SIX2 MEDIATES LATE-STAGE METASTASIS VIA DIRECT REGULATION OF SOX2 AND
INDUCTION OF A CANCER STEM CELL PROGRAM

By

Michael Uretz John Oliphant

B.S., University of Illinois Champaign-Urbana 2008
M.S., University of Illinois Champaign-Urbana 2010

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Integrated Physiology Program

2019
This thesis for the Doctor of Philosophy degree by

Michael Uretz John Oliphant

has been approved for the

Integrated Physiology Program

Jennifer Richer, Chair

Andy Bradford

James Costello

Aaron Johnson

Andrew Thorburn

Heide Ford, Advisor

Date: August 9, 2019
Abstract

The capacity for tumor cells to metastasize efficiently is directly linked to their ability to colonize secondary sites. Here we identify Six2, a developmental transcription factor, as a critical regulator of a breast cancer stem cell program that enables metastatic colonization. In several triple-negative breast cancer (TNBC) models, Six2 enhanced the expression of genes associated with embryonic stem cell programs. Six2 directly bound the Sox2 Srr2 enhancer, promoting Sox2 expression and downstream expression of Nanog, which are both key pluripotency factors. Regulation of Sox2 by Six2 enhanced cancer stem cell properties and increased metastatic colonization. Six2 and Sox2 expression correlated highly in breast cancers including TNBC, where a Six2 expression signature was predictive of metastatic burden and poor clinical outcome. Our findings demonstrate that a SIX2/SOX2 axis is required for efficient metastatic colonization, underscoring a key role for stemness factors in outgrowth at secondary sites.

The form and content of this abstract are approved. I recommend its publication.

Approved: Heide L. Ford
I dedicate this Thesis to my wife, Jen and my son, Lincoln. I would not be the person or scientist I am today without your love, support and inspiration. I love you more than I can ever imagine.
ACKNOWLEDGEMENTS

I want to start by thanking Heide for giving me the opportunity to join her lab and work with her. Even before I arrived for graduate school in 2013, I had already heard about how wonderful of a person and mentor she was. When I got to campus it became even more clear from speaking to anyone who knew her, that she was a fantastic scientist who was very invested in the success and education of her students. Because of this, I knew I wanted to join her lab and be a part of all of the great science she was doing. Her passion for science, patience and positivity is unmatched and inspires me to always strive to reach my full potential. I would not be where I am today without your guidance and insight both in research and in life in general! Importantly, your ability to somehow balance your whirlwind of a science life and family life, while being as successful as you are, gives me solace in the fact that you can truly have it all! Thank you so much for all of your time, patience and guidance! I will never be able to repay you for everything you’ve done for me!

I want to also especially thank Christie for being my student mentor during my rotation in the Ford Lab. You taught me so much in such a short period of time. Thank you so much for the time and effort you put into teaching me how to think about science, career paths and family life. You are such an awesome person, scientist and friend. Thanks for being there for me then, now and in the future. You have also been such a great inspiration and I can’t wait to see what you do next!

I know that my time here would be nowhere near as fun as it was without all of the past and present lab members that I have been lucky enough to call GREAT friends! During all of the highs and lows, all of you were there to celebrate, dance, sing, dab, throw tissues, listen and brainstorm. Michelle, Deepika, Mel, Taylor, Connors (2 and 3), Jessica, Hengbo, Deguang, Anna, David, Lina, Ritz, Jenean, Chuan and Aaron I can’t thank you enough for all of your suggestions, positivity and support during this journey. I really feel like I was in the best lab with the best people and I could not imagine being anywhere else. I know I could reach out to any of
you at a moment’s notice and you would be there to help. Here’s to more happy hours, lunches on the strip, intense sports debates and even MORE intense political discussions!

Lastly, I want to thank my family. Jon and Kathryn, I’m grateful for our friendship and your support during these crazy last few years. It has been really hard to not always be there with my bro and sis for the birthday parties and family gatherings but thank you for understanding. I know you will always be there for me and I can’t wait to see our little ones grow up and start playing together! I want to say thank you to my parents who taught me the values of hard-work, dedication and perseverance. Mom and Dad, you are truly my greatest heroes. You’ve taught me to never give up and to always see the bright side of things. Thank you both SO MUCH for all of the sacrifices you made to give us a better life and an opportunity to succeed and follow our dreams.

Last, but not least, I want to thank my wife, Jen. I can’t even begin to tell you how much you have meant to me during this journey. There is absolutely no way I would have made it through this without you. Thanks for sitting/sleeping/nodding through all of my practice talks and presentations and giving me a fresh perspective that allows me to think about things in a new way and communicate my science better. You have also sacrificed so much to support me and push me forward. You have been, and will always be, there to lift me up through so many tough times and be my biggest fan through all of the great times. I love you more than anything and I’m so thankful that I get to call you my wife and best friend. I can’t wait to continue on this crazy journey with you and I wouldn’t want to do this with anyone else xoxo! Lincoln, I want you to know that I’m already so proud of you. Your genuine curiosity and happiness shines so bright that it doesn’t matter what might have went wrong that day. I can look at you and know everything is going to be ok. I hope I can show you that anything is possible and that if you follow your passion, great things will happen for you and everything will fall into place. I love you so much son. Thank you for being you and being the greatest gift your mom and I will ever receive.
# TABLE OF CONTENTS

## I. INTRODUCTION

- Stem Cells During Mammary Gland Development and Breast Cancer Metastasis
  - The Role of Mammary Stem Cells During Mammary Gland Development
  - The Role of Stem Cells in Breast Cancer Metastasis

## Developmental Regulators and Signaling Pathways in Mammary Stem Cells and Breast Cancer Metastasis

- The Hedgehog (HH) Pathway
  - Roles of the HH Pathway in Mammary Stem Cells
  - Role of HH Pathway During Breast Cancer Metastasis
- The WNT Pathway
  - Role of WNT Signaling in Mammary Stem cells
  - Role of WNT Signaling During Breast Cancer Metastasis
- The Notch Pathway
  - Role of NOTCH Signaling in Mammary Stem Cells
  - Role of NOTCH Signaling in Breast Cancer Metastasis
- The Six Family of Homeobox Transcription Factors
  - Role of Six Family in Mammary Stem Cells
  - Role of Six Family During Breast Cancer Metastasis

## Pluripotency Factors in Mammary Stem Cells and Breast Cancer Metastasis

- Pluripotency Factors as Regulators of Mammary Stem Cells
- Pluripotency Factors as Regulators of Breast Cancer Metastasis

## II. MATERIALS AND METHODS

- Cell Lines and Culture
- RNA-sequencing
- RNA-sequencing and Data Analysis
III. SIX2 MEDIATES LATE-STAGE METASTASIS VIA DIRECT REGULATION OF SOX2 AND INDUCTION OF A CANCER STEM CELL PROGRAM ................................................................. 48

Introduction .................................................................................................................... 48

Results ............................................................................................................................. 52

SIX2 expression is enriched in TNBC where it promotes metastasis .................................. 52

SIX2/Six2 regulates stem cell-associated phenotypes in TNBC and mouse mammary carcinoma cells ........................................................................................................... 54

Six2 regulates a transcriptional program associated with stemness .................................. 55

Six2 regulates master stemness factors Sox2 and Nanog .................................................. 63

Sox2 mediates stem cell-associated phenotypes and late-stage metastasis downstream of Six2 .................................................................................................................................. 64

SIX2 and SOX2 expression correlate in human breast cancer, and are together associated with poor prognosis .......................................................................................... 68

Discussion ....................................................................................................................... 71

Acknowledgements ......................................................................................................... 79

IV. ESTABLISHING NOVEL EXPERIMENTAL MODELS TO EXAMINE LATE-STAGE METASTASIS AND DETERMINE OVERLAPPING AND DISTINCT OF SIX1 AND SIX2 ............................................................................................................ 81

Introduction ..................................................................................................................... 81
LIST OF TABLES

Table 2.1 List of Oligos used for qRT-PCR, shRNA and siRNA experiments ................................................................. 40
Table 2.2 List of Antibodies used for experiments ........................................................................................................ 41
Table 3.1 List of top 15 upregulated genes in response to Six2 OE (SIX2-Gene Signature) ......................................................... 73

LIST OF FIGURES

Figure 1.1 The NOTCH, WNT and HEDGEHOG signaling pathways regulate differentiation and stem cell fate during mammary gland development ........................................ 6
Figure 1.2 Differentiation and pluripotency factors play critical roles in the MIC population ...................................................... 9
Figure 1.3 Cellular plasticity as the core characteristic of MICs .............................................................................................. 10
Figure 3.1 SIX2 expression is enriched in triple-negative breast cancer (TNBC) where it promotes metastasis ....................... 51
Figure 3.2 SIX2/Six2 regulates stem cell-associated phenotypes in TNBC and mouse mammary carcinoma cells ........................................... 54,55
Figure 3.3 Six2 OE in 4T07 cells is comparable with multiple other human TNBC cell lines with relatively high endogenous SIX2 expression .......................................................... 59
Figure 3.4 Six2 OE leads to a change in the transcriptional program of 4T07 cells and differential expression of stemness-related genes and pathways ................................................. 60,61
Figure 3.5 SIX2/Six2 regulates master pluripotency factors SOX2/Sox2 and NANOG/Nanog .......................................................... 64
Figure 3.6 Sox2 is upstream of Nanog and is directly regulated by Six2 via its Srr2 enhancer .......................................................... 65,66
Figure 3.7 Sox2 mediates stem cell-associated phenotypes and late-stage metastasis downstream of Six2 ............................... 68,69
Figure 3.8 Six2 regulates Sox2 in multiple human TNBC models .............................................................................................. 71
Figure 3.9  *SIX2* and *SOX2* mRNA positively correlate in breast cancer patient samples and upregulates genes associated with stemness and poorer prognosis ................................................................. 72

Figure 3.10  *SIX2* and *SOX2* expression correlate in human breast cancer, and a *SIX2*-gene expression signature is associated with poor prognosis ................................................................. 74

Figure 3.11  *Six2* is more highly expressed in the fetal mammary stem cell population compared to adult mouse mammary cells ................................................................. 75

Figure 3.12  The *SIX2/SOX2* axis mediates metastatic colonization ........................................................................................................................................ 79

Figure 4.1  Inducible *Six2* expression recapitulates gene expression of downstream targets observed with stable *Six2* OE ........................................................................................................................................ 87

Figure 4.2  Schematic displaying the experimental design for testing whether *Six2* regulates the continued growth of metastatic lesions .................................................................................................. 88

Figure 4.3  66cl4 *Six1* KO results in upregulation of *Six2*, whereas siRNA KD of *Six1/SIX1* or *Six2/SIX2* results in no compensation .................................................................................................................. 92

Figure 4.4  Stable shRNA KD *SIX2/Six2* results in no consistent changes in *SIX1/Six1* ........................................................................................................................................ 93

Figure 4.5  3X-FLAG tagged *Six2* 66cl4 cells preferentially bind to the Srr2 enhancer region compared to 3X-FLAG tagged *Six1* 66cl4 cells .................................................................................................................. 95

Figure 4.6  The Six Family of Homeoproteins ........................................................................................................................................ 100

Figure 4.7  Regulation of *Six1/SIX1* and *Six2/SIX2* during embryonic kidney development ........................................................................................................................................ 10
CHAPTER I
INTRODUCTION

Embryonic development involves the intricate and dynamic coordination of cellular and molecular processes under strict spatial and temporal control, such that a unicellular zygote gives rise to progeny that not only multiply in number, but migrate and differentiate, ultimately resulting in a mature, multicellular organism at the time of birth. The mammary gland is unique in the context of embryogenesis, as it is the only organ that does not fully mature until later in adult life during pregnancy, lactation and involution. A diverse set of cell types drive the formation of this highly complex organ through interaction with, and regulation by, local and systemic hormones and growth factors. Because of the highly structured order of events and substantial availability of mouse models to study mammary gland development, we have gained vital knowledge of the molecular underpinnings that regulate this process. This, in turn, has provided critical insight into how cells transform to become tumorigenic, and how we might target dysregulation of mammary gland development therapeutically to inhibit disease progression.

Mammary gland development in both mice and humans can be described in three distinct stages: embryonic, pubertal and adult [1]. During the embryonic stage, there are three key events: establishment of the bilateral milk lines, placode development, and primitive mammary bud formation and branching [2]. Slight differences in human embryonic mammary gland development include earlier formation of mammary milk-lines relative to mice, the development of only one pair of mammary placodes, and the development of multiple ductal trees that concentrate at the nipple [2]. For both mice and humans, birth to the onset of puberty is considered a quiescent state where the growth of the ductal tree within the mammary gland is isomorphic with body growth [1]. This growth pattern changes dramatically at puberty when both the mouse and human mammary glands undergo a resurgence in ductal tree growth and branching. Ovarian- and pituitary-derived hormones mediate multiple rounds of initiation,
elongation and invasion of the maturing ductal tree into the mammary mesenchyme. In the case of mice, lobules do not appear until the onset of pregnancy. Of note, although there is significant expansion of the ductal tree network into the fat pad, space remains in order to support additional branching during diestrus and/or pregnancy [1]. Development of the post-pubertal/adult mammary gland is also very similar between mice and humans. The mammary gland continues to respond to secretion of ovarian hormones during each estrous cycle. This allows for the formation of alveolar and lateral buds. Pregnancy ushers in a dynamic and extensive remodeling of the mammary gland mainly involving proliferation and differentiation of the alveolar buds into cells capable of producing and secreting milk [1]. Although this process is largely similar between mice and humans, the exceptions include lobules being present in the human mammary gland prior to pregnancy and full maturity not being reached in humans until the end of the first full-term pregnancy [1]. Studying this highly organized and complex developmental process has provided critical insight into the molecular mechanisms that allow for proper mammogenesis. It has also highlighted the relationship between normal development and cancer, as tumor cells often hijack the proliferative potential and plastic nature of cells within the mammary gland to become aggressive and metastasize.

**Stem Cells During Mammary Gland Development and Breast Cancer Metastasis**

**The Role of Mammary Stem Cells During Mammary Gland Development**

Two main lineages make up the hierarchal organization of the mammary gland, where the basal population is mostly responsible for contraction and the luminal population is critical for milk production [3]. Over the years, many studies have identified a critical population of mammary stem cells (MaSCs) that sits atop of this hierarchy and can give rise to the progenitors and differentiated cells of both lineages. The first evidence of MaSCs was identified through transplantation studies, where segments of mammary epithelium were injected into the cleared mammary fat pads of recipient mice and assessed for their repopulating capability [4]. Subsequent studies further supported the theory of MaSCs by using flow cytometry to isolate
rare populations of cells based largely on surface marker expression of CD24, CD29 and CD49f, which could repopulate an entire, functional, mammary gland. For instance, Visvader and colleagues found that CD24\textsuperscript{med}/CD49\textsuperscript{fhi} expressing cells represented MaSCs or mammary repopulating units (MRUs) that could reliably repopulate and generate entire functioning mammary glands via repeated transplantations [5]. Recent studies using lineage-tracing and single-cell sequencing analyses have refined the identity and role of MaSCs by suggesting that instead of one multipotent population that exists throughout development, there are distinct populations of bipotent and unipotent MaSCs that are present and function at different times to ensure proper mammary gland development. Specifically, bipotent MaSCs are mainly present during embryonic development where they are critical for the expansion of unipotent basal and luminal stem cells [6]. After birth, these more differentiated progenitor cells serve to produce all of the necessary cell types needed throughout postnatal development, puberty, pregnancy, lactation and involution [6]. Although controversies still exist as to the origin, timing and designations of MaSCs, the fact still remains that these rare populations are critical for proper mammary gland development. Importantly, the molecular mechanisms and signaling pathways that regulate the functions of these populations are often shared with malignant cells, providing insight into how tumor cells undergo transformation, tumor progression and metastasis.

**The Role of Stem Cells in Breast Cancer Metastasis**

Like normal MaSCs, breast cancer stem cells (CSCs) are associated with increased self-renewal, survival and plasticity compared to their more differentiated counterparts. For instance, one of the first studies to identify a CSC population was through isolation of CD24\textsuperscript{-}/CD44\textsuperscript{+} cells from human breast tumors [7]. It was found that this population of cells could initiate tumors with heterogeneous progeny over multiple transplantations in vivo [7]. These studies were some of the first to propose that a rare population of cells within a tumor could hijack normal stem cell functions to promote tumor initiation and progression. Additionally, mouse models of spontaneous mammary carcinoma, such as the MMTV-PyMT model, have demonstrated that
during hyperplasia the normal mouse MaSC population is expanded and, when isolated, has increased tumorigenicity in vivo [8]. These data suggest that the MaSC population could serve as a driver for tumor initiation after oncogenic transformation.

There has been controversy around the idea that rare stem-like cells contribute to tumor initiation, because in other tumor types such as melanoma, relatively high populations of CSCs (as high as 25%) make up the bulk tumor [9]. However, it should be noted that patient samples used for the melanoma studies were advanced and high grade. Since aggressive and advanced tumors tend to be more undifferentiated compared to early stage, non-invasive tumors, this could account for the enrichment of the CSC population, and suggests a stemness phenotype contributes to more advanced disease. Indeed, a later study by Boiko and colleagues confirmed that when examining earlier stage melanomas, rare populations of CSCs with increased self-renewal and tumor-initiation could be found [10]. More recent studies using intravital imaging in combination with lineage-tracing [11] demonstrated that CSCs differentiate into heterogeneous progeny, but that non-CSC could also dedifferentiate into CSCs, supporting the importance of cellular plasticity in tumorigenesis [11].

Previous studies have also revealed that breast CSCs, like normal MaSCs, often rely on developmental signaling pathways to maintain and modulate stemness properties throughout tumor progression [12]. Pathways including Hedgehog, Wnt and Notch have been shown to not only be critical for normal MaSC function, but can also mediate important functions of breast CSCs such as expansion, tumor-initiation, and plasticity [3] (Fig 1.1). Importantly, because of their critical roles in both MaSCs and breast CSCs, and the association of stemness properties with aggressive disease, recent studies have focused on how to exploit developmental pathways to inhibit metastasis.
Figure 1.1: The NOTCH, WNT and HEDGEHOG signaling pathways regulate differentiation and stem cell fate during mammary gland development. Figure from Shared signaling pathways in normal and breast cancer stem cells by Malhotra et. al [13].
Thus, while developmental pathways such as HH, WNT and NOTCH are important for multiple aspects of mammary gland development, in this thesis I will discuss how key developmental pathways and mediators impact normal mammary stem cells, and how they are hijacked during tumor progression and metastasis to enhance their ability to proliferate, survive, migrate to and invade distant organs. I will also discuss how a better understanding of the relationship between developmental signaling pathways and normal MaSCs has provided critical insight into how stemness properties can influence metastasis, and how this additional knowledge has led to promising therapeutic approaches to combat breast cancer progression.

Intriguingly, it has recently been proposed that disseminated tumor cells (DTC) contain a unique population of metastasis-initiating cells (MICs) that possess the ability to generate detectable metastases at secondary sites [14]. Importantly, recent studies have suggested that the MIC population specifically regulates stemness and differentiation properties that allow for better survival and establishment of newly arriving tumor cells at secondary sites, and thus share many properties with CSCs [14]. In addition, given the increasingly important role that plasticity plays in critical aspects of metastasis [14,15], recent studies have also aimed to determine whether master embryonic pluripotency factors function in the MIC population to promote metastatic outgrowth (Fig 1.2). Although it still remains difficult to isolate, study and characterize MICs [14], recent advances in genomic sequencing, imaging and animal modeling have begun to provide insights into the molecular mechanisms that regulate MICs. Through these studies, MICs have been shown to display a diverse set of properties including immune evasion, dormancy exit, regulation of foreign niches and maintenance of tumor-initiating capabilities to survive the harsh conditions they encounter to colonize secondary sites [14]. It is proposed that central to all of these abilities is the fact that MICs have an extraordinary level of plasticity, allowing them to adjust to a plethora of different stimuli and microenvironments (Fig 1.3) [14].
Figure 1.2: Differentiation and pluripotency factors play critical roles in the MIC population. A) Roles of differentiation and pluripotency factors during development. B) Repurposed roles of differentiation and pluripotency factors in MICs. Figure from Distinctive properties of metastasis-initiating cells by Celià-Terrassa and Kang [14].
Figure 1.3: Cellular plasticity as the core characteristic of MICs. This diagram lists nine characteristics commonly observed in MICs. Figure from Distinctive properties of metastasis-initiating cells by Celià-Terrassa and Kang [14].
Thus, current studies examining MICs have mainly focused on the role of pluripotency and differentiation as a way to better understanding how we might target this population to inhibit metastasis.

Therefore, in addition to reviewing the roles of developmental pathways and regulators in both MaSCs and breast cancer metastasis, I will also discuss how pluripotency factors impinge on, not only tumor progression and metastasis overall, but also their roles in critical aspects of MICs. Lastly, I will discuss how my thesis work has added novel mechanistic insight to how MICs may facilitate metastatic colonization by demonstrating how tumor cells hijack the developmental roles of the homeoprotein transcription factor, SIX2, including its regulation of stemness/progenitor phenotypes, to enable tumor cells to better survive, self-renew and colonize secondary sites. Specifically, I found that SIX2 promotes metastatic outgrowth via direct regulation of the master stemness and pluripotency factor, SOX2, which leads to subsequent activation of an additional pluripotency regulator, NANOG. Using flow cytometry, RNA-seq analyses and in vivo limiting dilution assays, I demonstrated that SIX2 regulates a stemness program in multiple triple negative breast cancer (TNBC) models. Importantly, I identify mechanistically that Sox2 is the critical mediator downstream of Six2 in promoting stemness properties in vitro and late-stage metastasis in vivo. Overall, the studies in this thesis uncover novel mechanistic insights into how developmental regulators and pluripotency factors may influence normal stem cell biology, how they are repurposed to mediate critical aspects of tumor progression and metastasis and how this possibly occurs through regulation of MIC maintenance and function.
Developmental Regulators and Signaling Pathways in Mammary Stem Cells and Breast Cancer Metastasis

The Hedgehog (HH) Pathway

Roles of the HH Pathway in Mammary Stem Cells

Initiation of the canonical Hedgehog (HH) signaling cascade relies on the presence of the HH ligands: Sonic Hedgehog (SHH), Desert Hedgehog (DHH) and Indian Hedgehog (IHH) [16,17]. In the absence of HH ligands, the PATCHED (PTCH) receptor inhibits the SMOOTHENED (SMO) receptor by keeping it sequestered at vesicle membranes. Sequestration of SMO to vesicles results in phosphorylation (via Protein Kinase A (PKA), Glycogen Synthase Kinase 3 Beta (GSK-3β) or Casein Kinase 1 (CK1)) of members of the GLI transcription factor family (GLI1, GLI2 and GLI3). This results in forms of GLI (generally GLI2 and/or GLI3) that translocate to the nucleus and act as transcriptional repressors of HH target genes and suppression of the pathway. When HH ligand is present, it binds to PTCH receptors, resulting in a conformational change that allows for release of SMO. SMO goes on to inhibit phosphorylation of one or more GLIs (generally GLI1 and GLI2), allowing translocation away from the cilium and accumulation in the nucleus, where one or more GLI transcription factors can now act as transcriptional activators of target genes that are important for cellular processes including survival, proliferation, migration, invasion and stemness [16] (Fig 1.1).

Importantly, studies have demonstrated that the Hh signaling pathway can play a role in regulating stemness in the mammary gland [16]. For example, Ihh is expressed by mammary progenitor cells and functions in a feedback loop with Tp63 to promote the clonogenicity and expansion of the mammary progenitor cell population [18]. Similarly, addition of recombinant SHH in normal breast tissue collected from reduction mammoplasties results in increased mammosphere formation [19]. This effect was shown to be dependent on SMO because treatment with cyclopamine, a direct antagonist of SMO, blocks SHH-mediated increases in mammosphere formation [19]. Taken together, these data suggest that HH signaling can
promote mammary stem cell self-renewal in the normal mammary gland. However, these studies were conducted in vitro, with mammosphere formation as the only functional readout for stemness. In contrast, in vivo transplantation of Shh or Ihh deleted mammary epithelium display no mammary gland phenotypes when transplanted into cleared mammary fat pads, as shown by complete formation of a lactation competent ductal tree [20]. Studies examining the role of Dhh in MaSC function have yet to be performed. It is proposed that functional redundancy among the ligands is responsible for masking phenotypes in the mammary gland, but this hypothesis has not been fully explored.

Numerous molecules downstream of the Hh ligands, including PTCH1, SMO, GLI1 and GLI2, are more highly expressed in mammary stem/progenitor cells compared to differentiated mammary cells, suggesting a role for this pathway in stemness [19]. Functional in vivo studies targeting the downstream components of the Hh ligands have provided support for Hh signaling as a key regulator of mammary stem cell function. For instance, Li and colleagues demonstrated that Patch1+/− mice have an increased mammary progenitor population, as defined by Lin−/CD24+/CD29low cells [18]. In the aforementioned study, Patch1 was shown to be critical for mammary stem cell quiescence. Specifically, Ptch1+/− mice exhibit constitutively active Hh signaling in MaSCs, leading to an expansion of the mammary progenitor pool due to increased proliferation of mammary stem cells, as measured by BrdU uptake and retention [18]. Similarly, using an activated human SMO mouse mammary gland model (MMTV-SmoM2), Moraes and colleagues found that SMO activation resulted in an increase in mammosphere formation efficiency. However, SMO activation also negatively impacted the regenerative capacity of mammary stem cells in vivo [21]. These data suggest that while SMO functions to promote the survival of mammary stem cells, it does not regulate the expansion or proportion of the overall stem cell population.

