

1-4-2016

Function of the $\beta 4$ Integrin in Cancer Stem Cells and Tumor Formation in Breast Cancer: A Masters Thesis

Huayan Sun

University of Massachusetts Medical School

Follow this and additional works at: http://escholarship.umassmed.edu/gsbs_diss



Part of the [Cancer Biology Commons](#), and the [Neoplasms Commons](#)

Recommended Citation

Sun, H. Function of the $\beta 4$ Integrin in Cancer Stem Cells and Tumor Formation in Breast Cancer: A Masters Thesis. (2016). University of Massachusetts Medical School. *GSBS Dissertations and Theses*. Paper 814. DOI: 10.13028/M2588G. http://escholarship.umassmed.edu/gsbs_diss/814

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

**FUNCTION OF THE β 4 INTEGRIN IN CANCER STEM CELLS AND TUMOR
FORMATION IN BREAST CANCER**

A Masters Thesis Presented

By

Huayan Sun

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

January 4, 2016

CANCER BIOLOGY

**FUNCTION OF THE β 4 INTEGRIN IN CANCER STEM CELLS AND TUMOR
FORMATION IN BREAST CANCER**

A Masters Thesis Presented

By

Huayan Sun

The signatures of the Master's Thesis Committee signify
Completion and approval as to style and content of the Thesis

(Signature)

Arthur Mercurio, Thesis Advisor

(Signature)

Leslie Shaw, Chair of Committee

(Signature)

Mary Ellen Lane, Member of Committee

(Signature)

Michael Lee, Member of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences
signifies that the student has met all master's degree graduation requirements of
the school

(Signature)

Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

Molecular, Cell and Cancer Biology

January 4, 2016

ACKNOWLEDGEMENTS

This project would not have been possible without the support of many people. Many thanks to my adviser, Dr. Arthur Mercurio, who is very supportive and did his best to guide me along this process. Also many thanks to my TRAC committee members, Dr. Leslie Shaw, Dr. Stephen Jones, Dr. Michael Lee and Dr. Jaime Rivera, who gave me lots of helpful suggestions on my project. And finally, thanks to my parents and numerous friends who endured this process with me, always offering support and love.

Abstract

The integrin $\alpha6\beta4$ (referred to as $\beta4$) is expressed in epithelial cells where it functions as a laminin receptor. Integrin $\beta4$ is important for the organization and maintenance of epithelial architecture in normal cells. Particularly, $\beta4$ is shown to be essential for mammary gland development during embryogenesis. Integrin $\beta4$ also plays important roles in tumor formation, invasion and metastasis in breast cancer. However, the mechanism of how integrin $\beta4$ mediates breast tumor formation has not been settled. A few studies suggest that integrin $\beta4$ is involved in cancer stem cells (CSCs), but the mechanism is not clear. To address this problem, I examined the expression of $\beta4$ in breast tumors and its potential role involved in regulating CSCs. My data shows that $\beta4$ is expressed heterogeneously in breast cancer, and it is not directly expressed in CSCs but associated with a basal epithelial population. This work suggests that $\beta4$ can regulate CSCs in a non-cell-autonomous manner through the interactions between $\beta4^+$ non-CSC population and $\beta4^-$ CSC population. My data also shows that $\beta4$ expression is associated with $CD24^+CD44^+$ population in breast tumor. To further study the role of $\beta4$ in breast cancer progression, I generated a $\beta4$ reporter mouse by inserting a p2A-mCherry cassette before ITGB4 stop codon. This reporter mouse can be crossed with breast tumor models to track $\beta4^+$ population during tumor progression.

Table of Contents

Acknowledgement	III
Abstract	IV
List of Tables	VI
List of Figures	VII
List of Abbreviations	VIII
Chapter I: Introduction	1
Overview	1
Breast Cancer Overview	1
Tumor Heterogeneity and Breast Cancer Stem Cells	3
Integrins	5
Integrin $\alpha 6\beta 4$ in Mammary Epithelial Biology	7
Rational for the Thesis Work	9
Chapter II: Integrin $\beta 4$ Regulates Cancer Stem Cells in a Non-cell- autonomous Manner	12
Abstract	12
Introduction	12
Results and Discussion	14
Materials and Methods	23
Figures	26
Chapter III: Generate $\beta 4$ Reporter Mouse by CRISPR	44
Abstract	44

Introduction	44
Results and Discussion	45
Materials and Methods	47
Figures	49
Chapter IV: General Discussion	53
EMT and CSC	53
Intratumor Heterogeneity and Cell Cooperation	54
Integrin β 4 in Breast Tumor Metastasis	55
Integrin β 4 and PTHrP	56
Conclusion Remarks and Perspectives	57
Bibliography	60

LIST OF TABLES

Chapter II

- Table 1 Summary of RNA-seq of $\beta 4^{\text{hi}}$ and $\beta 4^{\text{lo}}$ populations in Src-transformed MCF10A system
- Table 2 Summary of RNA-seq of $\beta 4^+$ and $\beta 4^-$ populations in PDX tumor
- Table 3 Summary of RNA-seq of genes downregulated in $\beta 4$ knockout cells
- Table 4 Sequences of the primers used in this study

Chapter III

- Table 5 Sequences of the primers used in this study

LIST OF FIGURES

Chapter II

- Figure 1 Integrin β 4 is expressed heterogeneously in breast tumor.
- Figure 2 Integrin β 4 regulates PTHrP expression.
- Figure 3 Knocking down integrin β 4 affects CSC function.
- Figure 4 Integrin β 4 is associated with CD24⁺CD44⁺ population in breast tumor.

Chapter III

- Figure 5 Generate integrin β 4 reporter mouse using CRISPR.

LIST OF ABBREVIATIONS

AML	Acute myeloid leukemia
APC	Adenomatous polyposis coli
ATCC	American Type Culture Collection
BM	Basement membrane
BMP	Bone morphogenetic protein
CD	Cluster of differentiation
cDNA	Complementary DNA
CSC	Cancer stem cell
DCIS	Ductal carcinoma in situ
ECM	Extracellular matrix
EMT	Epithelial-to-Mesenchymal Transition
ER	Estrogen receptor
FACS	Fluorescence activated cell sorting
FN	Fibronectin
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GFP	Green Fluorescence Protein
HD	Hemidesmosome
HER2	Human epidermal growth receptor
LM	Laminin
MCF	Michigan Cancer Foundation
MDA	MD Anderson

MET	Mesenchymal-to-Epithelial Transition
MMP	Matrix metalloproteinase
NF- κ B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NMuMG	Normal murine mammary gland
NSG	NOD scid gamma
PDX	Patient-derived xenografts
PI3K	PIP3 kinase
PR	Progesterone receptor
qPCR	Quantitative realtime PCR
RANKL	Nuclear factor kappa B ligand
RNA-seq	RNA deep sequencing
SUM	The University of Michigan Human Breast Cancer Cell Lines
TGF- β	Transforming growth factor β
UMMS	University of Massachusetts Medical School
VEGF	Vascular endothelial growth factor

Chapter I

Introduction

Overview

This thesis focuses on the role of integrin $\alpha6\beta4$ in cancer stem cells and tumor initiation in breast cancer. This issue is significant because integrin $\alpha6\beta4$ as an adhesion receptor is important for organization and maintenance of epithelial architecture. It also can mediate a variety of signaling pathways to promote tumor progression (Soung et al., 2011).

Breast Cancer Overview

Breast cancer is a major health problem and the most commonly diagnosed cancer in women worldwide. The American Cancer Society estimates that about 231,840 new cases of female invasive breast cancer and 60,290 new cases of breast carcinoma in situ will be diagnosed in 2015 (Siegel et al., 2015). Breast cancer alone is expected to account for 29% of all new cancers among women in the United States (Siegel et al., 2015). It is also the second leading cause of cancer death in the United States, with about 40,290 deaths expected in 2015 for women representing 15% of all cancer-related deaths among women (Siegel et al., 2015).

It is believed that breast cancer develops stepwise from normal breast epithelial cells to atypical hyperplasia, ductal carcinoma in situ (DCIS), invasive ductal

carcinoma and metastatic breast cancer. Multiple genetic and epigenetic alterations may occur during this process. DCIS, though noninvasive, is greatly associated with invasive breast cancer and has a very high chance to develop into invasive breast cancer if at high grade (Allred, 2010; Ward et al., 2015). Studies have shown the mechanism of how DCIS turns into invasive breast cancer (Hu et al., 2008). However, early molecular alterations preceding DCIS have not been clearly demonstrated so far (Rivenbark et al., 2013).

Breast cancers are classified by immunohistochemistry to detect the expression of certain protein biomarkers, including estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth receptor 2 (HER2) and various cytokeratins. Based on this clinical classification, breast cancer can be classified as ER⁺/PR⁺/HER2⁻, ER⁺/PR⁺/HER2⁺, ER⁻/PR⁻/HER2⁺, and ER⁻/PR⁻/HER2⁻ breast cancers. Breast cancer can also be classified by gene expression (Perou et al., 2000; Sorlie et al., 2003; Sorlie, 2004) . Based on this molecular classification, breast cancer can be classified as six molecular subtypes, luminal A (ER⁺), luminal B (ER⁺/HER2⁺), HER2⁺, basal-like, claudin-low, and normal-like breast cancers (Rakha et al., 2008; Rivenbark et al., 2013).

ER⁺ and HER2⁺ breast cancers can be treated by effective targeted therapy of anti-estrogen treatment (tamoxifen) or anti-HER2 drug (trastuzumab), respectively, both in conjunction with chemotherapy (Mukai, 2010; Rimawi et al.,

2015). These targeted treatments have improved survival and the long-term outcomes for ER⁺ and HER2⁺ patients dramatically. However, basal-like and claudin-low molecular subtypes of breast tumors lack the expression of ER, PR as well as amplification of HER2 (ER⁻PR⁻HER2⁻), and are therefore also referred to as triple-negative breast cancers (TNBCs). TNBCs tend to be more aggressive compared to ER⁺ breast cancers, and there are no effective targeted therapies (Chacón and Costanzo, 2010; Dent et al., 2007; Li et al., 2013). Though they respond to some chemotherapies, the patients often have poor recurrence-free and overall survival outcomes (Kassam et al., 2009; Kennecke et al., 2010; Liedtke et al., 2008).

Tumor Heterogeneity and Cancer Stem Cells

A single tumor harbors a mixture of different subpopulations with distinct molecular properties (Meacham and Morrison, 2013; Skibinski and Kuperwasser, 2015). Two models are used to explain this intratumor heterogeneity, the clonal evolution model and the cancer stem cell model, which are not mutually exclusive but could also be integrated together (Campbell and Polyak, 2014; Kreso and Dick, 2014).

