epithelium rapidly expands from a rudimentary structure with few branches to form Terminal End Buds (TEBs) that drive epithelial growth into the fat pad. These large club-shaped structures have a leading edge composed of cap cells that rapidly proliferate. As the cells migrate back into the center of the TEB they undergo apoptosis, forming a hollow tube [1], while those on the periphery differentiate into the luminal and myoepithelial layers. Once the epithelial network completely fills the fat pad, the TEBs are lost and the gland becomes largely static with small estrous related alterations. Upon the initiation of pregnancy, there is rapid proliferation and differentiation with well-regulated transcriptional programs to generate a lactating mammary gland. After weaning of the pups, apoptosis and remodeling of the gland occurs to return the mammary ductal network to a state closely resembling the nulliparous gland. This change and return to a virgin like state is also reflected in transcriptional programs, readily seen through a principle component plot [2]. Transcriptional studies of genes differentially expressed between the TEB and the ducts revealed numerous drivers of mammary growth [3, 4]. A study using a shRNA knockdown approach in Mammary Stem Cells (MaSC) identified a series of novel genes that influence MaSC and their ability to function as mammary stem or progenitor cells [5]. These studies, along with other transcriptional studies of mammary development [6, 7]

have defined many of the transcriptional programs involved with mammary development and function.

In our previous analysis of mammary gland development, we predicted a role for the E2F family of transcription factors in mammary gland development [8]. The E2F transcription factors are commonly thought to regulate cell cycle, proliferation and apoptosis [9]. Individual developmental roles have been established with the knockouts of E2F1[10], E2F2[11], E2F3[12], E2F4[13] and E2F5[14]. More specifically, the role of E2F1-3 in development have been examined in the setting of the retina and small intestine using tissue specific knockouts [15, 16]. Interestingly, these E2F1-3 tissue specific triple knockout demonstrated that E2F1-3 are not crucial for normal cell proliferation but are needed for cell survival. However, in the setting of mammary gland development, E2F1 and E2F3 appear to have a role in proliferation and differentiation as evident by the delay in mammary gland outgrowth and branching defects. A delay in mammary gland outgrowth and branching defects was not observed in E2F2$^{-/-}$ mice [8]. Outgrowth in this study was examined at 4 and 8 weeks of development with effects observed at both timepoints. However, virgin adult glands were not significantly different. In addition, loss of one copy of E2F3 resulted in a slight delay in involution. However, this effect was not noted in the other activator E2F knockouts. No effects were noted during pregnancy or lactation for the knockout strains [8]. Importantly, the binding motif for individual E2Fs is not distinct [17], requiring other proximal regions [18], likely indicating that E2Fs function with other co-activators [19]. As such, the loss of individual E2Fs in mammary gland development may therefore be compensated by other E2F family members [20]. This

may occur both in developmental contexts as well as in cancer [21]. However, the extent of compensation in the mammary gland has not previously been explored.

Here we have integrated the development transcriptional studies [6, 7] with TEB gene expression [4] and the shRNA screen for MaSC genes [5] to predict a strong role for E2F activity in mammary gland proliferation and differentiation. Given the ability of the E2Fs to compensate, we then noted altered expression of E2Fs in individual E2F knockout backgrounds at various stages of mammary development. Examining mice lacking multiple E2Fs, we noted mammary gland outgrowth effects that were more extensive than for individual knockouts alone, indicating that there was significant compensation occurring. Combining our bioinformatic predictions with the individual and double E2F knockouts demonstrates the role of E2F compensation in mammary gland development.

**Results**

In order to explore the role of the E2F transcription factors in the regulation of mammary gland stem cells and progenitor cells, we have integrated the published shRNA screen data [5] with E2F signatures [23, 25] (Figure 1A). The shRNA screen was done in a CD29hi CD24+ subset of MaSC enriched basal cells which is a subpopulation of mammary cells that have the ability to reconstitute a complete mammary gland in vivo[5]. A customized mouse lentiviral library composed of 1,294 shRNAmirs targeted against genes involved in transcriptional regulation was used in the screen. The shRNA screen identified 73 genes that potentially regulate mammary stem and progenitor cell behavior, including genes with no previous implications in mammary gland development. The E2F

signature data was derived from human mammary epithelial cells (HMECs) that had either GFP control adenovirus or E2F1, E2F2 or E2F3 adenovirus expression (23). Differentially expressed genes relative to the GFP control HMECs were identified for each of the activator E2Fs (Figure 1B). The genes in the various E2F signatures and in the shRNA screen for mammary stem and progenitor cells were then compared through a Venn diagram, revealing 34 shRNA screen genes that were potentially regulated by E2F1, 1 by E2F2 and 37 by E2F3. In addition, only 21 of the 73 potential stem and progenitor genes were not contained in an E2F signature dataset. Given the prevalence of E2F targets in the list of genes that regulate mammary stem and progenitor cells, we examined a mammary gland developmental gene expression dataset [6] for E2F1 targets. To accomplish this, we used the E2F1 signature genes [25], a subset of those shown in Figure 1B, defined through a binary regression approach. We clustered the mammary gland development dataset using only the genes present in the E2F1 signature, thus only genes that are regulated by E2F1 were used. Strikingly, this revealed that E2F1 target genes alone were able to cluster the various developmental phases assayed by gene expression (Figure 1C). This illustrates that E2F1 regulated genes are differentially expressed throughout the mammary developmental stages. Given the overlap in differentially regulated E2F genes (Figure 1B), we noted which genes were also contained in the E2F2 and E2F3 signatures (Figure 1C, right). Based on the small overlap between E2F2 targets and MaSC and progenitor cells (Figure 1B), it was not surprising that the E2F2 signature genes did not resolve the stages of mammary differentiation as clearly as E2F1 (S1A Figure). E2F3 signature genes also stratified mammary development stages in a manner similar to E2F1 (S1B Figure). To determine

if a particular E2F is able to resolve the stages of mammary gland developments, we assessed how well late pregnancy, lactation and involution stages were able to be stratified.

With the noted roles for E2F1 and E2F3 in mammary stem cell and progenitor cells, we examined a TEB and duct gene expression dataset for E2F targets. Using a TEB vs duct gene expression dataset that was generated from microdissecting and enzymatic processing of TEBs and ducts (Figure 2A) [4], we identified a list of differentially expressed TEB genes with E2F gene enrichment. The TEB / duct gene list was compared to the gene expression from the E2F signatures (Figure 2B) and from E2F1, E2F2 and E2F3 ChIP-seq/ChIP-chip data (Figure 2C). Upon analysis of the genes that were shared between the differentially expressed TEB/duct genes and E2F signatures, we noted that E2F1 regulated genes were the most represented group (Figure 2B). These genes were equally split between being up regulated in the TEB and upregulated in ducts. In addition to key E2F signature genes, we also explored which direct targets from

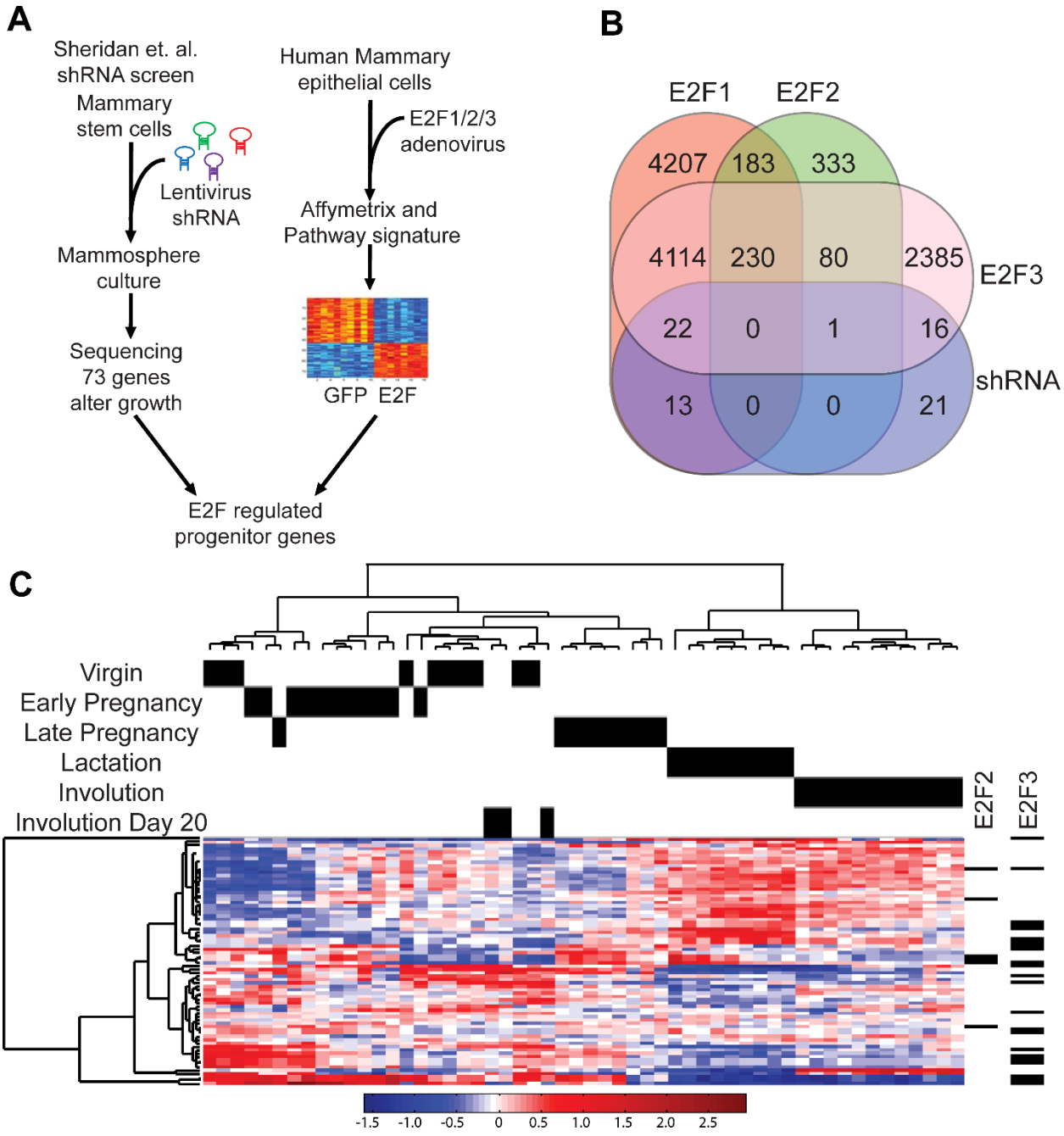


**Figure 2.1 E2F transcription factors define mammary developmental states. A.** Here we have combined a shRNA screen in primary mouse MaSC-enriched basal cells with E2F1-3 gene expression signatures to define progenitor genes regulated by the E2F transcription factors. **B.** Using a Venn diagram overlap between E2F1, E2F2 and E2F3

**Figure 2.1 (cont'd)** target genes as determined by CHIP-seq/CHIP-chip and microarray experiments, and the 73 unique genes identified from the pooled shRNA screen was noted. **C.** Unsupervised hierarchical clustering of E2F1 signature genes from mammary glands collected at various stages of mammary development is stratified into the indicated stages. The overlap between E2F1, E2F2 and E2F3 signature genes is depicted on the right of the of the heatmap.

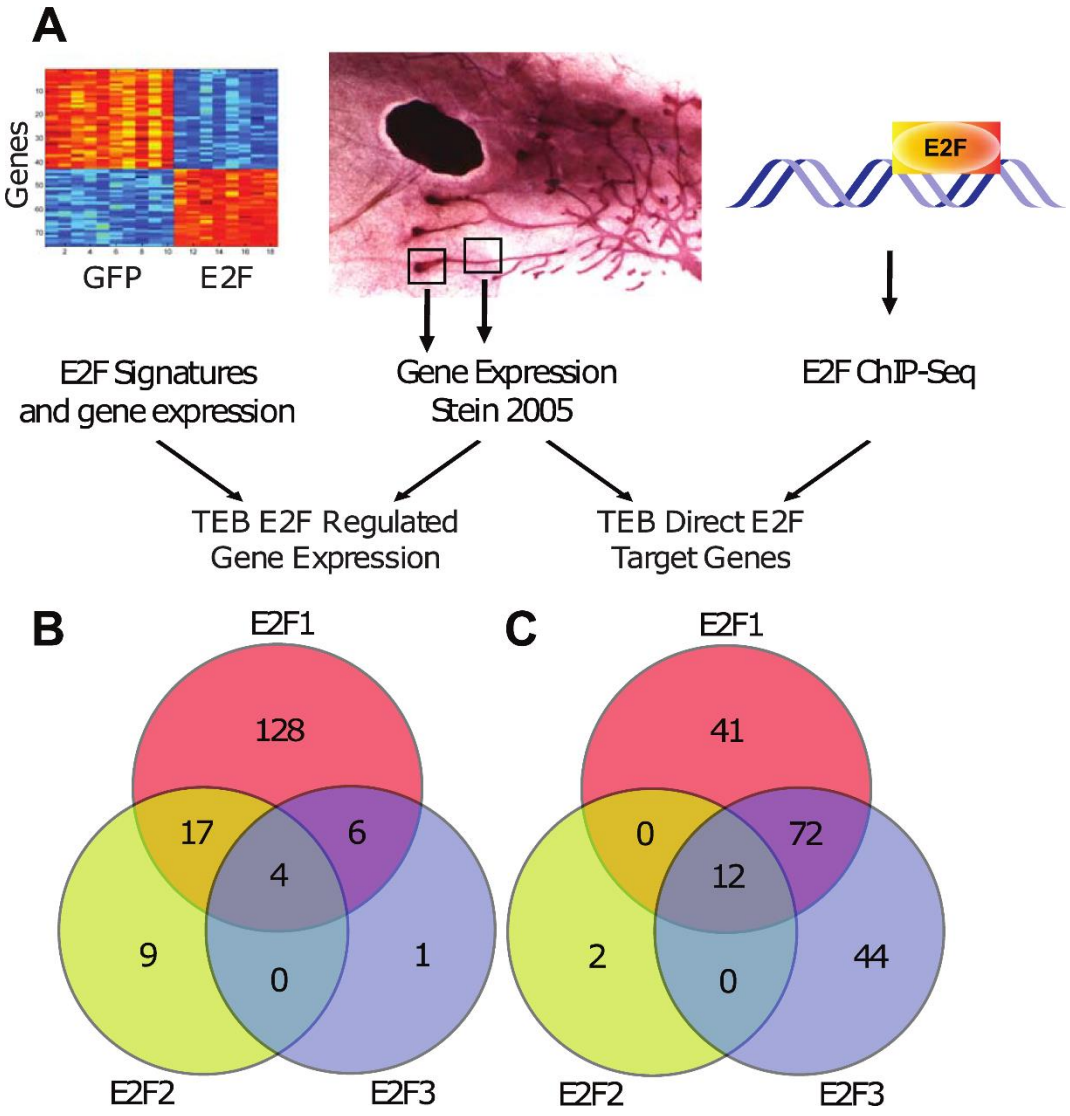


**Figure 2.2 E2F regulated genes in the terminal end bud. A.** Combining the E2F1, E2F2 and E2F3 signature genes and differentially expressed genes derived from microarray with a TEB / duct dataset revealed E2F regulated genes in the TEB. Further, E2F ChIP-seq and ChIP-chip data was combined with this TEB/duct dataset to predict direct E2F targets in the TEB. **B.** Using Venn Diagrams to compare TEB E2F1, E2F2 and

**Figure 2.2 (cont'd)** E2F3 signature genes/differentially expressed genes revealed which genes were regulated singly and by multiple E2F transcription factors. **C.** Similarly, direct TEB E2F1, E2F2 and E2F3 targets were analyzed.

E2F ChIP-Seq/ChIP-chip experiments were represented in the TEB / duct dataset. This analysis clearly demonstrated that again E2F1 and E2F3 target genes were largely shared in the TEB and duct genes. In addition, we also noted 41 TEB genes that were direct targets of only E2F1 and 44 TEB genes that were direct targets of only E2F3. These two sets of differentially regulated E2F1 or E2F3 targets include genes that are involved in cell proliferation and division (S1 Table) This list includes genes that have previously described roles in mammary gland development including *Wnt5a*[32], *Cebpd*[33] and *Pttg1*. For example, deletion of *Pttg1*, which is regulated by E2F1, was reported to result in a defect in mammary gland branching and progression[34].