The GLI transcription factors also have overlapping and distinct functions in the context of mammary stem cell maintenance and function. For instance, Liu and colleagues
demonstrated that overexpression (OE) of either GLI1 or GLI2 increases mammosphere formation and induces transcription of the master stem cell regulator BMI-1 [19]. Additionally, studies by Lewis and colleagues have uncovered a unique role for Gli2 in mammary stem cell activity. The authors found that embryonic Gli2Lost mammary glands (composed of both the stromal and epithelial compartments) transplanted into cleared mammary fat pads were hyperplastic and displayed morphological defects compared to wildtype counterparts [22]. Interestingly, transplantation of Gli2Lost epithelial cells only, resulted in normal formation of the mammary gland in recipient mice. More recent studies by Zhao and colleagues support these data by showing that stromal-specific conditional KO of Gli2 results in decreased colony formation in vitro and transplantation efficiency in vivo [23]. Collectively, these data suggest that, in addition to cell autonomous regulation of stemness in the mammary gland, Hh signaling also plays a role in promoting a supportive microenvironment for mammary stem cells.

Similar to Hh ligand studies, studies attempting to identify the underlying mechanisms by which individual Gli transcription factors regulate mammary gland development have been difficult. It is likely that compensation between the family members is responsible for the lack of phenotypes with single KOs of either the ligands or the Gli transcription factors. For instance, Bai and colleagues showed that replacement of Gli2 with Gli1 was sufficient to restore activation of the Hh signaling pathway in Gli2-null KO mice (which are embryonic lethal) [24]. Further, although Gli1 KO mice are viable, heterozygous KO of Gli2 downstream of Gli1 KO results in severe developmental defects that are perinatally lethal, suggesting dose-dependent compensation of the Gli transcription factors [24]. Taken together, additional studies focusing on the compensation mechanisms of Gli transcription factors may shed more light on their functional roles in mammary stem cell function, as currently the mechanistic roles of Gli transcription factors in this context remain poorly understood.
**Role of HH Pathway During Breast Cancer Metastasis**

Dysregulation of the HH signaling network is associated with multiple aspects of tumorigenesis including tumor initiation, progression and metastasis [20]. While activating mutations in multiple HH signaling components have been identified in medulloblastoma, basal cell carcinoma and rhabdomyosarcoma [20], such mutations are not often observed in breast cancer [25]. Still, active HH signaling has been associated with breast cancer, and with breast cancer stemness. For example, CD24-/CD44+ breast cancer cells, which are enriched for the breast cancer stem cell (BCSC) population, have higher expression of *PTCH1*, *GLI1*, and *GLI2* mRNA when compared to the bulk tumor [19]. Similarly, it was shown that Hoechst dye negative-associated stem-like cells (stem cells can efflux Hoechst dye more efficiently than differentiated cells) express higher levels of SHH and GLI1 compared to the bulk tumor [26]. Interestingly, these studies also demonstrated that shRNA KD of either *SHH* or *GLI1*, as well as treatment with the SMO antagonist, cyclopamine, can decrease this population. These data suggest that HH signaling is critical for tumor initiation, and that targeting SMO or GLI could be a promising therapeutic strategy to prevent expansion of the CSC population. For example, recent studies have also implicated Hh signaling in the epithelial-to-mesenchymal transition (EMT) and metastasis. Specifically, Hh pathway components such as Ptch1 and Smo have been implicated in breast cancer metastasis to the bone [27]. In these studies, the authors demonstrated that mice with heterozygous loss of *Ptch* exhibit increased bone resorption and tumor burden, whereas treatment with Smo inhibitors decreased osteoclast activity and tumor burden, suggesting that microenvironmental effects of Hh signaling likely impinge on metastasis and growth of the tumor cells at distant sites. This has important implications for the MIC population, as these data also indicate that Hh signaling crosstalk between the metastatic niche and newly arriving tumor cells could be important for metastasis initiation and colonization. Therefore, future studies investigating the role of Hh signaling specifically within the MIC
population, and how that influences the metastatic niche, would provide a better understanding of the underlying mechanisms that support growth of MICs at distant sites.

Of all the HH signaling components, studies suggest that the GLI transcription factors play the most critical role in tumor progression and metastasis, as they serve as the rate-limiting step in the HH signaling pathway. The GLI gene family is often aberrantly expressed in the context of tumorigenesis, with DNA copy-number gain alterations identified in all three GLI family members across multiple tumor types. Further, several studies have demonstrated that GLI transcription factors can be activated through a multitude of non-canonical tumor promotional ligands and regulators such as Phosphoinositide 3-kinase (PI3-K), Transforming Growth Factor Beta (TGFβ), RAS and Protein kinase C (PKC) [20]. Gli activation leads to upregulation of genes associated with multiple tumor progression and metastasis-associated phenotypes including proliferation, EMT, stemness, migration and invasion. Most studies examining the roles of GLI transcription factors in breast cancer have focused on GLI1, as it is the most upregulated gene of the GLI family relative to normal breast and contributes to multiple aspects of tumorigenesis. For instance, conditional transgenic OE of Gli1 is sufficient to induce tumor initiation and survival [28]. Both cell line and mouse models have demonstrated that GLI1/Gli1 function is critical for cancer stem cell function and maintenance. For example, Goel and colleagues demonstrated that GLI1 regulates an autocrine loop to induce expression of the key stem cell factor, BMI1 [29]. This regulatory axis is specifically activated in TNBC, one of the more aggressive subtypes of breast cancer, where it leads to increased tumorsphere formation in vitro and tumor initiation in vivo. GLI1 is also upregulated in the poorly differentiated claudin-low subtype of breast cancer, where it is associated with increased self-renewal, clonogenicity and EMT in breast cancer cells [25]. Further, Tao and colleagues identified both SMO and GLI1 as being significantly increased in TNBC patient samples compared to non-TNBC samples [30]. Additionally, GLI1 has recently been shown to be involved in non-cell autonomous effects that promote EMT and metastasis in breast cancer [31]. Our group demonstrated that EMT breast
cancer cells act via the homeoprotein SIX1 to upregulate GLI1 expression in neighboring non-EMT breast cancer cells through paracrine signaling [31]. Subsequent activation of GLI1 in non-EMT breast cancer cells results in increased migration, invasion, and anoikis resistance in vitro [31]. Interestingly, treatment with the downstream GLI inhibitor GANT61 inhibited metastasis-associated phenotypes in vitro and non-cell autonomous-mediated metastasis in vivo. Further, PDX tumor growth was also inhibited with GANT61 treatment compared to upstream inhibition of SMO, suggesting that GLI1 can act in a SMO-independent manner to promote tumor growth [31].

Given the above, it is not a surprise that HH family members are significantly associated with patient prognosis and survival. Indeed, along with SHH, GLI1 is positively associated with invasive breast cancer compared to non-invasive breast cancer [20]. In addition, activated GLI1, as defined by nuclear localization, is correlated with aggressive tumor grade, lymph node positivity, increased recurrence and decreased overall survival across all breast cancer subtypes [20]. Taken together, the data above support continued research into the molecular mechanisms that are regulated by GLI transcription factors, and that direct blockade of GLIs may be the most effective means to inhibit HH signaling in breast cancer. However, because of the various roles of HH signaling in the crosstalk between tumor cells and the surrounding microenvironment, therapeutic strategies should be tailored to a specific patient. This is evident in recent reports that suggest stromal-derived HH signaling can both promote and suppress tumor growth and metastasis, and that this is tumor type and stage dependent [17]. Therefore, further investigation is needed to better understand the roles of Hh signaling both in cancer stem cells, as well as in the microenvironment, to more effectively inhibit its tumor promotional functions.
The WNT Pathway

Role of WNT Signaling in Mammary Stem cells

WNT signaling plays a fundamental role in embryonic development. Genetic mouse models examining multiple Wnt signaling components such as Wnt ligands, and its downstream transcription factors Ctnnb1 (β-catenin), Lef1 and Tcf, have demonstrated that the Wnt pathway is critical for the morphogenesis of multiple organ systems such as the intestine, bone, lungs, kidneys and hematopoietic system [32]. Similar to HH signaling, WNT pathway activation relies on ligand binding to receptors in order to stabilize downstream transcription factors and activate target genes. In the absence of WNT signaling, the downstream regulator β-CATENIN is readily bound by a complex consisting of tumor suppressors (APC), and kinases (GSK3β) that target β-CATENIN for degradation through the proteasome. Upon binding of WNT to the FRIZZLED (FZD) receptor, the APC/GSKβ complex is dissociated, resulting in stabilization and translocation of β-CATENIN to the nucleus where it forms a complex with TCF to activate transcription of target genes that are involved in self-renewal, EMT and cell proliferation (Fig 1.1). However, in contrast to HH signaling, which is largely regulated by activating ligands, the WNT pathway can be regulated by both activating and inhibitory extracellular proteins [32], suggesting complex and tight control of the pathway is required to ensure proper signaling and expression of downstream target genes.

Due to its critical roles in cellular processes such as stemness, proliferation and EMT during embryonic development, it is not surprising that multiple components of the Wnt pathway also influence mammary stem cell maintenance and function [33]. For example, Zeng and colleagues demonstrated that in addition to its roles in regulating crucial aspects of embryogenesis such as cell fate determination, polarity, and cell migration, Wnt signaling is also critical for the self-renewal and expansion of the mammary stem cell population during mammogenesis [34]. More specifically, Wnt ligand expression marks the mammary stem cell population, and addition of purified Wnt protein, such as Wnt3a, is sufficient to expand and
maintain the population. Additionally, the Wnt target genes, Lgr5 and Axin2, are used as markers of mammary stem cells, with the latter also having a functional role in mammary gland development [34,35]. Using Axin2-LacZ reporter mice, Amerongen and colleagues demonstrated that Axin2+ cells are required for the outgrowth of the mammary epithelial network during puberty, and that these cells are capable of giving rise to both basal and luminal cell lineages [36]. Overall, the Wnt-pathway is essential for mammary gland development, with a significant and specific role in mammary stem cell maintenance and proliferation. Future studies identifying the origin of Wnt ligands and the role of the mammary stem cell niche in Wnt signaling would provide an even better understanding of the functions of the pathway during normal mammary gland development and neoplasia.

As a main downstream regulator of the WNT signaling network, β-CATENIN has been implicated in several aspects of mammary stem cell function including survival, proliferation and maintenance [33,37]. Transgenic models have been valuable in determining the roles of both transcriptional and cell-cell adhesion regulation by β-CATENIN. For example, when Woodward and colleagues isolated primary mouse mammary epithelial cells (MECs) from mice with gain-of-function, conditionally stabilized β-catenin, they found increased radioresistance of progenitor cells specifically. This indicated that activated β-catenin could enhance stem cell survival [38]. Although a potential role for β-CATENIN, as a measure of WNT pathway activation, in mammary stem cell maintenance and function is supported by data from other tissue types, few studies have examined its specific role in this context. Dissection of the functions of specific domains within the β-CATENIN protein would enable investigation of which aspects of β-CATENIN function (its role in adherens junctions, as a transcription factor, or both) are critical for the mammary stem cell population and niche.

Role of WNT Signaling During Breast Cancer Metastasis

The Wnt pathway is implicated in numerous cancer types including colon, breast, melanoma and hepatocellular carcinomas [39]. In the context of breast cancer, WNT signaling is
activated (as measured by β-CATENIN nuclear accumulation) in approximately 50% of all breast cancer patients and correlates with poor survival [40]. Interestingly, only a small proportion of breast tumors contain somatic mutations of WNT signaling components [41]. Instead, many studies have identified aberrant expression of pathway components as a prominent mechanism of oncogenic function [33]. Wnt-1 ligand was one of the first Wnt pathway components to be identified as an oncogene due to its activation in the MMTV mouse model of breast cancer [42]. Additionally, WNT signaling components such as CTNNB1, LEF1 and AXIN2 are often amplified in breast tumors or more highly expressed in breast cancer cell lines compared to their normal counterparts, and play a critical role in multiple aspects of tumor progression and metastasis. In the MMTV mouse model, aberrant Wnt signaling causes an expansion of mammary stem cells during tumorigenesis [43]. Since then, multiple studies have implicated Wnt signaling in breast cancer stemness, where it has remained an area of significant interest with respect to tumor initiation and recurrence [43]. Similar to development, WNT signaling can also regulate the conversion of epithelial breast cancer cells to a more metastatic, mesenchymal-like cell type [44,45]. For example, Wu and colleagues demonstrated that canonical WNT signaling promotes a mesenchymal phenotype in breast cancer cells through upregulation of EMT-inducing transcription factors such as SLUG and SNAIL [46].

In addition to regulating EMT and stemness of cancer cells, WNT signaling has also been implicated in promoting metastatic colonization at secondary sites, both through effects on the cancer cells themselves as well as the tumor microenvironment. Malanchi and colleagues demonstrated that the ECM protein, Periostin, produced by stromal fibroblasts, promotes Wnt signaling in the lung microenvironment to mediate crosstalk with newly arriving tumor cells [47]. This allows for high concentrations of Wnt ligands in the metastatic niche, thereby promoting survival and proliferation of MIC/CSCs. Similarly, Oskarsson and colleagues demonstrated that the lung metastasis associated protein, Tenascin-C (TNC), regulates WNT signaling to increase metastatic fitness and outgrowth of breast cancer cells in the lung parenchyma [48].
Studies identifying roles for individual TCF family members are limited, but diverse and multifaceted. For example, consistent with its role in maintaining normal mammary epithelial cell (MEC) sphere formation and stemness, TCF3 is highly expressed in poorly differentiated human breast cancers [49]. Indeed, silencing of TCF3 expression in basal-like breast cancer cell lines results in decreased tumorsphere formation \textit{in vitro} and tumor initiation and growth \textit{in vivo} [49]. Interestingly, although TCF1 plays a role, in concert with TCF3, to control embryonic stem cell function [50], there is no current evidence that this function is mirrored in its roles during breast cancer progression. Instead, TCF1 has identified roles in invasion and metastasis that are thought to be dependent on its regulation of \textit{MMP} expression, leading to ECM degradation/remodeling and infiltration of tumor cells into the blood stream [51]. TCF4 activity can also increase the invasiveness of ER+ and TNBC cells \textit{in vitro} via direct regulation of the immune modulator and key metastasis regulator, OSTEOPONTIN (OPN) [52].

In contrast to the TCF family, many studies have uncovered roles for the critical WNT downstream regulator, \textit{β}-CATENIN, in almost every aspect of tumorigenesis such as initiation, cancer stemness, EMT and metastasis [39]. Here, we will summarize the role of \textit{β}-CATENIN in the functions above and how this has led to promising therapeutic strategies to inhibit Wnt signaling to improve patient outcomes. Using the MMTV mouse mammary carcinoma model, Li et. al showed that transgenic OE of \textit{β}-catenin induces oncogenesis by targeting the luminal progenitor cell population (as marked by \textit{Keratin-6} expression) to increase proliferation [53]. Further, another report from Michaelson and Leder demonstrated that expression of a constitutively active form of \textit{β}-catenin in the mammary gland was sufficient to induce hyperplasia, as well as adenocarcinoma, and that this effect was mediated by transcriptional upregulation of oncoproteins such as \textit{cMyc} and \textit{CyclinD1} [54]. These data indicate that \textit{β}-catenin is a key factor in tumor survival and progression, and that a better understanding of its role in tumor progression could present as a promising therapeutic target.
During metastasis, reports have shown that β-CATENIN activity leads to decreased levels of the epithelial marker E-CADHERIN and increased levels of EMT transcription factors such as VIMENTIN and SNAIL [55]. For instance, by examining a phosphorylation-resistant (S33Y mutant) form of β-CATENIN, Yook and colleagues demonstrated that β-CATENIN mediates the transformation of weakly metastatic, epithelial MCF7 cells into more aggressive and invasive mesenchymal-like cells through increasing levels of the EMT-inducing transcription factor SNAIL1 [56]. Interestingly, this effect was not mediated through transcriptional activity of β-CATENIN, but rather through protein stabilization of the SNAIL1 protein. More recently, Jang and colleagues not only demonstrated the necessity of β-catenin in metastasis in vivo, but also demonstrated its therapeutic potential [57]. In their previous study, the authors observed that β-CATENIN activity was higher in breast cancer stem cells compared to the bulk tumor [58]. They went on to show that β-CATENIN-mediated expansion of the breast cancer stem cell population was required for tumorsphere formation in vitro, tumor growth in vitro and in vivo and migration in vitro and in vivo [57]. Importantly, they demonstrated that treatment of breast cancer cells with the small molecular inhibitor, CWP232228, which prevents β-Catenin/Tcf binding in the nucleus, preferentially decreased the growth of breast cancer stem cells and reduced breast cancer stem cell activity, as measured by tumorsphere formation [59]. Consequently, treatment with CWP232228 after tail vein injection of 4T1 cells led to decreased metastatic burden and increased overall survival [59]. These data suggests that therapies targeting β-Catenin could be effective at inhibiting metastasis in patients where Wnt signaling is active [57], and that this might be a result of decreasing the breast cancer stem cell population, as well as of numerous other pro-metastatic activities β-Catenin.

Importantly, WNT signaling correlates strongly with decreased patient survival, particularly in more aggressive subtypes of breast cancer. OE of WNT ligands is associated with decreased distant-metastasis free survival and increased recurrence in TNBC patients [60]. Additionally, using a β-CATENIN specific gene signature in a small breast cancer patient cohort,
Dey and colleagues found that not only is WNT pathway activation more associated with the TNBC subtype compared to other subtypes, but patients with activated WNT signaling were more likely to have increased lung metastasis [61]. These data emphasize the importance of developing therapeutic strategies to target the WNT signaling network to improve patient outcomes. However, since the WNT signaling axis also has a prominent role in normal homeostasis, particularly in tissues such as the colon, the development of therapeutics targeting WNT pathway components have been marginally effective due to toxicity in normal tissues. Restoration of WNT inhibitory ligands such as WNT5A [62], have shown promising results in vitro, but more physiologically relevant, context-dependent in vivo studies testing the functional role of WNT5a in metastasis are lacking. More recent development of WNT inhibitors have shown increased promise. For instance, Porcupine inhibitors (which prevent palmitoylation and transport of WNT ligands) and Frizzled inhibitor antibodies inhibit both canonical and noncanonical WNT signaling and cancer proliferation, and do so with reduced toxicity in mouse models [63].

Lastly, although expression data from patient samples have identified β-CATENIN/TCF/LEF1 activation as associated with prognostic indicators of tumor aggressiveness including higher grade, disease-free survival and poor differentiation [64], the association of individual components of the transcriptional complex with disease progression is more complicated. For example, high TCF3 is significantly correlated with decreased survival and poor prognosis in multiple patient datasets [49]. In contrast, TCF4 has been shown to be both positively or negatively associated with prognosis depending on the context [57,65]. However, these studies almost exclusively assessed mRNA expression. Therefore, IHC analysis in patient tumor samples should be performed to confirm these relationships.

Additionally, most of the studies examining the roles of TCF family members in breast cancer progression were only done in cell line models in vitro. With the emerging knowledge that other factors such as the tumor microenvironment, immune function and tumor heterogeneity all influence tumor progression, future studies should also incorporate immunocompetent mouse or
patient-derived xenograft models, where appropriate, to more accurately recapitulate disease progression in the clinical setting.

**The Notch Pathway**

*Role of NOTCH Signaling in Mammary Stem Cells*

The NOTCH signaling pathway is involved in the development of multiple organ systems including heart, pancreas, blood vessels, nervous system and the mammary gland [66]. In general, the pathway consists of NOTCH receptors (NOTCH1-4) interacting with surface bound or secreted ligands (known collectively as DSL ligands: DELTA, SERRATE and LAG-2). In mammals, the two main sets of Notch ligands are DELTA-LIKE (DELTA-LIKE-1, DELTA-LIKE-2, DELTA-LIKE-3 and DELTA-LIKE-4) and SERRATE-LIKE (JAGGED1 and JAGGED2) [67]. Activation of NOTCH signaling occurs when NOTCH receptors on one cell bind to ligand receptors on another cell. Following binding, the NOTCH receptor is cleaved by the γ-SECRETASE proteasome complex, leading to release of the NOTCH intracellular domain (NOTCH-ICD). NOTCH-ICD is then translocated to the nucleus where it interacts with cofactors such as RBP-JK, CSL and P300 to activate transcription of target genes involved in numerous cellular processes including proliferation, differentiation and self-renewal [66] (Fig 1.1).

In the mammary gland, Notch signaling has mainly been implicated in the function and maintenance of the mammary stem cell population and driving luminal lineage commitment. The expression of Notch pathway ligands is diverse, as demonstrated by the mammary stem cell-enriched populations mainly expressing Delta1, the luminal cell populations expressing Jagged1 and all subsets expressing varying levels of Jagged2, Delta3 or Delta4 [68]. This expression pattern largely defines their functions throughout mammary gland development, where the role of pathway members can both promote and inhibit various aspects of mammary stem cell function. For instance, a recent study by Chakrabarti and colleagues showed that Dll1 is enriched in mammary stem cells and is required for mammary stem cell survival and ductal morphogenesis [69]. Interestingly, they also found that Dll1 is a critical mediator of crosstalk
between mammary epithelial cells and macrophages in the stem cell niche to promote stem cell activity [69]. Further association with Notch ligand and stemness was identified by Dontu and colleagues where they demonstrated that activated Notch signaling results in mammary stem cell self-renewal and branching morphogenesis [70]. These effects could be reversed through treatment with antibodies targeting Notch4 receptor or inhibitors against y-Secretase activity. NOTCH pathway ligands are also involved in inhibiting the expansion of mammary stem cells through non-cell autonomous effects. In a study done by Sizemore and colleagues, they demonstrated that stroma-derived JAGGED1 was necessary for NOTCH3 receptor-mediated inhibition of stem cell proliferation and activity [71]. Mechanistically, JAGGED1 was downregulated in PTEN-null fibroblasts, leading to a decrease in NOTCH3 receptor-mediated signaling. This resulted in significant increases in stem cell activity and expansion of the population. However, reintroduction of JAGGED1 downstream of PTEN loss in fibroblasts restored inhibition of stem cell activity [71]. These data indicated that the PTEN-JAGGED1 axis was critical for controlling the mammary stem cell population through niche signaling. Animal models have also shown that OE of either Notch1 or Notch3 receptors is sufficient to block mammary gland development and induce tumors, suggesting that aberrant Notch pathway activity restricts mammary stem cells to a highly proliferative, undifferentiated state that is susceptible to transformation and inappropriate outgrowth [72]. Collectively, these data emphasize the importance of stromal-epithelial cell crosstalk for MaSC function, where Notch signaling plays a key role in ensuring the proper control of the mammary stem cell population.

In addition to NOTCH pathway ligands, NOTCH pathway receptors have also been identified as critical mediators of stemness properties in mammary epithelial cells. Both Notch1 and Notch4 have been the most extensively studied in this context. Previous studies have shown that NOTCH1 receptor specifically regulates asymmetric cell division in human breast epithelial stem cells [73]. Additionally, Notch1 receptor marks the ER- luminal progenitor population [73]. Interestingly, although these cells are unipotent in mice, they become highly
plastic when conducting transplantation experiments, suggesting that Notch1 receptor can mediate plasticity in mammary epithelial cells \textit{in vitro} [73]. A more recent study used two mouse models, one containing Confetti reporters to mark Notch receptor+ cells, and another that contained a gain-of-function \textit{Notch1} receptor expressed specifically in mammary epithelial cells, to demonstrate that Notch1 was critical for inhibiting the transdifferentiation of multipotent mammary stem cells to basal progenitors [74]. Further, late in luminal differentiation, Notch1 receptor also inhibits conversion of ER- cells to ER+ cells. Similarly, \textit{Notch4} receptor OE is also sufficient to inhibit differentiation and promote self-renewal of breast epithelial cells \textit{in vitro}. Specifically, mammospheres treated with Notch4 receptor blocking antibody prevent secondary mammosphere formation. Interestingly, this effect was time-dependent, as treatment with the Notch4 receptor blocking antibody after allowing the mammospheres to form for 48h had no effect. These data suggest that Notch4 receptor is a regulator of self-renewal, but only in the initial stages where the sphere-initiating cells are most critical for survival and proliferation in suspension [70]. Taken together, the data overwhelmingly show that Notch regulators are key mediators of stemness and plasticity in the mammary gland.

\textit{Role of NOTCH Signaling in Breast Cancer Metastasis}

Dysregulation of the NOTCH pathway is associated with several cancer types, and appears to have an especially critical role in breast cancer. OE of either \textit{Notch1} or \textit{Notch4} receptors in the mammary glands of mice results in mammary tumors [75]. Additionally, \textit{in vitro} soft agar colony formation studies demonstrated that OE of \textit{NOTCH1/Notch1} or \textit{NOTCH4/Notch4} receptors was sufficient to transform both mouse and human mammary epithelial cells [76]. Subsequent studies showed that Notch signaling can regulate tumor proliferation in stress and non-stress conditions through upregulation of \textit{CyclinA}, \textit{CyclinB} and \textit{CyclinD1} [77,78]. Notch signaling can also promote tumor growth and resistance to endocrine therapies through downregulation of the tumor suppressor Pten, leading to activation of the Akt pathway [79]. Notch signaling has also been implicated in promoting EMT through direct
transcriptional activation of the EMT-inducing transcription factor SLUG [77,80]. Additionally, recently published data from our group has shown that NOTCH1 is critical for stem cell expansion and activity in vitro and tumor initiation in vivo across multiple subtypes of breast cancer [81].