The clonal evolution model states that tumor cells acquire various mutations over time and cancer could initiate in any random single cell. Certain mutation could give cells growth advantage over other normal cells, and these cells will become

more aggressive and may leave a large number of offspring by chance. Thus, during tumor progression, any cell within the tumor could potentially be the one that drives tumor progression and results in therapy resistance and recurrence since mutations will occur in random cells over time.

The cancer stem cell model states that there is a small subpopulation of cells in the tumor that have stem cell-like properties, called cancer stem cells (CSCs). CSCs have the ability of self-renewal and can also differentiate into different cell types. They are thought to be responsible for cancer recurrence. Current therapies kill tumor cells but not CSCs, and CSCs can proliferate extensively. Thus, despite good therapeutic response by the bulk tumor, it shrinks but will grow back. Moreover, the recurred tumor will gain resistance to the previous therapies.

In this work, I focused on the CSC model. TNBCs harbor a relatively high frequency of CSCs compared with other breast cancer subtypes. This could be one of the reasons why TNBCs are more aggressive with poor recurrence-free and overall survival outcomes. Despite the evidence supporting the existence of CSCs, much less is known about the origin of these cells and the mechanisms that regulate their functions.

CSCs can be isolated by cell-surface markers, and a series of assays can be conducted to examine their stem-cell properties. The gold standard measure of CSCs is examining their ability to repopulate new tumors and maintain the long-term clonal growth by either transplantation into serial recipients or in situ tracking. Initial studies on CSCs were conducted in acute myeloid leukemia (AML), and the CD34⁺CD38⁻ subpopulation was identified as leukemia-initiating cells (Bonnet and Dick, 1997; Lapidot et al., 1994). After that many studies on CSCs in various solid tumors followed (Collins et al., 2005; Eramo et al., 2008; Hermann et al., 2007; Li et al., 2007; O'Brien et al., 2007; Patrawala et al., 2006; Prince et al., 2007; Ricci-Vitiani et al., 2007; Singh et al., 2004; Wu et al., 2007). The first identification of CSCs in a solid tumor was in human breast cancer (Al-Hajj et al., 2003). A CD24⁻CD44⁺ subpopulation was identified as CSCs in breast cancer by showing this subpopulation alone is sufficient to generate new tumors in immune-deficient mice while the other three subpopulations (CD24⁺CD44⁺, CD24⁻CD44⁻, and CD24⁺CD44⁻) cannot (Al-Hajj et al., 2003).

Integrins

In this context, we are interested in integrins in breast cancer. Integrins are transmembrane receptors that form heterodimers composed of α and β subunits. In mammals, the integrin family comprises 18 α and 8 β subunits that bind noncovalently, generating 24 unique integrins, and they bind to specific ligands including collagen, fibronectin, vitronectin and laminin (Hynes, 2002). Integrin

subunits span the cell membrane with a large extracellular domain, a single transmembrane domain and a short cytoplasmic tail of 40-70 amino acids. The exception is the $\beta 4$ subunit, which has a cytoplasmic domain of 1088 amino acids, one of the largest known cytoplasmic domains of any membrane protein (Campbell and Humphries, 2011; Srichai and Zent, 2010; Takada et al., 2007).

As adhesion receptors, integrins mediate cell-cell adhesions as well as cell adhesion to extracellular matrix proteins. All integrins except integrin $\alpha 6\beta 4$ are connected to actin, while the integrin $\beta 4$ subunit connects to keratin intermediate filaments. They can form adhesion structures such as focal adhesion and hemidesmosome which connect ligands in the ECM with the cytoskeleton. Other than functioning as adhesion receptors, integrins also mediate signaling transduction and play important roles in cell proliferation, division, apoptosis, survival, motility, and differentiation (Guo and Giancotti, 2004; Levental et al., 2009; Glukhova and Streuli, 2013). Each integrin plays distinct, nonredundant roles shown by the phenotypes of genetic knockout mice, such as in development, vasculogenesis and immune responses (Takada et al., 2007).

Unlike other transmembrane receptors, integrins mediate signaling in a bidirectional way (Hynes, 2002). They connect the extracellular and intracellular environments by binding ligands in the ECM and adaptor molecules in the cytoplasm through outside-in signaling or inside-out signaling. Outside-in

signaling involves the extracellular domain of integrins binding to their ligands to promote conformational changes first, which is followed by integrin clustering and the assembly of large intracellular adhesion complexes. Inside-out signaling involves an internal signaling molecule binding to the cytoplasmic tail of integrins first, which then induces conformational changes and influences the binding of integrins to their ligands. Moreover, integrins also participate in signaling crosstalk, where in addition to the activation of their own downstream signaling cascades, they interact with and regulate the activation of growth hormones and G-protein coupled receptor-induced signaling pathways (Mariotti et al., 2001; Trusolino et al., 2001; Guo et al., 2006).

Integrin $\alpha6\beta4$ in Epithelial Biology

Integrin $\alpha6\beta4$ is referred to as $\beta4$ integrin since its only partner is the $\alpha6$ subunit. Integrin $\beta4$ is special among all integrin subunits in that it has an extremely large cytoplasmic domain and it connects to keratin intermediate filaments instead of the actin cytoskeleton (Borradori and Sonnenberg, 1999; Dowling et al., 1996; Green and Jones, 1996). It is expressed primarily on the basal surface of most epithelia and in a few other cell types, preferentially binding to most of the known laminins (Lee and Streuli, 2014; Mercurio, 1995). In normal cells, $\beta4$ integrin is essential for the organization and maintenance of epithelial architecture by serving as an important component of the stable and rigid adhesive structures termed hemidesmosomes (HDs) that link the intermediate filament cytoskeleton

with laminins in the basement membrane (BM) (Borradori and Sonnenberg, 1999; Dowling et al., 1996; Green and Jones, 1996).

Homozygous knockout reveals the function of $\beta 4$ in maintaining the integrity of epithelia, especially the epidermis. In the absence of $\beta 4$ expression, the epidermis detaches from the dermis in response to mechanical stress, which leads to death of $\beta 4$ knockout mice shortly after birth (Dowling et al., 1996; van der Neut et al., 1996; 1999a). This critical role for $\beta 4$ derives from its ability to mediate the formation of HD on the basal cell surface (Borradori and Sonnenberg, 1999; Green and Jones, 1996). HDs are most pronounced in stratified epithelia, but rudimentary forms exist in most epithelia (Green and Jones, 1996).

Although many studies have been conducted on the contribution of $\beta 4$ to epithelial biology in the past two decades, its role in mammary gland development is unclear and confusing (Klinowska et al., 2001). During embryogenesis, integrin $\beta 4$ is expressed at least from embryonic day 12.5 (E12.5), restricted in the mammary epithelial bud which begins to elongate to form the rudimentary mammary tree after E14.5. The adult mammary glands are composed of two layers of cells: luminal cells and myoepithelial cells. Integrin $\beta 4$ is expressed primarily in myoepithelial cells and anchors them to the basement membrane (Gjorevski and Nelson, 2011). Seminal studies by Bissell and others

using three-dimensional cultures revealed the importance of laminins in mammary morphogenesis and the contribution of the $\alpha 6$ integrins ($\alpha 6\beta 4$ and $\alpha 6\beta 1$) to this process (Nisticò et al., 2014). Our previous work shows that integrin $\beta 4$ regulation of parathyroid hormone related protein (PTHrP) contributes to the ability of this integrin to regulate mammary bud development during embryogenesis (Li et al., 2015).

Rational for the Thesis Work

In breast cancer, $\beta 4$ has been shown to be important for tumor formation, invasion and metastasis (Soung et al., 2011). The role of $\beta 4$ in tumor migration and invasion has been extensively studied. In cancer cells, $\beta 4$ integrin switches from associating with HD to interacting with actin cytoskeleton and redistributes to cell protrusions, which can stimulate carcinoma migration and invasion (Lipscomb and Mercurio, 2005; Rabinovitz and Mercurio, 1997; Rabinovitz et al., 1999). Integrin $\beta 4$ also promotes tumor invasion by regulating several downstream effector molecules and activating various signaling pathways. Integrin $\beta 4$, together with growth factor receptors, initiates signaling cascades involving Shp2 and Src, which leads to activation of PI3K and Ras pathways (Shaw et al., 1997; Mercurio, 2003; Bertotti et al., 2006). As a result, integrin $\beta 4$ enhances expression of cancer related genes such as VEGF, ERBB3, NFAT, NF- κ B, STAT3 and c-Jun at the translational or transcriptional levels (Chung et

al., 2002; Folgiero et al., 2007; Gerson et al., 2012a; Jauliac et al., 2002; Yoon et al., 2006; Zahir et al., 2003).

Integrin $\beta 4$ has also been found to be essential for breast tumor formation. Our lab showed that $\beta 4$ is important for breast tumor growth by demonstrating that SUM-159 breast cancer cells with $\beta 4$ knocked down generate fewer and smaller tumors in vivo compared to control cells (Lipscomb et al., 2005). Moreover, this tumor growth phenotype could be rescued by overexpressing $\beta 4$ back into the knockdown cells. Another study also shows that $\beta 4$ promotes tumor formation by amplifying ERBB2 signaling (Guo et al., 2006). However, the mechanism of how integrin $\beta 4$ regulates tumor formation still needs to be further investigated.

The role of CSCs in tumor formation has arisen much interest in the field studying tumor initiation. Several integrins such as $\alpha 6$ and $\beta 1$ have been shown to be stem cell markers (Pontier and Muller, 2009). However, less is known about the role of $\beta 4$ in CSCs. Data in our lab suggests that $\beta 4$ may be involved in stem cell properties by showing that $\beta 4$ knockout mammary buds cannot develop into mature mammary glands when transplanted into immune-deficient mice (Li et al., 2015). Moreover, several studies suggest that $\beta 4$ is involved in CSCs of breast cancer and lung cancer in the past two years (Zheng et al., 2013; Vieira et al., 2014). In non-small-cell lung cancer, integrin $\beta 4$ has been identified as a CSC

marker by showing that a rare population of CD24⁺ITGB4⁺Notch^{hi} cells are able to form new tumors in serial transplantation assay (Zheng et al., 2013). Another study shows that P-cadherin signaling induces stem cell properties in basal-like breast cancer cells through integrin β 4 (Vieira et al., 2014). This study demonstrates that P-cadherin directly interacts with β 4 and regulates the expression of both integrin α 6 and β 4 subunits, which are essential for stem cell properties shown by mammosphere assay (Vieira et al., 2014). However, the mechanism of β 4 regulation on CSCs is still unclear. Given these data and the importance of CSCs in tumor initiation, I was interested in investigating the role of integrin β 4 in CSCs and tumor initiation in breast cancer.