In order to test the premise that the key mammary stem cell and progenitor genes were regulated by E2F transcription factors, we examined the expression of these genes throughout the mammary gland developmental cycle (Figure 3A). Filtering a mammary gland development dataset using only the list of 73 potential mammary stem cell and progenitor genes resulted in a list of 36 overlapping genes. Clustering based on these 36 genes alone stratified the mammary gland development gene dataset into various developmental stages. In addition, we demonstrated that many of these genes were either predicted (GATHER) or experimentally defined by ChIP-Seq/ChIP-chip to be E2F1 target genes. To further confirm E2F1 role in regulating mammary development, we performed a broad analysis of transcription factor binding (23) and isolated the genes that played a potential role in mammary stem cell development. This analysis demonstrated that amongst the group of various transcription factors, E2F1 was third in regulating the highest number of potential mammary stem cell genes (Fi g 3 B). Together, these findings

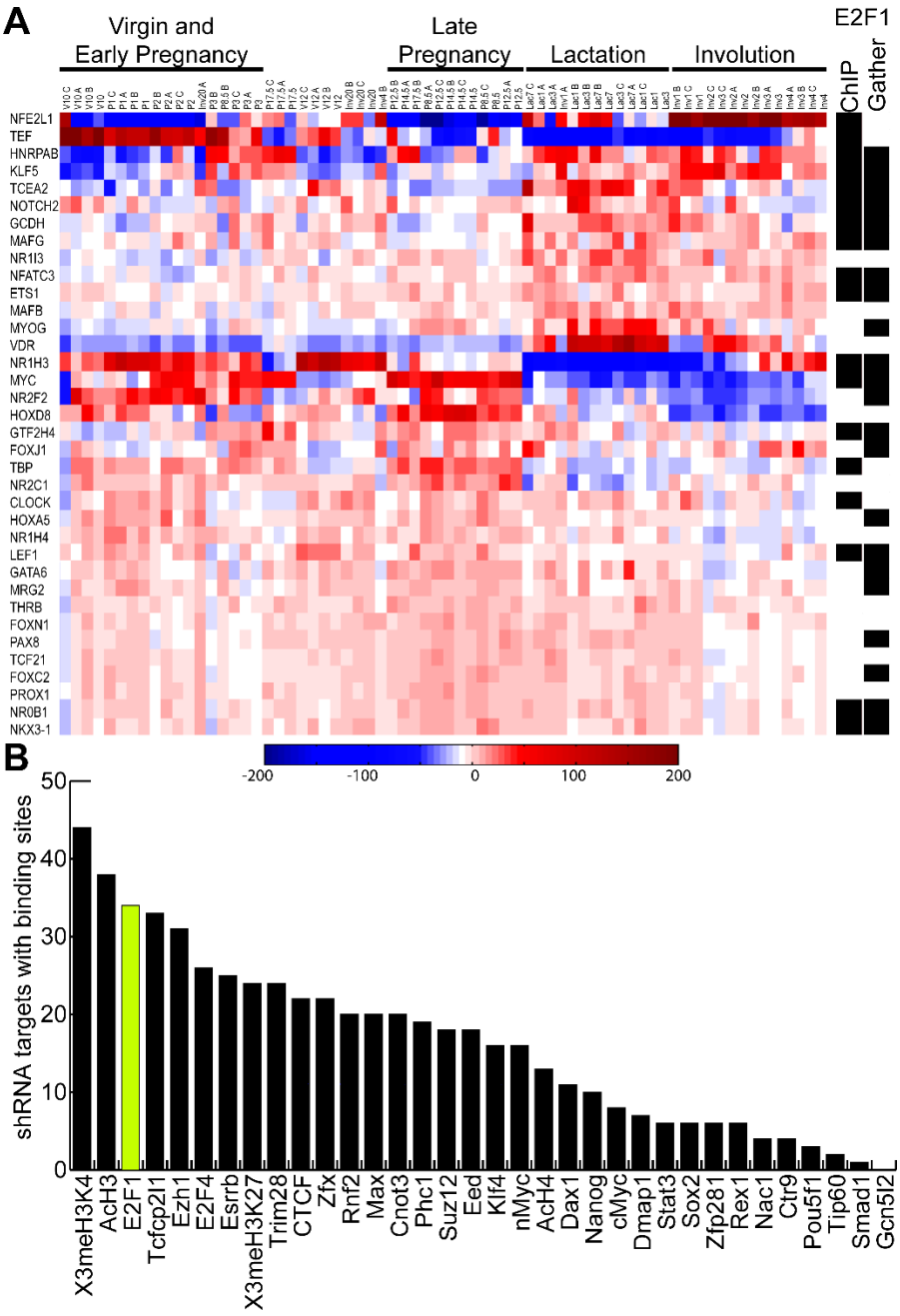


**Figure 2.3 Progenitor genes, mammary stages and transcription factor enrichment.**

**A.** Unsupervised hierarchical clustering of 17 mammary gland developmental timepoints using genes with a potential role in mammary gland development as determined by the pooled shRNA screen. Each of the 17 developmental timepoints have three replicates.

**Figure 2.3 (cont'd)** The overlap between potential mammary stem/progenitor cell genes and E2F1 direct and predicted target genes as determined by ChiP-ChIP and GATHER is depicted on the right side of the heatmap. The predicted E2F1 targets have a p-value of 0.002 and a Bayes factor of 3. **B.** Examining transcription factors with overlap of ChIP data in the various stages of mammary development highlights the importance of these genes in mammary gland function. The E2F1 transcription factor is highlighted (yellow) with over 30 of the 73 genes being regulated.

along with the mammary development clustering data suggest that E2F1 and E2F3 are of critical importance to mammary gland development and function.

In addition to the E2F transcription factors sharing a binding motif [18, 19], they are critical across all stages of mammary development (Figure 1) and are known to compensate for loss of other E2Fs [20, 21]. We therefore hypothesized that E2F transcription factors would compensate for each other during various stages of mammary development and function. Given that no major effects were previously noted [8] during pregnancy and lactation in E2F knockout mice, we hypothesized that the E2Fs were compensating for each other at this stage. To test this theory, mammary glands from control, E2F1$^{-/-}$ , E2F2$^{-/-}$ and E2F3$^{+/-}$ mice were collected and assayed by QRT-PCR for the other activator E2Fs. Relative to the wild type control, we observed a significant increase in E2F2 and E2F3a levels in the E2F1$^{-/-}$ mice at pregnancy day 14.5 ($p<0.05$ for both). Similarly, E2F3$^{+/-}$ mice had a large increase in only E2F2 levels during pregnancy (Figure 4A)($p<0.05$). In addition to pregnancy, mammary glands from the 5$^{th}$ day of lactation were examined. QRT-PCR at this timepoint demonstrated that there was a clear increase in E2F3b levels in the E2F2$^{-/-}$ mice (Figure 4B)($p<0.05$).

In a prior study of E2F function during development, it was noted that loss of individual E2Fs delayed mammary gland outgrowth [8] thus we hypothesized that the absence of multiple E2Fs would lead to greater outgrowth defect due to the disruption of a potential compensatory mechanism. In order to directly test the premise that E2Fs functionally could compensate for each other we interbred E2F1$^{-/-}$, E2F2$^{-/-}$ and E2F3$^{+/-}$ mice to generate double knockout strains. Wholemounts of the mammary glands (Figure 5A-D) revealed intriguing findings for each of the double knockout mice. Relative to wildtype

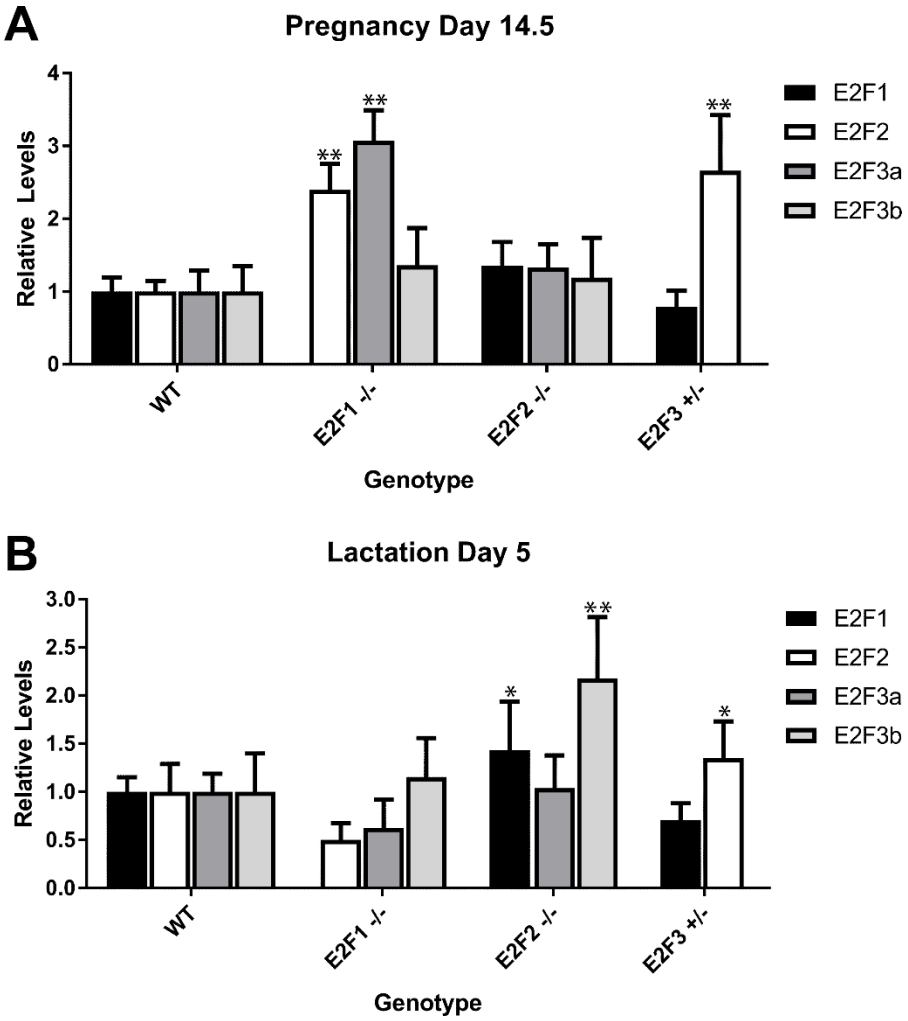


**Figure 2.4 Compensation during mammary gland function. A.** The levels of E2F1, E2F2, E2F3a and E2F3b were quantified in mammary gland RNA at pregnancy day 14.5 in wild type (WT) FVB (n = 3), E2F1-/- (n = 3), E2F2-/- (n = 3) and E2F3+/- (n = 3) mice. A significant increase in E2F2 and E2F3a was observed in E2F1-/- mice. A significant increase in E2F2 was also observed in E2F3+/- mice. **B.** The levels of E2F1, E2F2, E2F3a and E2F3b were quantified at lactation day 5 in WT FVB (n = 3), E2F1-/- (n = 3), E2F2-/- (n = 3) and E2F3+/- (n = 3) mice. A significant increase in E2F1 and E2F3b was observed in the E2F2 null mice. A significant increase in E2F2 was also noted in E2F3 heterozygous

**Figure 2.4 (cont'd)** knockout mice. This analysis revealed differential E2F specific compensatory gene expression dependent upon the developmental context of the mammary gland. A significant increase in E2F expression levels relative to WT control mice are depicted with an asterisk. * p value < 0.05 and ** p value < 0.006.

control, E2F1$^{-/-}$/E2F2$^{-/-}$ mice had a 40% delay in outgrowth (Figure 5E). This is a significantly greater delay in mammary outgrowth in comparison to the 20% reduction delay seen in E2F1$^{-/-}$ mice and lack of delay in E2F2$^{-/-}$ mice. These data indicate that E2F1 and E2F2 compensate for each other, resulting in a dramatic effect on the outgrowth of the double knockouts (Figureure 5E). The E2F1$^{-/-}$/ E2F3$^{+/-}$ mice had a mild additive effect (Figure 5F), despite sharing many potential targets (Figure 2). Finally, the E2F2 $^{-/-}$/E2F3$^{+/-}$ mice demonstrated that E2F2 was not able to compensate for loss of a copy of E2F3, with E2F3$^{+/-}$ mice having the same outgrowth as the E2F2$^{-/-}$/E2F3$^{+/-}$ mice (Figure 5G). In all of the double knockout crosses, the mammary epithelial network was fully formed at 16 weeks of age, indicating that the outgrowth was only delayed. Taken together, these data indicate the specificity of the compensatory mechanisms in place for the E2F transcription factor family.

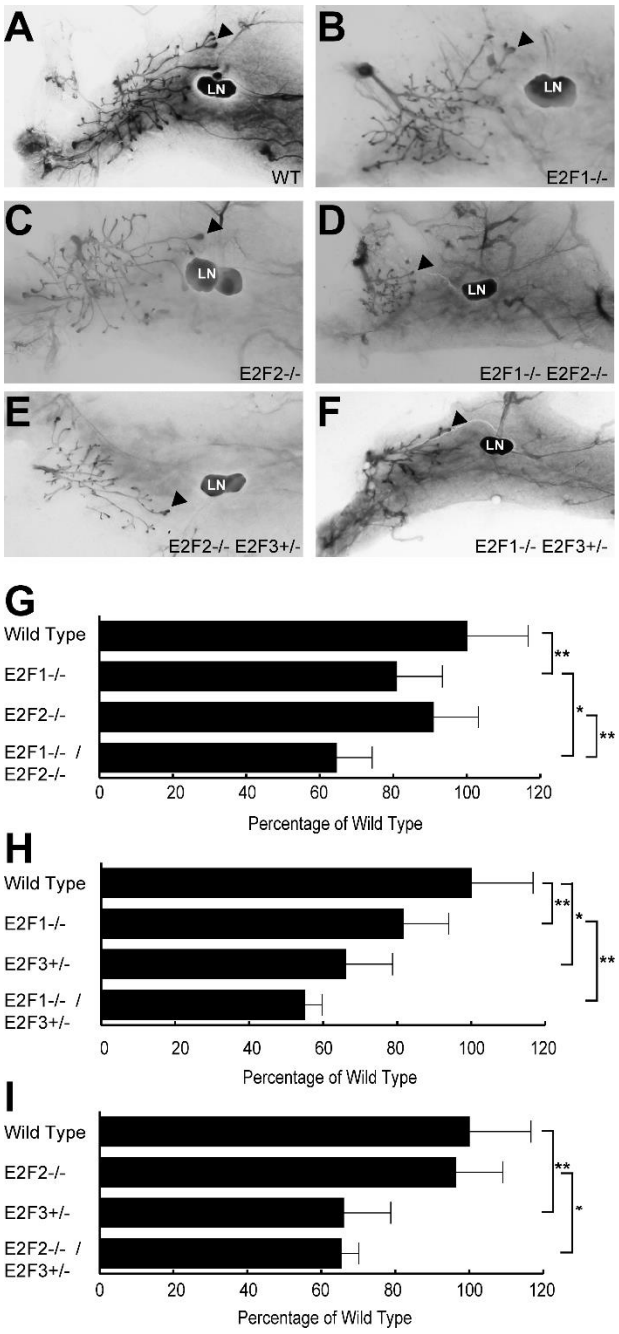


**Figure 2.5 Mammary gland outgrowth in double knockouts.** Representative wholemount images of mammary gland outgrowth at 4 weeks of development in wild type controls (A), E2F1$^{-/-}$ (B), E2F2$^{-/-}$ (C), E2F1$^{-/-}$/E2F2$^{-/-}$ (D) E2F2$^{-/-}$/E2F3$^{+/-}$ (E), E2F1$^{-/-}$ /E2F3$^{+/-}$ (F) mice are shown. Measuring from nipple to lymph node (LN) as well as from

**Figure 2.5 (cont'd)** nipple to the most distal TEB (arrowhead) allowed for quantification of mammary gland outgrowth. The results of the E2F1/E2F2 (E), E2F1/E2F3 (F) and E2F2/E2F3 (G) mammary outgrowth experiments are shown as a percentage of the wild type control growth. The mammary outgrowth was quantified in 12 WT FVB, 11 $E2F1^{-/-}$, 15 $E2F2^{-/-}$, 8 $E2F3^{+/-}$, 5 $E2F1^{-/-}/E2F2^{-/-}$, 5 $E2F1^{-/-}/E2F3^{+/-}$ and 3 $E2F2^{-/-}/E2F3^{+/-}$. There is a significantly greater delay in mammary outgrowth in the $E2F1^{-/-}/E2F2^{-/-}$ mice relative to the $E2F1^{-/-}$ mice. Differences in mammary outgrowth delay that are significant between two strains are depicted with an asterisk. * p value < 0.05 and ** p value < 0.006.

## Discussion

Here we have integrated shRNA screen data, gene expression signatures, developmental gene expression and ChIP-chip/ChIP-seq data to predict an overlapping role for E2F transcription factors in mammary gland development and function. This prediction was validated at a gene expression level in knockout mice during pregnancy and lactation and was experimentally tested in double knockout mice. The double knockout mice revealed specificity in the ability of E2Fs to compensate for loss of other family members. Together this integrative study underscores the importance of combining multiple large scale datasets to pose experimental questions.

The various bioinformatic analyses presented here strongly illustrates the importance of the E2F pathway in regulation of mammary gland development and function. This concept is reinforced by examination of the stratification of mammary gland developmental phases by genes from the E2F1 signature alone (Figure 1C). These effects are mediated through the regulation of genes within the TEB as well as later in functional states including lactation and involution. While E2F transcription factors are essential in this process, other transcription factors were also noted to potentially regulate the mammary stem cell and progenitor genes (Figure 3B). Interestingly, many of these promoters were noted to potentially regulate E2F activity through transcriptional repression and/or directly inhibiting its activity. This included transcription factors like Trim28 (Kap1), which has previously been shown to transcriptionally repress E2F1 and can also directly bind and inhibit its activity [35]. CTCF was identified in this assay and also has potential E2F co-factor activity [36]. Rnf2 was noted to regulate E2F1 activity

through transcriptional repression after binding E2F promoters [37]. Together, these data illustrate the importance of the E2F transcription factors in the mammary stem cells.

The role of E2Fs in mammary gland development has been observed in our prior work [8], but the integration of various datasets indicated the potential for compensatory activity based on shared targets. While E2F compensation has been previously reported in both a tissue culture setting [20] as well as in our tumor models [21], the ability for E2Fs to compensate during mammary developmental processes has not previously been examined. Additionally, the specificity of the collaborative potential by individual E2Fs has not been explored. Based on the overlap of ChIP-seq/ChIP-chip data, E2F1 and E2F3 appeared to have significant potential for overlap. Interestingly, the loss of E2F1 resulted in an increase in both E2F2 and E2F3a expression during pregnancy. Examining the TEB genes that are direct targets of E2Fs also revealed shared genes between E2F1 and E2F3. Despite this, E2F1$^{-/-}$ and E2F3$^{+/-}$ mice both demonstrated delay in mammary outgrowth suggesting incomplete compensation. Given that mammary outgrowth is controlled by epithelial cells in the TEB, we believe that the TEB genes that are direct targets of either E2F1 or E2F3 are potentially responsible for the incomplete compensation during ductal elongation (Figure 2C). Furthermore, this list of differentially regulated E2F1 or E2F3 targets include genes involved in cell proliferation and division. Future work needs to be done to functionally test if these differentially regulated genes are responsible for the incomplete compensation. In addition, we identified a list of shared E2F1, E2F2 and E2F3 TEB targets that are potentially responsible for ductal development. However, whether knockout of these specific targets genes will functionally

disrupt the compensation activity seen in the mammary gland of E2F individual knockout mice require further investigation.