Because NOTCH-ICD activity is synonymous with stemness, self-renewal and differentiation during development, subsequent studies focused on how these roles may influence breast cancer progression and metastasis. In the context of cancer stemness, multiple Notch transcription factors have been shown to play key roles. For instance, studies have shown that NOTCH1-ICD OE can increase tumorsphere formation in both human and murine mammary tumor cells, which is a distinct characteristic of cancer stem cells (CSC) [82]. Further, Simmons and colleagues showed that Notch1 ICD-mediated increases in breast cancer stem cell activity was due to direct transcriptional activation of the embryonic stem cell factor Nanog [83]. Recent studies from our group have supported a key role for NOTCH1 in promoting CSC activity in vitro and tumor initiation in vivo [81]. We show that the miRNA mir-106b-95 downregulates the E3 ubiquitin ligase, NEDD4L, to allow for stabilization of NOTCH1. This leads to increases in both tumorsphere formation and tumor initiation. Although less studied, NOTCH3 has been proposed to regulate tumor initiation and survival during the hypoxic response in breast tumors [84]. Additionally, NOTCH4 expression and activity is associated with the ESA+/CD44+/CD24low breast cancer stem cell population. This was confirmed through knockdown (KD) studies of NOTCH4, which result in decreases in the stemness-associated population and complete inhibition of tumor initiation in vivo [82].

The ability of cells to become resistant to treatment and outgrow at secondary sites are two key factors of tumor progression that can predict prognosis in breast cancer. It is hypothesized that stemness and self-renewal are indispensable for these processes [14]. As may then be anticipated, Notch transcription factors contribute to these critical aspects of tumor progression. For example, in the context of ER+ breast cancer, NOTCH4 has been shown to
drive resistance to targeted therapies by promoting cancer stem cell activity [85]. In the context of metastatic colonization, NOTCH3 is required for regulating the secretion of IL-6 from osteoblasts to promote outgrowth in the bone [85]. These mechanistic data are supported by the observation that OE of multiple NOTCH signaling components are associated with poor prognosis. For example, several studies have shown that higher levels of the JAGGED1 ligand and downstream regulator NOTCH1 correlate with poorer prognosis in breast cancer patients, with JAGGED1 also being associated with a more aggressive, basal phenotype and increased recurrence [86,87]. Overall, due to their roles in multiple aspects of tumor progression and metastasis, it is clear that targeting one component of the Notch pathway therapeutically will not be adequately effective. This is also supported by the fact that Notch receptors/transcriptional regulators have high sequence homology and can be activated by the same sets of ligands, and therefore can compensate for loss of other family members. Additionally, crosstalk between Notch signaling and other developmental pathways will make it difficult to target a metastatic process solely based on inhibition of the Notch transcription factor family.

The Six Family of Homeobox Transcription Factors

Role of Six Family in Mammary Stem Cells

The *sine oculis* (So/Six) family of genes were first discovered through both gain- and loss-of-function studies in drosophila, where they were found to be critical for eye development [88]. Subsequent reports then identified that these proteins functioned within a network involving multiple other developmental regulators including, *Eyes absent* (*Eya*), *Pax* and *dachshund* (*Dach*), allowing for precise regulation of important developmental processes including proliferation, stemness, differentiation and EMT [88]. The mammalian genome contains 6 homologous gene family members (*Six1-6*) to the single *So* gene found in Drosophila [88] and encode for homeobox transcription factors that can either activate or repress target genes based on context and interactions with specific co-factors. The protein structure of the *Six* family consists of two main functional domains: the Six domain, which has been shown to mediate
protein-protein interactions, such as binding with the transcriptional co-factors Eya and Dach, and the homeodomain, which is critical for DNA binding and specificity of transcriptional target genes [88]. Based on sequence similarity within these two domain regions, the Six family is further sectioned into 3 subgroups: Six1/2, Six4/5 and Six3/6 [88]. Interestingly, protein structure analysis identified that while Six2, Six4 and Six5 possess intrinsic activation domains, Six1, Six3 and Six6 do not, and thus require interactions with co-factors to regulate transcription [88]. Importantly, depending on the context and interaction with specific co-factors (mostly via the Six domain), Six family proteins can act as either transcriptional activators or repressors [88]. This underscores our understanding of their diverse roles in development and has provided critical insight into how their functions are hijacked during tumorigenesis and metastasis.

While much has been elucidated in terms of the developmental roles of the Six family during embryogenesis [88], very few studies have examined the roles of the Six family during mammogenesis and within the mammary stem cell population, and thus their roles in this aspect of development remain poorly understood. Still, studies by our group have provided some insight into the role of Six1 in the context of mammary gland development overall and mammary stemness. For example, our lab has shown that transplantation of Six1+/− embryonic mammary glands into the cleared fat pads of wildtype mice results in no significant morphological or functional abnormalities [89]. Interestingly, in response to this loss in Six1, increased expression of other family members such as Six2 and Six4 were observed, suggesting redundant functions of Six family members during mammary gland development [89]. Further, aberrant expression of Six1 in adult mammary glands also had no significant effects on mammary differentiation [89]. This result is especially surprising given SIX1 is often re-expressed in breast cancers and has been shown to be sufficient to initiate invasive carcinomas in mouse models [90], in part by promoting stemness properties such as anchorage-independent growth and self-renewal [91]. Taken together, these data indicated that although Six1 is capable of promoting breast cancer initiation, it may be dispensable for mammary gland development due to compensation
mechanisms by other Six family members. This could also somewhat explain why functional roles for the Six family members during mammary gland development may be difficult to identify. Although Six1 appears to not be required for mammary gland development overall, studies from our group have demonstrated that Six1 is sufficient to drive mammary stem cell phenotypes. For example, using an inducible, mammary-specific Six1 transgenic mouse model, our group showed that Six1 OE leads to expansion of the mammary stem cell population and increased stem/progenitor cell phenotypes, as measured by flow cytometry analysis of the mammary stem cell population and mammosphere formation [92].

While roles for Six1 in regulation of the mammary stem cell population have been identified, very few studies have examined the roles of other Six family members. However, data presented in this thesis has begun to suggest possible roles for Six2 in regulating the mammary stem cell population, as expression of Six2 is significantly increased in fetal mammary stem cells compared to their more differentiated, adult mammary epithelial cell counterparts (Chapter III and [93]). In addition, although expression of other Six family members (namely Six3, Six4 and Six5) have been found within the mammary gland [89,94], functional analyses focusing on other Six family members within mammary stem cells remains to be elucidated. Therefore, a more targeted analysis of expression and function of the Six family is needed to determine if they indeed play critical roles in the mammary stem cell population.

Role of Six Family During Breast Cancer Metastasis

Many studies have established that aberrant expression and/or function of the SIX family can lead to increases in both tumor progression and metastasis. In the context of breast cancer metastasis, the most widely studied SIX family member is SIX1. For instance, previous reports by our group have established that re-expression of SIX1 is significantly associated with tumor progression and metastasis [95], and that this is accomplished through multiple mechanisms including regulation of EMT, migration, invasion, survival and stemness [31,91,92,96–98]. Our group has demonstrated that, similar to its roles in development, Six1 plays a critical role in the
maintenance and survival of breast cancer stem cells, as measured by anchorage-independent growth in soft agar and tumorsphere formation \textit{in vitro} and tumor initiation at decreasing dilutions \textit{in vivo} \cite{90,91}. Importantly, studies from our group also showed that OE of SIX1 in ER+ MCF7 cells leads to increased metastatic capability, and that this was in part through upregulation of a specific cluster of miRNAs miR-106b, 93 and 25 and subsequent downstream activation of TGFβ signaling \cite{97,99}. Another report from our group showed that SIX1 could promote metastasis through cooperation with the angiogenesis factor, VEGFC to increase lymphangiogenesis and metastatic spread \cite{96}. Intriguingly, a recent report from our lab has also demonstrated that SIX1 functions downstream of the EMT-inducing transcription factors, TWIST and SNAIL, in metastatic cells to promote aggressive properties such as migration and invasion, in non-EMT, non-metastatic breast cancer cells through cooperation with HH signaling and GLI transcriptional activity \cite{31}. Taken together, these data underscore the diverse and critical functions of SIX1 in promoting cancer stemness and metastasis and highlight its potential as a therapeutic target to inhibit tumor progression.

While expression data from patient samples suggests other Six family members may be associated with tumorigenesis in multiple cancer types \cite{88}, few studies have characterized their functional roles in breast cancer metastasis. However, more recent studies from our group have identified that Six2, a closely related family member to Six1, not only appears to compensate for Six1 loss during mammary gland development \cite{89}, but also compensates for Six1 loss in mediating breast cancer metastasis \cite{96,100}. Specifically, our group showed that although Six1 KD results in a dramatic decrease in the metastatic capability of a highly aggressive mouse mammary carcinoma cell line model (66c14), the very few metastatic Six1 KD cells that survived and were retrieved from the lungs of mice had increased expression of Six2 \cite{96}. These data suggested that Six2 compensates for Six1 in mediating metastasis. Interestingly, subsequent studies showed that although Six2 KD in highly metastatic mouse mammary carcinoma cells did result in decreased metastasis overall \textit{in vivo}, phenotypic \textit{in vitro} assays demonstrated that Six2
KD had no significant effects on proliferation or lymphangiogenesis, functions associated with Six1-mediated metastasis. Moreover, Six2 OE in another mouse mammary carcinoma cell line that can reach secondary sites but cannot efficiently outgrow (4T07) resulted in no significant changes in migration or adhesion [100]. Instead, Six2 was shown to promote anoikis resistance and anchorage-independent growth in soft agar (two properties known to be important for initial survival and outgrowth at secondary sites [101,102]). Importantly, tail-vein injection experiments to test late-stage metastasis specifically showed that Six2 OE in 4T07 cells also lead to increased metastatic colonization and that this was in part dependent on downregulation of E-cadherin [100]. Taken together, these data determined that in contrast to Six1, which has been shown to regulate multiple aspects of the metastatic cascade, the functions of Six2 appear to be more specific to promoting the later stages of metastasis, specifically metastatic outgrowth and colonization [100].

Building on this work, data from this thesis suggests that similar to its roles in development, where it is critical for maintaining kidney progenitor populations and inhibiting differentiation [104], Six2 can promote metastatic colonization through regulation of stemness associated phenotypes in breast cancer cells. Specifically, I show that SIX2 is a critical mediator of cancer stem cell-associated phenotypes including tumorsphere formation in vitro and tumor initiation in vivo [93]. Mechanistically, I also demonstrate that Six2-mediated stemness phenotypes and late-stage metastasis is dependent on direct transcriptional upregulation of the master stemness factor, Sox2. This was supported by KD of Sox2 downstream of Six2 OE resulting in decreased stemness-associated phenotypes in vitro and metastatic outgrowth in vivo [93]. Thus, given the parallel roles and cooperation between Six1 and Six2 in mediating metastasis, additional studies examining their functional relationship could lead to more effective therapeutic strategies to inhibit metastatic burden and increase overall survival. Indeed, in this thesis I will discuss how using CRISPR/Cas9-mediated KO and KI strategies will allow for the identification of the distinct and overlapping functions of Six1 and Six2 in order to
better understand how they regulate metastasis and more effectively inhibit their functions to decrease the metastatic capabilities of breast cancer cells.

The prominent roles of SIX1 and SIX2 in mediating metastasis are also supported by patient expression data related to prognosis, where both family members are significantly associated with increased metastatic indicators and poorer survival. For instance, upregulation of SIX1 mRNA is significantly correlated with up to 90% of breast cancer metastatic lesions [105]. Additionally, SIX2 mRNA expression is more associated with patients diagnosed as M1 (metastatic disease at diagnosis) compared to those diagnosed as M0 (local disease only at diagnosis) [100]. Further, data in this thesis identified that SIX2 is also more associated with aggressive subtypes of breast cancer, such as the triple-negative subtype, compared to all other subtypes and that patients with high expression of a Six2-mediated gene signature have decreased distant metastasis-free survival, relapse-free survival and time to distant metastases compared to low expression of the signature [93]. Lastly, while few studies have investigated the functional consequences of the entire Six family in mediating metastasis, a recent report by Xu et al examined their expression in breast cancer patient samples to determine whether they could be correlated with patient outcomes. For example, they showed that SIX3 expression is significantly linked to high histological grade, with SIX4 being correlated with both poor differentiation and poor overall prognosis [106]. In addition, while expression of SIX5 was not shown to be associated with prognosis in this study, SIX6 expression was significantly associated with poor prognosis, albeit specifically in luminal breast cancer patients [106]. Still, these datamining analyses only suggest roles for other Six family members in breast cancer progression and metastasis. More functional studies are needed to definitively determine whether other Six family members, outside of SIX1 and SIX2, mediate breast cancer metastasis and whether they would also be attractive therapeutic targets.
Pluripotency Factors in Mammary Stem Cells and Breast Cancer Metastasis

Pluripotency Factors as Regulators of Mammary Stem Cells

The above sections underscore that key developmental stem pathways are critical in both tumor initiation and metastasis, and suggest that other stem regulators, such as key pluripotency factors, may also contribute to tumor initiation and progression. The transcription factors SOX2, OCT4, NANOG, KLF4 and cMYC are known as master regulators of pluripotency and stemness due to their critical roles in regulating target genes that are important for self-renewal, differentiation and survival [107]. Similar to embryonic development, where these core factors are essential for the function and maintenance of embryonic stem cells, studies have also identified important roles for pluripotency factors in regulating mammary stem cells over the course of mammary gland development and differentiation, with an especially strong association with the expansion of breast tissue during pregnancy and lactation [108]. For instance, Sox2, Nanog and Oct4 are all expressed in both luminal progenitors and adult mammary stem cells, and their expression decreases as the cells differentiate [109]. More specifically, upregulation of master pluripotency factors such as SOX2, NANOG, OCT4 and KLF4 is associated with the expansion of the mammary stem cell populations during pregnancy and lactation [110]. Their expression correlates with the onset of extensive remodeling and differentiation of the breast, in which integration of multiple cell types and lineages is needed for transformation of the breast into a milk-secretory organ [110]. Studies from Hassiotou and colleagues examined the expression of pluripotency factors during lactation and found that SOX2, NANOG, OCT4 and KLF4 were highly expressed in isolated breast milk-derived stem cells compared to non-lactating, non-pregnant breast cells [111]. Interestingly, these specialized cells were able to self-renew in anchorage-independent conditions, had multiple lineage potential and were also capable of differentiating into all three germ layers. Additionally, these cells could transport through the bloodstream to different organs, where they were shown to integrate and differentiate into functional cells of multiple target organs [112].
Functional studies support the above correlative studies indicating that pluripotency factors play a key role in mammary stem cell biology. For example, Wang and colleagues demonstrated that Sox2 is critical downstream of Lgr-mediated Wnt signaling to promote mammary stem cell activity [113]. Importantly, they identified Sox2 as a transcriptional target of the Lgr4/Wnt/β-catenin/Lef1 pathway and demonstrated that Sox2 OE could rescue the decrease in mammary stem cell repopulation capacity present in Lgr−/− mice [113]. Another pluripotency factor, cMyc, has been shown to be essential for mammary stem cell activity. Moumen and colleagues showed, using limiting-dilution and serial transplantation assays, that deletion of cMyc from mammary basal epithelial cells decreases self-renewal. Interestingly, although stimulation with Estrogen and Progesterone downstream of cMyc loss could partially rescue repopulation defects in vivo, this rescue was attributed to compensation by N-Myc. However, when cMyc-KO glands were transplanted a second time, the resulting mammary epithelium completely lacked stem and progenitor cells and failed to repopulate cleared fat pads, suggesting that cMyc regulates both the maintenance and self-renewal of mammary stem cells, and that N-Myc could not compensate for all activities of cMyc [114]. Overall, these studies have provided critical insight into the molecular mechanisms that underlie normal mammary stem cell functions and have been instructive into how this specialized population becomes susceptible to transformation during tumorigenesis and metastasis.

**Pluripotency Factors as Regulators of Breast Cancer Metastasis**

Although pluripotency factors have a well-established role in the activity and function of breast cancer stemness including transformation in vitro and tumor initiation in vivo [115–118], more recent studies have identified critical roles during tumor progression and metastasis, where they have been shown to regulate several critical properties of metastatic tumor cells including stemness, plasticity and differentiation. For instance, studies by Weinberg and colleagues demonstrated that Sox9, a pluripotency factor that has been shown to be critical for regulating the MaSC phenotype [119,120], regulates not only the tumorigenic, but also the
metastatic colonization abilities of human breast cancer cells [120] (Fig 1.2). Interestingly, Guo and colleagues found that Sox9, in cooperation with the stemness and EMT-inducing transcription factor Slug, were not only critical for the maintenance of MaSCs in normal mammary epithelial cells (MECs), but also that Sox9 and Slug cooperativity was sufficient to reprogram differentiated luminal epithelial cells into functioning MaSCs, as measured by organoid formation in vitro, competitive reconstitution assays in vivo and limiting dilution assays in vivo [120]. Further, using MCF7ras cells (transformed with RAS to allow for growth in the absence of estrogen), which could reach the secondary site and form micrometastases, but were unable to outgrow to macrometastases, they showed that combined induction of Sox9 and Slug resulted in a significant increase in outgrowth as measured by an increase in the number macrometastases compared to non-induced control or singly-induced cells [120]. Additionally, KD of either Sox9 or Slug in highly metastatic MDA-MB-231 cells resulted in decreased macrometastasis formation after tail vein injection [120]. Taken together, these data indicate that tumor cells, specifically CSCs, can hijack the stemness and differentiation functions of pluripotency factors to regulate multiple aspects of late-stage metastasis by promoting survival, self-renewal and metastatic outgrowth. Importantly, these data above also suggest that targeting MaSC factors to inhibit core CSC characteristics such as plasticity, stemness and differentiation may be an effective therapeutic strategy to inhibit metastatic burden and improve survival (Fig 1.3).

Additionally, a recent study by Cho and colleagues demonstrated that ubiquitin-mediated stability of OCT4 was critical for metastasis of TNBC cells in vivo [121]. Intriguingly, Lu and colleagues have shown that a transgenic mouse model of inducible OE of Nanog, in combination with Wnt-1 OE, was sufficient to induce mammary tumorigenesis and metastasis [122]. Interestingly, studies have shown that pluripotency factors can work in concert to promote metastatic phenotypes. For instance, Wang and colleagues showed that simultaneous OE of Oct4 and Nanog increased, whereas combined KD of Oct4 and Nanog decreased, EMT and
invasion of breast cancer cells [123]. These data suggest that pluripotency factors do not necessarily function in isolation to influence metastatic phenotypes. Thus, future studies should always consider the effects of the entire network of core pluripotency factors before determining any one factor’s role in metastasis.

Similar to its roles in promoting metastasis across multiple cancer types [124], cMYC has also been found to be a core regulator of metastatic progression in breast cancer [125,126], where it has mostly been shown to be required for migration, invasion and angiogenesis. Interestingly, cMYC is also critical for organ-specific metastatic colonization of breast cancer in that it regulates multiple cellular processes within the brain microenvironment to promote metastatic outgrowth including macrophage infiltration, GAP junction formation between tumor cells and astrocytes and invasive growth [127]. However, given that more recent studies have also identified metastasis inhibitory roles for cMyc [128], any clinical strategy to modulate its function must consider how context, timing and tissue type may influence drug responses.

Like cMYC, SOX2 is a key regulator of breast cancer metastasis. Studies have identified roles for SOX2 across multiple aspects of the metastatic cascade including EMT, migration, invasion, anoikis resistance and angiogenesis [129–132]. However, its most prominent role in tumor progression involves regulation of stemness and differentiation. This is especially evident in breast cancer, where SOX2 has been widely used to distinguish breast cancer stem cells from the bulk tumor population [133], and where its OE is associated with more aggressive subtypes of breast cancer [93]. Functionally, SOX2 has been shown to be essential for the maintenance and survival of cancer stem cells across multiple tumor types including Glioma [134], Head and Neck squamous cell carcinoma [135], osteosarcoma [136] and breast cancer [118]. Because of its role as a master pluripotency factor during embryonic development, SOX2 is controlled at the transcriptional, translational and post-translational levels to ensure proper expression and function. Since these mechanisms are also used by tumor cells to promote metastasis, it provides multiple opportunities to target and inhibit its functions to inhibit tumor
progression. For instance, recent examination of human breast cancer patient data has suggested that the threonine 116 phosphorylation site of SOX2 may be an attractive therapeutic target to inhibit its function [137]. Specifically, Gupta and colleagues showed that not only was increased phosphorylated-Sox2\textsuperscript{T116} associated with increased tumorigenicity in multiple breast cancer cell lines, but that it was also mostly found within breast cancer stem cells, suggesting that this phosphorylation state of SOX2 is important for breast cancer stem cell identity [137]. Consequently, when examining a breast cancer cohort, they found that phosphorylated-Sox2\textsuperscript{T116} levels correlated with increased histological tumor grade and invasion [137]. Collectively, these data provide critical insight into how we might target SOX2 to limit its functions and inhibit metastasis.

Recent data also suggests that the embryonic developmental roles of master pluripotency factors such as OCT3/4, SOX2 and NANOG in regulating stemness, differentiation and reprogramming, are also repurposed by tumor cells to promote multiple aspects of tumor progression, including metastatic initiation. Using intravital imaging, Liu and colleagues demonstrated that circulating tumor cells (CTCs) upregulate the breast cancer stem cell marker CD44 to promote cell clustering, increased metastasis and decreased overall survival [138]. Interestingly, it was also found that these CTC clusters had increased expression of stemness factors including OCT3/4 and that CD44 KO resulted in reduced OCT3/4, although the functional role of OCT3/4 downstream of CD44 was not examined. Nonetheless, these data suggested that CD44 depletion reduces stemness properties and that this could be mediated by OCT3/4 [138]. Additionally, SOX2 and NANOG have been shown to mediate survival and dormancy in DTCs and residual cancer cells via regulation of a genetic quiescence program downstream of the nuclear orphan receptor and cell fate regulator, NR2F1 [139]. Interestingly, in a metastatic dormancy model of squamous cell carcinoma, Sosa and colleagues showed the effects of SOX2 and NANOG induction resulted in no changes in tumor-initiating capacity, as measured by tumorsphere formation \textit{in vitro}, suggesting that the observed effects of SOX2 and
NANOG in this model are independent of their prominent roles in pluripotency and self-renewal and that the metastatic roles of SOX2 and NANOG are context dependent [139]. However, KD of NR2F1 allowed for DTCs to switch to a proliferative phenotype and resulted in corresponding downregulation of both SOX2 and NANOG [139]. Thus, while several associations between the regulation of stemness, pluripotency and differentiation and metastasis initiation have been identified, a better mechanistic understanding of the functional role of these properties and how they influence metastatic outgrowth remain to be elucidated. Importantly, a more recent analysis of breast cancer-associated DTCs showed that patients with ≥50% NR2F1^{high} DTCs had increased intervals of distant disease-free survival compared to patients with NR2F1^{low} DTCs [140]. Given that these data suggest that, similar to other cancer types, NR2F1 can also be used to predict prognosis in breast cancer, it would be interesting to pursue studies examining whether pluripotency factors such as SOX2 and NANOG are involved in NR2F1-mediated quiescence.

Many studies have examined how increased levels of pluripotency factors influence prognostic indicators such as disease-free survival, resistance and overall survival [141]. Studies examining either expression of OCT4, SOX2, NANOG, KLF4, or cMYC singly, or combined expression of OCT4 and NANOG specifically, showed that higher levels of mRNA expression in breast cancer patient samples across multiple subtypes is significantly correlated with decreased overall survival [141]. Additionally, higher levels of SOX2 mRNA are particularly predictive of decreased resistance- and disease-free survival [141]. Intriguingly, in addition to being associated with multiple prognostic markers including decreased overall survival, cMYC is also more highly expressed in patients that present with metastatic disease at diagnosis, suggesting that cMYC is a critical mediator of metastatic burden [142]. However, several studies have also identified that pluripotency factors can also be associated with improved prognosis [141]. These confounding data could be due to the influence of tumor heterogeneity, where it has been shown that uniform expression of factors does not exist within the bulk tumor.
Additionally, since pluripotency factors are more associated with the rare cancer stem cell population, their level of expression may be masked when assessing expression levels of the bulk tumor. Therefore, in the future it may be more effective to assess how a single gene may impact prognosis by examining it within a panel of genes that are representative of a specific tumorigenic process/phenotype. Further, if a putative gene is a transcription factor, it may be better to consider the expression of a panel of downstream targets to determine its prognostic value and better evaluate its potential as a therapeutic target.