Chapter II

Integrin β 4 Regulates Breast Cancer Stem Cells in a Non-cell-autonomous Manner

Abstract

Integrin β 4 is not only very important for the organization and maintenance of epithelial architecture in normal epithelia, but it also plays essential roles in cancer. It has been shown to be important for tumor formation, invasion and metastasis. However, the mechanism of how integrin β 4 is involved in tumor formation is still unclear. My study shows that integrin β 4 is heterogeneously expressed in breast cancer and the β 4⁺ population is critical for tumor formation. It further shows that β 4 affects CSCs in a non-cell-autonomous manner and I discuss the possibility of its regulation in CSCs through PTHrP paracrine signaling.

Introduction

Integrin β 4 has been shown to be important for tumor formation and survival in breast cancer (Guo et al., 2006; Lipscomb et al., 2005). However, the mechanism involved is still not very clear. Breast CSCs were identified as CD24⁻CD44⁺ by showing that it is necessary and sufficient to form new tumors in vivo while other three subpopulations (CD24⁻CD44⁻, CD24⁺CD44⁻, CD24⁺CD44⁺) cannot (Al-Hajj et al., 2003). Our lab showed that β 4 is important for mammary

gland development during embryogenesis through regulating parathyroid hormone-related protein (PTHrP) expression (Li et al., 2015). In this study, we knocked out integrin $\beta 4$ using MMTV-Cre and found that the mammary bud elongation was blocked, which is very similar to the phenotype of PTHrP knockout mice. Upon further investigation, we found that PTHrP expression was decreased in $\beta 4$ knockout mammary bud, and the mammary bud elongation phenotype could be rescued by adding additional PTHrP when $\beta 4$ knockout mammary bud was cultured in vitro. Moreover, in the study, the fact that $\beta 4$ knockout mammary buds cannot develop into mature mammary glands when transplanted into immune-deficient mice suggests that $\beta 4$ could be potentially involved in stem cell properties. Compared to integrin $\alpha 6$ and $\beta 1$, not much is known about $\beta 4$ in mammary stem cells (Pontier and Muller, 2009). Integrin $\beta 4$ has been indicated in progenitor cells (Chapman et al., 2011; Yoshioka et al., 2013). Moreover, $\beta 4$ is identified as a marker for CSCs in non-small-cell lung cancer (Zheng et al., 2013). In breast cancer, a study has shown that P-cadherin regulates CSC functions through integrin $\beta 4$ (Vieira et al., 2012; 2014). Given all these studies and that CSCs are important for tumor formation, I wanted to investigate whether integrin $\beta 4$ is involved in CSC functions and the mechanisms involved.

Previously, tumor formation is believed to be associated with epithelial-to-mesenchymal transition (EMT) (Thiery et al., 2009). It has been shown that EMT

generates cells with properties of stem cells, and CSCs express mesenchymal markers with decreased epithelial markers (Mani et al., 2008). However, integrin $\beta 4$ is a gene associated with epithelial cells, and our lab showed that $\beta 4$ is downregulated dramatically during EMT induced by TGF- β or Snail (Chang et al., 2013; Yang et al., 2009). In another system that our lab has been using, we sorted out two populations from Src-transformed MCF10A cells based on $\alpha 6$ staining (Goel et al., 2014). $\alpha 6^{\text{hi}}$ population has epithelial traits while $\alpha 6^{\text{lo}}$ population is mesenchymal based on both cell morphology and molecular markers (Goel et al., 2014). Moreover, $\alpha 6^{\text{lo}}$ population has higher CSC properties but lower $\beta 4$ expression compared to $\alpha 6^{\text{hi}}$ population. Based on these observations, it seems that $\beta 4$ expression is not directly expressed by CSCs. Since $\beta 4$ is important for tumor formation while it is not expressed in CSCs, I propose that $\beta 4$ is expressed heterogeneously in breast tumors and this $\beta 4^+$ non-CSC subpopulation is important for tumor formation.

Results and Discussion

2.1 Integrin $\beta 4$ is expressed heterogeneously in breast tumors.

We have a patient-derived xenograft (PDX) model in our lab. I used PDX tumors derived from two TNBC patients labeled as #1126 and #1258, respectively (Fig 1A). $\beta 4$ expression was first examined in these two tumor samples by immunofluorescence (IF) staining (Fig 1B). The result shows that $\beta 4$ is expressed

heterogeneously by a subpopulation of cells in the tumor. To examine the molecular difference between $\beta 4^+$ and $\beta 4^-$ populations, I sorted out $\beta 4^+$ and $\beta 4^-$ cells by FACS in PDX tumor #1126 (Fig 1C).

qPCR data shows that the $\beta 4^+$ subpopulation expresses a lower level of CSC marker OCT4 compared to $\beta 4^-$ subpopulation (Fig 1D), which confirms that $\beta 4$ is not directly expressed by CSCs. Since our lab showed that $\beta 4$ is associated with basal-like breast cancer (Lu et al., 2008) and it is downregulated during EMT together with epithelial markers (Chang et al., 2013; Yang et al., 2009), I checked the expression of several basal and epithelial markers. The $\beta 4^+$ subpopulation expresses higher levels of basal and epithelial genes compared to the $\beta 4^-$ subpopulation (Fig 1D).

The Ewald group showed that cell dissemination during invasion needs an epithelial program (Shamir et al., 2014). Moreover, a basal epithelial population is shown as leader cells at the invading front in breast cancer (Cheung et al., 2013). These studies indicate that not all cells undergo EMT during tumor progression, and there is a basal epithelial population driving tumor invasion. Interestingly, two back-to-back papers recently showed that EMT is not required for metastasis and epithelial cells can intravasate into the blood system and form metastasis (Cheung et al., 2013; Fischer et al., 2015; Shamir et al., 2014; Zheng et al., 2015). These results contradict with the previous model that cells need to

undergo EMT and become mesenchymal-like to break away from the primary site (Thiery et al., 2009). Overall, these studies address the important role of a basal-epithelial population within the tumor during tumor progression other than the cells that undergo EMT which are previously believed to drive tumor progression.

In cancer, integrin $\beta 4$ switches from associating with HD to connecting with actin and redistributing to cell protrusions (Lipscomb and Mercurio, 2005; Rabinovitz and Mercurio, 1997; Rabinovitz et al., 1999). Moreover, the invading front of several carcinomas are enriched in the expression of $\beta 4$ and its ligand laminin-332 (Kim et al., 2011a; Pyke et al., 1995; Xue et al., 2011). $\beta 4$ has been shown to drive tumor invasion by regulating several downstream effector molecules and signaling pathways (Gerson et al., 2012; Shaw et al., 1997). My data also shows that the $\beta 4^+$ population is associated with a basal epithelial program (Fig 1D). Based on all these findings, I propose that $\beta 4^+$ population may be the leader cells during invasion that Ewald group reported in their studies (Cheung et al., 2013; Shamir et al., 2014). To examine whether the $\beta 4^+$ population is present at the invasive front to drive tumor invasion and to track $\beta 4$ expression during tumor progression and metastasis, I made a $\beta 4$ reporter mouse using CRISPR (see Chapter III). In this work, I'm interested in investigating whether the $\beta 4^+$ subpopulation is important for tumor formation and the mechanism involved other than their potential role in tumor invasion.

2.2 Integrin $\beta 4$ regulates CSCs in a non-cell-autonomous manner through PTHrP paracrine signaling

I checked some genes that are important for CSC functions, and I found an interesting gene expression pattern that $\beta 4^-$ and $\beta 4^+$ cells have in certain important paracrine signaling pathways for example, the notch signaling pathway. qPCR data show that $\beta 4^+$ cells express higher level of JAG1 while $\beta 4^-$ cells express higher NOTCH3 as well as HEY1, which is a downstream target in Notch signaling (Fig 1D). Notch signaling is established to be important for CSC self-renewal (Bouras et al., 2008; Dontu et al., 2004; Harrison et al., 2010). One study showed that NF-kB regulates CSCs in a non-cell-autonomous manner through Notch paracrine signaling. CSCs express notch3 receptors, and NF-kB regulates JAG1 expression in non-CSCs (Yamamoto et al., 2013). This non-cell-autonomous regulation model is very similar with what my results indicate. Thus, I proposed that $\beta 4^+$ non-CSCs can interact with $\beta 4^-$ CSCs through paracrine signaling.

To show that $\beta 4^+$ and $\beta 4^-$ populations can interact with each other, I explored the Src-transformed MCF10A system in our lab where $\beta 4^{hi}$ and $\beta 4^{lo}$ populations were isolated (Goel et al., 2014). It has been shown that the $\beta 4^{lo}$ population has CSC properties compared to the $\beta 4^{hi}$ population (Goel et al., 2014). Part of the results from RNA-seq on $\beta 4^{hi}$ and $\beta 4^{lo}$ populations is shown in Table 1. I labeled $\beta 4^{hi}$

population with GFP and $\beta 4^{lo}$ population with mCherry. I mixed them together and plated them in single cells on Matrigel. The result shows that they grow into cell clusters together in 3D culture (Fig 2A). This indicates that $\beta 4^{hi}$ and $\beta 4^{lo}$ populations can interact with each other and grow together in 3D culture.

Our lab showed that integrin $\beta 4$ regulates PTHrP expression during embryonic mammary gland development and there are many signaling pathways during embryogenesis that have parallel signaling pathways in cancer, I checked the expression of PTHrP and its receptor PTH1R in $\beta 4^{+}$ and $\beta 4^{-}$ populations sorted from PDX tumors (Li et al., 2015). $\beta 4^{+}$ cells express higher PTHrP while $\beta 4^{-}$ cells express higher PTHrP as well as BMP4 which is a downstream target of PTH1R signaling (Fig 1D). It is known that PTHrP is expressed primarily by epithelial-derived cells, and its receptor PTH1R expression is restricted to mesenchymal cells. Thus PTH1R signaling is activated in the paracrine manner. PTHrP promotes increased tumor cell growth in breast cancer lines as well as in vivo (Dittmer et al., 2006). PTHrP also drives breast tumor progression and metastasis examined by mouse model (Li et al., 2011). However, no one has shown that PTHrP is involved in CSCs functions.