Moreover, in the individual knockout experiments, E2F2 loss normally had no effect on mammary outgrowth, but the dual loss of E2F1 and E2F2 revealed that E2F2 was partially compensating for the loss of E2F1 since the delay was more profound in the double knockout. However, E2F2 was not able to compensate for the reduction in function of E2F3. Together these data illustrate the nature of the specificity in the ability of the E2F family members to compensate in mammary development and ductal morphogenesis

The study of development is important to understanding other conditions, including cancer biology. In considering tumor biology, we have previously generated E2F knockout strains in MMTV-PyMT [38], MMTV-Neu [21] and MMTV-Myc transgenic mice [39]. In these experiments, we frequently noted altered gene expression of both direct and indirect E2F target genes. Additionally, increased metastasis with the loss of E2F2 in Myc induced tumors [40], and decreased metastasis with loss of E2F1 or E2F2 in Neu and PyMT induced tumors was observed. Consistent with these findings, E2F1 and E2F3 specific TEB targets includes genes that were previously reported to be involved with mammary tumor progression and metastasis in mice. Examples of these genes include *Loxl2*[41], *Klf4*[42], *Pdgfr*[43] and *Atf3*[44]. Importantly, previous studies have shown that elevated expression of upregulated genes in the E2F1 and E2F2 signature genes are associated with decreased time to distant metastasis free survival in breast cancer patients compared to those with low expression of these genes[21, 38, 40]. Based on these observations, we believe the biological relevance of the compensatory mechanism of E2Fs need to be carefully studied to understand their potential role in cancer biology.

## Material and Methods

### ***Animals***

All mice were bred and maintained according to guidelines and protocols approved by the Institutional Animal Care and Use Committee in Michigan State University. Euthanasia was performed as mandated and approved through CO2 followed by a secondary method including necropsy. E2F1$^{-/-}$, E2F2$^{-/-}$, and E2F3$^{+/-}$ mice were interbred to generate double knockouts in the FVB background. For wholemount generation, the inguinal mammary gland was excised and stained with Harris Modified Hematoxylin. To quantify the mammary epithelial outgrowth, the distance from the nipple to the leading edge of the epithelium and the distance from the nipple to the midpoint of the thoracic lymph node was measured. For the control, the ratio of the distance of outgrowth and distance between the lymph node and nipple was calculated. The ratio was set to 100% and used as standard to compare with various knockouts.

### ***Computational***

Genes for the shRNA screening experiment were obtained from a public dataset [5]. The pooled shRNA screen was done by infecting MaSC-enriched basal cells with a customized mouse lentiviral library consisting of 1,296 shRNAmirs. The study identified potential regulators of MaSCs by observing altered mammosphere growth in the non-adherent mammosphere formation assay. TEB / duct genes were extracted from published data [4]. The TEB and duct were isolated mechanically and enzymatically from microdissected mammary gland of pubertal Balb/C mice aged 5-6 weeks. E2F1, E2F2 and E2F3 gene expression data and binary regression signature methodology used to generate signatures was as previously described [9, 22-25]. ChIP-seq data for numerous

transcription factors was downloaded from public data [26, 27]. E2F1, E2F2 and E2F3 ChIP-Seq and ChIP-chip data was obtained from public sources [17, 28-30]. E2F1 ChIP-chip analysis was done in MCF10A cell lines. E2F2 ChIP-chip analysis was done in T lymphocytes isolated from 4 week old C57B16:129SV mice. E2F3 ChIP-chip and ChIP-seq analysis was completed in HCT116 cells and mouse myoblast and myotubes.

Unsupervised hierarchical clustering was completed with Cluster 3.0 using Euclidean distance and complete linkage and was visualized in Java Treeview. Heatmaps were altered to a blue/red color scheme using Matlab to ensure red-green color-blind viewers could distinguish the heatmap colors.

Transcription factor predictions based on motifs were generated using GATHER [31] and included the Bayes factor for statistical tests of enrichment. Other statistical tests were run using GraphPad Prism software and included Fishers two-tailed and t-tests.

***qRT-PCR***

Mammary glands from three WT FVB, three E2F1$^{-/-}$, three E2F2$^{-/-}$ and three E2F3$^{+/-}$ were excised on pregnancy day 14.5 and lactation day 5. Both of the number 4 inguinal mammary glands were collected from each mouse during necropsy. The mammary glands were snap frozen in liquid nitrogen and stored at -80ºC. Pregnancy was confirmed through observation. For lactation samples, litters were standardized to 6 pups and glands were excised 4 hours after pups were removed. RNA was extracted from flash frozen mammary glands with the Qiagen RNeasy midi kit. Quantitative RT-PCR was performed using a SYBR Green One-Step RT-PCR Kit (Qiagen). The following primers were used (5' to 3'): E2f1 forward, CGATTCTGACGTGCTGCTCT and reverse, CAGCGAGGTACTGATGGTCA; E2f2 forward, GCGCATCTATGACATCACCA and

reverse, CGGGTGGGGTCTTCAAATAG; E2f3a forward, CCAGCAGCCTCTACACCAC and reverse, GGTACTGATGGCCACTCTCG; E2f3b forward, CTTTCGGAAATGCCCTTACA and reverse, GGTACTGATGGCCACTCTCG; Gapdh forward, TCATGACCACAGTGGATGCC and reverse, GGAGTTGCTGTTGAAGTCGC. Relative change was calculated using the ΔΔCt method. Statistical analysis was performed using an unpaired T-test.

## WORKS CITED

## WORKS CITED


1. Hennighausen L, Robinson GW. Think globally, act locally: the making of a mouse mammary gland. Genes Dev. 1998;12(4):449-55. Epub 1998/03/21. PubMed PMID: 9472013.

2. Master SR, Hartman JL, D'Cruz CM, Moody SE, Keiper EA, Ha SI, et al. Functional microarray analysis of mammary organogenesis reveals a developmental role in adaptive thermogenesis. Mol Endocrinol. 2002;16(6):1185-203. Epub 2002/06/01. PubMed PMID: 12040007.

3. Kouros-Mehr H, Werb Z. Candidate regulators of mammary branching morphogenesis identified by genome-wide transcript analysis. Dev Dyn. 2006;235(12):3404-12. Epub 2006/10/14. doi: 10.1002/dvdy.20978. PubMed PMID: 17039550; PubMed Central PMCID: PMC2730892.

4. Morris JS, Stein T, Pringle MA, Davies CR, Weber-Hall S, Ferrier RK, et al. Involvement of axonal guidance proteins and their signaling partners in the developing mouse mammary gland. J Cell Physiol. 2006;206(1):16-24. doi: 10.1002/jcp.20427. PubMed PMID: 15920758.

5. Sheridan JM, Ritchie ME, Best SA, Jiang K, Beck TJ, Vaillant F, et al. A pooled shRNA screen for regulators of primary mammary stem and progenitor cells identifies roles for Asap1 and Prox1. BMC Cancer. 2015;15:221. doi: 10.1186/s12885-015-1187-z. PubMed PMID: 25879659; PubMed Central PMCID: PMCPMC4399223.

6. Stein T, Morris JS, Davies CR, Weber-Hall SJ, Duffy MA, Heath VJ, et al. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. Breast Cancer Res. 2004;6(2):R75-91. Epub 2004/02/26. doi: 10.1186/bcr753bcr753 [pii]. PubMed PMID: 14979920; PubMed Central PMCID: PMC400652.

7. Clarkson RW, Wayland MT, Lee J, Freeman T, Watson CJ. Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. Breast Cancer Res. 2004;6(2):R92-109. Epub 2004/02/26. doi: 10.1186/bcr754 bcr754 [pii]. PubMed PMID: 14979921; PubMed Central PMCID: PMC400653.

8. Andrechek ER, Mori S, Rempel RE, Chang JT, Nevins JR. Patterns of cell signaling pathway activation that characterize mammary development. Development. 2008;135(14):2403-13. Epub 2008/06/14. doi: dev.019018 [pii] 10.1242/dev.019018. PubMed PMID: 18550711.

9. Black EP, Hallstrom T, Dressman HK, West M, Nevins JR. Distinctions in the specificity of E2F function revealed by gene expression signatures. Proc Natl Acad Sci U S A. 2005;102(44):15948-53. Epub 2005/10/27. doi: 0504300102 [pii] 10.1073/pnas.0504300102. PubMed PMID: 16249342; PubMed Central PMCID: PMC1276052.

10. Field SJ, Tsai FY, Kuo F, Zubiaga AM, Kaelin WG, Jr., Livingston DM, et al. E2F-1 functions in mice to promote apoptosis and suppress proliferation. Cell. 1996;85(4):549-61. Epub 1996/05/17. doi: S0092-8674(00)81255-6 [pii]. PubMed PMID: 8653790.

11. Murga M, Fernandez-Capetillo O, Field SJ, Moreno B, Borlado LR, Fujiwara Y, et al. Mutation of E2F2 in mice causes enhanced T lymphocyte proliferation, leading to the development of autoimmunity. Immunity. 2001;15(6):959-70. Epub 2002/01/05. doi: S1074-7613(01)00254-0 [pii]. PubMed PMID: 11754817.

12. Danielian PS, Friesenhahn LB, Faust AM, West JC, Caron AM, Bronson RT, et al. E2f3a and E2f3b make overlapping but different contributions to total E2f3 activity. Oncogene. 2008;27(51):6561-70. Epub 2008/07/30. doi: onc2008253 [pii]10.1038/onc.2008.253. PubMed PMID: 18663357; PubMed Central PMCID: PMC2723773.

13. Gaubatz S, Lindeman GJ, Ishida S, Jakoi L, Nevins JR, Livingston DM, et al. E2F4 and E2F5 play an essential role in pocket protein-mediated G1 control. Mol Cell. 2000;6(3):729-35. Epub 2000/10/13. doi: S1097-2765(00)00071-X [pii]. PubMed PMID: 11030352.

14. Lindeman GJ, Dagnino L, Gaubatz S, Xu Y, Bronson RT, Warren HB, et al. A specific, nonproliferative role for E2F-5 in choroid plexus function revealed by gene targeting. Genes Dev. 1998;12(8):1092-8. Epub 1998/05/30. PubMed PMID: 9553039; PubMed Central PMCID: PMC316727.

15. Chong JL, Wenzel PL, Saenz-Robles MT, Nair V, Ferrey A, Hagan JP, et al. E2f1-3 switch from activators in progenitor cells to repressors in differentiating cells. Nature. 2009;462(7275):930-4. doi: 10.1038/nature08677. PubMed PMID: 20016602; PubMed Central PMCID: PMCPMC2806193.

16. Chen D, Pacal M, Wenzel P, Knoepfler PS, Leone G, Bremner R. Division and apoptosis of E2f-deficient retinal progenitors. Nature. 2009;462(7275):925-9. doi: 10.1038/nature08544. PubMed PMID: 20016601; PubMed Central PMCID: PMCPMC2813224.

17. Xu X, Bieda M, Jin VX, Rabinovich A, Oberley MJ, Green R, et al. A comprehensive ChIP-chip analysis of E2F1, E2F4, and E2F6 in normal and tumor cells reveals interchangeable roles of E2F family members. Genome Res. 2007;17(11):1550-

61. Epub 2007/10/03. doi: gr.6783507 [pii] 10.1101/gr.6783507. PubMed PMID: 17908821; PubMed Central PMCID: PMC2045138.

18. Rabinovich A, Jin VX, Rabinovich R, Xu X, Farnham PJ. E2F in vivo binding specificity: comparison of consensus versus nonconsensus binding sites. Genome Res. 2008;18(11):1763-77. Epub 2008/10/07. doi: gr.080622.108 [pii] 10.1101/gr.080622.108. PubMed PMID: 18836037; PubMed Central PMCID: PMC2577861.

19. Freedman JA, Chang JT, Jakoi L, Nevins JR. A combinatorial mechanism for determining the specificity of E2F activation and repression. Oncogene. 2009;28(32):2873-81. Epub 2009/06/23. doi: onc2009153 [pii] 10.1038/onc.2009.153. PubMed PMID: 19543322; PubMed Central PMCID: PMC2726897.

20. Kong LJ, Chang JT, Bild AH, Nevins JR. Compensation and specificity of function within the E2F family. Oncogene. 2007;26(3):321-7. Epub 2006/08/16. doi: 1209817 [pii] 10.1038/sj.onc.1209817. PubMed PMID: 16909124.

21. Andrechek ER. HER2/Neu tumorigenesis and metastasis is regulated by E2F activator transcription factors. Oncogene. 2015;34(2):217-25. Epub 2013/12/24. doi: 10.1038/onc.2013.540. PubMed PMID: 24362522; PubMed Central PMCID: PMC4067469.

22. Huang E, Ishida S, Pittman J, Dressman H, Bild A, Kloos M, et al. Gene expression phenotypic models that predict the activity of oncogenic pathways. Nat Genet. 2003;34(2):226-30. Epub 2003/05/20. doi: 10.1038/ng1167ng1167 [pii]. PubMed PMID: 12754511.

23. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006;439(7074):353-7. Epub 2005/11/08. doi: nature04296 [pii] 10.1038/nature04296. PubMed PMID: 16273092.

24. Andrechek ER, Cardiff RD, Chang JT, Gatza ML, Acharya CR, Potti A, et al. Genetic heterogeneity of Myc-induced mammary tumors reflecting diverse phenotypes including metastatic potential. Proc Natl Acad Sci U S A. 2009;106(38):16387-92. Epub 2009/10/07. doi: 0901250106 [pii] 10.1073/pnas.0901250106. PubMed PMID: 19805309; PubMed Central PMCID: PMC2752567.

25. Gatza ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, et al. A pathway-based classification of human breast cancer. Proc Natl Acad Sci U S A. 2010;107(15):6994-9. Epub 2010/03/26. doi: 0912708107 [pii] 10.1073/pnas.0912708107. PubMed PMID: 20335537.

26. Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. Cell. 2008;132(6):1049-61. Epub 2008/03/25. doi: S0092-8674(08)00328-0 [pii] 10.1016/j.cell.2008.02.039. PubMed PMID: 18358816.

27. Kim J, Woo AJ, Chu J, Snow JW, Fujiwara Y, Kim CG, et al. A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. Cell. 2010;143(2):313-24. Epub 2010/10/16. doi: S0092-8674(10)01058-5 [pii] 10.1016/j.cell.2010.09.010. PubMed PMID: 20946988; PubMed Central PMCID: PMC3018841.

28. Asp P, Acosta-Alvear D, Tsikitis M, van Oevelen C, Dynlacht BD. E2f3b plays an essential role in myogenic differentiation through isoform-specific gene regulation. Genes Dev. 2009;23(1):37-53. Epub 2009/01/13. doi: 23/1/37 [pii] 10.1101/gad.1727309. PubMed PMID: 19136625; PubMed Central PMCID: PMC2632163.

29. von Eyss B, Maaskola J, Memczak S, Mollmann K, Schuetz A, Loddenkemper C, et al. The SNF2-like helicase HELLS mediates E2F3-dependent transcription and cellular transformation. EMBO J. 2012;31(4):972-85. doi: 10.1038/emboj.2011.451. PubMed PMID: 22157815; PubMed Central PMCID: PMCPMC3280551.

30. Laresgoiti U, Apraiz A, Olea M, Mitxelena J, Osinalde N, Rodriguez JA, et al. E2F2 and CREB cooperatively regulate transcriptional activity of cell cycle genes. Nucleic Acids Res. 2013;41(22):10185-98. doi: 10.1093/nar/gkt821. PubMed PMID: 24038359; PubMed Central PMCID: PMCPMC3905855.

31. Chang JT, Nevins JR. GATHER: a systems approach to interpreting genomic signatures. Bioinformatics. 2006;22(23):2926-33. Epub 2006/09/27. doi: btl483 [pii] 10.1093/bioinformatics/btl483. PubMed PMID: 17000751.

32. Roarty K, Serra R. Wnt5a is required for proper mammary gland development and TGF-beta-mediated inhibition of ductal growth. Development. 2007;134(21):3929-39. doi: 10.1242/dev.008250. PubMed PMID: 17898001.

33. Thangaraju M, Rudelius M, Bierie B, Raffeld M, Sharan S, Hennighausen L, et al. C/EBPdelta is a crucial regulator of pro-apoptotic gene expression during mammary gland involution. Development. 2005;132(21):4675-85. doi: 10.1242/dev.02050. PubMed PMID: 16192306.

34. Hatcher RJ, Dong J, Liu S, Bian G, Contreras A, Wang T, et al. Pttg1/securin is required for the branching morphogenesis of the mammary gland and suppresses mammary tumorigenesis. Proc Natl Acad Sci U S A. 2014;111(3):1008-13. doi: 10.1073/pnas.1318124111. PubMed PMID: 24395789; PubMed Central PMCID: PMCPMC3903200.

35. Wang C, Rauscher FJ, 3rd, Cress WD, Chen J. Regulation of E2F1 function by the nuclear corepressor KAP1. J Biol Chem. 2007;282(41):29902-9. doi: 10.1074/jbc.M704757200. PubMed PMID: 17704056.

36. Julian LM, Liu Y, Pakenham CA, Dugal-Tessier D, Ruzhynsky V, Bae S, et al. Tissue-specific targeting of cell fate regulatory genes by E2f factors. Cell Death Differ. 2016;23(4):565-75. doi: 10.1038/cdd.2015.36. PubMed PMID: 25909886; PubMed Central PMCID: PMCPMC4986644.

37. Choi D, Lee SJ, Hong S, Kim IH, Kang S. Prohibitin interacts with RNF2 and regulates E2F1 function via dual pathways. Oncogene. 2008;27(12):1716-25. doi: 10.1038/sj.onc.1210806. PubMed PMID: 17873902.

38. Hollern DP, Honeysett J, Cardiff RD, Andrechek ER. The E2F transcription factors regulate tumor development and metastasis in a mouse model of metastatic breast cancer. Mol Cell Biol. 2014. Epub 2014/06/18. doi: MCB.00737-14 [pii] 10.1128/MCB.00737-14. PubMed PMID: 24934442.

39. Fujiwara K, Yuwanita I, Hollern DP, Andrechek ER. Prediction and Genetic Demonstration of a Role for Activator E2Fs in Myc-Induced Tumors. Cancer Res. 2011;71(5):1924-32. Epub 2011/01/20. doi: 0008-5472.CAN-10-2386 [pii] 10.1158/0008-5472.CAN-10-2386. PubMed PMID: 21245101.