Substantial progress has been made towards understanding the molecular underpinnings of developmental pathway networks and how they function to ensure proper mammary gland development. It is now understood that during breast cancer progression and metastasis, these developmental programs are activated out of context to promote processes such as proliferation, EMT, and most importantly for this thesis, stemness. Modulation of these programs is associated with drug response and clinical prognosis, supporting the continued examination of the underlying biology to develop therapeutic strategies. Importantly, this thesis highlights the similarities between development and breast cancer progression by identifying a novel role for the developmental transcription factor, SIX2, in regulating stemness properties in tumor cells, leading to increased survival and self-renewal to promote metastatic colonization. The goal of this work is to continue to explore how developmental regulators and networks promote breast cancer metastasis, in order to better understand the integration of context-dependent crosstalk between multiple developmental pathways and regulators. The ultimate goal of this work is to use such information to identify vulnerabilities in tumor cells to inhibit metastatic burden and increase patient survival.
CHAPTER II
MATERIALS AND METHODS

Cell Lines and Culture

The 4T07 and 66cl4 mouse mammary carcinoma cells were generously provided by Dr. Fred Miller[143]. Stable Six2 OE and KD 4T07 and 66cl4 cell lines were generated and grown in media as described previously[100]. MDA-MB-231 cells have been cultured long-term in the Ford laboratory, and were grown in MEM/EBSS, FBS, HEPES, non-essential amino acids, insulin, pen/strep and L-glutamine. Authentication of the human cell line MDA-MB-231 was performed using STR analysis by the University of Colorado Cancer Center Tissue Culture Shared Resource in March of 2017. All cells were regularly monitored for mycoplasma contamination every 3 months and only mycoplasma free lines were used for studies. All cell lines were cultured at 37°C in HEPA-filtered humidified air in 5% CO2. Stable Six2 KD in the MDA-MB-231 cells was performed by transduction of two different lentiviral shRNAs, or with a non-targeting (NT) shRNA. After transduction, cells were selected for 1 week using 3μg/ml of puromycin to obtain stable KD of Six2. Six2 levels in KD and OE cells were examined using RT-qPCR and Western blot analysis. Transient KD of SIX2 and Six2 was performed using ON-TARGETplus SMARTpool siRNA constructs (Table 2.1). A non-targeting siRNA pool was used as a negative control. KD levels were assessed 72h post transfection via RT-qPCR. For inducible Six2 expression, 4T07-Empty Vector (4T07-EV) and 4T07-Six2 (4T07-Six2i) were treated with Doxycycline or vehicle control for 24h and 48h. After each timepoint, cells were

harvested for Western blot and qRT-PCR analysis. To generate Six1 and Six2 knockout cell lines, Lipofectamine Transfection Reagent (ThermoFischer Scientific, 13778030) was used to transfect the pSpCas9(BB)-2A-GFP plasmid containing gRNAs targeting either Six1 or Six2 into 66cl4 cells. After 48h, cells were flow sorted based on GFP positivity into single cell clones and screened via western blot analysis for knockout. In order to introduce a 3X-FLAG tag to the C-terminal ends of both Six1 and Six2, similar conditions for generation of KO lines were used, with the addition of donor DNA templates containing 3X-FLAG and Neomycin flanked by homology arms consisting of either the C-terminal end of Six1 or Six2. After transfection, cells were selected with 500 ug/ml of Neomycin.

RNA-sequencing

4T07 CTL and Six2 cells were plated at 3x10^5 cells/ml in 10cm plates and allowed to attach and grow for 48h. Cells were collected and lysed by removing media, rinsing with PBS and adding 1ml of TriReagent directly to the plate. After vortexing to ensure homogeneity, total RNA was isolated using the Zymo Research RNA Isolation kit (Light Labs, R2053). RNA samples were evaluated for purity and concentration using a spectrophotometer and submitted to the Genomics and Microarray Core Facility at the University of Colorado Denver Anschutz Medical Campus. Sequencing sample preparation was performed using standard Illumina HiSeq protocols and reagents. Directional mRNA Seq was performed using Illumina HiSeq 4000 HT Mode 1X150 cycles.

RNA-sequencing and Data Analysis

The read quality and mapping were done as described previously[144]. Significantly differentially expressed genes were determined after using DESeq2 in R with batch correction, and a significance cutoff of adjusted p value < 0.1. MA plots, PCA plots and heatmaps were generated using Matplotlib from Python plotting library. To find overlap between significantly differentially expressed genes and stemness-associated gene signatures, DESeq genes were input into the StemChecker database (http://stemchecker.sysbiolab.eu).
### Table 2.1: List of Oligos Used for qRT-PCR, shRNA and siRNA Experiments

<table>
<thead>
<tr>
<th>Locus</th>
<th>Source</th>
<th>Primers and shRNA target sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIX2</td>
<td>This paper</td>
<td>F CCAATTCTACAGCAGGACAC &lt;br&gt; R GGCTGGATGATGAGTGGTCT</td>
</tr>
<tr>
<td>SIX2</td>
<td>This paper</td>
<td>F TCACCACCACGCAAGTCAGCAA &lt;br&gt; R ACCGACTTGCCACTGCCATTGA</td>
</tr>
<tr>
<td>Six2</td>
<td>This paper</td>
<td>F AGGATAAGTACACGCTGCCC &lt;br&gt; R TAACTGTCCATGCGCTGGTT</td>
</tr>
<tr>
<td>SOX2</td>
<td>This paper</td>
<td>F CATGAAGGAGCACCCGGATTA &lt;br&gt; R CGGGAAGCGTGTACTTATCCT</td>
</tr>
<tr>
<td>Sox2</td>
<td>This paper</td>
<td>F TGAGTGTGGGTCTTCCTGGT &lt;br&gt; R CAGGCATTGATGAGGCGTTC</td>
</tr>
<tr>
<td>NANOG</td>
<td>This paper</td>
<td>F CTGCTTATTCAGGACAGCCC &lt;br&gt; R TCTGTTTCTTGACCGGGACC</td>
</tr>
<tr>
<td>Nanog</td>
<td>This paper</td>
<td>F TGAGTGTGGGTCTTCCTGGT &lt;br&gt; R CAGGCATTGATGAGGCGTTC</td>
</tr>
<tr>
<td>NME7</td>
<td>This paper</td>
<td>F AGTGGCACGCACAGATGCTT &lt;br&gt; R CAACTCCATTTCTCTGGCCGCA</td>
</tr>
<tr>
<td>Nme7</td>
<td>This paper</td>
<td>F TTCTCCCAAGCGGTGGAGTTCT &lt;br&gt; R ATGCAACGCCGGCTGCTAAA</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>This paper</td>
<td>F ATGCTGTGAGCCTCCTTGCACA &lt;br&gt; R CGGCCATGTGGTTTGTGCATTT</td>
</tr>
<tr>
<td>Pik3ca</td>
<td>This paper</td>
<td>F TGGCGACGACTTACGGCAAGAT &lt;br&gt; R ACCCACACAGTCCCCAATGGAT</td>
</tr>
<tr>
<td>PKN3</td>
<td>This paper</td>
<td>F TGGTACCGCAGGGACTGCTTTT &lt;br&gt; R TCCTGGCCTCTGCGTTTAGAGA</td>
</tr>
<tr>
<td>Pkn3</td>
<td>This paper</td>
<td>F TCTGTCAGTGCAAGGGCTCGAA &lt;br&gt; R TGCCAGTTGGTGGTCCTGAAGA</td>
</tr>
<tr>
<td>COL18A1</td>
<td>This paper</td>
<td>F ACCATCTTCAGAGCGCATCAGC &lt;br&gt; R TGGCATCTGGCCCAAAGACGTA</td>
</tr>
<tr>
<td>Col18a1</td>
<td>This paper</td>
<td>F ACTTTCAGCCAGTGCTCCACCT &lt;br&gt; R AGCACTGGAAATCTGCTCCACG</td>
</tr>
<tr>
<td>NBEA</td>
<td>This paper</td>
<td>F AGCTCTGCCATGCACCTGTGTT &lt;br&gt; R AGTGTTGGCTGCCACATGGGTT</td>
</tr>
<tr>
<td>Nbea</td>
<td>This paper</td>
<td>F ACAGTGCCTGCGGCTAGTTTGT &lt;br&gt; R TCCTGAGGTTTGCTGCTTCCGT</td>
</tr>
<tr>
<td>TGIF2</td>
<td>This paper</td>
<td>F TCGCGCTCCGCACAAAGTTTA &lt;br&gt; R ATCTTCACCGACTCCTTGGGCA</td>
</tr>
<tr>
<td>Tgif2</td>
<td>This paper</td>
<td>F TGTTGACTGGCTTGGAGTGCCT &lt;br&gt; R TGCACAGCAAGGGATGGGTGAA</td>
</tr>
<tr>
<td>PPIB</td>
<td>This paper</td>
<td>F GGTCCGTCTTCTTCCTGCT &lt;br&gt; R CGGCCTACATCTTCATCTCC</td>
</tr>
<tr>
<td>Ppib</td>
<td>This paper</td>
<td>F AGCAAGTTCCATCGTGTCATC &lt;br&gt; R CTCTTTCCTCCTGTGCCATC</td>
</tr>
<tr>
<td>shRNA SIX2 KD1</td>
<td>Sigma Aldrich Cat# TRCN0000433131</td>
<td>CCGGGACGCCAGACCACTCATATCATCTCGAGGATGATGAGTGGTCTGGCGTCTTTTTT</td>
</tr>
<tr>
<td>shRNA SIX2 KD2</td>
<td>Sigma Aldrich Cat# TRCN000014684</td>
<td>CCGGCAACGAGAACTCCAATTCTAACTCGAGTTAGAATTGGAGTTCTCGTTGTTTTTT</td>
</tr>
<tr>
<td>shRNA SIX2 KD1</td>
<td>Wang et al. 2014; Clone ID: V3LMM_459347</td>
<td>TGAGCGAGGAAGCCAGCGG</td>
</tr>
<tr>
<td>shRNA SIX2 KD2</td>
<td>Wang et al. 2014; Clone ID: V2LMM_83091</td>
<td>TGAGTATGGAGTCCTGCAG</td>
</tr>
<tr>
<td>shRNA SOX2 KD1</td>
<td>Sigma Aldrich Cat# TRCN0000416106</td>
<td>CCGGCAAAGAGATACAAGGGAATTGCTCGAGCAATTCCCTTGTATCTCTTTGTTTTTTT</td>
</tr>
<tr>
<td>shRNA SOX2 KD2</td>
<td>Sigma Aldrich Cat# TRCN0000420955</td>
<td>CCGGACCAATCCCATCCAAATTAACCTCGAGGTTAATTTGGATGGGATTGGTTTTTTT</td>
</tr>
<tr>
<td>ChIP Col18a1 F</td>
<td>This paper</td>
<td>CAGAAAGCATGGGACAGAAGA</td>
</tr>
<tr>
<td>ChIP Col18a1 R</td>
<td>This paper</td>
<td>CCAACCTGCGTAAACCTAGAA</td>
</tr>
<tr>
<td>ChIP Nbea F</td>
<td>This paper</td>
<td>AGGAAATACGAAAGCCACCTAT</td>
</tr>
<tr>
<td>ChIP Nbea R</td>
<td>This paper</td>
<td>AGGCTAATCATTACCAGCTAACA</td>
</tr>
<tr>
<td>ChIP Nme7 F</td>
<td>This paper</td>
<td>TGAGTGATAGCAGCTTTGGG</td>
</tr>
<tr>
<td>ChIP Nme7 R</td>
<td>This paper</td>
<td>CCAAGTCCACGGAAGAAAGT</td>
</tr>
<tr>
<td>ChIP Pik3ca F</td>
<td>This paper</td>
<td>CCCTCCATGGGATGTGTTTAT</td>
</tr>
<tr>
<td>ChIP Pik3ca R</td>
<td>This paper</td>
<td>AGTTGGGATGGTAAGCAAGAG</td>
</tr>
<tr>
<td>ChIP Pkn3 F</td>
<td>This paper</td>
<td>AGCGTCAGTGGCTTGAAA</td>
</tr>
<tr>
<td>ChIP Pkn3 R</td>
<td>This paper</td>
<td>GTCGTCGTCATCATCATCTTCTT</td>
</tr>
<tr>
<td>ChIP Tgif2 F</td>
<td>This paper</td>
<td>CTCTTTGGGCTGATGTTCCT</td>
</tr>
<tr>
<td>ChIP Tgif2 R</td>
<td>This paper</td>
<td>CAGCTTCATGGTAGAAGGTAAGA</td>
</tr>
<tr>
<td>T27G F</td>
<td>This paper</td>
<td>TTAGAGTCTCTGCATTACTAGGAGGGAAAGGATGCC</td>
</tr>
<tr>
<td>T27G R</td>
<td>This paper</td>
<td>CGAGGTTCTCCCTTAAAGTACGAGAAGACTCTCA</td>
</tr>
</tbody>
</table>

### References

- Wang et al. 2014: Clone ID: V3LMM_459347
- Wang et al. 2014: Clone ID: V2LMM_83091
- Sigrist et al. 2006: TRC00000433131
- Sigrist et al. 2006: TRC000014684
- Sigrist et al. 2006: TRC0000146106
- Sigrist et al. 2006: TRC00000420955
- Dharmacon
- Dharmacon
- Dharmacon
- ON-TARGETplus Non-targeting Pool (D-001810-10-05)
- ON-TARGETplus Human SOX2 siRNA (L-011778-00-0005)
- ON-TARGETplus Human SIX1 siRNA (L-020093-00-00010)
- ON-TARGETplus Mouse SIX1 siRNA (L-092788-01-00020)
- ON-TARGETplus Mouse SOX2 siRNA (L-035689-00-0005)
- ON-TARGETplus Human MAX1 siRNA (L-020093-00-00010)
- ON-TARGETplus Human MAX1 siRNA (L-042788-01-00020)
Significantly enriched gene signatures were determined using a Hypergeometric test and adjusted by Bonferroni correction. Heatmaps were generated by converting reads per kilobase per million (RPKM) values to z-scores and using ComplexHeatmap and Circularize library R packages to visualize z-score normalized gene expression. Genes were grouped by hierarchical clustering of z-score normalized gene expression. The RNA-seq data was deposited into the GEO repository with the accession number GSE123489.

**GSEA**

Gene-Set Enrichment Analysis (GSEA) was performed by utilizing the GSEA Preranked Module of the GenePattern Webserver[145]. All significantly expressed genes were ranked by adjusted Log$_2$ fold change values (Six2 OE vs CTL) and tested against the C2: Curated gene sets and C5: Gene Ontology gene sets to test for enrichment.

**Western Blot**

Whole cell and nuclear lysates were isolated as done previously[100]. Protein concentration was determined using the Lowry Protein assay (BioRad). After gel electrophoresis of protein in 10% polyacrylamide gels, proteins were transferred to a polyvinylidene difluoride membrane at 100v for 1h 30min, blocked in 5% Milk-TBST and incubated at 4°C in primary antibodies overnight (see Table 2.2 for list of primary antibodies). Blots were imaged using either Film radiography or the OdysseyFc imaging system after HRP substrate chemiluminescence incubation and detection.

**RT-qPCR**

RNA was isolated using the Zymo Research RNA isolation kit (Light Labs, R2053). cDNA was made using the QuantaraBio cDNA synthesis kit (VWR Scientific, 95047-100). PCR was performed using the SYBR green master mix per manufacturer’s instructions and protocols (Qiagen, 330500). Primers used for this study are contained in Table 2.1 (IDT). Gene expression was normalized to peptidylpropyl isomerase B (cyclophilin B) (Ppib/ppib) mRNA expression and fold-change is relative to the CTL, NS or NT cells.
Significance was determined by t-test or one-way anova on means ± SEM of triplicate samples for a representative experiment (n=3).

**Tumorsphere Formation**

Tumorsphere assays were performed as previously described[91]. Briefly, cells were trypsinized and plated at either 300, 1000 or 20,000 cells/well in 6-well, ultra-low attachment plates (Corning) in 2 mls of serum-free DMEM/F12 media (Hyclone), supplemented with 20ng/ml bFGF (BD Biosciences), 20ng/ml EGF (BD Biosciences), 4 μg/ml heparin (Sigma), penicillin-streptomycin (Hyclone), and B27 (Gibco) for all cell lines tested. 500 μl of media was added every 3-4 days. On day 7, all cells/spheres were collected, washed with PBS, digested using 0.05% trypsin in HBSS with 0.2g/L EDTA (Fisher), and single cells were plated as above to perform secondary tumorsphere assays. After an additional 7 days, tumorspheres were imaged at 4x magnification and counted by the Incucyte Zoom Live-Cell Analysis System.

**Flow Cytometry**

Flow cytometry analysis was done as previously described[91]. Briefly, cells were incubated for 30 min in either CD24-biotin/CD49f-APC or Epcam-FITC/CD44-APC and in the case of CD24, were additionally incubated in Strep-FITC or Strep-V450 secondary for 30 min (Table 2.2). Positive control ultracomp beads were also stained and used as a positive control for consistent gating. Flow cytometry was performed by submitting samples to the Flow Cytometry Core Facility at the University of Colorado Anschutz Medical campus or using a C6 Flow Cytometer. Single-stained compensation controls (UltraComp Beads, eBioscience) were stained with individual antibodies (EpCAM-FITC, CD44-APC, CD24-biotin-strep-FITC, CD24-biotin-strep-V450 or CD49f-APC) to define gating for each cell-surface marker. Gates were set at the lowest edge of fluorescence for the positive stain populations relative to the negative stain populations.
Additionally, negative control (unstained) cell populations were checked to make sure they overlapped <1% with the single or double positive quadrants after gating with single-stained control beads. Flow cytometry analyses were performed using the FlowJo Analysis Software.

**Animal Studies**

Tail vein injections were performed in 6-8 week old female immunodeficient NOD-scid-gamma (for MDA-MB-231) and immunocompetent BALB/c mice (for 4T07) (Jackson Laboratories). For the MDA-MB-231 cells, 4 x 10^5 cells in 100ul of HBSS were injected. For the 4T07 cells, 1 x 10^6 cells in 100ul of DMEM medium were injected. Both the MDA-MB-231 and 4T07 cells were luciferase tagged to enable imaging of metastases in vivo. IVIS imaging was used to image the mice as described previously[100]. Quantification of luciferase signal was determined using Living Image Software. Limiting dilution assays were performed by injecting 5x10^5, 2.5x10^5 or 5x10^4 luciferase-tagged MDA-MB-231 cells in 100ul of a 1:1 Matrigel/PBS (Corning, 354234) underneath the nipple of 4th mammary fat pad of nude mice. Tumor formation was monitored weekly via IVIS imaging. TIC capacity was determined using the ELDA software website application (http://bioinf.wehi.edu.au/software/elda/) with a 95% confidence interval setting. All animal work was approved and performed in compliance with the Institutional Animal Care and Use Protocol at the University of Colorado Anschutz Medical Campus.

**Luciferase Assays**

Cells were plated in 12-well plates to achieve equal confluence (4T07: 4x10^4 cells/ml, 66cl4: 7.5 x10^4 cells/ml and MDA-MB-231: 1 x10^5 cells/ml) and then co-transfected with the indicated firefly-luciferase constructs (pGF-mSox2-ONL and pGF-mSox2-T27G-ONL) and renilla-luciferase constructs (pCNL-N1_3xNLS) using lipofectamine 2000 according to manufacturer instructions (Thermo Fisher Scientific). After 72h, cells were collected and analyzed for luminescence using the Dual-luciferase Reporter Assay System (Promega). To assess relative Srr2 activity, luciferase luminescence was normalized to renilla luciferase activity.
Table 2.2: List of Antibodies Used for Experiments

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Company</th>
<th>Product Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD44-APC</td>
<td>BD Pharmingen</td>
<td>559942</td>
</tr>
<tr>
<td>CD49f-APC</td>
<td>eBioscience</td>
<td>17-0495-82</td>
</tr>
<tr>
<td>CD24-Biotin</td>
<td>Invitrogen</td>
<td>13-0247-82</td>
</tr>
<tr>
<td>EpCAM-FITC</td>
<td>BD Biosciences</td>
<td>347197</td>
</tr>
<tr>
<td>Streptavidin-V450</td>
<td>BD Biosciences</td>
<td>560797</td>
</tr>
<tr>
<td>Streptavidin-FITC</td>
<td>BD Biosciences</td>
<td>554060</td>
</tr>
<tr>
<td>Six2</td>
<td>Proteintech</td>
<td>11562-1-AP</td>
</tr>
<tr>
<td>Six2</td>
<td>Sigma-Aldrich</td>
<td>SAB1105672</td>
</tr>
<tr>
<td>Sox2</td>
<td>Proteintech</td>
<td>66411-1-Ig</td>
</tr>
<tr>
<td>Nanog</td>
<td>Proteintech</td>
<td>14295-1-AP</td>
</tr>
<tr>
<td>Six1</td>
<td>Heide Ford Lab*</td>
<td>N/A</td>
</tr>
<tr>
<td>Hdac1</td>
<td>Santa Cruz Biotechnology</td>
<td>sc-7872</td>
</tr>
</tbody>
</table>

*NOTE: mouse monoclonal Six1 antibody was generated by the Tissue Culture Core Facility at the University of Colorado Anschutz Medical Campus*
ChIP-qPCR

Cells were plated in 15cm plates to similar confluency and then fixed with 1% formaldehyde to crosslink DNA and proteins. After terminating the crosslinking with glycine, cells were collected with ice cold PBS containing protease inhibitors and pelleted by centrifugation. Cell pellets were then processed to obtain ChIP DNA by using the Zymo Research ChIP kit (Zymo Research, D5210) according to manufacturer instructions. After obtaining ChIP DNA, qPCR was performed by diluting the samples 1:5 with nuclease-free water and using primers to amplify regions of SRR2 (Table 2.1). Ct (quantification) values from qPCR were normalized to input values for each condition. Results are presented as relative enrichment by subtracting the %input for IgG negative control from %Input for Six2 IP to adjust for background signal.

Datamining

Publicly available datasets were analyzed using cBioportal, Oncomine, SurvExpress or KMplotter. For survival data analyzed through KMPlotter, we used the 206510_at probe for SIX2 and the 228038_at probe for SOX2 (mean expression) to analyze the relationship between SIX2/SOX2 gene expression and Distant Metastasis-Free Survival and Relapse-Free Survival. After removal of biased array data (patient and clinical heterogeneity, different outcome measures and size effects) and redundant samples, results were split by the “auto select best cut off” and graphed to show the approximate low vs high expression, FDR of <20% and stated p-values.

Statistical Analysis

The experiments performed were analyzed using GraphPad Prism 6. Error bars were generated by Standard Error Mean (SEM) calculations. For experiments with two conditions, an unpaired one-tailed Student’s T-test was performed. For experiments with three or more conditions, one-way anova followed by a Bonferroni comparison was used. For the animal experiments measuring luciferase signal over time, fitting of a mixed effects model for testing
the main effect of KD/OE and KD/OE-time interactions were performed. Given the statistical significance of the overall mode (P<0.05), we proceeded to carry out the multiple comparisons (with Bonferroni correction) to determine significant differences in the increase of luciferase signal over time between the KD/OE groups and their corresponding CTL groups. Asterisks represent the significance of difference from control group *, P<0.05; **, P<0.01; ***, P<0.001 and ****, P<0.0001.
CHAPTER III

SIX2 MEDIATES LATE-STAGE METASTASIS VIA DIRECT REGULATION OF SOX2 AND
INDUCTION OF A CANCER STEM CELL PROGRAM

Introduction

The molecular mechanisms behind the metastatic progression of breast cancer remain poorly understood. Although advances in therapeutics are extending the lives of patients, metastatic spread remains the major cause of death. Understanding the mechanisms driving metastasis, particularly establishment and maintenance of metastasis, could lead to development of more effective therapeutics and increased survival.

The metastatic cascade is a multistep process involving escape from the primary tumor, dissemination, extravasation, survival/establishment and colonization. The rate-limiting step of metastasis is outgrowth at the secondary site, as newly arriving tumor cells must survive a harsh new microenvironment to establish and maintain proliferation at secondary sites[146]. Outgrowth of metastatic cells is complicated by the fact that these sites initially lack sufficient vasculature to provide nutrients needed by the cells[147]. Because the microenvironment at the secondary site presents challenges for newly arriving tumor cells, only a small percentage of these cells will go on to develop overt metastases[148]. Despite increased evidence around the importance of metastatic cell outgrowth at secondary sites, the vast majority of metastasis studies continue to focus on identifying molecular mechanisms that drive the earlier stages of metastasis, such as migration and invasion. However, recent studies have established that

tumor cells exit the primary site at very early stages of the disease[149]. Thus, targeting early stages of metastasis may be less effective, because by the time the primary tumor is diagnosed, many cancer cells would have likely already left the primary site. Since the later stages of metastasis are less efficient, and often not completed at the time of diagnosis, targeting this aspect of the metastatic cascade may be more effective[146].

Recently, advancements in imaging, mouse models and genomic and single cell genomics analyses have allowed for more accurate modeling of the later stages of metastasis. This technological progress has led to the identification of several proteins that specifically alter metastatic colonization. For example, tumor cells can indirectly activate osteoprotegerin (OPG), allowing for bone degradation and release of growth factors into the bone microenvironment, which provides signals to metastatic cells to proliferate[150]. Additionally, tumor cell-derived tenascin C (TNC) promotes the survival and outgrowth of lung metastases through regulation of stem cell signaling components musashi homolog 1 (MSI1) and leucine-rich repeat–containing G protein–coupled receptor 5 (LGR5)[48]. However, even with these promising discoveries, metastatic colonization remains the major cause of mortality in breast cancer patients. In part, this continued mortality can be attributed to a dearth of targeted therapies inhibiting metastatic colonization and/or the outgrowth of metastatic clones, which are often resistant to the chemotherapy regimens to which the primary tumor responded[146].