To better characterize $\beta 4^{+}$ and $\beta 4^{-}$ population within a breast tumor, RNA-seq was performed on $\beta 4^{+}$ and $\beta 4^{-}$ populations sorted from PDX tumor #1258 (Fig 1E). Part of the results was shown in Table 2. I first used mouse lineage markers

CD31, CD45 and Ter-119 to exclude the potential mouse stromal cells recruited to the PDX tumor, and then sorted out the $\beta 4^+$ and $\beta 4^-$ cells in Lin^- cells. I confirmed $\beta 4$ expression in $\beta 4^+$ and $\beta 4^-$ subpopulations before sending them for RNA-seq (Fig 1E). RNA-seq data is similar with the previous qPCR data in that $\beta 4^+$ population expresses lower level of CSC markers while is associated with a basal epithelial program (Table 2).

PTHrP expression is always associated with $\beta 4$. I checked this association in another two systems. In the Src-transformed MCF10A system, PTHrP is associated with $\beta 4^{\text{hi}}$ population (Table 1). I overexpressed VEGF promoter with GFP tag in MDA-MB-231 cells and sorted out GFP^+ cells which have higher VEGF activity (VEGF-high). Our lab has shown that VEGF-high cells have CSC properties (Goel et al., 2014). I checked $\beta 4$ and PTHrP expression in VEGF-low and VEGF-high cells by qPCR, and found that PTHrP expression is associated with $\beta 4$ in VEGF-low population (Fig 2B). I knocked down $\beta 4$ with siRNA in MCF10A cells, and found that PTHrP expression decreased at both mRNA and protein level (Fig 2C). When I ligated integrin $\alpha 6\beta 4$ by plating the cells on laminins, PTHrP expression increase (Fig 2D). This indicates that PTHrP expression is directly induced by ligation of integrin $\alpha 6\beta 4$. Based on these results, I propose that $\beta 4^+$ non-CSCs and $\beta 4^-$ CSCs interact with each other through

paracrine signaling, which potentially involves $\beta 4$ regulation on PTHrP expression.

2.3 Integrin $\beta 4$ is essential for CSC functions.

Our lab has shown that $\beta 4$ integrin is important for mammary tumor formation and growth by knocking down $\beta 4$ using shRNA in SUM-159 cells (Lipscomb et al., 2005). The sh $\beta 4$ cells form fewer and smaller tumors when injected into the mammary fat pad of immunocompromised mice compared to shGFP cells. Overexpressing $\beta 4$ back in to the sh $\beta 4$ cells can rescue the tumor formation phenotype. To investigate the mechanism involved, I studied the same system by knocking down $\beta 4$ in SUM-159 cells using shRNAs (Fig 3A). Since CSCs are known to be very important for tumor formation, I first examined if the tumor formation phenotype we got after $\beta 4$ knocking down was due to decrease in the amount of CSC population. Interestingly, I found that knocking down $\beta 4$ doesn't decrease the CD24⁻CD44⁺ population which is identified as CSCs in breast cancer (Al-Hajj et al., 2003), but it dramatically reduces CD24⁺CD44⁺ population (Fig 3B).

To further investigate the mechanism, I sorted out CD24⁻CD44⁺ and CD24⁺CD44⁺ populations in SUM-159 cells, and checked gene expression in these two populations by qPCR. The data show that CD24⁻CD44⁺ CSCs express lower levels of integrin $\beta 4$ but higher PTH1R. The CD24⁺CD44⁺ population

expresses higher level of basal and epithelial markers as well as PTHrP (Fig 3C). I cultured these two populations separately for 1 week and checked CD24 and CD44 expression by FACS again. The CD24⁻CD44⁺ population can give rise to the cell profiles in the parental SUM-159 cells, both CD24⁻CD44⁺ and CD24⁺CD44⁺ populations (Fig 3D). However, CD24⁺CD44⁺ population can only give rise to itself (Fig 3D). As one major property of stem cells, they should be able to differentiate into different types of cells. This result confirms that CD24⁻CD44⁺ population is CSCs while CD24⁺CD44⁺ population doesn't have this stem cell property. Moreover, this result also indicates that integrin β 4 is essential for CSC functions. As it is discussed above, I expected that intact CSC population (CD24⁻CD44⁺) in SUM-159 cells should be able to differentiate into all cell subtypes in SUM-159 cells, just as shown in Fig 3D. However, after knocking down β 4, there was a dramatic reduce in CD24⁺CD44⁺ population (Fig 3B), which means that CD24⁻CD44⁺ cells in sh β 4 cells failed to give rise to CD24⁺CD44⁺ population in the parental SUM-159 cells. This indicates that CSC function of CD24⁻CD44⁺ population is defective without β 4. Thus, I proposed that integrin β 4, though it is not directly expressed by CSCs, is essential for maintaining CSC functions.

Since shRNAs don't fully knock down β 4, I used CRISPR to target β 4 in MDA-MB-231 cells. I sorted out single cells and grew them into colonies before selecting the clones with β 4 completely knocked out. I selected one single cell

clone from each of the two guide RNAs sg3 and sg4, and performed RNA-seq (Fig 3E). $\beta 4$ expression was examined in control and two knockout clones before RNA-seq (Fig 3F). Part of the RNA-seq results is shown in Table 3.

2.4 Integrin $\beta 4$ is associated with $CD24^+CD44^+$ population in breast tumor.

I checked $\beta 4$ expression by FACS in a Luminal A type human breast tumor.

There are $\beta 4$ -low and $\beta 4$ -high populations, and $\beta 4$ -low cells are mostly $CD24^-CD44^+$ while $\beta 4$ -high cells are mostly $CD24^+CD44^+$ (Fig 4A). If the tumor cells are analyzed first by $CD24$ and $CD44$ expression, there are two subpopulations $CD24^-CD44^+$ and $CD24^+CD44^+$ (Fig 4A). Interestingly, $CD24^+CD44^+$ subpopulation has higher level of $\beta 4$ expression compared to $CD24^-CD44^+$ subpopulation.

The model I propose from my data is described in fig 4B. $\beta 4$ is expressed heterogeneously in breast tumor with $\beta 4^+$ and $\beta 4^-$ populations. The $\beta 4^+$ population is $CD24^+CD44^+$ expressing basal and epithelial markers, and the $\beta 4^-$ population is $CD24^-CD44^+$ with CSC properties. $\beta 4^+$ non-CSCs are important for and interact with $\beta 4^-$ CSCs by providing the ligands for paracrine signaling. Together, $\beta 4$ regulates CSC functions in a non-cell-autonomous manner, and both $\beta 4^+$ and $\beta 4^-$ populations are needed for tumor formation.

Materials and Methods

Antibodies. $\beta 4$ (505, for western blot) (Rabinovitz et al., 1999); $\beta 4$ (4399B, for FACS, IF and IHC); actin (Sigma-Aldrich); phycoerythrin-conjugated anti-CD44 (BD Bioscience); Alexa Fluor 647 anti-CD24 (BD Bioscience); PTHrP (a gift from Dr. Jiarong Li).

Cell lines. SUM-159 cells were provided by Dr. Steve Ethier (Medical College of South Carolina). MCF10A cells were obtained from the Barbara Ann Karmanos Cancer Institute. MDA-MB-231 cells were purchased from American Type Culture Collection. ER-Src-transformed MCF10A cells were provided by Dr. Kevin Struhl (Harvard Medical School). Isolation of the CD24⁻CD44⁺ population from these Src-transformed MCF10A cells and characterization of the $\alpha 6^{\text{hi}}$ and $\alpha 6^{\text{lo}}$ populations have been described (Goel et al., 2014). The SUM-159 shGFP and $\beta 4$ -shRNAs (sh68 and sh70) human breast carcinoma cell lines were generated and maintained as described previously (Chung et al., 2004). MDA-MB-231 cells stably expressing GFP under control of the VEGF promoter were generated and maintained as described previously (Goel et al., 2014). MCF10A siCtrl and si $\beta 4$ cells were generated as described previously (Gerson et al., 2012b). A6hi and a6low cells stably expressing GFP and mCherry, respectively were generated by PCDH-GFP and PLV-mCherry (Addgene) lentiviral plasmids and selected on puromycin. MDA-MB-231 cells stably expressing sgRNAs were generated by cloning indicated sgRNA sequences into v2 lentiviral backbone

plasmid (Addgene). Single-cell sorting was performed to select clones with $\beta 4$ knockout.

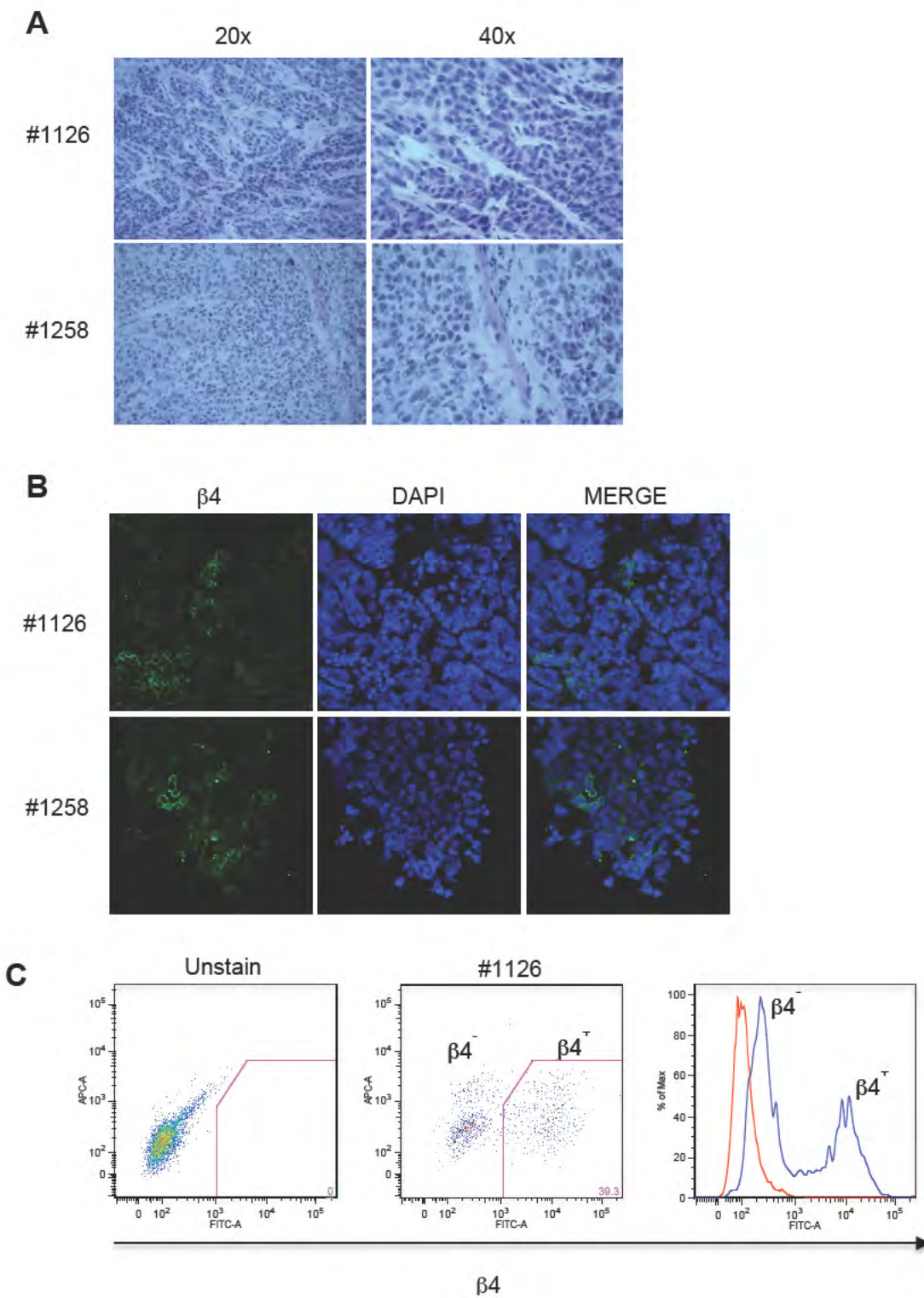
RNA-seq. RNA was extracted from the indicated cells and sent to Applied Biological Materials (ABM) for quantification and sequencing. Sequencing was performed on the paired ends of the mRNA fragment with a reading length of 75bps. Each sample had a sequencing depth of 8 million.