40. Yuwanita I, Barnes D, Monterey MD, O'Reilly S, Andrechek ER. Increased metastasis with loss of E2F2 in Myc-driven tumors. Oncotarget. 2015. Epub 2015/10/17. doi: 10.18632/oncotarget.5690. PubMed PMID: 26474282.

41. Barker HE, Chang J, Cox TR, Lang G, Bird D, Nicolau M, et al. LOXL2-mediated matrix remodeling in metastasis and mammary gland involution. Cancer Res. 2011;71(5):1561-72. doi: 10.1158/0008-5472.CAN-10-2868. PubMed PMID: 21233336; PubMed Central PMCID: PMCPMC3842018.

42. Yori JL, Seachrist DD, Johnson E, Lozada KL, Abdul-Karim FW, Chodosh LA, et al. Kruppel-like factor 4 inhibits tumorigenic progression and metastasis in a mouse model of breast cancer. Neoplasia. 2011;13(7):601-10. PubMed PMID: 21750654; PubMed Central PMCID: PMCPMC3132846.

43. Jechlinger M, Sommer A, Moriggl R, Seither P, Kraut N, Capodiecci P, et al. Autocrine PDGFR signaling promotes mammary cancer metastasis. J Clin Invest. 2006;116(6):1561-70. doi: 10.1172/JCI24652. PubMed PMID: 16741576; PubMed Central PMCID: PMCPMC1469776.

44. Wang A, Arantes S, Yan L, Kiguchi K, McArthur MJ, Sahin A, et al. The transcription factor ATF3 acts as an oncogene in mouse mammary tumorigenesis. BMC Cancer. 2008;8:268. doi: 10.1186/1471-2407-8-268. PubMed PMID: 18808719; PubMed Central PMCID: PMCPMC2564979.

# CHAPTER 3:

# THE ROLE OF E2F5 IN NORMAL MAMMARY GLAND DEVELOPMENT AND BREAST CANCER

**Abstract**

The development of breast cancer has been observed as a result of dysregulated mammary gland developmental processes. Thus, a better understand of the normal mammary gland development can reveal possible mechanism in how normal cells are re-programmed to become malignant cells. E2F1-4 are part of the E2F transcription factor family whose role in mammary gland development have been previously described. However, little is known about the role of E2F5 in mammary gland development. Using a mammary-specific E2F5 knockout mouse model we demonstrate that loss of E2F5 resulted in modest mammary gland development changes. Strikingly, E2F5CKO mice developed mammary tumors after a prolonged latency. In this study, we characterize E2F5CKO mammary tumor model and investigate the mechanism of tumorigenesis. Through bioinformatic analysis and in vitro studies we identified two oncogenes, KRas and Cyclin D1, that are dysregulated in E2F5CKO tumors. Based on these findings, we propose that loss of E2F5 leads to dysregulation of KRas and Cyclin D1, which facilitates the development of mammary tumors.

## Introduction

The mammary gland is a complex organ that undergoes dynamic changes during different stages of development from puberty to menopause. Analysis of transcriptional profiles at each developmental stage, including pregnancy, lactation and involution, have revealed unique gene expression changes [1–3]. One of the family of transcription factors that regulate these intricate transcriptional changes are the E2F transcription factors. The E2F transcription factor family consist of 9 members that can be divided into two groups based on their roles as a transcriptional activator or repressor. E2Fs are best known for their role in cell cycle progression [1]. However, they are a functionally diverse group of transcription factors with numerous studies highlighting their role in apoptosis, cell differentiation, metabolism and development [4]. The role of E2Fs in development have been established through the characterization of single and compound E2F knockout mice [2–8]. Specifically, the role of E2Fs in mammary gland development were first characterized in E2F1, E2F2, E2F3 and E2F4 knockout mice [5]. Loss of E2F1, E2F3 and E2F4 resulted in mammary outgrowth delay and branching defects [5]. However, these changes were not observed in in E2F2KO mice. In addition, E2F3 heterozygous mice demonstrated slight delay in involution. Given the high functional redundancy observed among E2Fs [4,6,7], our lab investigated the extent of compensation between the activator E2Fs in the mammary gland developmental processes [8]. Using double knockout mice, we revealed that E2F2 can partially compensate for the loss of E2F1 as evident by the greater delay in mammary outgrowth in the double knockout. In contrast, E2F2 was not able to compensate for decreased E2F3 function. Outcomes from this study demonstrate that E2F compensation can vary by degree and specificity. Although studies

have characterized the roles of activator E2F1-3 and repressor E2F4 in mammary development and ductal morphogenesis, little is known whether the other repressor, E2F5, has a role in mammary development. Similar to E2F4, E2F5 is considered to be a transcriptional repressor and canonically functions to repress cell cycle progression. Moreover, E2F4 and E2F5 share the most structural similarities among all the E2F members and have demonstrated functional redundancy [9]. In this study, we investigate the role of E2F5 in mammary gland development through bioinformatic analysis and the generation of mammary-specific E2F5 knockout mice.

## Results

### ***E2F5 during normal mammary gland development***

Given that terminal end buds play a critical role in mammary gland morphogenesis, we examined a possible role of E2F5 in the regulation of gene expression in TEB. First, we identified genes that are regulated by E2F5 by overexpressing E2F5 or GFP in Human Mammary Epithelial Cells (HMEC) followed by microarray analysis. Differential gene analysis between E2F5-HMEC and GFP-HMEC revealed E2F5 regulated gene expression changes. Although, E2F5 is canonically known as a transcriptional repressor, our analysis revealed upregulated and downregulated genes with E2F5 overexpression. Furthermore, when examining differentially expressed genes in TEB and duct [3], we see that 137 TEB and 7 duct genes are E2F5 regulated targets (Figure 1A). To further investigate the relationship between E.02F5 and TEB/ducts, we ran single sample Gene Set Enrichment Analysis (ssGSEA) 0on a TEB and duct gene expression dataset [3] using a geneset consisting of genes whose promoter region contain binding sites for

E2F5. The results demonstrate that E2F5 target genes are more enriched in the TEB than in the duct (Figure 1B). Taken together, this data suggest that E2F5 may play a potential role in TEB development.

The mammary gland is a dynamic tissue that undergoes remodeling and structural changes in response to varied levels of hormones and growth factors during different development stages. To determine if E2F5 plays a role in mediating these changes we performed supervised hierarchical clustering on a mammary gland developmental gene expression dataset [9] using E2F5 activation signature genes. The results demonstrate that E2F5 signature genes can stratify the different stages of mammary gland development (Figure 2A). This suggest a role of E2F5 in regulating each of the major stages of mammary function. To further explore the role of E2F5 in mammary gland development, we examined its expression in various developmental stages using a previously generated single cell RNA-seq data set [10]. This dataset was generated from different stages of mammary gland development including virgin, gestation, lactation and post-involution. Interestingly, this analysis revealed that E2F5 expression is detected at all stages of mammary gland development. However, its expression appears to be highest in the lactation and post-involution cell populations (Figure 2B).This is in contrast to E2F1 whose expression is mostly limited to gestation and post-involution populations (Figure 2B). To determine if E2F5 activity is enriched in any of the development stages, we performed ssGSEA on a development gene expression dataset using a E2F5 activation signature. The results suggest that E2F5 activity is higher in the lactation and involution stages (Figure 2C). Given that E2F5 has higher expression and activity during

lactation and involution, we hypothesized that E2F5 may have a functional role in these developmental stages.

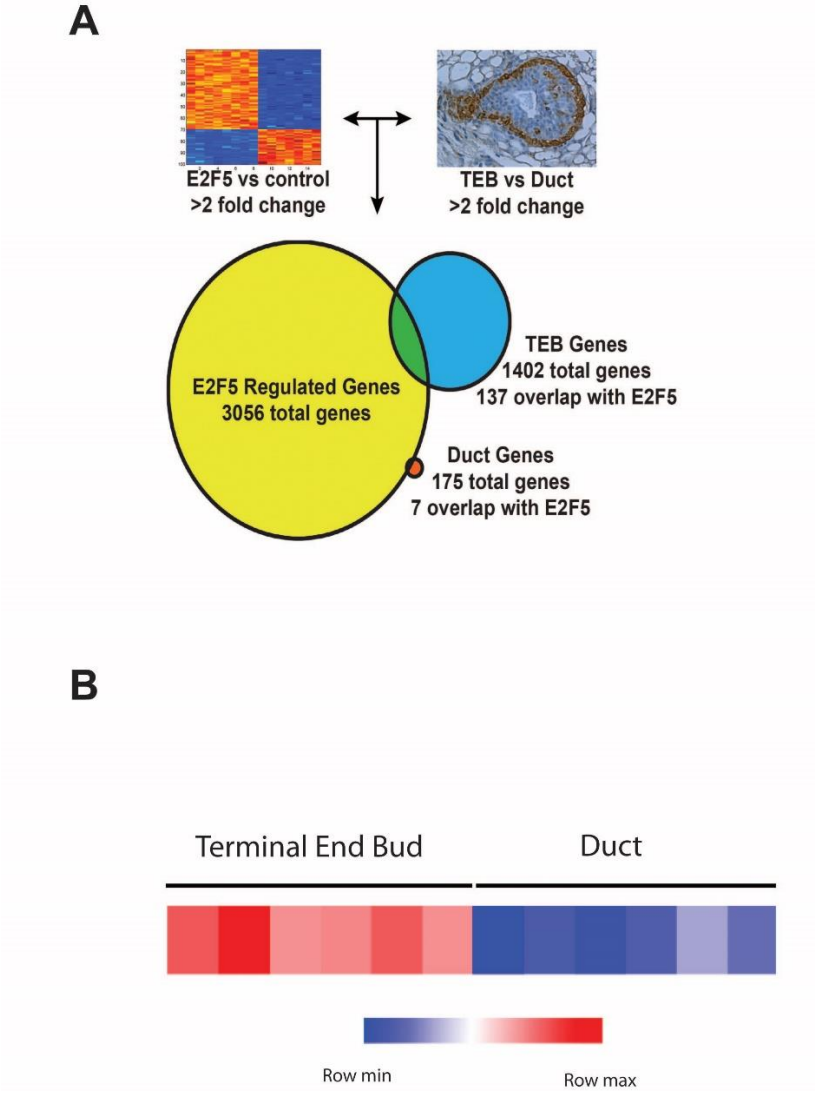


**Figure 3.1 E2F5 regulated genes in the TEB A.** Differentially expressed genes with 2-fold change in expression were identified in E2F5-HMEC vs GFP-HMEC. Likewise, TEB and duct differentially regulated genes were identified using the same methodology and cut-off. E2F5 regulated genes were overlapped with TEB-specific genes and Duct-specific genes. **B.** Single Sample Gene Set Enrichment Analysis was performed on a gene expression dataset from microdissected terminal end buds and mature ducts using a E2F5 target gene set.

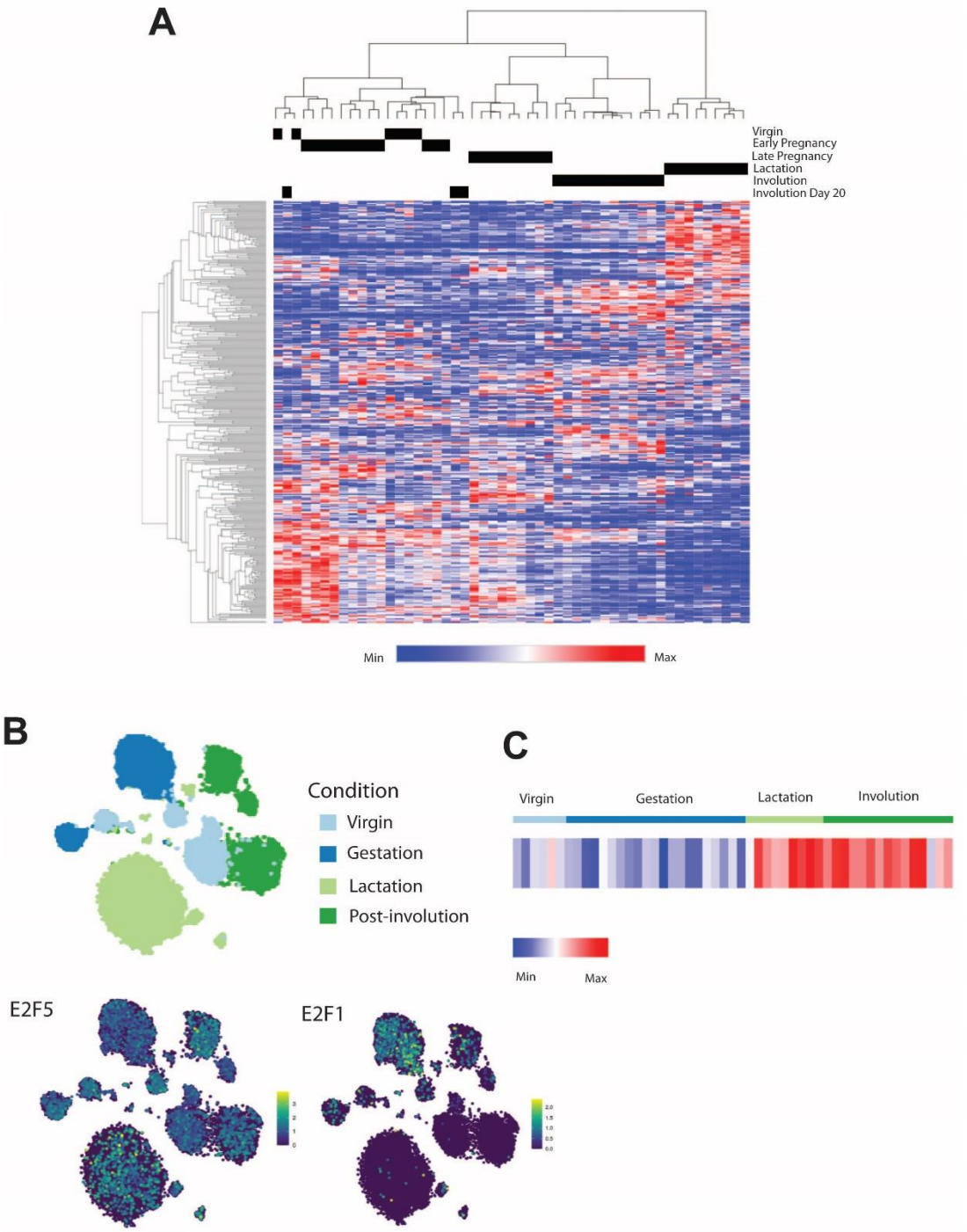


**Figure 3.2 E2F5 activity in mammary gland developmental stages A.** Unsupervised clustering of mammary developmental dataset based on E2F5 activation signature genes stratified the samples by developmental stage. **B.** E2F5 and E2F1 expression in single cell RNA-seq data generated from mammary gland derived from various developmental

**Figure 3.2 (cont'd)** stages. **C.** Single Sample Gene Set Enrichment Analysis was performed on a gene expression dataset from various stages of mammary gland development using a E2F5 activation signature.

***Characterizing mammary gland development in E2F5CKO mice***

Our bioinformatic predictions suggest that E2F5 may regulates stages of mammary gland development. To further characterize E2F5's role in the mammary gland, we generated a mouse model with mammary-specific deletion of E2F5 using the Cre-LoxP recombinase system. LoxP sites were inserted between exons 2 and 3 of E2F5 (Figure 3A). After 12 backcrosses to the FVB background, these floxed mice were interbred with MMTV-Cre mice to generate a mammary-specific E2F5 conditional knockout (E2F5CKO) model. Excision of exon 2 and 3 of E2F5 was confirmed with PCR (Figure 3B). Wholemount staining technique was used to examine the gross morphology and the mammary outgrowth. Compared to wildtype controls (Figure 3C), mammary glands from E2F5CKO mice demonstrated slight but significant delay in mammary outgrowth during week 4 of pubertal development (Figure 3D-E). However, this delay was not observed at 8 weeks of development, suggesting that the delay was transient. No branching defects were observed. Given that bioinformatic analysis predicts a more prominent role for E2F5 during lactation and involution, we examined these stages for abnormalities. Based on E2F5CKO dams' ability to rear pups as well as normal histological lactation phenotype, we concluded that loss of E2F5 did not affect lactation. Our analysis of the mammary gland on the fourth day of involution in E2F5KO mice revealed a slight delay in remodeling but no other major defects (Figure F). Following our analysis of early mammary gland development, we assessed aged virgin mammary gland to examine long term effects of E2F5 loss. Interestingly, in comparison to MMTV-Cre controls, aged (>9 months) E2F5CKO mammary glands have numerous

enlarged alveoli, a reduction in adipocytes and enlarged ducts, phenotypes that are reminiscent of lactating glands (Figure G).