SIX2 is a homeodomain-containing transcription factor that plays a key developmental role in the kidney. KO of Six2 in mice demonstrates a critical role for the gene in maintenance of the mesenchymal phenotype, self-renewal, and survival of progenitor cells that will eventually give rise to all epithelial cell types of the developing nephron[151]. Not surprisingly then, Six2 loss in mice leads to postnatal lethality due to defects caused by premature differentiation of cells in the kidney as well as kidney hypoplasia[151]. Importantly, mutations in SIX2 have been found in humans with congenital abnormalities of the kidney, such as renal hypoplasia[151], underscoring an important role for this gene in human kidney development.
Tumor cells often hijack developmental processes to promote metastasis[88]. Recent studies have thus begun to examine the role of SIX2 in tumor progression. OE of SIX2, as compared to normal tissue, has been observed in esophageal and lung carcinomas[100]. SIX2 also plays a critical role in tumor cell invasion and drug resistance in colorectal cancer[152], and regulates proliferation and EMT in hepatocellular cancer[153]. Most recently, whole exome sequencing of Wilms tumor patient samples identified that almost a fifth (18.1%) of patients with the most aggressive blastemal subtype of Wilms tumor possess a gain-of-function mutation (Q177R) in the homeodomain of SIX1 or SIX2 that correlates with high proliferation[154]. This mutation has been shown to enhance the SIX1-mediated Warburg effect, suggesting that a similar function could be attributed to SIX2[155]. Importantly, SIX2 marks the cancer stem cell population in Wilms tumor[156], suggesting conservation of its developmental function in cancer. Collectively, these data support a critical role for SIX2 across multiple different cancer types and underscore the need to better understand the molecular mechanisms by which SIX2 promotes tumor progression and metastasis.

Previous studies from our lab have demonstrated that Six2 is critical for late-stage metastasis in triple-negative mouse mammary carcinoma models[100]. KD of Six2 in the triple-negative, metastatic 66cl4 mouse mammary carcinoma line results in decreased metastatic burden when injected orthotopically[100]. Interestingly, Six2 KD does not affect primary tumor growth or tumor-associated lymphangiogenesis, contributors to early stage metastasis that are known to be regulated by a related Six family member, Six1[88]. Conversely, Six2 OE in the 4T07 mammary carcinoma cell line results in increased lung metastases after tail vein injection in BALB/c mice when compared to control cells[100]. This result is of particular interest, as the 4T07 cells, which are syngeneic to the 66cl4 cells, can reach secondary sites but are unable to efficiently colonize them[143]. These data demonstrate that Six2 promotes the ability of cancer cells to colonize secondary sites. Further analysis demonstrated that Six2 OE leads to a decrease in E-cadherin (Cdh1) expression, suggesting that Six2 may mediate at least a partial
EMT [100]. Importantly, restoration of E-cadherin in 4T07 cells downstream of Six2 OE inhibited metastasis induced by Six2[100]. However, KD of Six2 inhibited metastasis of 66cl4 cells without restoration of E-cadherin expression, likely due to epigenetic silencing of Cdh1[100]. Taken together, these data suggest that Cdh1 repression is necessary downstream of Six2 to mediate metastasis, but that additional genes regulated by Six2 are also required.

Herein, we demonstrate that SIX2 plays a critical role in metastatic colonization by promoting stemness-associated properties in triple-negative breast cancer (TNBC) and TN mouse mammary carcinoma cells. We show that SIX2 is more highly expressed in patients with TNBC as compared to other subtypes, and that it regulates metastasis not only of mouse mammary carcinoma lines, but also in a human TNBC model. Using three different models of TNBC, we demonstrate that SIX2 enhances cancer stem-cell associated phenotypes and regulates a genetic stem cell program. We identify Six2 as a direct transcriptional regulator of a master pluripotency factor, Sox2, further resulting in upregulation of a second master pluripotency factor, Nanog. Importantly, we show that Sox2 is the critical mediator downstream of Six2 in promoting stemness-associated phenotypes in vitro and late-stage metastasis in vivo. Finally, we demonstrate for the first time that SIX2 and SOX2 expression positively correlate in human breast cancer, including TNBC, and that a Six2-mediated gene signature is associated with significantly shortened distant metastasis free survival as well as relapse and recurrence-free survival. Collectively, our data suggest that a novel SIX2/SOX2 axis may promote stem cell characteristics in newly arriving breast cancer cells at secondary sites, providing insight into the regulation of metastatic colonization and outgrowth.
Results

**SIX2 expression is enriched in TNBC where it promotes metastasis**

Our previous studies demonstrated that SIX2 is more highly expressed in human breast carcinomas when compared to normal breast tissue, and that its expression is also increased in invasive ductal carcinomas when compared to ductal carcinoma in situ (DCIS) [100]. Building on these results, analysis of SIX2 mRNA expression from multiple patient datasets, including TCGA and Tabchy datasets [157], reveals that it is more highly expressed in TNBC compared to all other subtypes of breast cancer (Fig 3.1A). Patients with TNBC present with higher tumor grade at diagnosis, have the highest rate of recurrence and possess the poorest prognosis overall compared to the other subtypes, making it one of the more aggressive breast cancer subtypes [158]. Given that we previously established a role for Six2 in late-stage metastasis of murine triple-negative mouse models [100] and its high expression in human TNBC, we next sought to understand the molecular mechanism through which SIX2 regulates TNBC metastasis. However, as Six2 has previously only been shown to regulate murine triple negative mammary carcinoma metastasis, we first examined whether SIX2 can also increase metastasis of human TNBC. To this end, we knocked down (KD) SIX2 in luciferase tagged triple-negative MDA-MB-231 cells (Fig 3.1B). Tail-vein injections of non-targeting control (NT) and SIX2 KD MDA-MB-231 cells into immunodeficient NOD-scid-gamma (NSG) mice demonstrate that mice injected with SIX2 KD cells have significantly reduced metastatic burden when compared to mice injected with NT control cells (Fig 3.1C and D). These data, in conjunction with previous data in mouse mammary triple-negative models [100], demonstrate a critical role for SIX2 in the later stages of metastasis in human TNBC.
Figure 3.1: *SIX2* expression is enriched in triple-negative breast cancer (TNBC) where it promotes metastasis. A) *SIX2* mRNA expression in TNBC vs all other subtypes. Expression levels of *SIX2* obtained from the TCGA Cell 2015 (cbioportal) (Gao et al. 2013; Cerami et al. 2012) and Tabchy breast datasets (Oncomine) (Rhodes et al. 2004) B) Levels of *SIX2* mRNA (left) and protein (right) in non-targeting (NT) Control MDA-MB-231 cells and in *SIX2* KD MDA-MB-231 cells. Gene expression was normalized to *PPIB* mRNA expression and fold-change is relative to the NT cells. P-values were calculated using one-way anova followed by Bonferroni multiple comparisons test of triplicate samples in a representative experiment (n=3). For western blot analysis, nuclear extracts are shown with HDAC1 used as a loading control. C) Luciferase-labeled MDA-MB-231 NT, KD1 and KD2 cells were injected into NSG mice through the tail vein. Metastatic burden was measured by IVIS imaging. Luciferase images are of mice at the week 6 timepoint (n=5 mice/condition). D) Quantification of average whole-body luciferase signal (total luciferase counts) for each condition over the duration of the animal experiment, (n=5 mice/condition). P-values to determine differences in luciferase signal over time were calculated using mixed model effects interaction analysis followed by Bonferroni correction.
SIX2/Six2 regulates stem cell-associated phenotypes in TNBC and mouse mammary carcinoma cells

Because Six2 plays a critical role in the maintenance and self-renewal of nephron progenitor cells during kidney development, and because it promotes metastasis of a cell line (4T07) that is capable of reaching secondary sites, but cannot efficiently colonize those sites[143], we hypothesized that SIX2 mediates late-stage metastasis, particularly the survival and establishment of tumors at secondary sites, partly due to its role in stem/progenitor-like phenotypes. To test this hypothesis, we first performed flow cytometry to determine if there are differences in the stemness-associated population in response to the presence of SIX2.

Because MB-MDA-231 cells express little to no CD24, we measured cell surface expression of EPCAM and CD44[159], which is known to more accurately enrich for cells with tumor-initiation and metastatic capacity[160]. Interestingly, SIX2 KD results in a decrease in the EPCAM+/CD44+ population in MDA-MB-231 cells (Fig 3.2A). We additionally examined whether Six2 alters the cancer stem cell (CSC) population within triple-negative mouse mammary carcinoma cells where we had previously found it to mediate metastasis[100]. In this case, flow cytometry was performed to examine the CD24+/CD49f+ population, which enriches for cells with the ability to repopulate the mouse mammary gland[161], in 4T07 cells overexpressing Six2 (4T07-Six2), and in 66cl4 cells with KD of Six2, as compared to their control counterparts. Our data show that Six2 loss decreases, while Six2 OE increases, the percentage of CD24+/CD49f+ cells within the population (Fig 3.2B and C).

To assess the role of SIX2 in stem cell-associated phenotypes using a more functional assay, we next performed tumorsphere formation experiments in all three models. The ability of cells to form secondary tumorspheres after dissociation and dilution to single cells is a measure of self-renewal capability, and thus serves as a surrogate CSC assay[162]. Similar to what was observed in our flow cytometry analysis, in both the MDA-MB-231 and 66cl4 cells, SIX2/Six2 KD significantly decreases the ability of cells to form secondary tumorspheres (Fig 3.2D and E). In
line with these findings, Six2 OE in the 4T07 cells significantly increases secondary
tumorsphere formation (Fig 3.2F).

To further test whether SIX2 regulates cancer stem cell frequency and functionality, we
performed limiting dilution assays in vivo using the MDA-MB-231 cells with SIX2 KD. We found
that compared to NT control cells, SIX2 KD cells have decreased tumor-initiating cell (TIC)
capacity, as demonstrated by a significant decrease in TIC frequency from 1:146,096 for NT
control cells to 1:813,444 and 1:357,622 for SIX2 KD1 and KD2 cells, respectively (Fig 3.2G).
Taken together, these data suggest that SIX2 may play a similar role in mediating stemness
characteristics during tumor progression as it does during development[151], and that this role
may allow it to mediate late-stage metastasis by enhancing the ability of cells to self-renew and
survive at secondary sites.

**Six2 regulates a transcriptional program associated with stemness**

Since SIX2 is a transcription factor, we hypothesized that it directly transcriptionally
regulates downstream effectors to promote CSC-like phenotypes and metastatic burden. We
have previously shown that 4T07 cells express very little Six2 endogenously and that the OE
levels are similar to other murine mammary triple-negative cell lines with endogenously high
Six2 expression, such as the 66c14 cells [100]. Additionally, although we observe a range of
SIX2 expression in human TNBC cell lines, we found that Six2 OE in 4T07 cells is comparable
to, and does not artificially exceed, the levels of SIX2 observed in multiple TNBC cell lines (Fig
3.3). In addition, although we attempted to perform RNA-seq in the 66c14 Six2 KD cells, the
resulting expression data was very inconsistent (most likely due to varied KD levels). Therefore,
we chose the 4T07 system to perform RNA-seq analysis comparing 4T07 control (4T07-CTL)
and 4T07-Six2 cells to examine how Six2 regulates downstream targets across the genome.
Six2 OE leads to a robust transcriptional response as evidenced by the significant increased
expression of 1,997 genes and decreased expression of 2,002 genes (FDR q-value<0.1 and
log2 Fold Change ± 1.0) (Fig 3.4A).
Figure 3.2: *SIX2/Six2* regulates stem cell-associated phenotypes in TNBC and mouse mammary carcinoma cells. A) *SIX2* knockdown in MDA-MB-231 cells decreases the EPCAM⁺/CD44⁺ population, as measured using flow cytometry. Quantification on the right is of triplicate samples for a representative experiment (n=3). Error bars represent the mean +/- SEM. P-values were calculated using one-way anova followed by a Bonferroni multiple comparisons test. B) *Six2* KD in 66cl4 cells decreases the CD24⁺/CD49f⁺ population, as measured using flow cytometry. Quantification and p-values were done as described above. C) *Six2* overexpression in 4T07 cells increases the CD24⁺/CD49f⁺ population, as measured using flow cytometry. Quantification was performed as described above. P-values were calculated using an unpaired two-tailed t test.
Figure 3.2 (Continued): **SIX2/Six2** regulates stem cell-associated phenotypes in TNBC and mouse mammary carcinoma cells. D) **SIX2** knockdown in MDA-MB-231 cells decreases tumorsphere formation when compared to control (NT) cells. Quantification on the right is of triplicate samples for a representative experiment (n=3). Error bars represent the mean +/- SEM. P-values were calculated using one-way anova followed by a Bonferroni multiple comparisons test. E) Six2 knockdown in 66c14 cells decreases tumorsphere formation when compared to control (NS) cells. Quantification and p-values were performed as described above. F) Six2 overexpression in 4T07 cells increases tumorsphere formation when compared to control cells. Quantification was performed as described above. P-values were calculated using an unpaired two-tailed t test. G) **SIX2** knockdown in MDA-MB-231 cells decreases tumor-initiating capacity compared to control (NT) cells. Cells were transplanted into the 4th mammary fat pad of nude mice at limiting dilutions. The graph displays the percentage of positive tumors at week 5 post tumor cell injection (n=6). Mean tumor-initiating cell (TIC) frequency was determined using the ELDA software program, with estimated ranges of TIC frequency being 72,956-292,560 for NT, 326,264-2,028,088 for KD1 and 175,059-730,575 for KD2.
Additionally, Principal-Component Analysis (PCA) based on the top 500 most-variable genes from each sample revealed a stark difference between 4T07-CTL and 4T07-Six2 cells, indicating that Six2 OE in 4T07 cells drives distinct changes in the transcriptional landscape (Fig 3.4A). Given the well-established role of Six2 as a critical transcriptional regulator of progenitor and stem cell populations[151], we used the web-server StemChecker[163] to examine whether differentially expressed genes identified from the RNA-seq analysis overlap with previously established and curated stemness-associated gene signatures. We first queried the 1,997 significantly upregulated and 2,002 significantly downregulated genes in response to Six2 OE (Adjusted p-value < 0.1 and log2 Fold Change ± 1.0) from the RNA-seq analysis to determine the overlap with gene signatures from StemChecker that represent specific normal stem cell types. We found that 77% of the Six2-upregulated genes from the RNA-seq analysis overlapped with genes from StemChecker signatures representing various stem cell types, whereas the Six2-downregulated genes contained a 72% overlap (Fig 3.4B). Interestingly, gene signatures representing embryonic stem cells (ESCs) were amongst the most highly enriched and represented signatures in the analysis when examining all significantly differentially expressed genes together (Fig 3.4C).

We then determined whether SIX2/Six2 KD or Six2 OE (Fig 3.4D shows expression levels of SIX2/Six2 in all three cell lines as determined in the same RNA samples used in panels 3.4E and F) could affect the expression of genes from the ESC-associated signatures. For instance, when examining a dataset that used a shRNA functional screen to identify crucial factors for embryonic stem cell proliferation and pluripotency[164], we found that the genes contained in this signature were dramatically differentially expressed between the 4T07-CTL and 4T07-Six2 cells (Fig 3.4E, ESC_Wang dataset[164]). The same was true for a signature containing genes that are differentially expressed between mouse embryonic stem cells and adult tissues[165] (Fig 3.4F, SC Wong dataset[165]), as well as signatures associated with maintaining pluripotency [166,167] (Fig 3.4G). To validate that SIX2 mediates gene expression
alterations that are linked to stem-associated phenotypes, we performed RT-qPCR for stem cell genes identified in the different signatures on independent sets of RNA from all three cell line models. These genes were chosen based on their established roles in normal embryonic/cancer stem cell function, tumor progression, and/or metastatic colonization\[164,168–172\], and were amongst the top quartile of upregulated genes after DESeq analysis. This analysis confirmed that SIX2/Six2 KD in both MDA-MB-231 and 66cl4 cells leads to a corresponding decrease in the mRNA expression of numerous genes associated with embryonic stem cells (Fig 3.4E and F), whereas Six2 OE leads to an increase in many of the same genes (Fig 3.4E and F).

To further confirm that SIX2 plays a role in a genetic stem cell program, we also performed Gene-Set Enrichment Analysis (GSEA) after converting mouse gene symbols to their human equivalents\[173\]. Notably, GSEA analysis using the C2-Curated gene sets revealed that Six2 OE results in negative enrichment of pathways associated with E-cadherin function (ONDER CDH1 TARGETS 2 DN) and cell focal adhesion (KEGG FOCAL ADHESION) (FDR q-value < 0.25 and Normalized p-value < 0.05) (Fig 3.4H and I). This data supports previous studies from our lab demonstrating that Six2 OE in 4T07 cells results in increased metastasis in part through downregulation of E-cadherin and increased anchorage-independent growth\[100\]. Interestingly, we also observed enrichment of pathways associated with stemness. For instance, pathways that are upregulated to promote stemness (BENPORATH PROLIFERATION, WONG EMBRYONIC STEM CELL CORE, MUELLER PLURINET) were among the most highly positively significantly enriched pathways across all pathways analyzed in the C2-Curated gene sets (Fig 3.4H and I). Additionally, we identified significantly negatively enriched pathways that contain genes that are downregulated in ESCs (BENPORATH ES WITH H3K27ME3) (Fig 3.4H and I). When performing GSEA analysis using the C5-Gene Ontology gene sets, Six2 OE led to negative enrichment of pathways involved in adhesion and differentiation (3.4J and K).

Together, these data demonstrate that SIX2 expression broadly enhances a stem-cell gene program in both triple negative breast and mammary cancer cells.
Figure 3.3: Six2 OE in 4T07 cells is comparable with multiple other human TNBC cell lines with relatively high endogenous SIX2 expression. Western blot analysis showing the level of Six2 OE in 4T07 cells compared to other human TNBC cell lines. Lines were determined to be TNBC cell lines via the ATCC database HDAC1/Hdac1 was used as a loading control.
Figure 3.4: Six2 OE leads to a change in the transcriptional program of 4T07 cells and differential expression of stemness-related genes and pathways. A) MA plot displaying the number of differentially expressed genes in response to Six2 OE in 4T07 cells after RNA-seq analysis (Fold-change > 1.0 relative to 4T07-CTL cells, FDR q-value <0.1) (Left). Principle Component Analysis of 4T07-Six2 and 4T07-CTL cells (n=2) after batch correction (Right). B) Area-proportional Venn Diagram displaying percent overlap of significantly upregulated (left) or downregulated (right) genes with normal stem cell type signatures. C) Genes from RNA-seq analysis that overlap with genes from signatures that represent different normal stem cell types. D) RT-qPCR analysis measuring SIX2/Six2 mRNA expression after KD (MDA-MB-231 and 66cl4) and OE (4T07). Gene expression was normalized to PPIB (human) or Ppib (mouse) mRNA expression and fold-change is relative to the NT cells. Heatmap and mRNA expression of significantly expressed genes from RNA-seq analysis that overlap with the E) Mm_ESC_Wang and F) Mm_SC_Wong gene signature. P-values were calculated using either one-way anova followed by Bonferroni multiple comparisons test or unpaired T-test on means +/- SEM of triplicate samples for a representative experiment (n=3).
Figure 3.4 (Continued): Six2 OE leads to a change in the transcriptional program of 4T07 cells and differential expression of stemness-related genes and pathways. G) Heatmaps displaying differential expression of 4T07-Six2 and -CTL cells from the Mm_ESC_Yang and PluriNetWork gene signatures (Pinto et al., 2015). H) Volcano plot of false discovery rate (FDR q-val) versus normalized enrichment score (NES) after GSEA from RNA-seq data (Pre-ranked GSEA module). Red dots denote significantly positively enriched pathways and blue dots denote significantly negatively enriched pathways (0.25 FDR cutoff). I) GSEA enrichment plots for EMT-, adhesion- and stemness-associated pathways selected from the C2-Curated gene set. Black bars represent genes contained within the specific pathway and are ranked from positively expressed (left) to negatively expressed (right) based on log2 fold change (Six2 OE vs CTL). J) Volcano plot of FDR q-values and normalized enrichment scores (NES) from GSEA analysis using the C5-Gene Ontology gene set of all significantly expressed genes in response to Six2 OE (Blue dots, significantly negatively enriched pathways; Red dots, significantly positively enriched pathways). K) Enrichment plots, NES and FDR q-values of a subset of significantly enriched gene set pathways.
**Six2 regulates master stemness factors Sox2 and Nanog**

Based on the large number of stem-related genes altered by SIX2, we asked whether SIX2 promotes the expression of master regulators of stemness. To address this question, we first examined the mRNA and protein expression of well-established master regulators of stemness and pluripotency. Intriguingly, KD of SIX2/Six2 in the MDA-MB-231 and 66cl4 cells resulted in corresponding decreases in both SOX2/Sox2 and NANOG/Nanog mRNA as measured by RT-qPCR (Fig 3.5A and B). In contrast, Six2 OE in the 4T07 cells resulted in upregulation of mRNA expression of both pluripotency factors (Fig 3.5C). Additionally, KD of SIX2/Six2 leads to a decrease in both SOX2/Sox2 and NANOG/Nanog protein (Fig 3.5D and E), whereas OE of Six2 leads to corresponding increases in Sox2 and Nanog levels (Fig 3.5F). These data suggest that SIX2 may promote stemness-associated phenotypes through regulation of SOX2 and NANOG.

Previous studies have shown that in the context of breast cancer, SOX2 regulates cancer stem cell properties upstream of NANOG[141]. To determine whether Sox2 is upstream of Nanog in Six2 OE cells, we established stable KD of Sox2 in the 4T07-Six2 cells (Fig 3.6A). Indeed, we observed that Sox2 KD in 4T07-Six2 cells resulted in a corresponding decrease in Nanog mRNA and protein levels (Fig 3.6B), suggesting that Six2 may act more directly on Sox2 to mediate alterations in Nanog levels.

Sox2 transcriptional regulation is, in part, under the control of two enhancer regions, Srr1 and Srr2. The Srr2 domain, located ~3kb downstream of the Sox2 gene, functions as a transcriptional enhancer of Sox2 expression in ESCs[174]. Intriguingly, further analysis of the sequence of the Srr2 enhancer revealed the presence of a Six2 binding region[175]. Therefore, we performed luciferase reporter assays containing the Srr2 regulatory region[176] to determine whether Six2 may regulate Sox2 through the Srr2 domain (Fig 3.6C shows the reporter construct used in assays). After transfection of a Srr2 luciferase construct (pGF-Sox2-ONL, 3.6C)[176], along with a Renilla-luciferase plasmid (pCNL-N1_3xNLS) to normalize for
transfection efficiency, we found that SIX2/Six2 KD decreases, whereas Six2 OE increases, activation of the Srr2 region (Fig 3.6D). To determine if Six2 directly regulates the Srr2 enhancer via this binding region, we introduced a single point mutation (T27G) within the putative Six2 binding region of the Srr2 enhancer based on a Six2 binding site that was previously identified and characterized in the Gdnf gene[175] (Fig 3.6E). Interestingly, although introduction of the wildtype srr2 region resulted in a similar increase in Srr2 enhancer activity in 4T07-Six2 cells compared to 4T07-CTL cells, introduction of the T27G mutation into the Srr2 enhancer-luciferase construct inhibited the activation of luciferase activity (Fig 3.6F). We next performed ChIP-qPCR and found that Six2 binding at the Srr2 enhancer region is enriched in 4T07-Six2 cells compared to control cells (Fig 3.6G). In contrast, the enrichment of Six2 observed in the Sox2 regulatory regions is lost when performing a pulldown at a region approximately 6kb downstream of the Srr2 region that contains no known Six2 binding sites (Fig 3.6G). Together, our results suggest that Six2 plays a critical role in altering CSC characteristics through direct transcriptional activation of Sox2, resulting in upregulation of Nanog and induction of a stem cell-like program.