Cell-Based Assays. Flow cytometry was used to analyze surface expression of the $\beta 4$ integrin, CD44, CD24 and lineage markers (CD31, CD45, Ter-119) in breast cancer cell lines and breast tumors. Cells in culture were detached, stained with antibodies and analyzed by FACS. PDX tumors and primary human breast tumors were dissociated into single cell suspensions as described before staining and FACS analysis (Goel et al., 2013). For 3D co-culture, $\alpha 6^{hi}$ -GFP and $\alpha 6^{lo}$ -mCherry cells were suspended in single cells and plated on Matrigel-coated plates for 1 week. For integrin $\alpha 6\beta 4$ ligation assay, MCF10A cells were cultured on FN, LM111 and LM511-coated plates for 4 hours before collecting for RNA isolation.

Tumor samples. For PDX tumors, triple-negative breast tumor tissue samples #1126 and #1258 were obtained from breast cancer surgery patients, through Brigham and Women's Hospital and Massachusetts General Hospital. These

breast tumor samples were cut into small pieces, which were then surgically implanted into the fourth mammary fat pad of NSG mice. Tumor xenografts were expanded into multiple recipient NSG mice for several passages. For primary human tumors, two Luminal A type breast tumors were obtained from Tissue Bank at UMass Medical School.

Immunoblotting and qPCR. Cells were extracted in RIPA (50mM Tris-HCL [pH 7.4], 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1mM PMSF, and protease inhibitors). The proteins were separated by SDS-PAGE and immunoblotted using Abs as specified in the figure legends. For qPCR, RNA isolation was done with the NucleoSpin RNA kit (Macherey-Nagel) and 1ug of total RNA was used to produce cDNAs using Roche First strand synthesis kit (Roche). qPCR was performed using a SYBR green (Applied Biosystems) master mix as described by the manufacturer. Sequences for qPCR primers used are provided in Table 4.

Figures

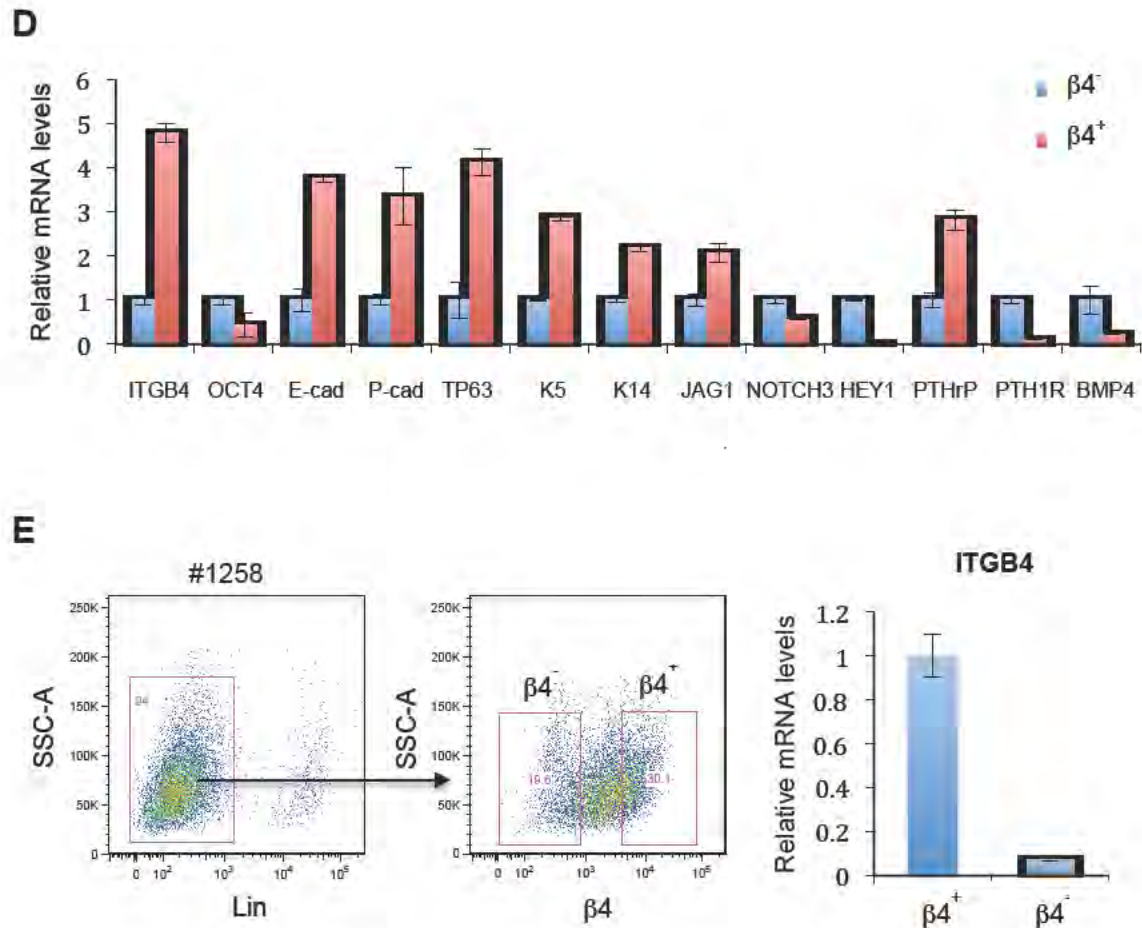
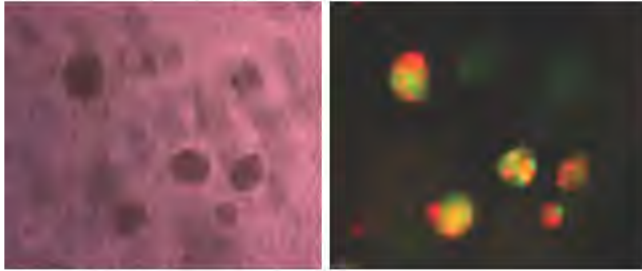
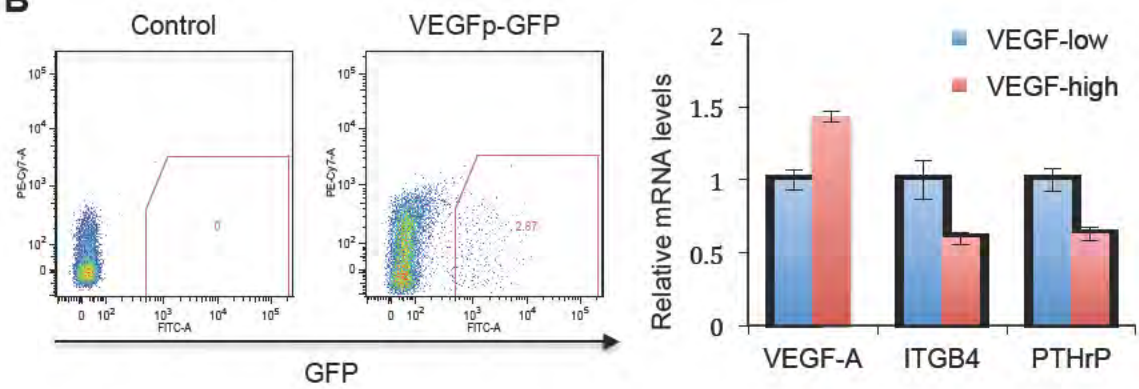
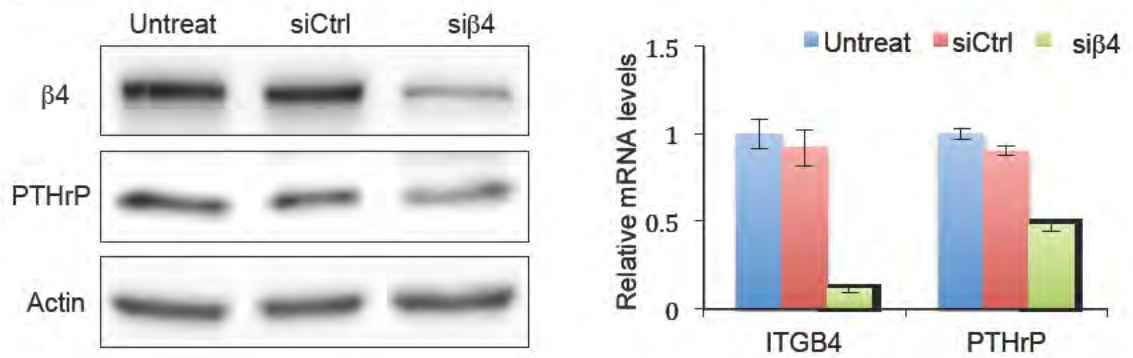


Figure 1. Integrin $\beta 4$ is expressed heterogeneously in breast tumor.