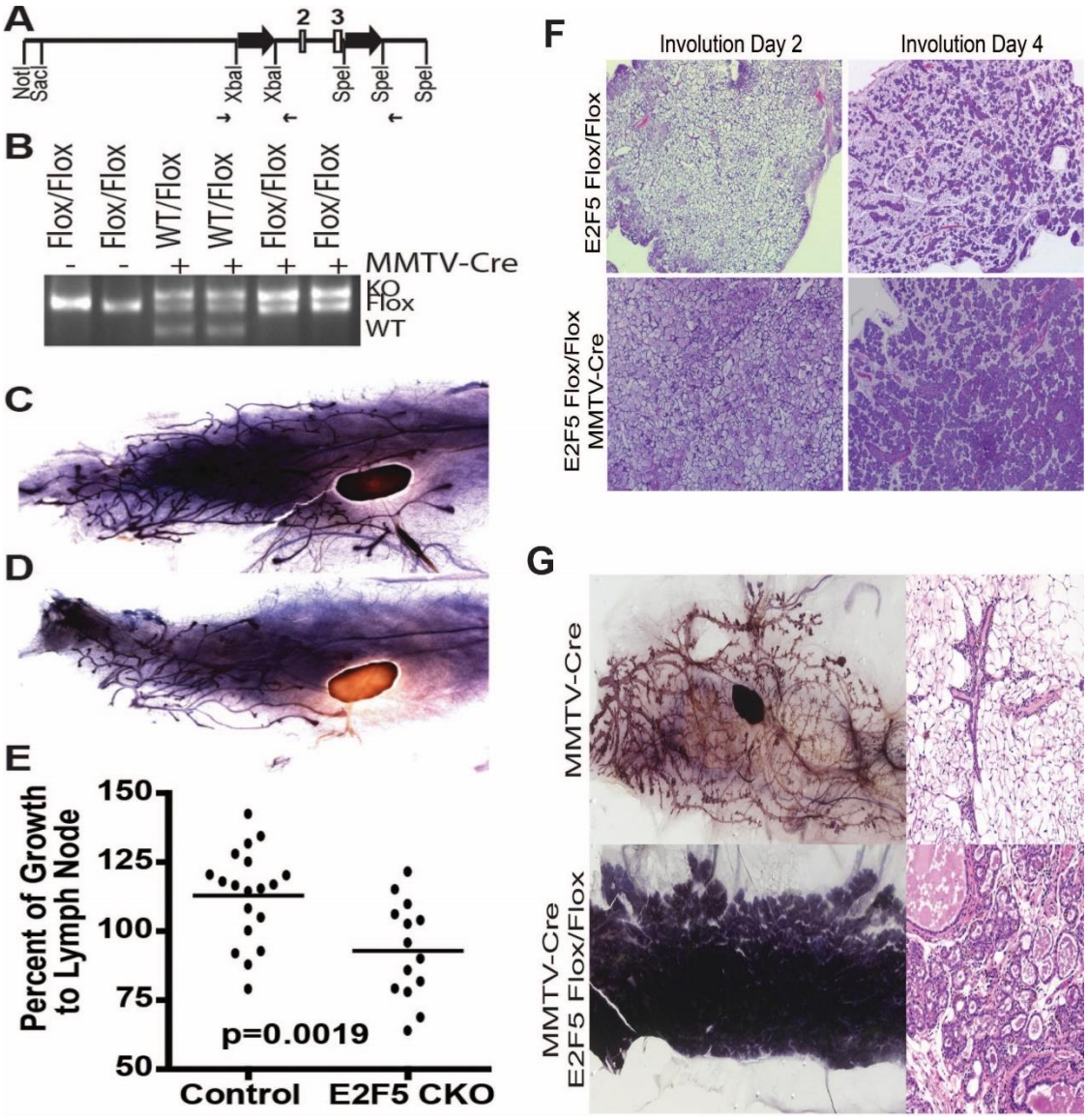


**Figure 3.3 Generation of E2F5 Conditional Knockout Mice A.** LoxP sites were inserted between exon 2 and 3 of E2F5. After backcrossing to FVB, E2F5 Flox/Flox mice were interbred with MMTV-Cre mice. **B.** Excision of exons 2 and 3 were confirmed with PCR. **C.** In comparison to control, mammary glands from conditional knockout mice demonstrated a slight but significant delay in mammary outgrowth at 4 weeks. **D.** Quantification of mammary gland outgrowth delay. **E.** Assessment of involution at day 2

**Figure 3.3 (cont'd)** and day 4 revealed a slight delay in E2F5CKO relative to controls. **F.** Analysis of wholemounts and histology from aged (> 9 months) virgin E2F5KO mammary glands revealed that mammary glands lacking E2F5 in the mammary epithelium resemble lactating glands with numerous alveoli, a reduction of adipocytes and enlarged ducts.

***E2F5CKO mice develop spontaneous mammary tumors***

In a previous study, we applied a copy number prediction method to 26 major mouse models to identify distinct and conserved changes [11]. When comparing basal-like mouse models with luminal- like mouse models, we noted that E2F5 was deleted across the basal-like models, with up to 17% in the BRCA p53 model. Interestingly, this deletion was not observed in the luminal-like models (Figure 4A). Based on this observation as well as E2F5 role as a cell cycle repressor, we monitored E2F5CKO mice for possible tumor development by palpation. Strikingly, mammary tumors were observed in E2F5CKO mice after a prolonged latency. The multiparous group developed tumors at a median latency of 19 months while the virgin group developed tumors at a median latency of 21 months (Figure 4B). Importantly, mammary tumors were not observed in the MMTV-Cre control mice. Tumors derived from E2F5CKO mice demonstrated diverse morphology including papillary, adenocarcinoma, EMT, mixed and adenocarcinoma (Figure 4D). In addition, distant metastasis, primarily to the lungs, were observed in 60% of tumor bearing mice.

Due to the prolonged latency, we generated a syngeneic transplantation model. E2F5CKO tumors were implanted into the abdominal mammary fat pad of MMTV-Cre mice. Interestingly, we observed a mass in the axillary region ipsilateral to the transplantation site in a small percentage of animals (Figure 5A). Histological examination of the axial lymph node revealed both lymph tissue and metastatic tumor (Figure 5B). To confirm that the axial lymph node contains metastatic breast tumor cells, we labeled for pan-cytokeratin markers by immunohistochemistry (Figure 5C). Importantly, control MMTV-Cre axial lymph nodes are negative for tumor cell labeling (Figure 5D). As a

control, we transplanted other tumor models (MMTV-Neu, MMTV-Myc and MMTV-PyMT) into the abdominal mammary fat pad. Lymph node metastasis was not observed in any of the other tumor models tested, demonstrating that this increase in metastatic potential is driven by loss of E2F5. Since the lymph node metastasis had low penetrance in the transplanted mice, we enriched for lymph node metastasis in order to further study this phenomenon. To enrich for mammary tumors that metastasize to the lymph node, we used a serial transplantation technique (Figure 5E), a strategy which has been previously used in human breast cancer cell lines to enrich for lung [12] and brain metastasis [13]. Transplants of the original E2F5CKO mammary tumors into the mammary fat pad of MMTV-Cre mice resulted in less than 5% of mice developing lymph node metastasis. An axillary lymph node tumor derived from the 1$^{st}$ generation of transplants was transplanted into the mammary fat pad of a second group of mice. This resulted in 30-40% of the transplanted mice developing lymph node metastasis. This strategy was repeated with the 2$^{nd}$ generation axial tumors being transplanted into a third group of animals, leading to >80% enrichment for lymph node metastasis. Each round of transplantation also resulted in shorter latency to tumor-end stage.

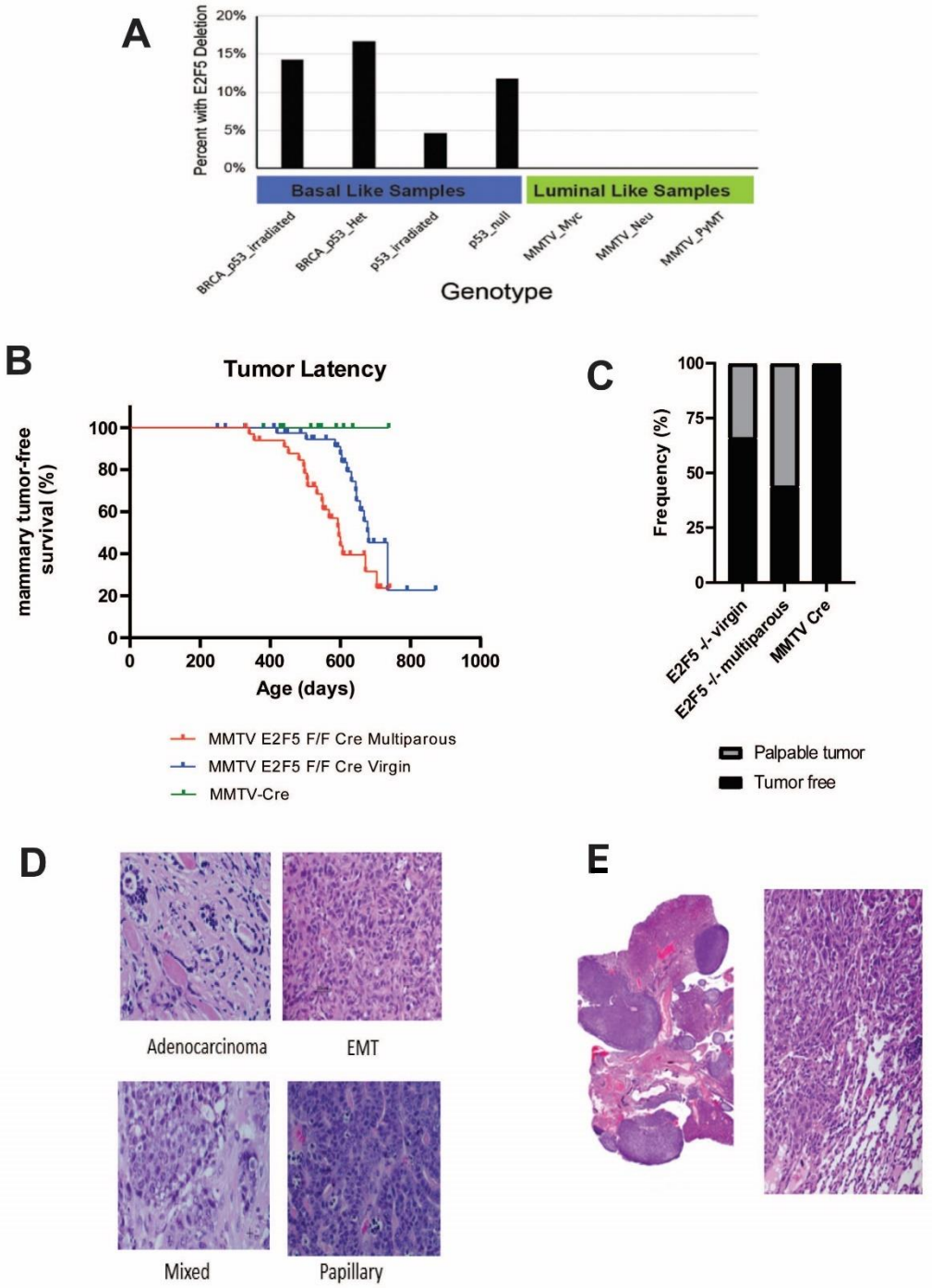


**Figure 3.4 Loss of E2F5 in the mammary epithelium results in tumor formation A.** Screening basal and luminal sub-populations of mouse model tumors revealed loss of E2F5 in the basal-like tumors. **B.** Virgin (blue) and multiparous (red) mice lacking E2F5

**Figure 3.4 (cont'd)** in the mammary gland developed tumors after a long latency. MMTV-Cre control shown in green. **C.** 56% of mice developed mammary tumors in the multiparous group and 34% of mice developed mammary tumors in the virgin group. **D.** A variety of histological features were identified in the E2F5cKO mammary tumors. **E.** Metastasis was observed in 83% of tumor bearing mice. Pulmonary metastasis is shown at 5x and 20x

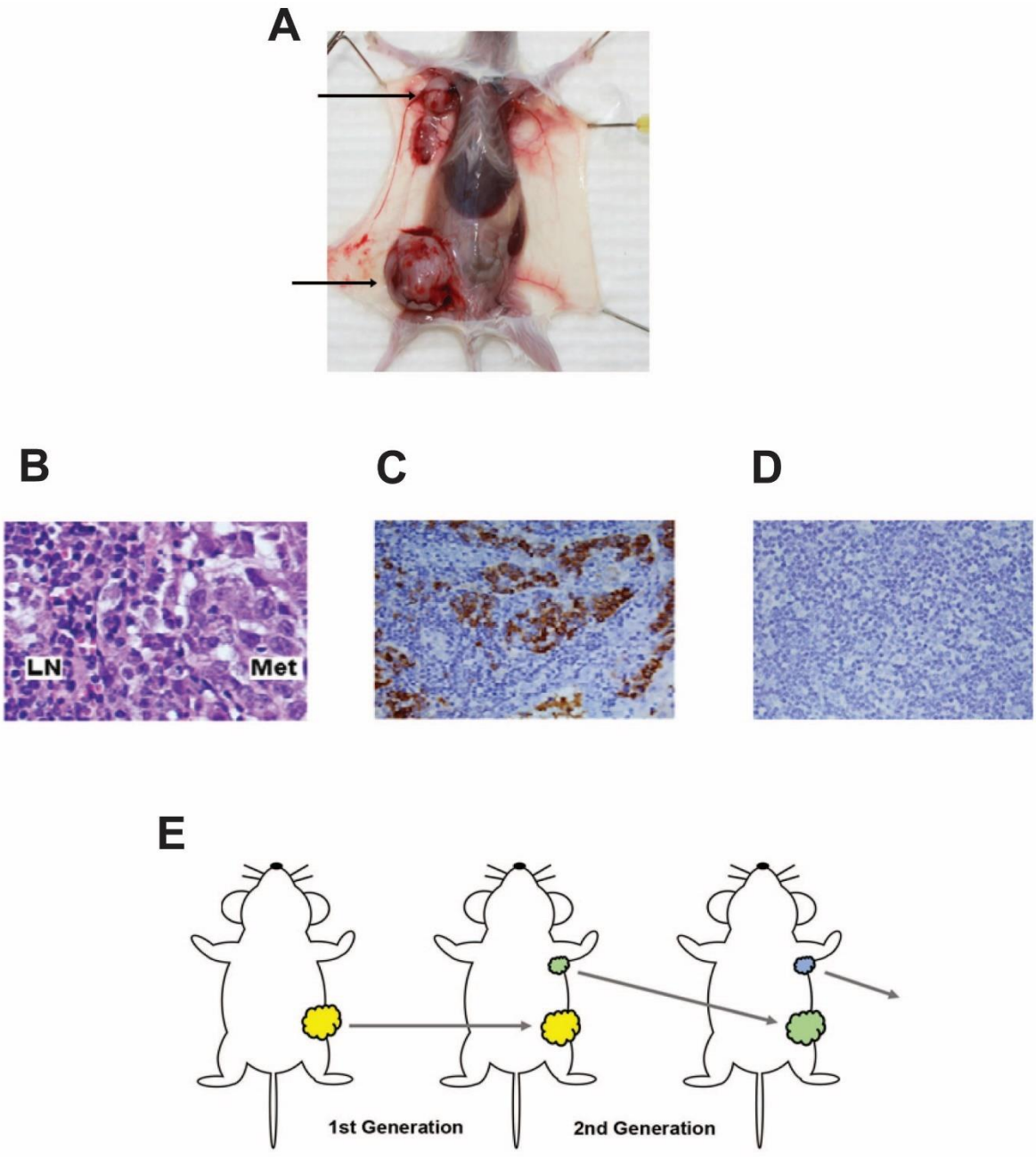


**Figure 3.5 Lymph node metastasis** **A**. Transplantation of an E2F5CKO tumors resulted in a mass in the axillary region ipsilateral to the surgical site**. B.** Histology of the lymph node showed both lymph (LN) and metastatic (Met) tumor tissue . **C.** Pan-cytokeratin staining of lymph node from E2F5CKO transplant mouse. **D.** Pan-cytokeratin

**Figure 3.5 (cont'd)** staining of lymph node from MMTV-Cre control mouse. **E.** Schematic illustrating serial transplantation technique to enrich for lymph node metastasis in E2F5CKO syngeneic transplantation model.

***Increased KRas activity in E2F5CKO mammary tumors***

To begin to elucidate the mechanism of tumorigenesis in E2F5CKO tumors, we examined E2F5 regulated genes. Differentially expressed genes in E2F5-overexpressed HMECs were analyzed. The list of differentially expressed genes was filtered several selection criteria including fold change, known roles in cancer based on current literature, and percent alteration in human cancers. One of the leading candidates from this analysis was the gene KRas. Overexpression of E2F5 in HMECs resulted in a significant decrease in KRas expression (Figure 6A). Given that deregulation of the Ras signaling pathway is associated with increased cell survival and proliferation, we further assessed this potential relationship between E2F5 and KRas. First, we tested Ras activity by applying a Ras signature to gene expression data from E2F5 and GFP overexpressed HMECs [14,15] using Bayesian regression modeling. The results illustrate that predicted Ras activity is decreased in samples with E2F5 overexpression relative to samples with GFP overexpression (Figure 6B). In addition, we determine that Ras, Akt and Erk pathways were less enriched in the E2F5 overexpressed HMECs (Figure 6C) using ssGSEA [16,17]. This indicates an inverse relationship between E2F5 activity and Ras activity. Given this inverse relationship between E2F5 and KRas, we postulate that E2F5 deletion in the mammary gland of mice leads to increased KRas signaling. To test this hypothesis, we examined KRas activity by measuring activation of its downstream pathway. As a control, we compared KRas activity in E2F5CKO tumors to Myc tumors with and without a known KRas activating mutation. First, we looked at the activation of the PI3K-Akt pathway by immunoblotting, which revealed high phospho-Akt levels in E2F5cKO tumors (Figure 6D). In addition, we examined the MAPKKK pathway, another downstream

pathway of KRas activation. Immunoblotting using a phospho-Erk antibody revealed elevated phospho-Erk levels in a majority of E2F5cKO tumors (Figure 6D). To further confirm if there is increased KRas activity in E2F5CKO tumors, we performed a Ras-GTP activation assay on tumor cell lines derived from E2F5CKO tumors. Increased levels of Ras activation were detected in the E2F5CKO tumor cell lines relative to MYC tumor cell line controls (Figure 6E). This data further suggest that E2F5CKO tumors have increased Ras activity. Furthermore, a binding site predictor, GATHER, revealed that KRas has a predicted E2F binding site [18]. This further suggests that KRas expression may be directly linked to E2F5 activity. Collectively, our data suggests that E2F5 negatively regulates KRas expression and E2F5 activity is negatively correlated with Ras activity.