**Sox2 mediates stem cell-associated phenotypes and late-stage metastasis downstream of Six2**

We then asked whether Sox2 is required downstream of Six2 to mediate CSC characteristics and late-stage metastasis. To this end, we performed flow cytometry to measure the mouse mammary stem cell enriched population in response to Sox2 KD downstream of Six2 in the 4T07 cells. We found that Sox2 KD in 4T07-Six2 cells leads to a significant decrease in the CD24+/CD49f+ enriched CSC population when compared to control cells (4T07-Six2 shNT) (Fig 3.7A). Similarly, secondary tumorsphere formation was also decreased in response to Sox2 KD in 4T07-Six2 cells (Fig 3.7B). Additionally, to better understand why we did not observe a complete reversal in the *in vitro* stemness-associated phenotypes after Sox2 KD, we performed ChIP-qPCR in 4T07-CTL and -Six2 OE cells to examine whether Six2 was directly bound to,
Figure 3.5: SIX2/Six2 regulates master pluripotency factors SOX2/Sox2 and NANOG/Nanog. SOX2/Sox2 and NANOG/Nanog mRNA expression in the A) MDA-MB-231, B) 66cl4 and C) 4T07 cells. Gene expression was determined by RT-qPCR and normalized to PPIB/Ppib mRNA expression. Fold-change is relative to the NT (MDA-MB-231 cells), NS (66cl4 cells) or CTL (4T07 cells). P-values determined using an unpaired T-test or one-way anova followed by Bonferroni multiple comparisons test of means +/- SEM of triplicate samples for a representative experiment (n=3). SIX2/Six2 (top), SOX2/Sox2 (middle) and NANOG/Nanog (bottom) protein expression in D) MDA-MB-231, E) 66cl4 and F) 4T07 cells. Protein expression after nuclear extraction was determined by Western Blot Analysis with HDAC1/Hdac1 being used as a loading control (n=3).
Figure 3.6: Sox2 is upstream of Nanog and is directly regulated by Six2 via its Srr2 enhancer. A) Sox2 mRNA expression in 4T07-CTL, 4T07-Six2 and 4T07-Six2/Sox2 KD (KD1 and KD2) cells measured by RT-qPCR (left). Western blot analysis of Sox2 protein in 4T07-CTL, 4T07-Six2, and 4T07-Six2/Sox2 KD cells (right). Gene expression was determined using RT-qPCR and normalized to Ppiib mRNA expression. Fold-change is relative to 4T07-CTL cells. P-values were determined using a one-way anova followed by Bonferroni multiple comparisons test of means +/- SEM of triplicate samples for a representative experiment (n=3). Sox2 protein was determined using western blot analysis after whole cell extraction with Hdac1 used as a loading control. B) Nanog mRNA expression in 4T07-CTL, 4T07-Six2 and 4T07-Six2/Sox2 KD (KD1 and KD2) cells measured by RT-qPCR (left). Gene expression was determined by RT-qPCR and normalized to Ppiib mRNA expression. Fold-change is relative to 4T07-CTL cells. P-values were determined as described above of triplicate samples for a representative experiment (n=3). Nanog protein was determined as described above (right).
Figure 3.6 (Continued): Sox2 is upstream of Nanog and is directly regulated by Six2 via its Srr2 enhancer. C) Diagram of pGF-Sox2-ONL luciferase reporter plasmid used in luciferase assays. Darker rectangle denotes putative Six2 binding region (T2A, encodes self-cleaving peptide sequence; ONL, Orange NanoLantern). D) pGF-Sox2-ONL luciferase assays performed in MDA-MB-231 +/- SIX2 KD, 66cl4 +/- Six2 KD and 4T07 +/- Six2 OE cells 72h after transfection. P-values determined using an unpaired T-test or one-way anova followed by Bonferroni multiple comparisons test of means +/- SEM of triplicate samples in a representative experiment (n=3). E) Sequence of WT Six2 binding region (underlined) within Srr2 enhancer region. Bottom sequence depicts the mutant version of the Srr2 enhancer (mutation in larger font) (T27G) used in the luciferase experiment. F) Luciferase reporter assay using WT and T27G Sox2-ONL luciferase reporter plasmids in 4T07-CTL and Six2 cells. P-values were determined as described above of triplicate samples in a representative experiment (n=3). G) ChIP-qPCR analysis of 4T07-CTL and -Six2 cells. ChIP-enriched DNA was quantified using RT-qPCR with primers surrounding the predicted Six2 binding region within the Srr2-enhancer. In a separate pulldown, Six2 binding at a downstream region that does not contain known Six2 binding sites was also examined. P-values were determined using an unpaired T-test of means +/- SEM of triplicate samples in a representative experiment (n=3).
and therefore could transcriptionally regulate, the promoter regions of the putative targets identified in the StemChecker analysis. We found that Six2 was bound to a subset of these targets, suggesting that Six2 regulates multiple genes that play a critical role in stemness (Fig 3.7C). This data suggests that in order to completely reverse the in vitro stemness-associated phenotypes regulated by Six2, multiple downstream targets, in addition to Sox2, may need to be inhibited.

To examine whether Sox2 is a critical mediator of Six2 in promoting late-stage metastasis, we performed tail-vein injections using 4T07-Six2/Sox2 KD cells (as compared to corresponding controls). As had been previously observed[100], Six2 OE in 4T07 cells enables metastatic outgrowth in vivo (Fig 3.7D, upper panels). However, KD of Sox2 in 4T07 cells dramatically decreases Six2-induced metastatic outgrowth (Fig 3.7D, lower panels and E for quantitation), suggesting that in vivo Sox2 is the key target downstream of Six2 in promoting late-stage metastasis. Thus, our data demonstrate that Six2 regulates metastatic outgrowth in a Sox2-dependent manner.

**SIX2 and SOX2 expression correlate in human breast cancer, and are together associated with poor prognosis**

On the basis of our experimental findings that Sox2 KD decreases Six2-mediated metastatic outgrowth, we next examined whether this relationship may be relevant to human breast cancer. We first performed RT-qPCR on multiple human TNBC cell lines to determine if the SIX2-SOX2 axis is present in additional human TNBC cells. Indeed, we find that in both BT549 and SUM159 cells (TNBC cell lines with high endogenous SIX2 expression[100]), transient KD of SIX2 leads to a corresponding decrease in SOX2 (Fig 3.8). These data suggest that similar to what we observe in the MDA-MB-231 cell line model, SIX2 also regulates SOX2 in other human TNBC models. We next performed correlation analysis of SIX2 and SOX2 mRNA expression obtained from cbioportal, Oncomine and R2 Cancer Databases and demonstrate that SIX2 and SOX2 positively correlate in breast cancer datasets encompassing
Figure 3.7: Sox2 mediates stem cell-associated phenotypes and late-stage metastasis downstream of Six2. A) KD of Sox2 in 4T07-Six2 cells decreases the CD24+/CD49f+ mouse mammary stem cell-like population when compared to 4T07-Six2 scramble control (shNT) cells as measured by flow cytometry. Quantification on the right is of triplicate samples for a representative experiment (n=3). Error bars represent the mean +/- SEM. P-values were calculated using one-way anova followed by a Bonferroni multiple comparisons test. B) Sox2 KD in 4T07-Six2 cells decreases tumorsphere formation compared to 4T07-Six2 non-targeting shRNA (4T07-Six2 shNT) cells. Quantification and calculation of p-values were done as described above. C) ChIP-qPCR analysis of 4T07-CTL and -Six2 cells. ChIP-enriched DNA was quantified using RT-qPCR with primers surrounding predicted Six2 binding sites within the promoter regions of validated genes from the StemChecker analyses.
Figure 3.7 (Continued): Sox2 mediates stem cell-associated phenotypes and late-stage metastasis downstream of Six2. D) Luciferase-labeled 4T07-CTL shNT, -Six2 shNT, -Six2/Sox2 KD1 and -Six2/Sox2 KD2 cells were injected into Balb/c mice through the tail vein. Metastatic burden was measured using IVIS imaging. Representative pictures of animals injected (n=16 per condition at Week 2 timepoint). E) Quantification of average +/- SEM lung luciferase signal for each condition over the duration of the animal experiment (n equals at least 15 per condition, as one 4T07-Six2 shNT mouse died after week 2). P-values to determine differences in luciferase signal over time were calculated using two-way anova followed by Bonferroni correction of average luciferase signal in each condition.
all breast cancer subtypes (Fig 3.9A); a relationship that is also observed when specifically examining TNBC datasets (Fig 3.9B). These data suggest that the relationship between SIX2 and SOX2 in our cell line models is recapitulated in human breast cancer.

We then generated a SIX2-gene expression signature compiled of the top 15 significantly upregulated genes in response to Six2 OE after converting mouse gene symbols to their human equivalent (Table 3.1). Of interest, the top 15 SIX2-regulated genes (based on highest expression) are all associated with roles in normal and CSC function and/or tumor progression[177–188]. Importantly, high expression of the SIX2-gene signature significantly correlates with a decrease in distant metastasis-free and relapse-free survival across all breast cancer patient samples as well as in TNBC-specific samples (Fig 3.10A and B). Additionally, the SIX2-gene signature significantly correlates with a decrease in time to distant metastasis, distant metastasis-free, and recurrence free survival across all breast cancer (we were unable to query TNBC specifically in this dataset) (Fig 3.10C). These data demonstrate that the relationship between SIX2 and metastatic burden is not limited to our cell line models, and that the SIX2 transcriptional program is predictive of poor outcomes in human breast cancer.

**Discussion**

Previous data from our group, using two different triple negative mouse mammary carcinoma xenograft models, established Six2 as critical for late stage metastasis[100]. In particular, the fact that Six2 results in outgrowth of 4T07 cells, which are known to reach the lungs after orthotopic injection into the mammary gland, but do not efficiently colonize the lungs[143], demonstrates a role for Six2 in the colonization step of the metastatic cascade. Further, data within this manuscript demonstrate that SIX2 also regulates late stage metastasis in human TNBC.

Although our data suggest that a main function of SIX2 is to mediate outgrowth of cells that have newly arrived at a secondary site, given its described role in mediating mesenchymal phenotypes and anoikis resistance[100,156], SIX2 may in fact affect several additional aspects
of late-stage metastasis, although this remains to be formally tested. If SIX2 regulates multiple steps of late stage metastasis (such as survival in the bloodstream and exit from the vasculature in addition to colonization at secondary sites), this finding would further emphasize the importance of developing novel means to target its function, as significant data suggests that metastases can seed additional metastatic sites[189]. In the future, it will be important to determine whether SIX2 is required for maintenance of metastatic lesions (addressed in detail in Chapter IV). This question is critical to address, as patients that already present with significant metastatic burden would benefit from a therapy that inhibits the continued growth of cells at the secondary site. Nonetheless, even if SIX2 is only critical for the establishment and initial stages of metastatic outgrowth, many breast cancer patients do not present with overt metastases, and thus inhibiting this step could be valuable. However, as SIX2 is a transcription factor, its direct inhibition may prove difficult. Thus, understanding the mechanism by which SIX2 mediates late stage metastasis may enable the identification of novel downstream targets or pathways that can potentially be targeted in the future.

Previous studies have shown that Six2 maintains the kidney progenitor pool during nephrogenesis, in part through impinging on stem cell-associated pathways[151]. Based on our RNA-seq analysis, we find that similarly, in the context of breast cancer, SIX2 regulates a genetic stem-cell like program. We demonstrate that Six2 OE increases, whereas SIX2/Six2 KD decreases, the expression of multiple genes that are enriched in normal stem cells. These data suggest that SIX2 may regulate stem cell-like properties across multiple tissue types and across normal and disease states. Interestingly, although it has mostly been studied in the context of nephrogenesis[151], previous studies from our laboratory have shown that during mammary gland development, Six2 expression is increased in the mammary gland of Six1 KO mice[89]. Because Six1 OE in the mammary gland results in the expansion of mammary stem/progenitor cells[88], yet Six1 loss does not affect mammary gland development[89], it is possible that upregulation of Six2 compensates for Six1 loss to maintain mammary stem cell function.
Figure 3.8: Six2 regulates Sox2 in multiple human TNBC models. Transient KD of Six2 in both BT549 and SUM159 TNBC cells results in a corresponding decrease in Sox2 mRNA expression as measured by RT-qPCR. Gene expression was measured 48h post-transfection of siRNAs targeting Sox2. A non-targeting siRNA (si0) was used as a negative control.
Figure 3.9: SIX2 and SOX2 mRNA positively correlate in breast cancer patient samples and upregulates genes associated with stemness and poorer prognosis. A) Linear regression analyses demonstrate significant positive correlation between SIX2 and SOX2 in Tumor Breast EXPO - 351 datasets and Hatzis via R2 Database. B) Linear regression analysis demonstrating that SIX2 and SOX2 mRNA significantly correlate in TNBC patient samples using both the TCGA Cell 2015 dataset and the Tabchy dataset (obtained from Oncomine).
Table 3.1: List of top 15 upregulated genes in response to Six2 OE (SIX2-Gene Signature)

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>FC (Six2/CTL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSPB1</td>
<td>302.3</td>
</tr>
<tr>
<td>NOS1</td>
<td>25.7</td>
</tr>
<tr>
<td>CHRM3</td>
<td>24.5</td>
</tr>
<tr>
<td>MID2</td>
<td>17.1</td>
</tr>
<tr>
<td>CDH19</td>
<td>12.5</td>
</tr>
<tr>
<td>EFS</td>
<td>12.3</td>
</tr>
<tr>
<td>TPS3I11</td>
<td>11.7</td>
</tr>
<tr>
<td>TSPAN12</td>
<td>11.6</td>
</tr>
<tr>
<td>PLCL1</td>
<td>11.4</td>
</tr>
<tr>
<td>COL9A1</td>
<td>10.8</td>
</tr>
<tr>
<td>NDP</td>
<td>10.7</td>
</tr>
<tr>
<td>SLC28A2</td>
<td>10.5</td>
</tr>
<tr>
<td>PPARGC1A</td>
<td>9.4</td>
</tr>
<tr>
<td>CYP7B1</td>
<td>9.0</td>
</tr>
<tr>
<td>GRIK2</td>
<td>8.5</td>
</tr>
</tbody>
</table>
Figure 3.10: SIX2 and SOX2 expression correlate in human breast cancer, and a SIX2-gene expression signature is associated with poor prognosis. A) Distant-metastasis free and relapse-free survival in patients with high or low combined expression of the SIX2-gene signature in all subtypes of breast cancer (KMPlotter). B) Distant-metastasis free and relapse-free survival of TNBC patients with high or low combined expression of the SIX2-gene signature (KMPlotter, auto select was used for cutoff). C) Time to distant metastasis, Distant metastasis-free survival and recurrence-free survival from the Survexpress meta-analysis dataset of the SIX2-Gene Signature (median cutoff).
Figure 3.11: *Six2* is more highly expressed in the fetal mammary stem cell population compared to adult mouse mammary cells. mRNA expression of *Six2* in normal adult mouse mammary cells and normal fetal mouse mammary stem cells (fMaSC). RNA-seq data was obtained from GSE27027(Spike et al., 2012).
Similar to SIX1, SIX2 is rarely expressed in adult tissues, but it is re-expressed in the context of breast cancer[100]. The role of SIX2 in nephrogenesis, coupled with its re-expression in human breast cancers, suggests that SIX2 may be a key regulator of both mammary stemness and mammary tumorigenesis and/or metastasis. However, the functional role of SIX2 in the normal mammary stem cell compartment and in mammary gland development has yet to functionally explored. Analysis of mRNA expression from isolated mouse fetal mammary stem and adult mouse differentiated mammary epithelial cells[161] shows that Six2 mRNA is more highly expressed in the fetal mammary stem cell population compared to the adult mammary cell population, suggesting that SIX2 may play a role in fetal mammary stem cells (Fig 3.11). Thus, investigation into whether SIX2 contributes to normal mammary development may be warranted.

Previous studies using ChIP-seq analysis in the developing kidney identified Gdnf, Eya1 and Six2 itself as direct transcriptional Six2 targets[151,175]. Interestingly, our RNA-seq analysis in 4T07 cells, identify Eya1 as one of the top upregulated genes in response to Six2 OE, suggesting conservation of specific downstream targets of Six2 between normal development and cancer. As outlined above, we have identified Sox2 as a novel direct transcriptional target of Six2 in breast cancer cells that is critical for Six2-mediated metastasis. We further demonstrate that Six2 directly regulates Sox2 transcription by binding to, and activating, the Srr2 enhancer region of the Sox2 gene. Thus, it will be of interest to determine whether Six2 similarly regulates Sox2 in the context of normal embryonic development. While it is important to continue to understand how SOX2, as a key stem cell factor, mediates SIX2-regulated late-stage metastasis, unfortunately, SOX2, like SIX2, is a transcription factor and thus not an ideal drug target. Moving forward, it will be important to identify more targetable downstream effectors of SIX2 that are critical in mediating its ability to promote metastatic burden.
Overall, our studies indicate that SIX2 plays an important role in promoting cancer stem cell-associated properties through a novel SIX2/SOX2 axis that is critical for metastatic colonization (Fig 3.12). Importantly, the relationship between SIX2 and SOX2 is relevant in human breast cancer, and high expression of a SIX2-mediated gene signature significantly correlates with poor metastatic prognosis. Although it is well established that targeting cancer stem cells at the primary tumor site can increase therapeutic efficacy[190], our findings support growing evidence that these strategies should be considered when targeting metastatic tumor cells, in which there may be different regulators of tumor outgrowth that provide an advantage in a new microenvironment[148]. Our results support this emerging concept by providing critical knowledge regarding how stem cell properties in newly arriving tumor cells promote late-stage metastasis.

**Acknowledgements**

I want to thank Sadie Doran and Katherine Johnson for their help with examining how Six2/SIX2 KD affected stemness in breast cancer cells. I thank Melanie Vincent for her help with the animal experiments. I also thank Matt Gailbrath and Ahwan Pandey for their help with RNA-seq analyses. Thanks to James Costello as well for helping interpret RNA-seq data and GSEA analyses. Vadym thank you for your help with tail vein injections to study late-stage metastasis. Finally, thank you to Pratyka for your help with statistical analysis for the animal experiments. We thank Yasushi Okada for providing the pGF-mSox2-ONL and pCNL-N1_3xNLS plasmids for use in the SRR2 luciferase reporter assays. We also thank the University of Colorado Cancer Center Animal Imaging (P30CA046934), Flow Cytometry (P30CA046934), Functional Genomics (P30CA046934), and Genomics Shared Resources (P30CA046934) for aiding in experimental analysis used within this work. This work was supported by R01CA180175 (JD), R01CA095277 (HLF), The Cancer League of Colorado - 25A5976 (HLF), R01CA117907 (JME), the NCI F31 Ruth L. Kirschstein National Service Research Award 5F31CA210622-02 (MUJO), and the NCI F99/K00 Fellow Transition Award F99CA223023 (MUJO).
Figure 3.12: The SIX2/SOX2 axis mediates metastatic colonization. In the absence of SIX2, breast cancer cells also have lower levels of SOX2. This results in decreased stemness properties such as survival and self-renewal. The ultimate result is that these cells are unable to efficiently outgrow at secondary sites (Left). In contrast, cells with higher levels of SIX2 now have increased SOX2 expression, allowing them to acquire stemness phenotypes including survival and self-renewal. This leads to an increased ability of newly arriving tumor cells to colonize distant sites (Middle). When SOX2 is decreased downstream of increased levels of SIX2, tumor cells cannot acquire stemness properties and outgrow efficiently at secondary sites. This suggests that SOX2 is the critical mediator downstream of SIX2 in promoting metastatic colonization through increases in stemness-associated phenotypes.
CHAPTER IV

ESTABLISHING NOVEL EXPERIMENTAL MODELS TO EXAMINE LATE-STAGE METASTASIS AND DETERMINE OVERLAPPING AND DISTINCT OF SIX1 AND SIX2

Introduction

An estimated 30% of breast cancer patients with an early stage diagnosis will develop metastatic disease, and virtually all deaths from breast cancer can be attributed to metastatic burden. Thus, there is a critical need to develop anti-metastatic therapeutics. Because tumor cells have already disseminated from the primary tumor at the time of diagnosis, research focusing on the early stages of metastasis, such as migration and invasion, may be limited in its potential to identify targets whose inhibition would effectively decrease metastatic burden. In contrast, the identification of molecules required for the establishment, survival, and expansion of cells at secondary sites (later stages of metastasis) is critical for developing impactful anti-metastatic therapies.

Six2 is a homeoprotein that plays a key role in kidney development [104,191–196]. In mice, Six2 maintains the mesenchymal phenotype of progenitor cells that will eventually give rise to all of the epithelial cell types of the developing nephron. Six2 expression in the developing nephron is associated with self-renewal, survival and resistance to apoptosis, and its loss results in premature epithelialization of the kidney and renal hypoplasia[192,193]. The importance of Six2 in overall development has been demonstrated by the fact that Six2 null mice die shortly after birth with significant kidney defects[193]. Taken together, these studies demonstrate that Six2 is a critical factor in maintaining progenitor cell populations and mesenchymal phenotypes, and leads one to ask whether its functions during development could also be important when misexpressed in tumors (as is seen with other homeoproteins) [197].

Recent studies have begun to examine the role of Six2 in tumor progression [93,100,152,154]. In contrast to its related family member, Six1, much less is known about the role of Six2 in cancer. Six2 is thought to act in a different manner than Six1, as Six1 influences...
primary tumor growth as well as metastasis, whereas the effects of Six2 are isolated to late stage metastasis [93,100]. We have previously demonstrated that Six2 KD in TN, metastatic 66cl4 mouse mammary carcinoma cells, when compared to a non-silencing control (NS), decreases lung metastases when injected orthotopically [100]. Additionally, Six2 OE in TN 4T07 mammary carcinoma cells results in significantly increased lung metastases in BALB/c, immune-competent mice compared to control cells [100]. This latter model defines a critical late-stage metastatic function for Six2, as 4T07 cells are capable of reaching the lungs after orthotopic or tail vein injection, but are unable to colonize the lungs [143], and the addition of Six2 to these cells allows colonization [100]. Intriguingly, although Six2 enables colonization in the lungs, we have not observed any effects of Six2 on proliferation in culture or in the primary tumor, nor have we seen effects of Six2 on adhesion or invasion, properties that are associated with tumor growth and earlier events in metastasis (data not shown and [100]).

Studies from this thesis have shown that upon Six2 OE in 4T07 cells, we observe a dramatic change in the transcriptome, resulting in a shift towards a genetic stem cell program [93]. Indeed, we found that not only were multiple pathways associated with stemness affected by Six2, but master stemness factors, Sox2 and Nanog, were also upregulated downstream of Six2 [93]. Our studies (described in Chapter III) demonstrated that Six2 directly regulates Sox2, through the Sox2 enhancer region (Sr2), and that Sox2 is the critical mediator downstream of Six2 in promoting stemness properties in vitro and late-stage metastasis in vivo. Overall, these data indicate that Six2 plays an important role in the later stages of metastasis (establishment and/or outgrowth and survival of the tumor cells at the secondary site), and as outlined above, our data suggest it does so via regulating a master regulator of stemness, Sox2. Our discovery has clinical implications because at the time of diagnosis, tumor cells have often already escaped the primary tumor. Thus, targeting earlier stages of metastasis may be ineffective, whereas targeting later stages of metastasis, such as establishment, outgrowth, and/or survival of tumor cells at the secondary site may lead to more effective anti-metastatic treatments.
Because we have shown that 4T07 cells overexpressing Six2 are highly metastatic [93,100], and because 4T07 cells without Six2 OE are capable of intravasation and extravasation in orthotopic metastasis models, but not of colonizing secondary sites [103], we hypothesize that the major role of Six2 during later stages of metastasis is in mediating establishment/outgrowth at secondary sites. However, whether cells remain dependent on SIX2 to mediate continued growth at the secondary site remains unknown. Therefore, in this chapter I will discuss how we plan to develop inducible cell line models to tightly control the timing and expression level of Six2 in order to better understand whether metastatic lesions remain dependent on Six2 for continued growth. Addressing this question will be critical for understanding the full spectrum of how Six2 functions (such as transcriptional regulation) during late-stage metastasis, as such information could ultimately lead to more effective therapeutic strategies to improve patient outcomes.

The Six family of homeobox transcription factors have been shown to play critical roles during normal development and tumor progression[88,91,96,100,195,198–201]. Specifically, in the context of kidney development, the closely related family members Six1 and Six2 are crucial to the expansion and maintenance of kidney progenitor cells[198]. Despite their overall high sequence (83%) and structural homology (conserved Homeo- and DNA binding domains), Six1 and Six2 have differing functions during nephrogenesis. For instance, previous studies have shown that Six1 is essential for the onset and continuation of nephrogenesis, whereas Six2 is more critical for the maintaining later progenitor self-renewal after the onset of nephrogenesis[198].

This phenomenon also extends to breast cancer metastasis, where Six1 appears to regulate multiple aspects of the metastatic cascade, but Six2 specifically affects the later stages of metastasis to promote metastatic outgrowth[91,96,100]. Additionally, Six1 and Six2 differentially impact distinct subtypes of breast cancer, suggesting their functions are context dependent. Interestingly, given their distinct roles in development and metastasis, they both are
correlated with poor prognosis in breast cancer overall, both promote stemness properties during development and breast cancer progression (thus influencing the MIC population) and compensate for each other in some contexts[96,100,199], suggesting the presence of shared pathways regulated by both proteins to promote cancer stem-like properties and tumor progression. However, despite these data, very few common or distinct downstream pathways/targets of Six1 and/or Six2 have been identified in breast cancer. Thus, it is of interest to identify the common/distinct downstream targets/pathways that are regulated by Six1 and Six2 so that we can better understand their roles in regulating stemness properties and breast cancer metastasis. Therefore, I hypothesize that the transcription factors Six1 and Six2 regulate mutual and distinct pathways to promote stemness properties and breast cancer metastasis. Additionally, I hypothesize that understanding how they regulate targets across the genome can ultimately lead to the development of more effective therapeutics to either simultaneously or individually target their ability to promote the properties of MICs and metastatic progression. In this chapter I will discuss the different experimental models and bioinformatics approaches we plan to use in order to address these hypotheses and determine how SIX1 and SIX2 regulate overlapping and distinct genes to promote metastasis. Understanding the downstream pathways that mediate Six1 and Six2 effects on metastasis are critical when considering therapeutics, as these proteins are transcription factors and will thus not be easy to target directly.

Results

4T07 murine mammary carcinoma cells are known to perform all steps of the metastatic cascade efficiently, except for colonization once they reach a secondary site [143]. We have shown that Six2 OE in these cells leads to metastatic outgrowth in the lungs. However, we do not yet know whether Six2 is required for continued growth and survival at the secondary site. This issue is critical, as targeting of Six2 (or downstream effectors) to inhibit metastasis will be more effective if it is required not only for colonization, but also for maintained growth and
survival at the secondary site. In this chapter we begin to test the hypothesis that Six2-mediated increases in Sox2 result in a CSC phenotype that enables both metastasis initiation and maintained growth at the secondary site, by developing a novel Six2 inducible mammary carcinoma system. The system I developed is an inducible 4T07-Six2 expression system (referred to as 4T07-Six2i) (Fig 4.1A). Importantly, induction of Six2 with doxycycline (dox) leads to both a decreased expression of the epithelial marker E-cadherin and increased expression of stem cell-associated markers including Sox2 (Fig 4.1B-E), suggesting that this system is recapitulating what I observed with stable OE of Six2 in 4T07 cells [93,100]. Unfortunately, I will not be able to test this in vivo before I leave, but I will outline here the experimental plan. To ensure that the 4T07 cells with induction of Six2 harbor the same properties as 4T07 cells with stable Six2 OE, we will first treat 4T07-EV and 4T07-Six2i cells with increasing doses of doxycycline over time and assess the minimal time needed for induced Six2 OE to promote increases in Zeb2, Sox2 and Nanog levels, as well as downregulation of E-cadherin levels and increased methylation at the E-cadherin promoter. This will serve as a genetic readout for proper induction of Six2 that recapitulates stable Six2 OE. In addition, we will perform the same induction experiments above followed by tumorsphere assays as a functional/phenotypic readout. Once optimization of doxycycline pre-treatment is performed, we will examine the role of Six2 in maintained metastatic growth in vivo, by using luciferase-tagged Six2 dox-inducible 4T07 cells and empty vector (EV) control cells that are pre-treated in culture with doxycycline (at the minimal dose and time needed to promote the genetic and phenotypic changes outlined above) to induce Six2 expression (in Six2i cells, Six2 expression should not be induced in EV cells). The cells will then be injected into the tail vein of BALB/c mice to isolate the experiment to the later stages of metastasis (Fig 4.2). At the time of injection, the mice will receive doxycycline (2mg/ml, as described in [202–204]) in their drinking water (4T07-EV mice will not express Six2, even in the presence of dox, but 4T07-Six2i mice will express Six2). Once metastases are detected by IVIS imaging (expected to be almost undetectable in the 4T07-EV injected mice, but
highly detectable in the 4T07-Six2i mice based on prior experiments [93,100]), doxycycline will be withdrawn in a subset of the 4T07-Six2i mice, after which metastases will be monitored twice a week by IVIS imaging to examine whether stasis and/or regression occurs. Completion of this work will be integral to better understand how SIX2 regulates late-stage metastasis, and could lead to novel therapeutic strategies to inhibit metastatic burden.