(A) H&E staining of PDX tumor #1126 and #1258. (B) Expression of integrin $\beta 4$ in PDX tumor #1126 and #1258 was examined by IF staining. (C) Integrin $\beta 4$ expression in PDX tumor #1126 is analyzed by FACS, and $\beta 4^+$ and $\beta 4^-$ populations were sorted. (D) Relative mRNA levels of indicated genes were quantified in $\beta 4^+$ and $\beta 4^-$ populations by qPCR. (E) $\beta 4^+$ and $\beta 4^-$ populations were sorted in PDX tumor #1258 by FACS, and $\beta 4$ expression was examined by qPCR.

A**B****C**

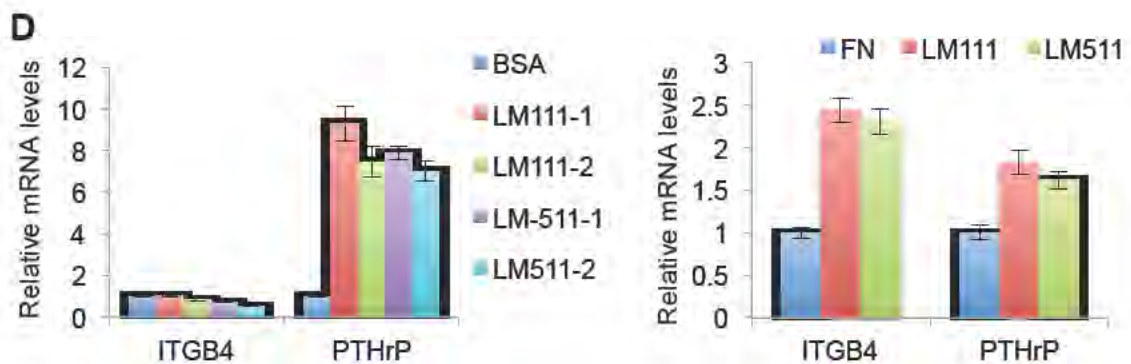
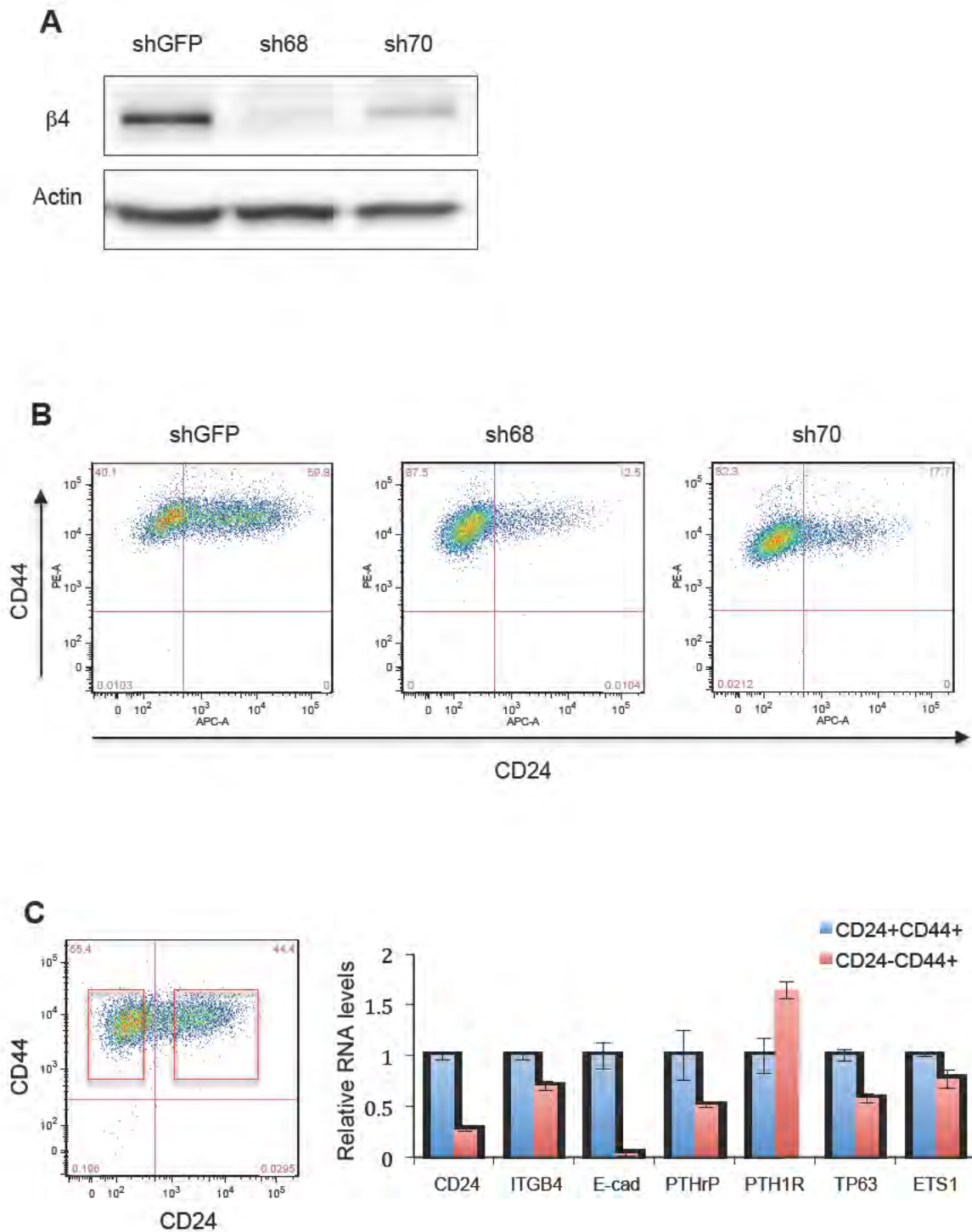


Figure 2. Integrin β 4 regulates PTHrP expression.

(A) $\alpha 6^{hi}$ -GFP and $\alpha 6^{lo}$ -mCherry cells were mixed in single-cell suspension and cultured on Matrigel for 1 week. (B) GFP⁺ cells were sorted from MDA-MB-231 cells stably expressing GFP under control of the VEGF promoter. Relative mRNA levels of VEGF-A, ITGB4 and PTHrP are examined by qPCR in VEGF-low (GFP⁻) cells and VEGF-high (GFP⁺) cells. (C) Expression of β 4 and PTHrP were examined by western blot and qPCR in MCF10A cells transfected with siCtrl or si β 4. (D) MCF10A cells were plated on BSA, fibronectin (FN) or indicated laminins (LM) for 4hr before collected for qPCR.



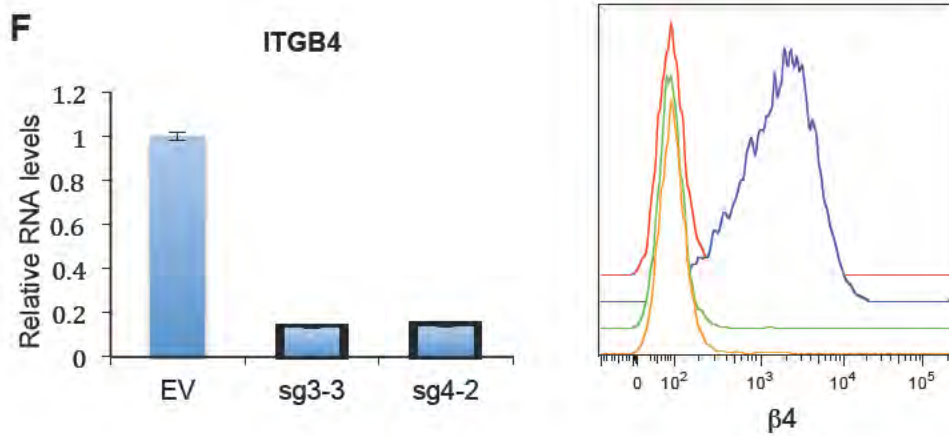
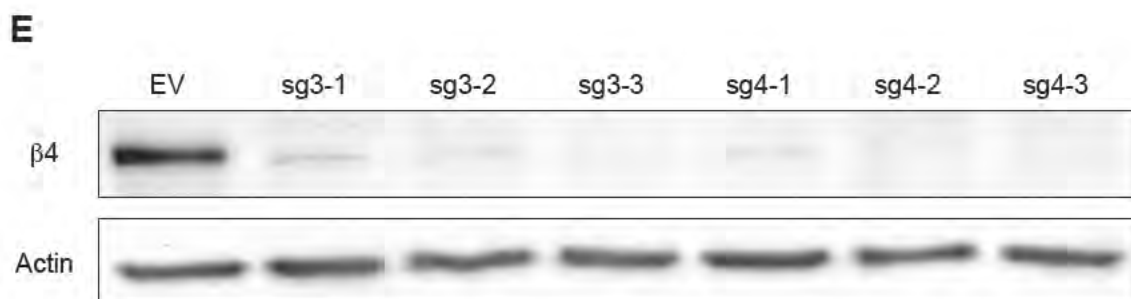
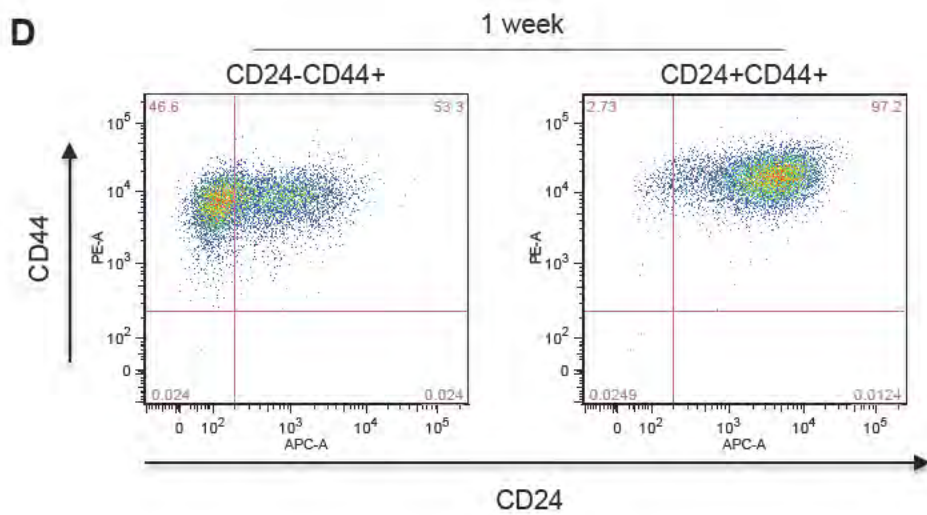


Figure 3. Knocking down integrin β 4 affects CSC functions.