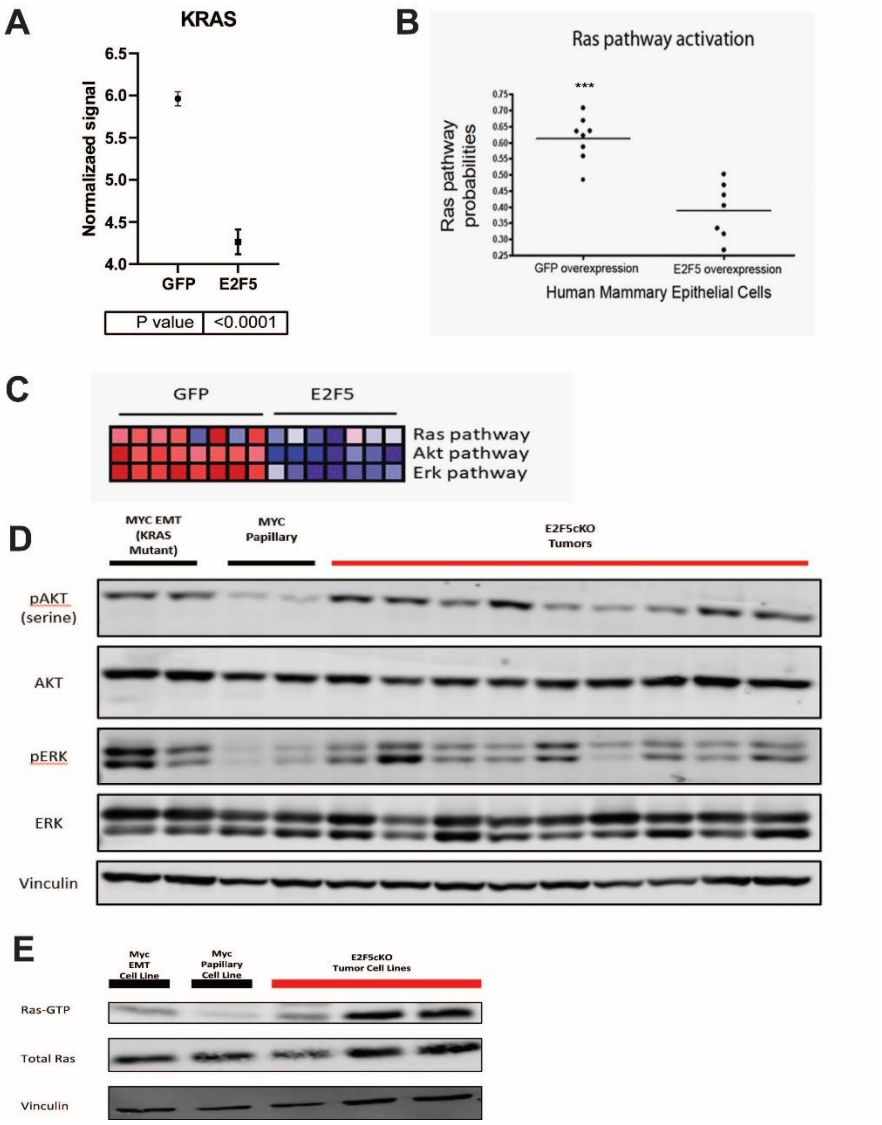


**Figure 3.6 Increased Ras activation in E2F5KO tumors A.** Expression of KRas in HMECs overexpressing E2F5 relative to GFP. **B.** E2F5 overexpressed HMECs have lower predicted Ras activity relative to GFP overexpressed HMECs. **C.** Ras, Akt, and ERK pathway enrichment in HMEC with E2F5 overexpression relative to GFP overexpression (blue indicates low enrichment, red indicates high enrichment). **D.** Phospho-Akt and Phospho-Erk levels in E2F5cKO tumors. MMTV-Myc mammary tumors with confirmed mutant KRas and WT KRas were used as controls. **E.** Ras activation in MMTV-Myc tumor (with and without KRas activating mutation) and E2F5KO tumors.

***Cyclin D1 was upregulated in E2F5CKO mammary gland and tumor***

To better understand transcriptional profiles of E2F5CKO tumors, RNA-sequencing was performed on MMTV-Cre mammary glands, E2F5CKO mammary glands, E2F5CKO mammary tumors and tumor cell lines. Differential gene expression analysis was performed between MMTV-Cre and E2F5CKO mammary glands and MMTV-Cre mammary glands and E2F5CKO tumors/cell lines. To identify gene expression changes driven by E2F5 loss, we composed a list of genes that are differentially regulated in both E2F5CKO mammary glands and tumors relative to MMTV-Cre mammary glands (Figure 7A). Next, we filtered this list down using several criteria including fold change, percent alteration in human breast cancer and known E2F targets. Based on these factors, 4 candidates (Rad51, Sphk1, Kif20a and Cyclin D1) were chosen for validation with qRT-PCR. In line with the differential gene expression analysis, all four genes demonstrated increased expression in E2F5CKO tumors relative to MMTV-Cre mammary glands (Figure 7B). However, the most striking difference was seen in Cyclin D1 where there was a 15-fold increase in E2F5CKO tumors relative to control. Furthermore, qRT-PCR also confirmed that Cyclin D1 levels were elevated in E2F5CKO mammary glands relative to MMTV-Cre mammary glands (Figure 7D). Consistent with these findings, gene set enrichment analysis revealed an enrichment for genes involved in cell cycle progression and G1/S check point transition (Figure 8A).

Previous studies have demonstrated that Cyclin D1 overexpression in the mouse mammary gland can initiate tumorigenesis after a prolonged latency. Moreover, data suggests that Cyclin d1 can mediate tumor development in the MMTV-Ras and MMTV-Neu mouse models. Loss of cyclin D1 in the MMTV-Ras and MMTV-Neu mouse models

inhibited tumor development while having minimal effects in the MMTV- Wnt and MMTV-Myc models. Interestingly, when comparing Cyclin D family members in these mouse models, Cyclin D1 levels were relatively similar while Cyclin D2 was elevated only in MMTV-Wnt and MMTV-Myc. Based on these findings, the authors suggest that, unlike the MMTV-Wnt and MMTV-Myc models, the MMTV-Ras and MMTV-Neu exclusively depend on Cyclin D1, thus loss of Cyclin D1 function have significant impact on tumorigenesis. Given that we observed increased expression of Cyclin D1 in E2F5CKO mammary gland and a greater upregulation in E2F5CKO tumors and tumor-derived cell lines, we hypothesize that Cyclin D1 dysregulation is contributing to tumor progression in E2F5CKO mice. To begin investigating the role of Cyclin D1 in E2F5CKO tumors, we examined the expression pattern of the three D-type cyclins. In comparison to MMTV-Wnt and MMTV-Neu mammary tumors, E2F5CKO tumors have relatively similar levels of Cyclin D1 expression to both tumor types. However, like MMTV-Neu tumors, majority of E2F5CKO tumors have decrease levels of Cyclin D2 relative to MMTV-Wnt tumors (Figure 7B).

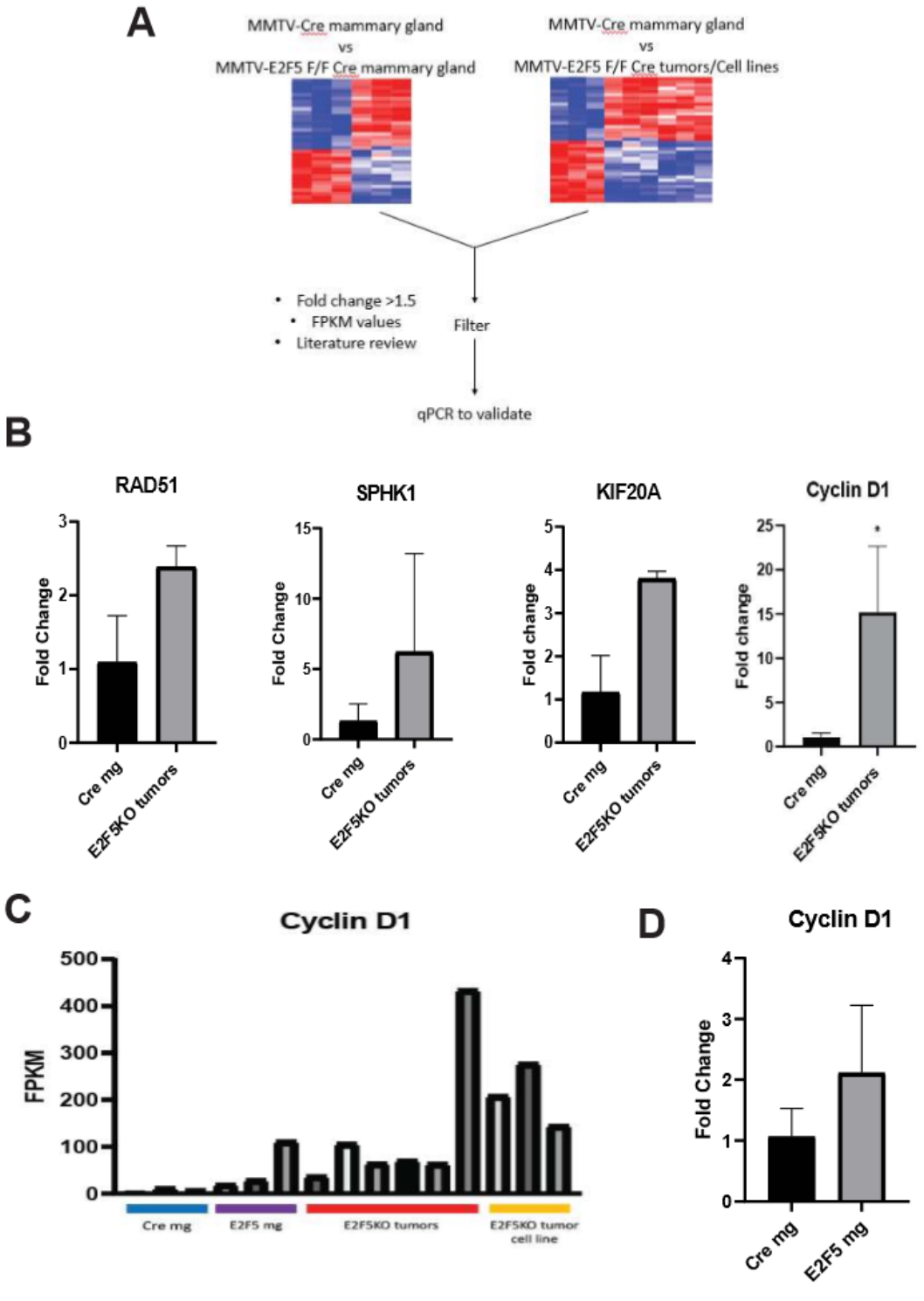


**Figure 3.7 Cyclin D1 expression in E2F5CKO tumors A.** Schematic demonstrates how genes were selected for qRT-PCR validation. Differentially expressed genes in E2F5CKO

**Figure 3.7 (cont'd)** mammary gland and tumors relative to MMTV-Cre mammary gland were identified. Target list were filtered based on selective criteria. **B.** Expression of four candidate targets were examined in MMTV-Cre mammary gland versus E2F5CKO tumors using qRT-PCR. **C.** FPKM (normalized RNA-seq counts) values of Cyclin D1 across samples. **D.** Expression of Cyclin D1 in MMTV Cre mammary gland versus E2F5CKO mammary gland were quantified using qRT-PCR.

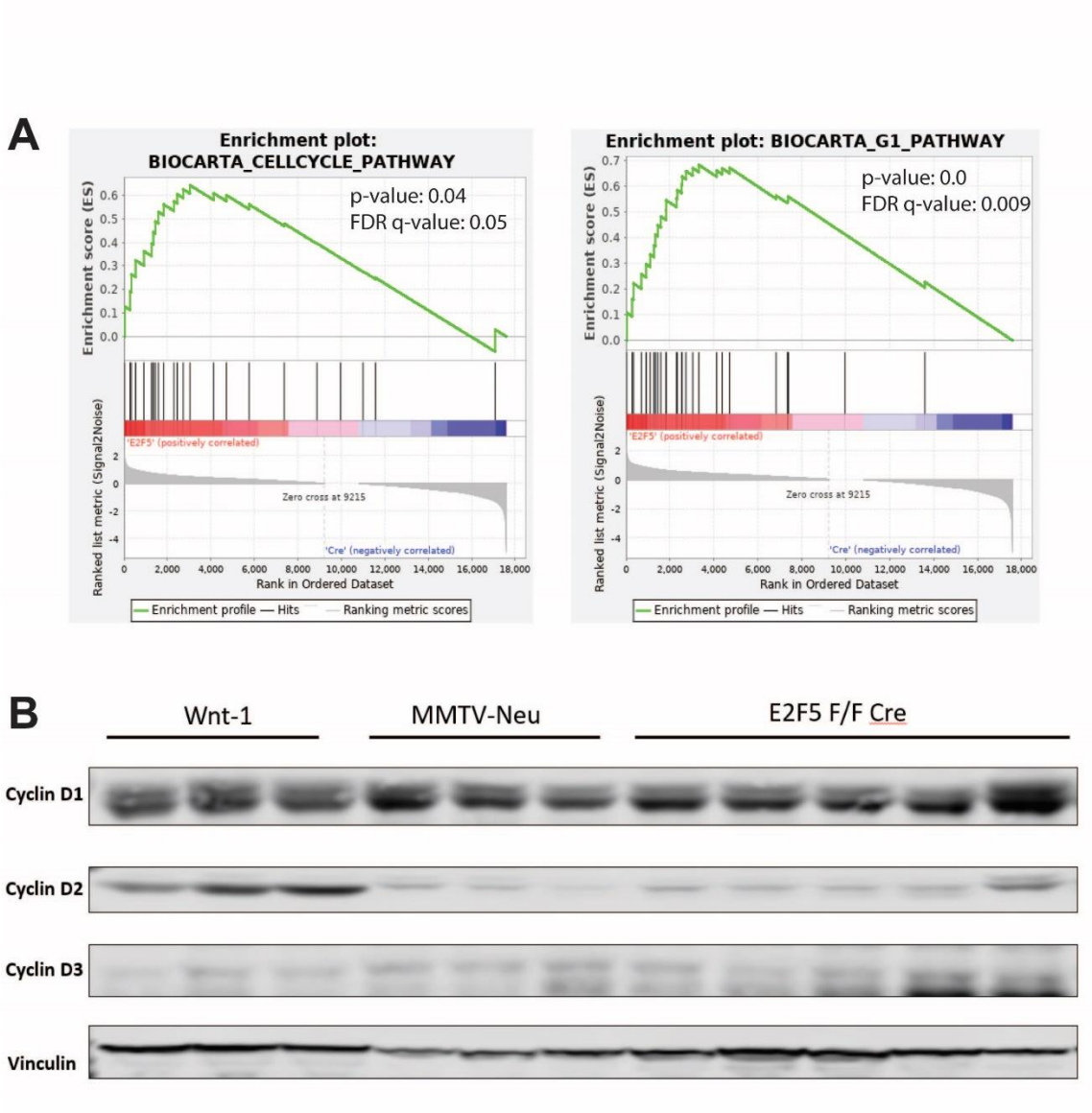


**Figure 3.8 D-type cyclin expression in E2F5CKO tumors A.** E2F5CKO tumors are enriched for genes involved in cell cycle progression and G1/S check point transition. **B.** Western blot analysis of D-type Cyclin in MMTV-Wnt, MMTV-Neu and E2F5CKO tumors.

***Role of E2F5 in human breast cancer***

Since E2F5 was shown to be deleted in basal-like mouse models, we wanted to determine which intrinsic subtype E2F5CKO tumors most closely resembles. Using the PAM50 genes [19], gene expression data from E2F5CKO tumors and TCGA breast cancer, genes were clustered together. Interestingly. E2F5CKO tumors generally clustered with luminal A and luminal B samples (Figure 8A). However, E2F5CKO tumors did not express ER or PR based on immunohistopathology (data not shown).

To further investigate the role of E2F5 in human breast cancer we examined its expression and copy number in human breast cancer datasets. To our surprise, E2F5 is amplified and/or overexpressed in 27% of patients in the Metabric and TCGA human breast cancer dataset [20,21]. Since E2F5 is located in chromosome 8q21, a commonly amplified region in breast cancer, we hypothesize that E2F5 is being co-amplified [22]. Indeed, when we examine genes neighboring E2F5, we see that these genes are also being co-amplified, suggesting that E2F5 is located within an amplicon. Therefore, it begs the question whether E2F5 is just located in the wrong place or whether its amplification indicate their role as an oncogene. To begin to answer this question, we examined if patients with high E2F5 expression have better or worst prognosis. We found that in basal and Her2+ patients, high E2F5 expression was associated with better overall and relapse-free survival (Figure 8B) [23]. This suggests that E2F5 may have a protective role in the basal and Her2+ subtypes. To better understand the role of E2F5 in human breast cancer, we deleted E2F5 in two triple negative breast cancer cell lines, BT549 and MDA-MB-231, using CRISPR-Cas9 (Figure 9A). Deletion of E2F5 was confirmed with sequencing and immunoblotting. There was no significant change in proliferate rate in E2F5KO cells

relative to wildtype clones. In addition, E2F5CKO cells did not demonstrate any change in their ability to migrate in a wound-healing assay.