Previous studies from our laboratory have shown that in the context of metastasis, Six2 may compensate for loss of Six1 [96]. Indeed, the very few 66cl4 cells with Six1 KD that are recovered from the lungs have upregulation of Six2 [96]. This prompted future studies examining the role of Six2 in metastasis. Interestingly, in contrast to Six1, which regulates multiple aspects of the metastatic cascade including proliferation, EMT, migration and invasion [31,92,96], Six2 was demonstrated to play a more isolated role in the later stages of metastasis [100]. Specifically, our group showed that Six2 was a critical mediator of metastatic outgrowth. Interestingly, similar to Six1, the data also indicated that Six2 regulated EMT and anoikis resistance, through downregulation of E-cadherin, to promote metastasis. Conversely, unlike Six1, the functions of Six2 in this context did not include effects on proliferation, migration or adhesion [100]. Lastly, work from my thesis added mechanistic insight into the role of SIX2 in late-stage metastasis by showing that the developmental roles of Six2 in maintaining stemness and survival were hijacked by metastatic tumor cells to colonize distant sites [93]. Taken together, our data indicate that SIX1 and SIX2 have both overlapping (EMT, Anoikis resistance and stemness) and distinct (Six1 regulates proliferation, migration and invasion exclusively) functions that promote metastasis. However, past studies from our group on the roles of Six1 and Six2 in metastatic breast cancer have not examined the function of one gene in isolation of the other. Importantly, knowing the context by which each or both proteins regulate metastasis will not only give us a better understanding of the metastatic cascade, but may also lead to the development of more effective therapeutic strategies to target and suppress the functions of Six1 and/or Six2, which can inhibit metastatic burden and increase overall survival.
Figure 4.1: Inducible Six2 expression recapitulates gene expression of downstream targets observed with stable Six2 OE. Inducible Six2 cells treated with increasing doses of doxycycline show A) increased Six2 mRNA and protein expression and B) decreased Cdh1 mRNA and E-cad protein expression (1ug/ml dox for 48h). qRT-PCR measuring mRNA expression of C) Six2, D) Sox2 and E) Nanog in 4T07-Six2i cells -/+ doxycycline. Total RNA was harvested 48h post treatment with 1ug/ml of doxycycline. Hdac1 or Gapdh was used as loading control for western blot analysis. Ppib was used as a loading control for qRT-PCR analysis.
Figure 4.2: Schematic displaying the experimental design for testing whether Six2 regulates the continued growth of metastatic lesions. Pre-treated 4T07-EV or 4T07-Six2i will be injected into the tail vein of BALB/c mice to isolate the experiment to the later stages of metastasis. At the time of injection, the mice will receive 2mg/ml of dox in their drinking water. Once metastases are detected by IVIS imaging, doxycycline will be withdrawn in a subset of the 4T07- Six2i mice, after which metastases will be monitored twice a week by IVIS imaging to examine whether stasis and/or regression occurs.
Previous work from our laboratory aimed to identify the relationship between Six1 and Six2 in metastasis. Given that Six1 and Six2 have high sequence homology overall (> 80%) and are almost 100% conserved across the Six and Homeodomains, it is anticipated that they will have overlapping gene targets. However, when microarray analysis was performed on 66cl4 cells with either Six1 or Six2 KD, no overlapping targets were identified (data not shown). Still, this could have been due to the different efficiencies of KD of Six1 vs Six2, the fact that this was done in a pooled population of cells with different KD efficiencies, and/or the fact that overlapping targets may not be unmasked if only one of the two transcription factors is knocked down, as they may compensate for each other. To address at least some of these concerns, we decided to generate complete KO of either Six1 or Six2 using CRISPR/Cas9 technology. Using this technology would allow us to dissect the function of one gene in isolation of the other and may provide insight into their distinct transcriptional targets. Additionally, we would be able to identify how each gene endogenously is responsible for advancing metastasis without the other present because we will be using the 66cl4 mouse mammary carcinoma cell line where both genes are already known to influence metastasis. Specifically, previous data from our group has shown that loss of either Six1 or Six2 in 66cl4 cells significantly affects metastasis [93,96,100].

To this end, I attempted to generate Six1 and Six2 KO cells in the 66cl4 system using CRISPR/Cas9 approaches. I designed gRNAs that would target the Cas9 nuclease to Exon1 of either Six1 or Six2. I then cloned the gRNAs into a non-lentiviral plasmid (pSpCas9-2A-GFP) that allows for the gRNA and Cas9 to be expressed from one plasmid [205]. Additionally, expression of GFP downstream of Cas9, and under the control of a single CMV promoter, allowed me to single cell sort transfected cells to screen for KO clones. After the clones grew up, I combined sequencing and western blot analysis to identify 66cl4 cells with successful knockout of either Six1 or Six2 (Fig 4.3A and B). However, in the process of generating these KO cell lines we discovered yet another issue with trying to study these genes in isolation. Although I was able to achieve Six1 KO in the 66cl4 cells (Fig 4.3A and B), I observed a
corresponding increase in $Six2$ mRNA (Fig 4.3C), likely compensating for $Six1$ loss. This suggests that unlike stable shRNA-mediated $Six1$ KD, where we do not observe significant upregulation of $Six2$ in vitro [96], there is a compensatory increase in $Six2$ when $Six1$ is completely lost in vitro, and that there may be a threshold level of $Six1$ loss that triggers upregulation of $Six2$. Because of this issue, we decided to instead use siRNAs to target either $Six1$ or $Six2$ transiently as an alternative way to better isolate the functions of each protein in the absence of the other. By introducing transient loss, we reasoned that we would avoid compensation such as that observed with complete gene knockout. In addition to using the 66c14 system, we also introduced siRNAs targeting either $SIX1$ or $SIX2$ in MDA-MB-231 cells to identify overlapping and distinct transcriptional regulation by SIX1 or SIX2 in a human context. In our transient KD system, although we observe significant loss of $Six1/SIX1$ and $Six2/SIX2$ without compensation from the other (Fig 4.3D-G) (in that the other does not become overexpressed as we see in our KO systems), western blot analysis did show that either transient $SIX1$ or $SIX2$ KD caused a corresponding decrease in the protein levels of the other family member (Fig 4.3F and G). These data suggest that in a transient setting, SIX1 and SIX2 positively regulate each other, and that only longer-term loss leads to compensation. Of interest, these data are also similar to human kidney development, where there is reciprocal regulation of SIX1 and SIX2 [198]. Interestingly, this is in contrast to what we observe with stable shRNA KD $Six2/SIX2$, where we do not see downregulation of $Six1$ or $SIX1$ in the 66c14 and MDA-MB-231 cells, respectively (Fig 4.4). Thus, alternative methods may need to be considered to ensure that when we have loss of one family member and analyze changes in downstream gene expression to identify distinct transcriptional targets, we do not observe corresponding changes in the other family member, as these data suggest that $Six1/SIX1$ and $Six2/SIX2$ can regulate each other differently depending on specific conditions and contexts.

Since Six1 and Six2 are both transcription factors, we also decided to couple our siRNA studies examining gene expression changes with ChIP-seq analysis to achieve a more
Figure 4.3: 66c4 Six1 KO results in upregulation of Six2, whereas siRNA KD of Six1/SIX1 or Six2/SIX2 results in no compensation. A) Alignment of DNA sequencing for Exon 1 of the Six1 gene in WT versus Six1 KO cells. Six1 KO results in a deletion at the beginning of Exon 1 leading to loss of Six1 protein as measured by western blot (right). CTL cells were transfected with the same plasmid as the Six1 KO cells, but without gRNAs targeting Six1. Hdac1 was used as loading control. B) qRT-PCR measuring Six2 expression in CTL and Six1 KO cells. Ppib was used as a loading control. qRT-PCR measuring mRNA expression of Six1 and Six2 in C) 66c4 and D) MDA-MB-231 cells with siRNA mediated knockdown of either Six1/SIX1 or Six2/SIX2. Ppib/PPiB was used as a loading control. Western blot analysis of E) Six1/SIX1 and F) Six2/SIX2 in 66c4 and MDA-MB-231 cells. HDAC1/Hdac1 was used as a loading control. For all siRNA experiments, cells were harvested 48h post transfection.
Figure 4.4: Stable shRNA KD SIX2/Six2 results in no consistent changes in SIX1/Six1. qRT-PCR analysis of SIX2/Six2 and SIX1/Six1 mRNA expression in A) MDA-MB-231 and B) 66cl4 cells. Gene expression was normalized to PPIB (human) or Ppib (mouse) mRNA expression and fold-change is relative to the NT/NS cells.
comprehensive assessment of Six1 and Six2 transcriptional regulation across the genome.
Measuring and comparing genome-wide endogenous transcription factor binding using ChIP-seq requires high-quality, non-cross-reactive ChIP-grade antibodies [206]. Unfortunately, few ChIP-grade Six1- or Six2-specific antibodies exist, and the antibodies that are available may bind with different affinity when measuring protein levels via western blot. Therefore, we decided to use CRISPR/Cas9 technology to introduce 3X-FLAG tags to the C-terminal ends of either Six1 or Six2 to more accurately compare their occupancy across the genome using ChIP-seq.

Indeed, using the CRISPR epitope tagging ChIP-seq (CETCH-seq) system, I efficiently introduced 3X-FLAG tags to the C-terminal ends of both Six1 and Six2 (Fig 4.5A). After validation by PCR, gDNA sequencing and western blot analyses, I performed ChIP-qPCR to test whether there was differential binding of Six1 or Six2 at a previously identified downstream target [93]. Preliminary data suggests that Six2 is more preferentially bound to the Srr2 enhancer region compared to Six1 (Fig 4.5B), suggesting that despite their high sequence homology across their DNA-binding domains, Six1 and Six2 preferentially bind target genes.

**Discussion**

The targeting of established metastases remains an immense challenge in the field of cancer biology. This can be mainly be attributed to their diffuse location in different organs (or location in vital organs), evolution of heterogeneity/genetic mutations and resistance to cytotoxic agents [207]. Thus, current treatments for metastatic breast cancer mainly focus on quality of life, rather than curing the disease. Because metastatic cells have been shown to dramatically diverge genetically and phenotypically from primary tumor cells, it is imperative that studies continue to identify vulnerabilities specific to metastatic cells, as opposed to primary tumor cells, to identify more effective therapeutics to inhibit metastasis. In this thesis, we demonstrate that the homeoprotein, SIX2, is specifically associated with tumor cells that can survive and outgrow at secondary sites.
Figure 4.5: 3X-FLAG tagged Six2 66cl4 cells preferentially bind to the Srr2 enhancer region compared to 3X-FLAG tagged Six1 66cl4 cells. A) Western blot analysis of 66cl4 Parental and FLAG-tagged Six1 and Six2 cells. Hdac1 was used as a loading control. B) ChIP-qPCR analysis of the Srr2 enhancer region in 66cl4 FLAG-tagged Six1 and Six2 cells. Primers were designed to flank a Six2 binding motif with the Srr2 enhancer region.
Additionally, our data suggest that this is, in part, reliant on a shift towards a more stemness phenotype [93]. Although previous data has suggested that stemness programs play a critical role in tumor initiation and progression, few studies have focused on the role stemness plays specifically during metastatic initiation and outgrowth. Indeed, the data from this thesis indicates that stemness is necessary to allow tumor cells to better survive and initiate tumors at secondary sites [93]. Importantly, future studies examining the role of SIX2 in the maintenance of metastatic lesions will be critical for identifying novel therapeutic strategies to combat metastatic disease, as this remains the most difficult aspect of breast cancer progression to treat, and thus is responsible for the vast majority of breast cancer related deaths.

In this chapter, I have established an inducible Six2 OE 4T07 model to test whether Six2 is required for the continued growth of metastatic lesions. Addressing this question is critical, as patients that present with metastasis would benefit from a therapy that inhibits the continued growth of the metastatic lesions. Nonetheless, even if Six2 is only important for establishment of metastases, inhibiting its action will still be of use to inhibit the formation of new metastases. Importantly, although my thesis work has determined that Sox2 is required for Six2-mediated metastatic outgrowth, I have not tested whether Sox2 is also required for maintenance of metastasis. Therefore, inducible systems of Sox2 expression downstream of Six2 OE can be developed to address this question. Additionally, development of an inducible SIX2 KD system in MDA-MB-231 cells, to address whether loss of SIX2 in a setting where it is normally expressed will be valuable in determining whether targeting SIX2 functions would be a promising therapeutic strategy to reverse metastatic outgrowth and maintenance. Lastly, our initial in vivo experiments will be done through tail vein injection and assessment of metastatic outgrowth to the lung. Therefore, in the future it would also be interesting to test whether the effects we observe with Six2 OE and metastatic outgrowth are organ-specific. Since Six2 is a critical mediator of late-stage metastasis overall, and significant data suggest that metastases can seed additional metastatic sites [189], it will be important to determine whether Six2 impacts
metastatic outgrowth at sites other than the lung. Thus, future experiments where we perform intracardiac injections will allow us to test the role of Six2 in metastatic outgrowth at additional organs such as the bone or liver. Determining whether Six2 affects outgrowth at additional sites will underscore its value as a therapeutic target, not only for preventing initial metastases, but also for patients who present with metastatic disease at diagnosis, as targeting Six2 would inhibit the ability of secondary metastases to form. Overall, determining whether Six2 is important for the maintenance of metastases could have major implications for improving outcomes for patients with metastatic disease. Unfortunately, the vast majority of cancer deaths are still attributed to the metastases and not the primary tumor. Despite this, few drugs still exist that specifically treat metastases directly. Therefore, additional mechanistic insight into how metastatic cells initiate, and more importantly maintain, growth at secondary sites will lead to the development of more targeted therapeutic strategies to treat metastatic disease.

The SIX family is an evolutionary conserved group of transcription factors that play pivotal roles in fundamental cellular processes that give rise to a myriad of tissues including eye, kidney, muscle, head and gonads [88]. Although all SIX genes come from a common ancestor, they evolved during a series of duplication events resulting in pairs of subclasses of Six genes. SIX1 and SIX2 make up the sine oculis subclass of the protein family and have high overall sequence homology (>80%), with almost 100% homology across their SD and HD [88]. The fact that these genes can compensate for each other is supported through previous studies from our laboratory demonstrating that Six1−/− mammary glands exhibit upregulation of Six2, suggesting functional redundancy and overlapping functions during mammary gland development [89]. However, despite this similarity in overall homology, they differ greatly in their C-terminal domains. Indeed, SIX1 lacks its own activation domain and therefore requires co-factors, such as EYA, to modulate transcription [199]. In contrast, SIX2 contains an intrinsic activation domain at the C-terminal end that can regulate transcription without co-expression of EYA (this is usually required in conjunction with SIX1 OE) [88] (Fig 4.6). Nonetheless, Six2 can bind to Eya
family members and may utilize the Eyas to boost their transcriptional output, as well as bind to additional factors via its activation domain including histone acetyl transferases (HATs), mediator proteins and other transcription-associated machinery. These data could in part explain how SIX1 and SIX2 may have both overlapping and distinct functions in tumor progression, though an elucidation of their direct target genes would greatly facilitate our understanding of how these proteins regulate metastasis and how we might develop more effective therapeutics to target their functions. Nonetheless, an alternative way to better understand the similarities and differences between Six1 and Six2 would be to perform Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins (RIME) using our Six1 and Six2 FLAG tagged 66c14 cell line model to identify transcriptional co-factors and other interacting-proteins that associate with Six1 or Six2 [208]. Briefly, 66c14 Six1 and Six2 FLAG tagged cells will be harvested and treated with formaldehyde to promote crosslinking of proteins. We will then isolate nuclei, sonicate DNA/Protein-protein complexes and immunoprecipitate Six1 and Six2 using a ChIP-grade FLAG antibody. After tryptic digestion of proteins, samples can be sent for mass spectrometry to identify proteins from each sample. Combined with RNA-seq and ChIP-seq, RIME will provide us with a more complete and integrated transcriptional profile of Six1 and Six2 regulation across the genome. In the context of interacting partners, this could have a significant clinical impact, as data from our group has demonstrated that the disruption of Six1 binding to its co-factor Eya2 with small molecular inhibitors can significantly inhibit metastasis in vivo. Thus, identifying interacting partners may lead to the development of additional therapeutic strategies to inhibit Six1- and/or Six2-mediated metastasis, as we could also target co-factors or other transcriptional machinery for which Six1 and/or Six2 are dependent to facilitate transcriptional regulation.

Previous studies have demonstrated that in the context of the developing mouse kidney, Six1 KO results in reduced expression of Six2 [209], suggesting that Six1 is upstream of Six2 during mouse kidney development. Interestingly, in the context of breast cancer,
Figure 4.6: The Six Family of Homeoproteins. Schematic of protein domains across Six family members. Figure from The Six Family of homeobox genes in development and Cancer by Christensen et. al [88].
data from my thesis showed that stable shRNA-mediated Six2/SIX2 KD in 66cl4 and MDA-MB-231 cells results in no change of Six1/SIX1 (Fig 4.4). These data suggest that during development and breast cancer progression, SIX1 is upstream of SIX2. However, when I generated stable Six1 KO in 66cl4 cells (4.3A and B), I observed a corresponding increase in Six2 mRNA (4.3C). This prompted us to generate siRNA-mediated transient KD cells to avoid compensation mechanisms. Still, although we did not observe protein compensation (meaning upregulation of the other family member when one is lost) in the context of transient Six1/SIX1 or Six2/SIX2 KD (Fig 4.3D-G), we did observe that transient loss of either SIX1 or SIX2 in the MDA-MB-231 cells resulted in downregulation of the other family member (Fig 4.3F and G). These data indicate that the regulatory relationship between Six1/SIX1 and Six2/SIX2 may be dependent on the timing of loss, a threshold level of expression and/or origin of species of each gene.

In line with this complex regulatory relationship between Six1/SIX1 and Six2/SIX2, recent developmental studies have shown that regulation between Six1/SIX1 and Six2/SIX2 is context and species dependent. Specifically, O'brien and colleagues demonstrated that during mouse kidney development, Six1 is indeed upstream of Six2 in the metanephric mesenchyme [198]. However, as nephrogenesis progresses, not only does Six1 continue to activate Six2 expression, but also Six2 feeds back on Six1 to maintain the nephron progenitor population throughout differentiation [198]. Interestingly, in the developing human kidney, SIX1 and SIX2 regulate each other and also regulate themselves to maintain kidney progenitor populations [198] (Fig 4.7). Given that the preliminary data from my thesis also suggests a complex mechanism of regulation between Six1/SIX1 and Six2/SIX2, targeting both SIX1 and SIX2 simultaneously may be the most effective therapeutic strategy to inhibit tumor progression and metastasis. Nevertheless, future studies examining the molecular mechanisms of compensation and transcriptional regulation between SIX1 and SIX2 would be useful in better understanding their relationships during development and metastasis.
Figure 4.7: Regulation of Six1/SIX1 and Six2/SIX2 during embryonic kidney development.

Schematic of how Six1/SIX1 and Six2/SIX2 are regulated and regulate each other in the metanephric mesenchyme and nephron progenitor cells. Figure from Differential regulation of mouse and human nephron progenitors by the Six family of transcriptional regulators by O'Brien et. al [198].
Interestingly, when comparing microarray gene expression data from 66cl4 Six1 KD and 66cl4 Six2 KD cells, no overlapping genes were identified. This could be due to the fact that shRNA KD does not completely abolish expression of the target gene. This is important in the context of Six1 and Six2 because, as previously shown, Six2 can compensate for Six1 loss during breast cancer metastasis [96]. Importantly, although their highly divergent C-terminal domains suggest Six1 and Six2 likely regulate a diverse set of downstream targets, possibly through c-terminal-mediated interactions that change their DNA binding, their high overall sequence homology indicates that they also likely regulate many overlapping target genes [31,90,93,96,100]. Given this, we used CRISPR-Cas9 technology to KO Six1 or Six2 in 66cl4 cells. Although we were able to generate Six1 KO cell lines (Fig 4.3A and B), unfortunately, we were unable to establish Six2 KO cells.

Alternatively, I moved to another approach before repeating attempts to generate Six2 KO cells, due to the observed upregulation of Six2 in response to Six1 KO. Thus, I instead introduced siRNAs targeting either Six1 or Six2 for two reasons: 1) By abolishing Six1 or Six2 transiently, I hoped to avoid compensation of the other family member to better assess the loss of Six1 or Six2 in the absence of the other gene, and 2) In contrast to stable long-term KO, transient KD will allow for identification of the most critical mediators downstream of Six1 or Six2. Overall, I anticipated that this experimental strategy may more definitively determine transcriptome profiles of SIX1 and SIX2, allowing for a more specific assessment of putative overlapping and distinct downstream targets. Unfortunately, even in this context, I still observed an effect on the other family member when I transiently KD either SIX1 or SIX2 in MDA-MB-231 cells (Fig 4.3 F and G). To address this, we may need to develop an inducible KD system for Six1/SIX1 and Six2/SIX2. Using this system, we can induce and assess gene expression changes at earlier time points where we would hopefully not observe compensation or changes in the other family member. However, it is also important to consider that the compensation and redundancy in the Six family is not isolated to Six1 and Six2. For example, previous data has
shown that Six4 can compensate for Six1 loss during limb development [210]. Therefore, any study examining loss of Six family members, should always consider compensation by other family members. This data suggests that a more ideal therapeutic strategy to inhibit Six1 or Six2 during metastasis may be to target the entire Six family to ensure inhibition of compensation between family members. However, since Six1 and Six2 appear to be the most critical regulators of metastasis within the family, targeting them individually or together may still be the most effective strategy to inhibit metastasis.

Intriguingly, studies have started to independently identify putative overlapping targets of Six1 in the context of tumorigenesis, although in a different cancer type. A recently published report, which came out just after my Cancer Research paper reporting the link between Six2 and Sox2, identified SOX2 as a downstream target of SIX1 in the context of glioma [211]. Specifically, they found that SIX1-mediated upregulation repressed cellular senescence, in part, through regulation of SOX2. Still, it is important to note that the antibodies used in this published study to measure SIX1 levels were generated using an immunogen that could cross react with SIX2. Thus, it is possible that their study could be mistakenly identifying SIX1 in regulating SOX2, when in fact SIX2 is the main mediator of this effect in glioma, similar to what we have shown in breast cancer. With the newly developed C-terminal FLAG tagged Six1 and Six2 66cl4 cells we have generated, we can begin to better address whether both proteins regulate the same transcriptional targets [96,100]. Using this system will allow us to avoid confounding results after ChIP-seq because we will utilize one antibody to perform immunoprecipitation (IP) instead of different antibodies that may have different binding efficiencies. Importantly, we plan to generate a similar cell line model system in the MDA-MB-231 cells to assess the human relevance of the transcriptomic profile of SIX1 and SIX2. Overall, a combination of RNA-seq, RIME and ChIP-seq analyses will enable us to perform a comprehensive, genome-wide, assessment of transcriptional regulation by SIX1 and SIX2, allowing for identification of
overlapping and distinct processes or pathways that can be targeted therapeutically to inhibit SIX1- and/or SIX2-dependent tumor progression and metastasis more effectively.