(A) Integrin β 4 was knocked down in SUM-159 cells by two shRNAs, sh68 and sh70. (B) CD24⁺CD44⁺ population was reduced dramatically after β 4 was knocked down. (C) CD24⁻CD44⁺ and CD24⁺CD44⁺ populations in SUM-159 cells were sorted, and the expression of indicated genes were examined by qPCR. (D) CD24⁻CD44⁺ and CD24⁺CD44⁺ populations were cultured separately for 1 week before CD24 and CD44 expression were analyzed by FACS again. (E) MDA-MB-231 cells stably expressing sgRNAs (sg3 and sg4) targeting integrin β 4 were generated. Three clones derived from single cells for each sgRNA were selected, and β 4 expression was examined by western blot. (F) Expression of integrin β 4 is examined by qPCR and FACS in two selected clones sg3-3 and sg4-2.

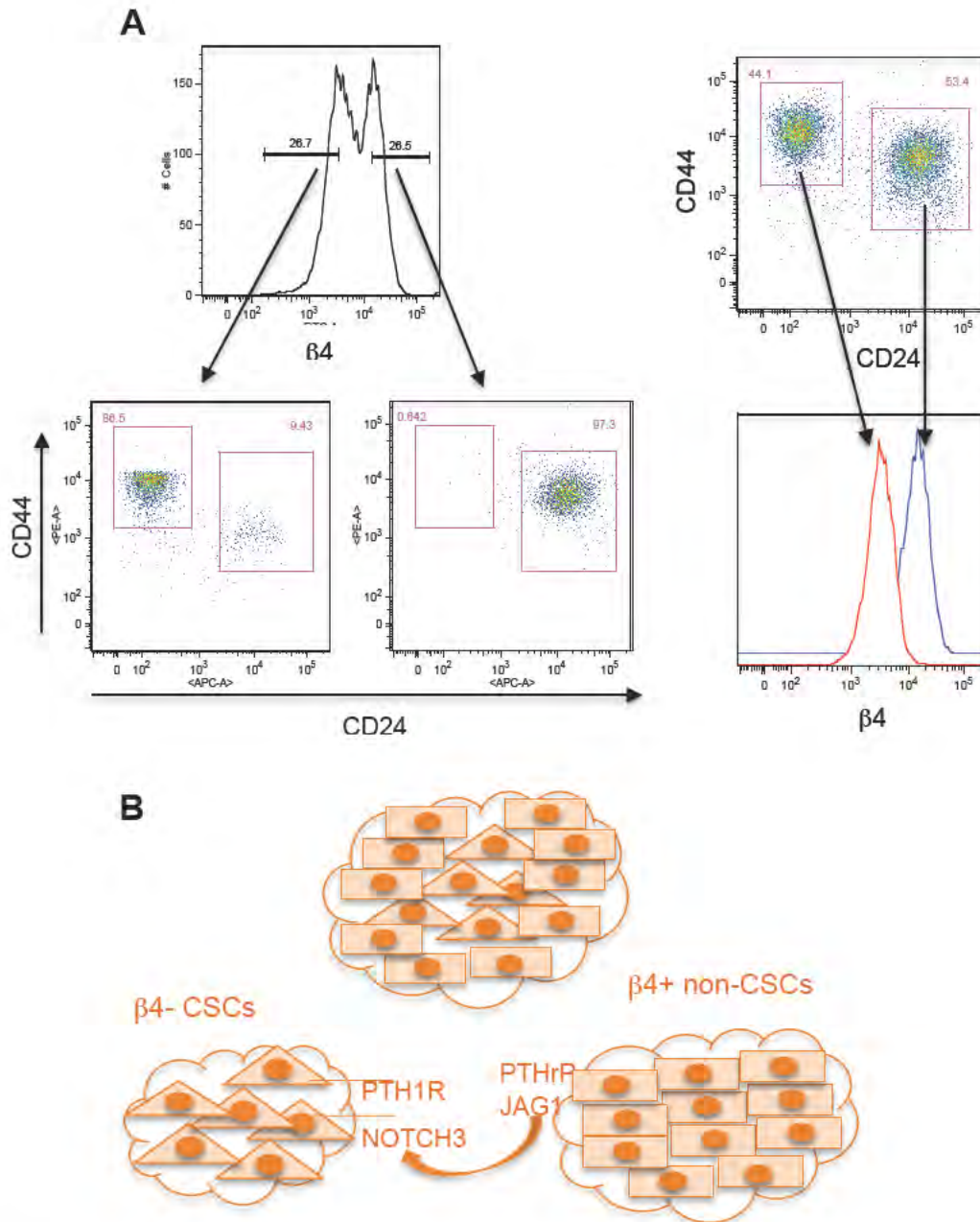


Figure 4. Integrin $\beta 4$ is associated with CD24+CD44+ population in breast tumor.

(A) Expression of integrin $\beta 4$, CD24 and CD44 were analyzed by FACS in Luminal A type human breast tumor. (B) Model of the study.

Table 1. Summary of RNA-seq of $\beta 4^{hi}$ and $\beta 4^{lo}$ populations in Src-transformed MCF10A system

Top genes in $\beta 4^{hi}$ population	\log_2 ratio $\beta 4^{hi} / \beta 4^{lo}$	Top genes in $\beta 4^{lo}$ population	\log_2 ratio $\beta 4^{lo} / \beta 4^{hi}$
CDH1	4.3402873	POSTN	6.093040102
TP63	3.743782825	BMP4	3.157352242
PTHrP	3.211881568	NOTCH3	3.13051063
LAMA3	2.986767935	WNT5A	2.973761667
ALDH1A3	2.304530856	WNT5B	2.543005934
CD24	2.267123678	TGFB3	1.832512552
KRT14	2.201647226	PVRL3	1.45199799
CDH3	1.879796858	CXCR4	1.451083445
PVRL1	1.762919545	TGFB2	1.224635659
ITGA6	1.351495691	WNT3	1.157352242
LAMC2	1.326901226	TGFB1	0.919192505
LAMB3	1.123110904	TWIST1	0.902375611
JAG1	1.119903245	ERBB2	0.838043658
PVRL4	0.889953473	COL4A2	0.815580486
KRT18	0.809841409	COL4A1	0.779803164
KRT5	0.746395911	SNAI1	0.731534904
NOTCH1	0.594431214	PVRL2	0.584585038
IGF1R	0.412021182	ITGB1	0.476975791
CD44	0.362532557	NOTCH2	0.416051599

Table 2. Summary of RNA-seq of $\beta 4^+$ and $\beta 4^-$ populations in PDX tumorTop 100 genes in $\beta 4^+$ population

Rank	gene_id	fold.change. $\beta 4^- / \beta 4^+$	genename
1	ERP27	0.055810447	endoplasmic reticulum protein 27
2	EDN2	0.068085924	endothelin 2
3	GABRP	0.07806187	gamma-aminobutyric acid (GABA) A receptor, pi
4	ITGB4	0.078433946	integrin, beta 4
5	SCGB2A2	0.08094378	secretoglobin, family 2A, member 2
6	PROM2	0.081264473	prominin 2
7	SOSTDC1	0.086105138	sclerostin domain containing 1
8	KRTAP19-5	0.088997743	keratin associated protein 19-5
9	LINC00617	0.089003394	
10	CAPS	0.090619562	calcyphosine
11	EEF1E1- MUTED	0.091320952	
12	ITGB6	0.100665881	integrin, beta 6
13	DEFB109P1	0.105029219	defensin, beta 109, pseudogene 1
14	KRTAP5-2	0.105787866	keratin associated protein 5-2
15	SLC34A2	0.108548112	solute carrier family 34 (type II sodium/phosphate cotransporter), member 2
16	SNORA58	0.10884567	small nucleolar RNA, H/ACA box 58
17	CLDN8	0.111889128	claudin 8
18	MIR205HG	0.112553915	MIR205 host gene
19	ZNF20	0.114010716	zinc finger protein 20
20	S100A14	0.114327068	S100 calcium binding protein A14
21	EPHA1	0.115049589	EPH receptor A1
22	AZGP1	0.115545491	alpha-2-glycoprotein 1, zinc-binding
23	EPN3	0.116073033	epsin 3
24	PROM1	0.116701246	prominin 1
25	SCARNA10	0.119531437	small Cajal body-specific RNA 10
26	PRODH	0.121185626	proline dehydrogenase (oxidase) 1
27	FXYD3	0.121581886	FXYD domain containing ion transport regulator 3
28	KRT17	0.125540397	keratin 17, type I
29	KRTAP21-2	0.127566641	keratin associated protein 21-2
30	HIST1H2AB	0.128848047	histone cluster 1, H2ab
31	S100P	0.132093822	S100 calcium binding protein P
32	VIPR1	0.132249714	vasoactive intestinal peptide receptor 1
33	LOC93432	0.138217499	maltase-glucoamylase (alpha-glucosidase)
34	EFNA3	0.139139826	ephrin-A3
35	DEFB109P1B	0.141477134	defensin, beta 109, pseudogene 1B
36	LOC642423	0.142028508	golgin A2 pseudogene
37	TRIM29	0.142518236	tripartite motif containing 29
38	LCN2	0.142842554	lipocalin 2
39	TMPRSS13	0.142981928	transmembrane protease, serine 13
40	KRTCAP3	0.143176989	keratinocyte associated protein 3