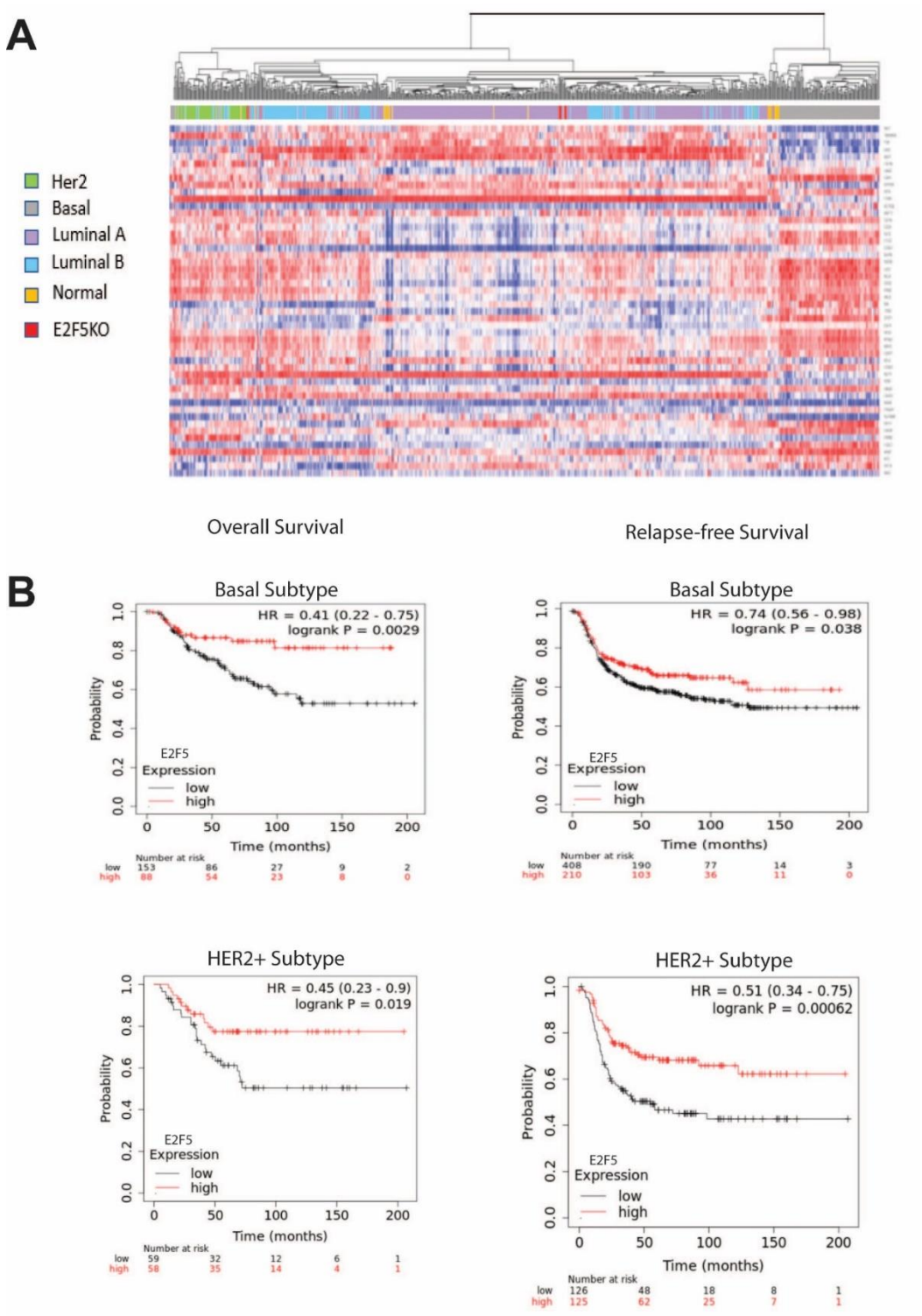


**Figure 3.9 E2F5 expression correlate with better overall survival A.** Clustering of TCGA human breast cancer samples with E2F5CKO tumors based on PAM50. **B.** Overall

**Figure 3.9 (cont'd)** survival and relapse-free survival in between with high E2F5 (red) versus low E2F5 (black) expression in basal and Her2+ subtypes.

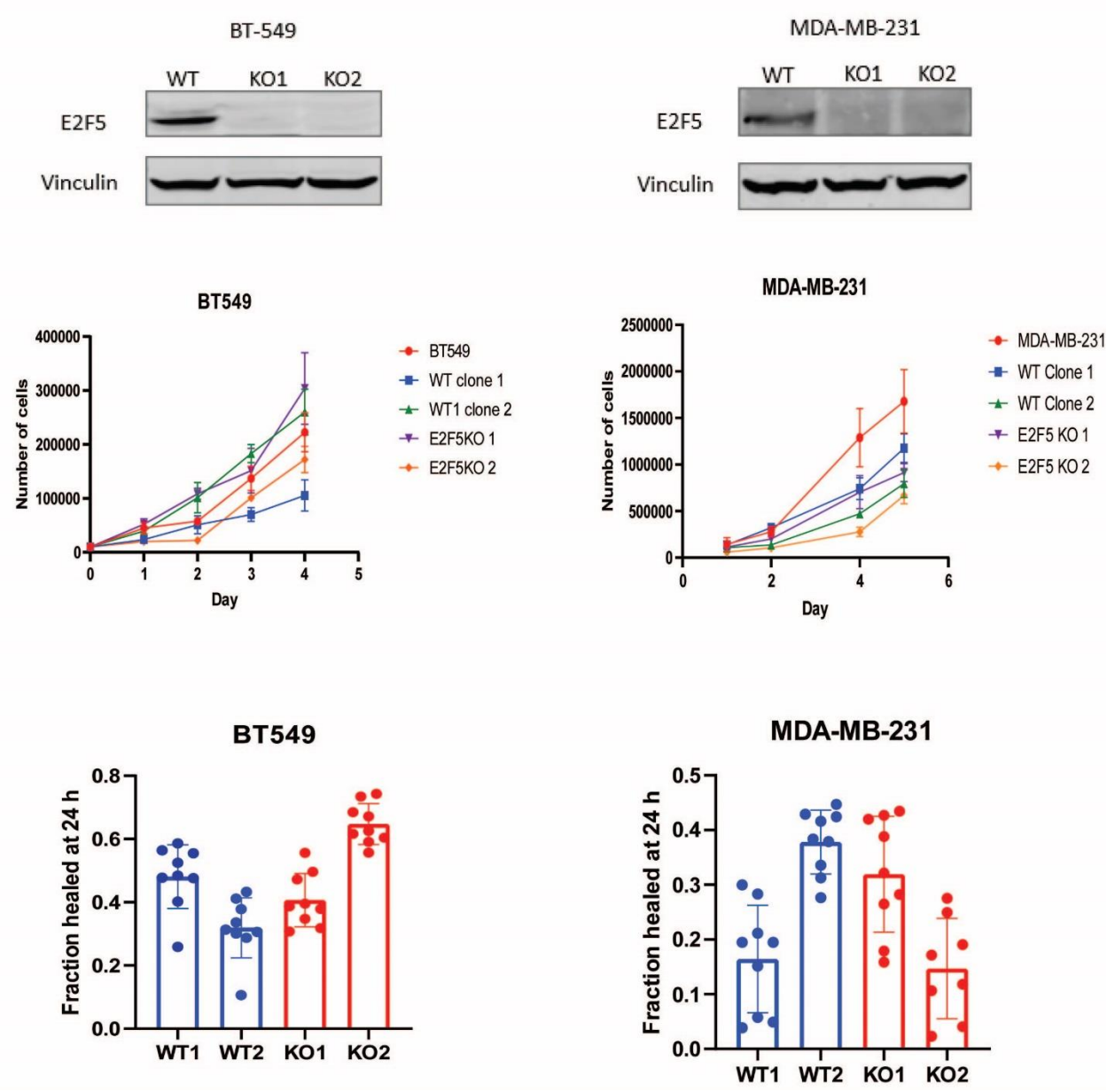


**Figure 3.10 Characterization of E2F5KO cell lines A.** Western blot showing loss of E2F5 in BT549 and MDA-MB-231 cells. **B.** Growth curve analysis with BT549 and MDA-MB-231 E2F5KO cells. **C.** Quantification of wound healing 24 hours after scratch in BT549 and MDA-MB-231KO cells.

## Discussion

Contrary to our bioinformatic predictions, loss of E2F5 resulted in only modest mammary gland developmental defects. However, it is likely that E2F4 can compensate for E2F5 loss. A previous study has demonstrated that whole body double knockout of E2F4 and E2F5 mice results in neonatal lethality while deletion of E2F4 or E2F5 alone produces a viable embryo [10]. This suggest that E2F4 and E2F5 may mediate overlapping functions during early mouse development. Furthermore, Kong *et al*. have demonstrated the ability of E2F members to functionally compensate for one another [7]. Importantly, a study from our lab has shown that compensation can occur among E2Fs during mammary gland development.

Although there were no remarkable developmental changes, we found that E2F5CKO mice spontaneously developed mammary tumors. Similar to human breast cancer, mammary tumors arising in E2F5CKO mice demonstrate diverse morphology. Furthermore, the prolonged tumor latency in the E2F5CKO model is similar to human breast cancers, where the majority occur in older postmenopausal women [11]. Given that majority of the transgenic mouse models develop tumors with at a shorter latency [11], E2F5CKO mice may be a useful for model for studying age-related changes in breast cancer [12]. The prolonged tumor latency also suggests that, in addition to the loss of E2F5, other genetic events need to accumulate prior to tumor initiation.

In addition, E2F5CKO mice develop metastasis, most commonly to the lungs. Interestingly, we discovered that E2F5CKO mammary tumors, when transplanted into the abdominal mammary fat pad, have a propensity to metastasize to the axillary lymph node. Given that axillary lymph nodes are most commonly the first site of metastasis in human

breast cancer, we wanted to enrich for E2F5CKO tumors that can metastasize to the axillary lymph node. Using a serial transplantation technique to re-transplant the axillary tumor into the abdominal mammary fat pad, we generated a syngeneic transplantation model that develop lymph node metastasis within one month of transplant and with >80% penetrance. Current mouse models of breast cancer rarely metastasize to the lymph node. Therefore, this model of enriched lymph node tumors is unique and can provide further insight behind the mechanism driving lymph node metastasis.

Integrating our bioinformatic analysis and in vitro studies, we identified two targets, KRas and Cyclin D1, that may play a role in tumor development and progression in E2F5CKO tumors. Ras mutation is only detected in 3.2% of breast cancer. However, the Ras/ERK pathway is hyperactivated in >50% of breast cancer, suggesting that other mechanisms may facilitate Ras activation [24]. Cyclin D1, on the other hand, is commonly amplified and/or overexpressed in human breast cancer [13]. Moreover, overexpression of Ras or Cyclin D1 in the mouse mammary gland results in mammary tumor develop, supporting their role in tumorigenesis [25,26]. Based on gene expression analysis, E2F5 expression is inversely correlated with both KRas and Cyclin D1, suggesting that E2F5 may be negatively regulating these two genes. In line with this finding, we demonstrate that E2F5CKO tumors demonstrates increased Ras activation. To investigate the role of Cyclin D1 in E2F5CKO tumors, we examined Cyclin D1, D2 and D3 levels in E2F5CKO tumors compared to MMTV-Wnt and MMTV-Neu tumors. Our analysis revealed similar levels of Cyclin D1 expression across all three models. However, Cyclin D2 and D3 levels were lower in the MMTV-Neu and MMTV-E2F5KO models compared to the MMTV-Wnt model. A study by Yu *et al.* have shown that Cyclin D1 loss does not affect MMTV-Wnt

tumor development because it also expresses Cyclin D2, thus it is able to compensate for Cyclin D1 loss [27]. In contrast, since MMTV-Neu mainly express Cyclin D1, the loss or inhibition of Cyclin D1 inhibited tumor development as demonstrated in several studies [27–29]. However, Zhang *et al.* demonstrate that loss of Cyclin D1 in MMTV-Neu tumors only delayed tumor latency [15]. Upon further investigation, they discovered that Cyclin D3 is able to compensate for Cyclin D1 loss, allowing for tumor initiation to occur. Thus, there is contradicting results in whether tumor development still occurs in Cyclin D1 deficient MMTV-Neu mice as a result of Cyclin D3 compensation. However, these discrepancies do not undermine the fact that Cyclin D1 is critical for tumor development and under normal circumstances is the main member of D-type cyclins to initiate tumorigenesis in MMTV-Neu. Furthermore, Cyclin D1 expression in MMTV-Neu is mediated by E2F1 [30]. Other studies have also demonstrated that E2F1 and E2F4 can directly bind to and regulate Cyclin D1 expression [31]. Given the functional redundancy and shared binding motif between E2F family members, it is likely that E2F5 can also regulate Cyclin D1 expression [32,33]. Taken together, we propose that loss of E2F5 in the mammary gland leads to deregulation of KRas and Cyclin D1, contributing to tumor development and progression. Although there is evidence suggesting that E2F5 may be directly regulating KRas and Cyclin D1 expression, it is also possible that disruption of E2F5 leads to dysregulation of other unknown targets that can result in KRas activation and Cyclin D1 expression. Further studies using chip-based technologies to identify direct E2F5 target genes may help further elucidate the mechanism of tumor development in E2F5CKO mouse model.

Based on current dogma, the development of mammary tumors from the loss of E2F5 suggest it is behaving as a tumor suppressor. However, in human breast cancer, E2F5 appears to be amplified and/or overexpressed, suggesting it may be behaving as an oncogene. Interestingly, in humans, E2F5 is located on chromosome 8q21, a commonly amplified region in breast cancer [21]. Furthermore, examining copy number changes in that region reveals that E2F5 becomes co-amplified with other genes in its vicinity, suggesting that E2F5 is part of the amplicon. Importantly, high E2F5 expression in basal and HER2+ subtype is associated with better survival outcomes, suggesting it may have a protective role. To better understand the role of E2F5 in human breast cancer, we characterized the effects of E2F5 loss in human breast cancer cells. Our analysis revealed that loss of E2F5 had no effect on cell growth or migration. Although these results are surprising, it suggests that E2F5 may not be directly involved in tumor progression given that BT549 and MDA-MB-231 are already transformed. Furthermore, migration is only one of the hallmarks of metastasis [34]. Thus, it is possible that loss of E2F5 promotes other stages of the metastatic cascade. Future experiments are needed to better characterize the role of E2F5 in tumor initiation and metastasis.

In this study, we have identified the novel role of E2F5 as a tumor suppressor. E2F5CKO mice develop histologically diverse mammary tumors and metastatic lesions after a prolonged latency. Given the lack of lymphatic metastasis models, E2F5CKO syngeneic transplantation model can be significant resource to studying the mechanism of lymphatic metastasis. We have identified the dysregulation of KRas and Cyclin D1 as two potential mechanism in tumor initiation and progression in E2F5CKO mice, but further studies are needed to elucidate the specific mechanisms by which KRas and Cyclin D1

are dysregulated. Given the prolonged tumor latency and diverse morphology of E2F5CKO tumors, we postulate that multiple mechanisms and pathways are responsible for driving tumorigenesis. Thus, E2F5CKO can be favorable model when trying to recapitulate the heterogeneity in human breast cancer as well as studying age-related cancer changes.

## Material and Methods

### ***Animal generation***

All animal husbandry and use was in compliance with local, national and institutional guidelines. Ethical approval for the study was approved by Michigan State University Animal Care & Use Committee (IACUC) under AUF 06/18-084-00. E2F5CKO mice was generated by floxing exon 2 and 3 of E2F5 gene with loxP sites. E2F5 flox/flox mice were interbred with MMTV-Cre mice (a gift from Dr. William Muller). Mammary glands at different stages of mammary development were excised for histology. Mice were monitored weekly for tumor development. The endpoint for primary tumor was at 2000 $mm^3$.

### ***Cell culture***

Human mammary epithelial cells were cultured in Mammary Epithelial Cell Basal Medium (ATCC, Manassa, Virginia #PCS-600-030) supplemented with rH-insulin (5 ug/mL), L-glutamine (6 mM), epinephrine (1 µM), apo-transferrin (5 µg/ml) , rH-TGF-α (5 ng/ml), extractP (0.4%) and hydrocortisone hemmisuccinate (100 ng/ml). BT549 were cultured in RPMI supplemented with 10% FBS (Gibco #10437028) and 1% Antibiotic- Antimycotic reagent (Thermo Fisher, Waltham, MA, USA #15240062). MDA-MB-231 were cultured in

DMEM supplemented with 10% FBS (Gibco # 10437028) and 1% Antibiotic- Antimycotic reagent (Thermo Fisher, Waltham, MA, USA #15240062).

***E2F5-regulated genes***

Human Mammary Epithelial cells were infected with adenovirus expressing E2F5 or GFP. Cells were collected eighteen hours after infection. Total RNA was extracted using Qiagen RNeasy Mini kit. RNA was used with Affymetrix Human Genome U133 chip to generate gene expression data. RMA algorithm was used to normalized microarray dataset. Significance Analysis of Microarray was applied to the dataset to identify differentially expressed genes in HMEC-E2F5.

***Pathway analysis***

Single Sample Gene Set Enrichment Analysis and Gene Set Enrichment Analysis were performed on Broad Institute Genepattern interface. Normalized RMA data was used as input for microarray data and normalized TPM data was used as input for RNA-seq data. E2F5 activation signature was generated from genes that were upregulated and downregulated in HMEC overexpressing E2F5. A fold change of 2 was used for cut-off. E2F5 target geneset and other genesets used were derived from mSigdb.

***Histology***

For wholemount analysis, abdominal mammary fat was excised and place on glass slide. The slide was incubated in acetone for 24 hours. Next, the slide was placed in Harris Modified Hematoxylin for 24 hours. The slide was washed repeatedly in alcohol acid until the appropriate amount of stain is left. Following several hours in 70% ethanol, the slide was place in 100% ethanol for at least 3 hours. The slide was then incubated in xylene for 24 hours. Finally, permount was carefully added to the mammary tissue followed by a

glass cover slip. To evaluate mammary outgrowth, the distance from the nipple to the leading edge of the epithelium and the distance from the nipple to the midpoint of the thoracic lymph node were measured. Samples for histology were fixed in 10% formalin and submitted to Michigan State University Pathology lab.

***Microarray analysis***

Significant Analysis Microarray was used for differential gene expression in microarray data. The following published microarray datasets were used for analysis: terminal end bud and duct (GSE2988) and mammary gland developmental stages (GSE12247).

***RNA-sequencing***

Flash frozen tumor pieces were homogenized using Fisher Homogenizer 150 (Thermo Fisher, Waltham, MA, USA). Total RNA was isolated using QIAGEN RNeasy Midi Kit (Hilden, Germany #75142) with the manufacturer's protocol. RNA concentration was measured by Qubit and Agilent 2100 Bioanalzyer. RNA samples with RIN >7 was used for library preparation using the Illumina Tru-Seq stranded total RNA kit. RNA library was sequenced to a depth of >20M reads/sample with paired end 150 base paired reads on Illumina NovaSeq 6000. Adaptors were removed from reads using Trimmomatic v0.33. Quality control was performed using FastQC v0.11.5. Reads were aligned and mapped using STAR [35]. RSEM was used to quantify and normalize reads [36]. Differential gene expression analysis was performed using EdgeR [37].