Acknowledgements

I would like to thank Taylor Hotz, Connor Purdy and Uyen Thai for their help with generating and screening for Six1 and Six2 KO cells. I would also like to thank Christina Towers, Darya Wodetzki and Jackie Thorburn for their help with RNP-mediated CRISPR/Cas9 KO of Six1 and Six2 in the MDA-MB-231 cells. Finally, I want to thank Moli Joshi for her advice with generating CRISPR-mediated Six1/Six2 KO and 3X-FLAG tagging of Six1/Six2 in 66cl4 cells.
CHAPTER V
DISCUSSION AND FUTURE PERSPECTIVES

The Role of Stemness In Metastatic Colonization of Breast Cancer

Innovations in research technologies such as mouse modelling, cell lineage tracing and advanced microscopy has provided powerful tools for understanding how tumor cells metastasize to distant organs. Using these technologies have demonstrated that although tumor cells enter circulation and reach secondary sites early-on in tumorigenesis, only a small fraction of these cells will be capable of seeding and outgrowth at distant organs [212]. In order to colonize secondary sites, tumor cells from the primary site must locally invade normal adjacent tissue, intravasate into the blood stream, survive in circulation, extravasate out of the blood stream and survive as either single cells or small clusters of cells in anchorage independent conditions until they can outgrow and become the clinically detectable metastatic lesions diagnosed in patients [212]. Although most studies have identified EMT as a hallmark of metastasis via its regulation of migratory and invasive phenotypes, subsequent studies have also identified critical roles for EMT in promoting stem-like properties such as self-renewal, pluripotency and survival [213]. Indeed, a seminal study by Weinberg and colleagues demonstrated that induction of an EMT in immortalized human mammary epithelial cells leads not only to the induction of mesenchymal properties, but also an increase in stem cell like characteristics [214]. Additionally, a recent study using single-cell RNA-seq analysis from patient-derived xenografts demonstrated that metastasis-initiating tumor cells possess EMT features along with expression of a stem cell program [215]. Thus, since tumor cells have likely already left the primary tumor at diagnosis, this suggests that rather than targeting EMT phenotypes associated with migration and invasion (steps associated with the earlier steps of metastasis), a better understanding of how EMT regulates stemness could lead to the development of more effective therapeutics to inhibit metastatic colonization.
My work has also revealed a link between EMT and stemness that is mediated by the developmental transcription factor Six2. In a previous report from our group, we demonstrated that Six2 OE in a murine mammary carcinoma cell line that can reach secondary sites, but cannot efficiently outgrow, results in increased metastatic burden in the lung after tail vein injection [100]. Mechanistically, this study demonstrated that Six2 can mediate late-stage metastasis through promoting EMT via upregulation of the EMT-inducing transcription factor Zeb2 and subsequent downregulation and promoter methylation of the epithelial gene Cdh1 (E-cadherin) [100]. These data would suggest that similar to its roles in development, where it represses the epithelial phenotype during nephrogenesis [193], Six2 promotes mesenchymal properties to enhance the ability of tumor cells to outgrow at secondary sites. This finding is counterintuitive though, as most studies suggest that cells need to undergo a MET to grow out at the metastatic site [216–218]. Recent data has further suggested that EMT is not required for metastasis at all [219,220], although there are significant flaws with these studies including the fact that the results come from fate mapping using a single marker of mesenchymal cells, and the fact that crosstalk between EMT and non-EMT cells in epithelial tumors was not considered [31,221]. Nonetheless, these recent reports also suggest that although EMT may be dispensable for metastasis, it is clearly important for drug resistance and survival [219,220] at the secondary site. Interestingly, recent studies have also demonstrated that in addition to properties such as self-renewal and survival, stemness also impacts metastasis through increased plasticity. This has been supported by mathematical modelling studies that have identified a “window of stemness” that links a hybrid EMT/MET state with plasticity [222]. Additionally, more recent studies have provided compelling evidence that circulating tumor cells that have undergone EMT have the unique ability to increase plasticity, survival and tumor-initiating capabilities to overcome the challenges they will face to establish tumors at secondary sites [212], and that this could be targeted therapeutically. For example, a recently published study by Ishay-Ronen and colleagues demonstrated that metastasis could be inhibited by
transdifferentiating tumor cells into adipocytes, and that cells that had undergone an EMT, and thus had increased plasticity, were especially susceptible to transdifferentiation treatment [15]. Interestingly, data from our group showed that although Six2 did promote expression of EMT markers and repression of epithelial markers, OE of Six2 did not affect EMT-associated phenotypes such as migration or invasion. Thus, we hypothesized that Six2 promotes metastatic outgrowth through alternative mechanisms aside from EMT. Indeed, in this thesis, I have demonstrated that Six2 also regulates stemness properties in both murine mammary and human breast cancer cells through direct transcriptional activation of the master pluripotency factor Sox2, with subsequent upregulation of another core pluripotency regulator, Nanog. Further, I show that Sox2 is the critical mediator downstream of Six2 to induce metastasis and that the SIX2/SOX2 axis is relevant in the more aggressive TNBC subtype, as shown in both cell line models and patient samples. Taken together, these data suggest that Six2 activity culminates in a more plastic phenotype, as opposed to an EMT and stemness phenotype, in tumor cells that allow them to better adapt, outgrow and produce heterogeneous progeny at secondary sites. Still, the connection between Six2, pluripotency and mesenchymal properties has yet to be fully examined as an integrated process. Therefore, future studies examining the molecular mechanisms by which Six2 could regulate plasticity and differentiation, in addition to self-renewal and survival, may provide a link between its roles in promoting mesenchymal properties and stemness, and importantly would provide additional critical insight into how to target Six2-mediated metastasis therapeutically.

It is possible, that while the major effect of Six2 is on establishment, it may also enhance extravasation, particularly based on its role in inducing a more mesenchymal phenotype [100,104,191]. Therefore, Six2 OE could result in enhancement of several late stage metastasis characteristics including increased transendothelial migration (perhaps due to its EMT inducing ability) and metastatic outgrowth (due to its induction of stem cell programs). However, because 4T07 cells can already extravasate to secondary sites, but cannot establish metastases, and
because we have previously shown that introduction of Six2 into these cells leads to enhanced metastases [100], we expect that the major function of Six2, as seen both in my in vitro and in vivo assays, is to allow establishment of metastases at the secondary site. Indeed, in my thesis work I demonstrate that this ability is in part due to the role of Six2 in stem/progenitor-like phenotypes [93]. However, it is unclear whether Six2 will be required for maintenance of metastatic lesions. To address this aspect of Six2-mediated metastasis, I have developed an inducible cell line model of Six2 OE. This will allow for precise control of Six2 at any given point in the metastatic cascade. By inducing Six2 OE with doxycycline to establish tumors followed by removing Six2 through withdrawal of doxycycline, we will determine if Six2 is required for the continued growth of metastases at distant sites. The question of whether Six2 regulates maintenance of metastasis has clinical implications, as patients that already present with metastasis would benefit from a therapy that inhibits the continued growth of the metastatic lesions. Nonetheless, even if Six2 is only important for establishment but not maintenance, many breast cancer patients do not present with overt metastases, where inhibiting establishment is most valuable.

Since a significant rate of attrition (<0.02% of disseminated tumor cells can successfully outgrow at secondary sites) occurs during metastatic progression, more recent studies have focused on the unique population of cells that has been hypothesized to be responsible for successful outgrowth at secondary sites [14]. Indeed, genome-wide comparisons between matched primary and metastatic tumor samples, single-cell analyses and lineage-tracing experiments have found that within a bulk metastatic tumor, cells termed metastasis-initiating cells (MICs), possess enhanced stemness characteristics including self-renewal, plasticity and survival. This allows for better adaptability and eventual outgrowth at harsh new microenvironments compared to their more differentiated counterparts. Given the data in my thesis work demonstrating that Six2 plays a critical role in stemness during late-stage metastasis, Six2 might be a defining feature of MICs, as it appears to regulate the later stages
of metastasis (during the arrival and attrition of newly arriving tumor cells). Because MICs appear to be the bottleneck for metastatic initiation and outgrowth, it will be important that future studies aim to identify the molecular mechanisms that are critical for the function and activity of this cell population (and how Six2 may influence this population), as this presents as an ideal therapeutic window to inhibit metastatic burden.

**Regulation of Pluripotency Factors by Six2 During Development and Metastatic Outgrowth**

Six2 is a homeoprotein that plays a key role in kidney development [193,194,223]. In mice, Six2 maintains the mesenchymal phenotype of progenitor cells that will eventually give rise to all epithelial cell types of the developing nephron. Six2 expression in the developing nephron is also associated with self-renewal and survival, and its loss results in premature epithelialization of the kidney and renal hypoplasia. Additionally, Six2 is important for overall development, demonstrated by the fact that Six2 null mice die shortly after birth with significant kidney defects [193]. These studies demonstrate that Six2 is a critical factor in maintaining progenitor cell populations in the kidney, as well as mesenchymal identity in the kidney, and have led others to examine whether Six2 could regulate stemness and differentiation in additional aspects of embryonic development. For example, studies by He and colleagues demonstrated that inactivation of Six2 via creation of a KO mouse results in abnormal craniofacial development. Specifically, they revealed that Six2 loss leads to premature chondrocyte differentiation. Additionally, gain-of-function experiments showed that Six2 was sufficient to induce cartilage development and growth in multiple parts of the body, suggesting that Six2 plays a general role in chondrocyte cell fate determination [224]. More recent studies by Zhou and colleagues showed that Six2 serves as a marker of heart progenitor cells, and that its expression was dependent on Hh signaling [225]. Interestingly, although ablation of Six2 progenitor cells caused no apparent heart defects at birth, cardiac hypertrophy and dysfunction
did occur later on in adulthood [225]. Taken together, these data highlight the essential role of Six2 in mediating stemness and differentiation during development.

In line with the data presented in this thesis, other groups have shown that Six2 expression and function is associated with master factors of stemness in multiple contexts during normal development and tumorigenesis. For example, during kidney development Six2 has been shown to be colocalized with Sox9 in neural crest cells [226]. Additionally, a report by Xu and colleagues demonstrated that Six2 regulates the stability of the master pluripotency factor, cMyc through its interaction with the Six family co-factor Eya1 [223]. Interestingly, recent studies by Pode-Shakked and colleagues showed that undifferentiated human fetal kidney and Wilms Tumor cells (isolated based on CD133 expression) co-expressed SIX2 and multiple core regulators of pluripotency including OCT4, NANOG and KLF4. In the context of breast cancer metastasis, my thesis work has identified the pluripotency factor, Sox2 as a direct transcriptional target of Six2 during late-stage metastasis. Additionally, upregulation of Sox2 also leads to increased levels of another master factor of pluripotency, Nanog. Taken together, these data suggest that Six2 regulates master stemness factors in both normal development and tumorigenesis. However, despite growing evidence that SIX2 regulates master factors of stemness, few studies have examined the molecular mechanisms by which SIX2 controls their expression and/or function. Therefore, future studies involving ChIP-seq, RNA-seq and Methyl-seq (as Six2 has been shown to regulate promoter methylation [100]) in these different contexts should be pursued to better understand how SIX2 may regulate pluripotency, and thus provide critical insight into how we might target this important function of SIX2 to inhibit metastatic burden.

Given the critical role of SOX2 as a master regulator of stemness and pluripotency, its expression is tightly controlled at many levels to ensure proper function and cell fate determination during development [107]. For example, in addition to basal promoter control of SOX2 gene expression, post-transcriptional control of SOX2 by miRNAs and IncRNAs has also
been reported in both embryonic stem cells and across multiple tumor types [130]. Additionally, SOX2 function can also be impacted post-translationally via phosphorylation, glycosylation, sumoylation and ubiquitination, although this has been almost exclusively examined in the context of mouse embryonic stem cells [130]. Similar to my thesis work, an additional level of transcriptional control of the SOX2 gene is through binding and interaction with enhancer regions that control SOX2 promoter activity [227]. Studies by Tomioka and colleagues were one of the first to identify two such enhancer regions, Srr1 and Srr2 in mouse ESCs [228]. Interestingly, further studies by Zhou and colleagues found that although Srr1 was active using reporter assays in mouse ESCs, its deletion had no significant impact on Sox2 gene expression [229]. Conversely, deletion of the Srr2 region downstream of Sox2 gene completely abolished transcriptionally activation and expression, suggesting that Srr2 was the critical enhancer region that controls Sox2 transcriptional activation. Subsequent studies have also identified distinct roles for the Srr2 enhancer in pluripotency. For example, Hotta and colleagues found that introduction of a lentiviral vector containing the Srr2 enhancer along with enhancers for Oct3/4 upstream of EGFP, was sufficient to isolate and enrich for induced pluripotent stem cells [230]. These data suggest that transcriptional control via the Srr2 enhancer is a critical mechanism for precise regulation of Sox2 and for the induction of stemness.

In my thesis, I identified that Six2 directly transcriptionally regulates the master stemness factor, Sox2. I demonstrated that Six2-mediated upregulation of Sox2 leads to subsequent upregulation of an additional master stemness regulator, Nanog. Additionally, using a luciferase reporter assay I showed that Six2/SIX2 can activate the Srr2 enhancer region in multiple TNBC cell line models. Interestingly, further analysis of the Srr2 enhancer sequence revealed the presence of a putative Six2 binding motif. Upon mutation of this region, Six2 binding was abolished and activation of the Srr2 was significant inhibited. Performing ChIP-qPCR analyses also revealed that Six2 binds the endogenous Srr2 enhancer in murine triple-negative mammary carcinoma cells, suggesting that Six2 directly regulates Sox2 via its Srr2 enhancer region to
promote stemness properties in metastatic breast cancer cells. Interestingly, in the context of breast cancer, SOX2 gene amplification has been identified in only a small proportion of patient tumors [130]. Instead, increased gene expression is the prominent mechanism of aberrant expression seen in patients. Importantly, SOX2 OE is significantly correlated with disease progression and poor survival [231], suggesting that identification of therapeutic strategies targeting SOX2 gene expression could have a significant impact on patient outcomes. Given my novel finding that Six2 regulates Sox2 transcription in the context of metastatic burden, and that both SIX2 and SOX2 expression are significantly positively correlated in breast cancer patient tumors, including the highly aggressive TNBC subtype specifically, further studies examining how Sox2 is controlled at the transcriptional level would provide additional insight into how we can better target its functions to inhibit metastatic colonization and improve survival. Due to the promising results our group has observed with small molecule inhibitors that target the Six1-Eya2 complex and that Six2 also interacts with the Eya proteins to regulate downstream targets [223,232], it would be interesting to determine if either Sox2 transcription or stemness phenotypes regulated by Six2 are dependent on interactions with the Eya protein family. If so, then future studies testing our small molecule inhibitors should be performed to examine whether therapeutic targeting of Six2 could inhibit metastatic outgrowth through inhibition of its ability to upregulate Sox2.

**Relationship Between Six1 and Six2 in Breast Cancer Metastasis**

Previous data from our lab and others have identified aberrant functions of SIX1 throughout multiple stages of the metastatic cascade [31,88,90,96,233], establishing SIX1 as a critical mediator of metastasis. Indeed, *in vitro* and *in vivo* studies from our group have shown that SIX1 activity impinges on metastatic properties such as proliferation, EMT, migration, invasion and lymphangiogenesis [31,88,90,96,233]. Further, *in vivo* studies using orthotopic, tail vein and intracardiac (the latter two being models for experimental metastasis) injection models have shown that SIX1 is both necessary and sufficient to promote metastasis [31,90,96,97].
Interestingly, although shRNA KD of Six1 in a highly metastatic murine mammary carcinoma cells (66cl4) significantly reduced ability to metastasize to the lung in vivo [96], the very few Six1 KD cells that were able to reach and colonize the lung had a dramatic increase in Six2 expression [96]. This led to studies examining the role of Six2 in metastasis. However, unlike SIX1, which plays a role in multiple aspects of metastatic progression, Six2 appears to act primarily at later steps of metastasis, specifically establishment and outgrowth at secondary sites such as the lung [93, 100]. To identify overlapping and distinct downstream targets of Six1 and Six2, which could explain their respective functions in metastasis, our lab performed microarray analysis in 66cl4 cells with either Six1 or Six2 KD. Surprisingly, despite Six1 and Six2 having almost identical sequence homology across their homeodomain and Six domain, no overlapping targets were identified (non-overlapping targets have not been extensively studied). This could be due to different KD efficiencies and the remaining presence of the other family member, thus masking any discernable differential or overlapping targets. One way to address this issue, is to generate Six1 and Six2 double knockout 66cl4 cells. In combination with our ChIP-seq analyses used as a reference for Six1 and Six2 baseline occupancy and putative co-regulated targets, downregulated genes in double KO cells compared to control cells (using RNA-seq) would be suggestive of a co-regulated gene. However, when a previous member of the lab attempted to generate stable Six1 and Six2 double KD 66cl4 cell lines, cells with loss of both genes did not survive, suggesting that the presence of either Six1 or Six2 is required for cell proliferation/survival.

As an alternative approach, I attempted to generate 66cl4 Six1 and Six2 Crispr/Cas9-mediated KO cells, and to couple gene expression information in these cells with Chip-Seq data from cells we made that expressed tags endogenously at the C-terminal ends of the Six1 and Six2 loci to identify non-overlapping downstream targets of Six1 or Six2 (Chapter IV). In theory, the KO environment should allow us to assess the function of one family member without the other. Using KO cells, we thought we could perform RNA-seq to identify targets specific to either
Six1 or Six2. However, upon KO of Six1, we observed compensatory upregulation of Six2. This suggests that generation of stable KOs might not be the best method to identify the most direct transcriptional targets because many redundant systems exist that could be activated when a gene is completely lost, as we observed when Six2 was upregulated in response to Six1 KO in the 66cl4 cells (Fig 4.3C). Therefore, we decided to perform siRNA KD of Six1 or Six2 in 66cl4 cells to more transiently remove activity, thus identifying the most critical and direct downstream targets of each protein. Additionally, we decided to also transiently KD SIX1 and SIX2 in the MDA-MB-231 cells to consider the human relevance when comparing and contrasting SIX1 and SIX2 function across the genome. Interestingly, although we were able to achieve significant KD of Six1 and Six2 mRNA and protein levels in the 66cl4 cells without significant effects on the other family member, this was not the case in the context of MDA-MB-231 cells where it appears that transient KD of one family member results in a corresponding decrease in the other (Fig 4.3 F and G). Taken together, these data suggest that depending on the context, Six1/SIX1 and Six1/SIX2 are differentially regulated by each other. These context-dependent, multilayered, aspects of regulation further support the need for these transcription factors to be tightly controlled to ensure proper development, and importantly, to inhibit transformation and tumor progression.

Previous studies have identified upstream regulators of Six2 during development. For instance, Gong and colleagues discovered that the Pax2-Hox11-Eya1 transcriptional complex directly regulates Six2 expression during nephrogenesis [234]. However, whether this complex, or other additional upstream regulators, are relevant in Six2-mediated metastasis requires further investigation. Additionally, studies examining the reciprocal relationship at the transcriptional level between Six1 and Six2 have yet to be fully explored. Understanding this aspect of regulation will provide additional insight into what phenotypes are impacted together or separately by SIX1 and SIX2 and how we might target them to inhibit metastasis. Importantly,
given the critical roles for Six1 and Six2 in tumor progression and metastasis in breast cancer, it would be interesting to see if this relationship is also relevant in other tumor types.

Importantly, although I have shown that SIX2 is critical for metastatic outgrowth of TNBC in the MDA-MB-231 cell line model [93], we have not tested the role of SIX1 in this cell line model. Therefore, we also plan to stably KD SIX1 in the MDA-MB-231 cells to confirm that, similar to SIX2, SIX1 plays a critical role in metastasis in this cell line model. If we determine that SIX1 also plays a role in metastasis in the MDA-MB-231 cells, then we also plan to generate double KO MDA-MB-231 cells to uncover shared downstream targets between SIX1 and SIX2 in the human context, where it is more clinically relevant. Overall, by combining RNA-seq analysis from singly KD (inducible KD) and double KO 66cl4 and MDA-MB-231 cells with ChIP-seq of endogenously FLAG-tagged Six1/SIX1 and Six2/SIX2 in 66cl4 cells and MDA-MB-231 cells, we should obtain a more comprehensive assessment of how SIX1 and SIX2 regulate their targets across the genome to mediate metastatic phenotypes (and which targets are shared vs not shared).

**Targeting Six2 to Inhibit Metastatic Burden**

Despite its importance in kidney development, few downstream targets of Six2 have been identified, particularly downstream targets that may serve as drug targets. As outlined in my thesis, I have identified Sox2 as critical for Six2-mediated metastatic outgrowth. This has clinical implications as metastasis is initiated early on in disease progression, suggesting that at diagnosis tumor cells have already disseminated. Therefore, therapeutic strategies targeting critical mediators of late-stage metastasis may be more effective at inhibiting disease progression. However, while it is important to understand how Sox2 mediates the effects of Six2 on late-stage metastasis, thus implicating activation of a stem cell axis, unfortunately, Sox2, like Six2, is a transcription factor and thus not an ideal drug target. Therefore, using a technique such as an in vivo shRNA screen downstream of SIX2 OE would be ideal to identify druggable downstream effectors of the SIX2/SOX2 axis that are required for the survival of metastatic cells.
and their subsequent outgrowth at secondary sites. Additionally, performing these experiments from cells isolated from an in vivo environment, rather than from in vitro tissue culture conditions, will account for the cooperativity between the lung microenvironment and tumor cells, as this has been shown to be critical for survival and outgrowth in the lungs [235].

Another therapeutic strategy for targeting SIX2-mediated metastasis may be through targeting upstream regulators such as the closely related family member, SIX1. For instance, the SIX/EYA complex serves as a core transcriptional regulator of target genes that are critical for both development and tumorigenesis [88,236]. Specifically, multiple studies by our group and others have shown that the SIX/EYA complex impacts important processes and pathways during tumor progression including proliferation, EMT and TGFβ signaling [95]. Additionally, multiple phenotypes induced by SIX1 during cancer progression and metastasis, have been shown to require parallel upregulation of EYA2 [88,95], suggesting that this complex in particular is critical for SIX1-mediated disease progression. Due to the critical role of the SIX1/EYA2 complex in breast cancer progression, our group has aimed to identify therapeutic strategies to disrupt the complex and thereby inhibit metastasis. Indeed, studies from our group for an upcoming publication, have identified a class of small compounds that disrupt the SIX1/EYA2 interaction and reverses SIX1-mediated phenotypes including EMT and TGFβ signaling. Importantly, the lead compound has been shown to be well-tolerated when administered to mice, and can significantly reduce metastases in vivo. Still, previous work from our laboratory has shown that although Six1 KD does reduce metastasis, Six2 upregulation occurs in the few cells that do metastasize. Therefore, it will be important not only to determine how our SIX1/EYA2 inhibitors affect SIX2, but also how Six2 compensation occurs in the context of loss of Six1 expression versus loss of Six1 transcriptional activity impacts metastasis. Additionally, Eya proteins have been shown to play critical roles together with Six2 during kidney development [223]. Therefore, it will also be important to determine whether the same is
true during late-stage metastasis. Still, because compensation is a common mechanism within the entire Six family, targeting the entire family may be the better way to inhibit metastasis.

A recent report from our group has identified SIX1 as playing a critical role in mediating non-cell autonomous effects between EMT and non-EMT cells to promote metastasis [31]. Specifically, SIX1 acts downstream of EMT-inducing transcription factors, TWIST and SNAIL1, in metastatic EMT cells to increase Hedgehog signaling in neighboring non-metastatic non-EMT cells. The increase in Hh signaling results in increased migratory and invasive phenotypes in the non-EMT cells, and in some cases these cells do not themselves convert to an EMT phenotype (based on marker expression) despite becoming more migratory and invasive. Since SIX1 and SIX2 have overlapping functions during metastasis including promoting stemness and mesenchymal phenotypes [31,44,93,100], it would be interesting to pursue whether SIX2 plays a role in any of the SIX1-mediated non-cell autonomous effects involving EMT. Importantly, studies examining non-cell autonomous effects of SIX1 were by mixing EMT and non-EMT cells in the primary tumor. Since SIX2 has been shown to be a critical mediator of late-stage metastasis specifically, future studies investigating how SIX2 mediates non-cell autonomous effects in the metastatic microenvironment could lead to valuable insight into how cell-cell interactions mediate survival and outgrowth at distant sites. Although SIX2 mainly acts cell autonomously to maintain nephron progenitor populations during kidney develop [104], a recent report has shown that the stromal compartment can act non-cell autonomously through Hedgehog signaling to restrict expansion of the SIX2+ nephron progenitor pool, and that this is SMO, GLI and TGFβ-dependent [237]. Importantly, in the context of tumorigenesis, stromal cells adjacent to the primary tumor site, and more recently in the distant sites, have been shown to provide a supportive niche for survival and tumor progression [238]. Therefore, a better understanding of the relationship between SIX2-expressing metastatic cells and the metastatic stromal compartment may provide a therapeutic window to inhibit survival at distant sites and reduce metastatic outgrowth.
Concluding Remarks

Advances in imaging technologies, genomic sequencing, mouse models of metastasis and, more recently, single-cell analysis have greatly improved our understanding of how metastatic tumor cells colonize secondary sites. This has led to substantial progress in the development of novel therapeutic strategies, such as mTOR and CDK4/6 inhibitors, to specifically treat advanced/metastatic breast cancer [239]. However, survival of patients with metastatic disease has remained largely marginal. This is mostly due to established metastases being inherently more resistant to current treatments compared to primary tumor cells. Given the clinical implications of metastatic outgrowth, it is unfortunate that current studies mainly focus on preventing the earlier steps of metastasis such as migration and invasion. It is well known that cancer cells enter the circulation long before a patient is diagnosed. Therefore, a better understanding of the molecular underpinnings of metastatic establishment and outgrowth are needed to develop more effective therapeutics to inhibit metastasis. The studies in this thesis further support this concept and provide novel mechanistic insight into how tumor cells successfully colonize distant sites to increase metastatic burden. Overall, the goal of this work is to not only better understand the molecular mechanisms that regulate this critical aspect of metastasis, but to also stimulate new avenues of investigation in order to identify additional processes and pathways that can be exploited to more effectively target metastatic outgrowth and improve patient survival.
REFERENCES


34. Zeng YA, Nusse R. Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. Cell Stem Cell. 2010;


126


209. Xu P-X. Six1 is required for the early organogenesis of mammalian kidney. Development. 2003;

210. Grifone R. Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. Development. 2005;


human iPS cells using EOS lentiviral vectors to select for pluripotency. Nat Methods. 2009;