***Clustering***

Unsupervised clustering was performed using Broad Institute's Morpheus interface.

***Immunoblotting***

To extract RNA from tissue, samples were homogenized using mortar and pestle in liquid nitrogen. Sample were lysed in TNE lysis buffer (0.05 M Tris HCl pH 8.0, 0.15 M NaCl, 2 mM EDTA, 0.01 N NaF and 1% NP40) with proteinase inhibitor (1 M $Na_3VO_4$, 58 µM PMSF, 10 µg/ml aprotinin and 10 µg/ml leupetin) for 1 hour on ice with constant agitation. Protein was quantitated using BCA (Thermo Scientific, Waltham, MA, USA #23225) and then boiled at 100°C for 5 min. Samples were loaded onto a 8-12% polyacrylamide gel. Separated protein was transferred onto a Immobilon- FL PVDF membrane (Millipore Sigma, Burlington, MA, USA #IPFL00010). Membranes were blocked in 5% milk in TBS with 0.1% Tween-20 (TBS-Tween) for 1 hour and then incubated in primary antibody overnight at 4°C. Following three washes in TBS-Tween the membrane was incubated in the appropriate antibody at a dilution 1:10,000 in 5% milk in TBS-Tween for 1 hour at room temperature. Membranes were washed 3x in TBS-Tween and imaged on LI-COR Odyssey imaging system (LI-COR Biosciences, Lincoln, NE, USA). The following antibodies were used: 1:1000 ERK (C-9), 1:100 E2F5 (C-8) from Santa Cruz Biotech (Santa Cruz, CA, USA), 1:1000 AKT, 1:1000 phospho- AKT, 1:4000 Vinculin from Cell Signaling Technology (Boston, MA, USA), 1:1000 phospho-ERK, 1:2000 Cyclin D1, 1:2000 Cyclin D3 from Thermo Fisher (Waltham, MA, USA), 1:1000 Cyclin D2 from Proteintech (Rosemont, IL, USA).

***Ras Activation assay***

Cells were grown to 50-80% confluency in 10-cm plate. After rinsing plate in cold PBS, lysis buffer (50 mM Tris pH 7.6, 500 mM NaCl, 0.1% SDS, 0.5% DOC, 1% Triton-X100, 0.5 mM $MgCl_{2)}$ with protease inhibitor (Roche, Indianapolis, IN #11836153001) was

added to plate. Cells were detached using cell scraper. The cell mixture was clarified at maximum speed at 4°C for 1 min. The lysate was mixed with Raf-RBD beads (gift from Dr. Sean Misek) and incubated at 4C for 1 hour with constant agitation. To pellet the beads, the mixture was centrifuge at 5000g at 4°C for 1 min. The beads were washed in 500 µl of wash buffer (50 mM Tris pH 7.6, 150 mM Nal, 1% Triton-X100, 0.5 mM $MgCl_2$). Protein was eluted off the beads using 2x laemmli buffer. Samples were boiled for 5 min at 100°C. The amount of activated Ras is determined by standard western blot procedure.

***Quantitative RT-PCR***

Flash frozen tumor pieces were homogenized using Fisher Homogenizer 150 (Thermo Fisher, Waltham, MA, USA). Total RNA was isolated using QIAGEN RNeasy Midi Kit (#75142; Hilden, Germany) with the manufacturer's protocol. Quantitative RT-PCR was performed using Luna Universal One-Step RT-qPCR kit (New England Biolabs, Ipswich, MA, USA) according to manufacturer's protocol using Agilent Mx3000P instrument. Primers were designed using Primer Bank tool (https://pga.mgh.harvard.edu/primerbank/). The following primers were used (5' to 3'): Rad51 forward, *TGTTGCTTATGCACCGAAGAA*; Rad51 reverse, *GCTGCCTCAGTCAGAATTTTGT*; KIF20A forward, *CAGCGGGCTTACTCTCTGATG*; KIF20A reverse, *GTCTGACAACAGGTCCTTTCG*; Sphk1 forward, *ACTGATACTCACCGAACGGAA*; Sphk1 reverse, *CCATCACCGGACATGACTGC*; CCND1 forward, *TGACTGCCGAGAAGTTGTGC*; CCND1 reverse, *CTCATCCGCCTCTGGCATT*; Gapdh forward, *AGGTCGGTGTGAACGGATTTG*; Gapdh reverse, *TGTAGACCATGTAGTTGAGGTCA*. Primer efficiency was 90-110% for all primers used. Delta-delta CT method was used for fold change analysis.

***CRISPRKO generation***

Guide RNA targeting E2F5 was inserted into CRISPR/Cas9 plasmid PX458, obtained from Addgene (Watertown, MA, USA #48138) as a gift from Dr. Feng Zhang, using the BBSI insertion site. This plasmid was transfected into BT549 and MDA-MB-231 using Lipopfecamine 3000 (Thermo Scientific, Waltham, MA, USA). 48 hours post-transfection, GFP- positive cells were sorted using FACs technology into 96 well plates. Clones were screened for the presences of INDELS using Sanger Sequencing. Knockouts were confirmed with western blot. For control, parental cell lines were transfected with empty (no sgRNA) and sorted into single- cell clones. The following guide RNA was used (5' to 3') forward: CACCGTCGAGTTCATCTAAGCCCG; reverse: AAACCGGGCTTAGATGAACTCGAC.

***Growth curve***

Wildtype and E2F5KO BT549 were seeded at 10,000 cells per well in a 6 well plate. Wildtype and E2F5KO MDA-MB-231 were seeded at 100,000 cells per well in a 6 well plate. Three wells were seeded per each time point. Cell were counted using Countess II (Thermo Scientific, Waltham, MA, USA) with 1:1 trypan blue. Duplicate counts were performed for each well.

***Wound-healing assay***

Wildtype and E2F5KO cells were seeded into a 6 well plate in triplicates and grown to 90-100% confluency. An optimized concentration of Mitomycin C was added prior to wound formation. A wound was created using a p200 pipette tip. Images were taken at 0 hours and 24 hours after the wound was created. The amount of wound healing at 24 hours was quantified using ImageJ software.

***Mammary fat pad transplantation***

E2F5CKO mammary tumors were harvested and stored in DMEM with 20% FBS and 10% DMSO at -80°C. Tumors were orthotopically implanted into the abdominal mammary gland of 6-to-10-week old MMTV Cre female mice. Mice were palpated 2x a week for mammary tumor formation. When the tumor size reached 2000 $mm^3$, samples were harvested for further analysis.

***Statistical analysis***

All statistical comparisons are performed with an unpaired students two-tailed, unpaired t-test.

## WORKS CITED

## WORKS CITED


1. Wu L, Timmers C, Malti B, Saavedra HI, Sang L, Chong GT, et al. The E2F1-3 transcription factors are essential for cellular proliferation. Nature. 2001;414:457–62.

2. Yamasaki L, Jacks T, Bronson R, Goillot E, Harlow E, Dyson NJ. Tumor induction and tissue atrophy in mice lacking E2F-1. Cell. 1996;85:537–48.

3. Murga M, Fernández-Capetillo O, Field SJ, Moreno B, R.-Borlado L, Fujiwara Y, et al. Mutation of E2F2 in mice causes enhanced T lymphocyte proliferation, leading to the development of autoimmunity. Immunity. 2001;15:959–70.

4. Humbert PO, Verona R, Trimarchi JM, Rogers C, Dandapani S, Lees JA. E2f3 is critical for normal cellular proliferation. Genes Dev. 2000;14:690–703.

5. Humbert PO, Rogers C, Ganiatsas S, Landsberg RL, Trimarchi JM, Dandapani S, et al. E2F4 is essential for normal erythrocyte maturation and neonatal viability. Mol Cell. 2000;6:281–91.

6. Lindeman GJ, Dagnino L, Gaubatz S, Xu Y, Bronson RT, Warren HB, et al. A specific, nonproliferative role for E2F-5 in choroid plexus function revealed by gene targeting. Genes Dev. 1998;12:1092–8.

7. Li FX, Zhu JW, Tessem JS, Beilke J, Varella-Garcia M, Jensen J, et al. The development of diabetes in E2f1/E2f2 mutant mice reveals important roles for bone marrow-derived cells in preventing islet cell loss. Proc Natl Acad Sci U S A. 2003 ;100:12935–40.

8. Panagiotis Zalmas L, Zhao X, Graham AL, Fisher R, Reilly C, Coutts AS, et al. DNA-damage response control of E2F7 and E2F8. EMBO Rep. E; 2008;9:252–9.

9. Gaubatz S, Lindeman GJ, Ishida S, Jakoi L, Nevins JR, Livingston DM, et al. E2F4 and E2F5 play an essential role in pocket protein-mediated G1 control. Mol Cell. 2000;6:729–35.

10. Bach K, Pensa S, Grzelak M, Hadfield J, Adams DJ, Marioni JC, et al. Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. Nat Commun. 2017;8:1–11.

11. Rennhack J, To B, Wermuth H, Andrechek ER. Mouse Models of Breast Cancer Share Amplification and Deletion Events with Human Breast Cancer. J Mammary Gland Biol Neoplasia. 2017;22:71–84.

12. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, et al. Genes that mediate breast cancer metastasis to lung. Nature. 2005;436:518–24.

13. Paula D. Bos, Xiang H.-F. Zhang, Cristina Nadal, Weiping Shu RRG and DXN. Genes that mediate breast cancer metastasis to the brain. Nature. 2009;459.

14. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006;439:353–7.

15. Gatza ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, et al. A pathway-based classification of human breast cancer. Proc Natl Acad Sci U S A. 2010;107:6994–9.

16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–50.

17. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. NIH Public Access; 2009 [cited 2018 Nov 22];462:108–12.

18. Chang JT, Nevins JR. GATHER: a systems approach to interpreting genomic signatures. Bioinformatics. Oxford University Press; 2006;22:2926–33.

19. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. J Clin Oncol. 2009;27:1160–7.

20. Pereira B, Chin SF, Rueda OM, Vollan HKM, Provenzano E, Bardwell HA, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun. 2016;7.

21. Choschzick M, Lassen P, Lebeau A, Marx AH, Terracciano L, Heilenkötter U, et al. Amplification of 8q21 in breast cancer is independent of MYC and associated with poor patient outcome. Mod Pathol. 2010;23:603–10.

22. Byrne JA, Chen Y, Martin la Rotta N, Peters GB. Challenges in Identifying Candidate Amplification Targets in Human Cancers: Chromosome 8q21 as a Case Study [Internet]. Genes and Cancer. 2012; 87–101.

23. Győrffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat. 2010;123:725–31.

24. McLaughlin SK, Olsen SN, Dake B, De Raedt T, Lim E, Bronson RT, et al. The RasGAP Gene, RASAL2, Is a Tumor and Metastasis Suppressor. Cancer Cell. 2013;24:365–78.

25. Wang TC, Cardiff RD, Zukerberg L, Lees E, Arnold A, Schmidt E V. Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. Nature. 1994;369:669–71.

26. Sinn E, Muller W, Pattengale P, Tepler I, Wallace R, Leder P. Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: Synergistic action of oncogenes in vivo. Cell. 1987;49:465–75.

27. Yu Q, Geng Y, Sicinski P. Specific protection against breast cancers by cyclin D1 ablation. Nature. 2001;411:1017–21.

28. Bowe DB, Kenney NJ, Adereth Y, Maroulakou IG. Suppression of Neu-induced mammary tumor growth in cyclin D1 deficient mice is compensated for by cyclin E. Oncogene. 2002;21:291–8.

29. Jeselsohn R, Brown NE, Arendt L, Klebba I, Hu MG, Kuperwasser C, et al. Cyclin D1 Kinase Activity Is Required for the Self-Renewal of Mammary Stem and Progenitor Cells that Are Targets of MMTV-ErbB2 Tumorigenesis. Cancer Cell. 2010;17:65–76.

30. Andrechek ER. HER2/Neu tumorigenesis and metastasis is regulated by E2F activator transcription factors. Oncogene. 2015;34:217–25.

31. Wang D, Russell JL, Johnson DG. E2F4 and E2F1 have similar proliferative properties but different apoptotic and oncogenic properties in vivo. Mol Cell Biol. 2000;20:3417–24.

32. Xu X, Bieda M, Jin VX, Rabinovich A, Oberley MJ, Green R, et al. A comprehensive ChIP-chip analysis of E2F1, E2F4, and E2F6 in normal and tumor cells reveals interchangeable roles of E2F family members. Genome Res. 2007;17:1550–61.

33. Kong L-J, Chang JT, Bild AH, Nevins JR. Compensation and specificity of function within the E2F family. Oncogene. 2007;26:321–7.

34. Welch, D.R. and Hurst, D.R., 2019. Defining the hallmarks of metastasis. *Cancer Research*, *79*(12), pp.3011-3027.

35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

36. Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.

37. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.

# CHAPTER 4:

# FUTURE DIRECTIONS

**E2F5 as a tumor suppressor in human breast cancer**

To test if E2F5 has a role in tumor initiation in human breast cancer, we can study the effects of E2F5 deletion in HMEC-hTERT expressing SV40 large T antigen (LT) and small t antigen (st). A previous report has demonstrated that expressing LT and st, which can inactivate p53 and pRb, along with HRas was suffice to transformed HMEC-hTERT [1]. Thus, we can examine if E2F5 loss is enough to induce transformation in HMEC-hTERT expressing LT and st. To evaluate transformation, we can perform soft agar assay to assess if the modified cells demonstrate anchorage-independent growth, a hall mark of cell transformation.

**Mechanism of tumorigenesis in E2F5CKO model**

Based on our findings, we hypothesize that Cyclin D1 can facilitate tumor initiation in E25CKO mice. To test this hypothesis, we can generate a cyclin D1-deficient E2F5CKO model. To examine whether Cyclin D1 is required tumor initiation in E2F5CKO mice, we will cross Cyclin D1KO FVB/NJ mice, from Jackson Lab, with E2F5CKO mice. After confirming loss of Cyclin D1 and E2F5 with PCR and western blot, mice will be monitored 2x weekly for mammary tumor development by palpation.

**E2F4 compensation in mammary gland of E2F5CKO mice**

To determine if E2F4 can compensate for E2F5 loss during mammary development, we can generate a E2F4/E2F5 double knockout mice. E2F4f mice from Jackson lab will be backcrossed to FVB background. E2F4f FVB mice will be interbred with E2F5KO mice. Loss of E2F4 and E2F5 will be confirmed with PCR and western blot. Wholemounts will be generated from mammary glands at week 4 and week 8 to examine mammary outgrowth and branching. At various developmental stages including virgin,

pregnancy, lactation and involution, mammary tissue will be collected for histology and flash frozen. To examine if E2F5 loss results in increased E2F4 expression, qRT-PCR will be performed on RNA extracted from frozen mammary tissue at each developmental stage. To assess for lactation defect, we will observe if E2F4/E2F5KO mice can nurse their pups effectively as indicated by body weight of pups at weaning.

**-Omic studies**

Given the morphological heterogeneity of E2F5CKO tumors, we postulate that different mechanism may be mediating tumor development and progression. To capture the heterogeneity between E2F5CKO tumors, we would perform RNA-seq on more E2F5CKO tumors (n= 15). In addition, we would include commonly used mouse models including MMTV-PyMT, MMTV-Neu and MMTV-MYC when generating our RNA-Seq in order to compare these models to E2F5CKO tumors. Pathway analysis, unsupervised hierarchical clustering and differential gene expression analysis will be performed on the RNA-seq data. To examine E2F5 direct targets in the mammary epithelium, we would perform ChIP- seq analysis of E2F5 in HMECs. Integrating the E2F5 ChIP-seq data with E2F5CKO tumor RNA-seq data will help identify direct E2F5 targets that may be dysregulated in E2F5CKO tumors.

**Identifying lymphatic metastasis genes**

To identify genes that potentially mediate lymph node metastasis, we will perform differential gene analysis on enriched axillary lymph node metastasis versus non-enriched primary mammary tumors. To prioritize which target genes to focus on, we will first rank order according to fold change. We will then screen for genes whose expression in breast cancer patients is associated with lymph node involvement. This will be

accomplished using a cohort of clinically annotated breast tumors. The function of target genes will be evaluated to identify those which may function in metastasis. This will be achieved using GATHER to determine gene ontology [2]. Top targets will be validated through qRT-PCR and western blot using enriched tumors compared to mammary tumors that did not metastasized to the lymph node. The top two validated candidates will be knocked out using CRISPR-Cas9 in tumor cell lines that have the propensity to metastasize to the axillary lymph node. We will use a cell line generated from enriched E2F5CKO axillary tumors as well as MDA-MB-231 human breast cancer cell line. MDA-MB-231 cell lines were chosen because previous studies demonstrate the ability of orthotopically injected MDB-MB-231 cells to form lymph node metastases [3]. Knockouts will be confirmed using Western blot and targeted sequencing. Following deletion of the target genes using CRISPR-Cas9, we will evaluate migration and invasion in these modified cell lines using wound healing and trans-well assays. To determine if the modified cell lines have altered ability to metastasize to the axillary lymph node, the knockout cell lines will be orthotopically injected into the mammary fat pad of 10 MMTV-Cre mice or NSG mice. As a control, WT E2F5CKO axillary tumor cell lines and MDA-MB-231 cell lines will be injected. Using the control cell lines, a pilot study will be performed to determine the required number of cells to inject.

## WORKS CITED

## WORKS CITED


1. Zhao JJ, Gjoerup O V., Subramanian RR, Cheng Y, Chen W, Roberts TM, et al. Human mammary epithelial cell transformation through the activation of phosphatidylinositol 3-kinase. Cancer Cell [Internet]. Cancer Cell; 2003 [cited 2020 Nov 24];3:483–95. Available from: https://pubmed.ncbi.nlm.nih.gov/12781366/

2. Chang JT, Nevins JR. GATHER: a systems approach to interpreting genomic signatures. Bioinformatics [Internet]. Oxford University Press; 2006 [cited 2018 Nov 22];22:2926–33.

3. Iorns E, Drews-Elger K, Ward TM, Dean S, Clarke J, Berry D, et al. A New Mouse Model for the Study of Human Breast Cancer Metastasis. PLoS One. Public Library of Science; 2012;7:e47995.