41	MB	0.145079255	myoglobin
42	NAALADL2	0.147429106	N-acetylated alpha-linked acidic dipeptidase-like 2
43	INHBB	0.147610419	inhibin, beta B
44	BSPRY	0.1478731	B-box and SPRY domain containing
45	LIPG	0.149503414	lipase, endothelial
46	AQP3	0.151534351	aquaporin 3 (Gill blood group)
47	RASAL1	0.152332362	RAS protein activator like 1 (GAP1 like)
48	CAMSAP3	0.152843564	calmodulin regulated spectrin-associated protein family, member 3
49	EHF	0.153213829	ets homologous factor
50	LAD1	0.153312724	ladinin 1
51	KRT14	0.153942955	keratin 14, type I
52	LGALS7	0.15404238	lectin, galactoside-binding, soluble, 7
53	PADI2	0.156071582	peptidyl arginine deiminase, type II
54	SUSD4	0.156225299	sushi domain containing 4
55	UNC13D	0.156339417	unc-13 homolog D (C. elegans)
56	LOC100128338	0.156403451	
57	SLPI	0.1564616	secretory leukocyte peptidase inhibitor
58	VGLL1	0.158340308	vestigial-like family member 1
59	GUCY1B3	0.158370436	guanylate cyclase 1, soluble, beta 3
60	VTCN1	0.158429337	V-set domain containing T cell activation inhibitor 1
61	KRT15	0.160475562	keratin 15, type I
62	DLX3	0.16236596	distal-less homeobox 3
63	JMJD7-PLA2G4B	0.163768749	JMJD7-PLA2G4B readthrough
64	TMEM54	0.166422176	transmembrane protein 54
65	ESRP1	0.167872286	epithelial splicing regulatory protein 1
66	B4GALNT3	0.167928138	beta-1,4-N-acetyl-galactosaminyl transferase 3
67	EFCAB4A	0.167970125	
68	PVRL4	0.168173739	poliovirus receptor-related 4
69	FMOD	0.169938592	fibromodulin
70	MUC5B	0.170962449	mucin 5B, oligomeric mucus/gel-forming
71	AGR2	0.171148032	anterior gradient 2
72	ATP2C2	0.172121778	ATPase, Ca ⁺⁺ transporting, type 2C, member 2
73	CALB2	0.172161605	calbindin 2
74	IGFL1	0.172251795	IGF-like family member 1
75	MUC1	0.172274959	mucin 1, cell surface associated
76	IGFL2	0.174560115	IGF-like family member 2
77	PTPRU	0.174636105	protein tyrosine phosphatase, receptor type, U
78	FLVCR2	0.175418597	feline leukemia virus subgroup C cellular receptor family, member 2
79	RTN4RL1	0.176413331	reticulon 4 receptor-like 1
80	LOC728463	0.178605874	
81	RBP1	0.178629336	retinol binding protein 1, cellular
82	DSC3	0.179223577	desmocollin 3
83	KRT6B	0.179990217	keratin 6B, type II
84	FREM2	0.180121264	FRAS1 related extracellular matrix protein 2
85	KRT6C	0.180811098	keratin 6C, type II
86	SPINT1	0.180864618	serine peptidase inhibitor, Kunitz type 1

87	HIST1H4D	0.181680765	histone cluster 1, H4d
88	TMC6	0.181806274	transmembrane channel-like 6
89	C6orf15	0.183153963	chromosome 6 open reading frame 15
90	PDGFA	0.183711689	platelet-derived growth factor alpha polypeptide
91	ALDH3B2	0.184243618	aldehyde dehydrogenase 3 family, member B2
92	MAP7	0.184346294	microtubule-associated protein 7
93	CLDN23	0.18458237	claudin 23
94	NEBL	0.187068659	nebulette
95	TNNI2	0.187391417	troponin I type 2 (skeletal, fast)
96	HSD11B2	0.188092237	hydroxysteroid (11-beta) dehydrogenase 2
97	CLIC3	0.188237351	chloride intracellular channel 3
98	S100A16	0.189391975	S100 calcium binding protein A16
99	KRT8	0.190759774	keratin 8, type II
100	CTSL2	0.19103296	

Top 100 genes in $\beta 4^-$ population

Rank	gene_id	fold.change. $\beta 4^- / \beta 4^+$	genename
1	RPL31P11	39.56084	ribosomal protein L31 pseudogene 11
2	KRTAP20-2	25.62712	keratin associated protein 20-2
3	SUMO4	19.8716	small ubiquitin-like modifier 4
4	RAB11A	19.22263546	RAB11A, member RAS oncogene family
5	SNORA28	16.47679023	small nucleolar RNA, H/ACA box 28
6	ZEB2-AS1	14.29556367	ZEB2 antisense RNA 1
7	POTEM	11.42675277	POTE ankyrin domain family, member M
8	RMRP	11.027	RNA component of mitochondrial RNA processing endoribonuclease
9	SCARNA7	10.50824	small Cajal body-specific RNA 7
10	CXCL14	9.965201093	chemokine (C-X-C motif) ligand 14
11	POTEJ	9.793759966	POTE ankyrin domain family, member J
12	CSNK1A1L	9.7768	casein kinase 1, alpha 1-like
13	SPON2	9.673812413	spondin 2, extracellular matrix protein
14	SNORA14B	8.895367534	small nucleolar RNA, H/ACA box 14B
15	FAM78B	8.836337781	family with sequence similarity 78, member B
16	POTEF	8.173411643	POTE ankyrin domain family, member F
17	HGF	7.977776299	hepatocyte growth factor (hepapoietin A; scatter factor)
18	SERPINE1	7.704713281	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1
19	MSC	7.662533796	musculin
20	ZEB2	7.239536869	zinc finger E-box binding homeobox 2
21	MIR3661	7.153655035	microRNA 3661
22	CXCL12	6.877907642	chemokine (C-X-C motif) ligand 12
23	ZNF804A	6.832372958	zinc finger protein 804A

24	FAM90A2P	6.813	family with sequence similarity 90, member A2 pseudogene
25	CUZD1	6.469748699	CUB and zona pellucida-like domains 1
26	FLNC	6.447612684	filamin C, gamma
27	HIST1H4B	6.4476	histone cluster 1, H4b
28	HS3ST3A1	6.427331867	heparan sulfate (glucosamine) 3-O-sulfotransferase 3A1
29	TGM2	6.396667265	transglutaminase 2
30	EBF1	6.37322519	early B-cell factor 1
31	TMEFF2	6.365964799	transmembrane protein with EGF-like and two follistatin-like domains 2
32	POTEE	6.359760119	POTE ankyrin domain family, member E
33	MIR3682	6.320083257	microRNA 3682
34	PIN4P1	6.307404114	protein (peptidylprolyl cis/trans isomerase) NIMA-interacting, 4 pseudogene 1
35	SERPINF1	6.196399774	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1
36	PLAC8L1	6.089834984	PLAC8-like 1
37	CREB3L1	6.011548445	cAMP responsive element binding protein 3-like 1
38	TRPC6	5.880595105	transient receptor potential cation channel, subfamily C, member 6
39	PCDH7	5.857565254	protocadherin 7
40	MMP3	5.701584203	matrix metalloproteinase 3
41	PTENP1	5.67822539	phosphatase and tensin homolog pseudogene 1 (functional)
42	SFRP2	5.671937194	secreted frizzled-related protein 2
43	DCN	5.59704733	decorin
44	TFPI2	5.457923828	tissue factor pathway inhibitor 2
45	FOXP1	5.434430939	forkhead box P1
46	COL5A2	5.394802843	collagen, type V, alpha 2
47	SNORA51	5.394565631	small nucleolar RNA, H/ACA box 51
48	DACT1	5.371712803	dishevelled-binding antagonist of beta-catenin 1
49	SCARNA11	5.314531105	small Cajal body-specific RNA 11
50	BMPER	5.305486231	BMP binding endothelial regulator
51	HEG1	5.304122148	heart development protein with EGF-like domains 1
52	PRAP1	5.24076	proline-rich acidic protein 1
53	GFAP	5.189984324	glial fibrillary acidic protein
54	LINC00235	5.051334656	long intergenic non-protein coding RNA 235
55	FEZ1	4.949002127	fasciculation and elongation protein zeta 1 (zygin I)
56	PRPS1L1	4.934884	phosphoribosyl pyrophosphate synthetase 1-like 1
57	DENND5B-AS1	4.91194	DENND5B antisense RNA 1
58	CDRT15	4.88256	CMT1A duplicated region transcript 15
59	TFAP2B	4.873186417	transcription factor AP-2 beta (activating enhancer binding protein 2 beta)
60	GPR176	4.869509648	G protein-coupled receptor 176
61	INO80B-WBP1	4.842572365	INO80B-WBP1 readthrough (NMD candidate)
62	MMP1	4.823947431	matrix metalloproteinase 1

63	LAYN	4.813852191	layilin
64	NNMT	4.762956969	nicotinamide N-methyltransferase
65	DDX3Y	4.740717564	DEAD (Asp-Glu-Ala-Asp) box helicase 3, Y-linked
66	COL3A1	4.661040209	collagen, type III, alpha 1
67	LOXL1	4.645667536	lysyl oxidase-like 1
68	DPT	4.640355521	dermatopontin
69	ADAM19	4.591757985	ADAM metallopeptidase domain 19
70	ESM1	4.590916332	endothelial cell-specific molecule 1
71	FHL1	4.549355217	four and a half LIM domains 1
72	PTPRN	4.513944	protein tyrosine phosphatase, receptor type, N
73	RPS10-NUDT3	4.485972758	RPS10-NUDT3 readthrough
74	EXT1	4.475825975	exostosin glycosyltransferase 1
75	C9orf53	4.45826	
76	MARCH4	4.456227523	membrane-associated ring finger (C3HC4) 4, E3 ubiquitin protein ligase
77	LOC728875	4.417964747	
78	HIC1	4.417683196	hypermethylated in cancer 1
79	DACT3	4.394819526	dishevelled-binding antagonist of beta-catenin 3
80	METRN	4.354512	meteorin, glial cell differentiation regulator
81	KCNE4	4.348465048	potassium channel, voltage gated subfamily E regulatory beta subunit 4
82	NPTX1	4.331684174	neuronal pentraxin I
83	HBEGF	4.277256745	heparin-binding EGF-like growth factor
84	HAS2	4.242831962	hyaluronan synthase 2
85	MIR25	4.213524347	microRNA 25
86	ARL2-SNX15	4.206012874	ARL2-SNX15 readthrough (NMD candidate)
87	NOVA1	4.205469071	neuro-oncological ventral antigen 1
88	PRG1	4.176908	p53-responsive gene 1
89	IL13RA2	4.158880367	interleukin 13 receptor, alpha 2
90	TSNAX-DISC1	4.158744	TSNAX-DISC1 readthrough (NMD candidate)
91	NPPB	4.141444389	natriuretic peptide B
92	HS3ST3B1	4.136352538	heparan sulfate (glucosamine) 3-O-sulfotransferase 3B1
93	PTEN	4.114100491	phosphatase and tensin homolog
94	RUNX1T1	4.103495281	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)
95	CNRIP1	4.103266866	cannabinoid receptor interacting protein 1
96	TOP1P1	4.100549285	topoisomerase (DNA) I pseudogene 1
97	STEAP1B	4.076097027	STEAP family member 1B
98	RNF216-IT1	4.040784	RNF216 intronic transcript 1
99	FMN2	4.025966549	formin 2
100	EEF1DP3	4.008811645	eukaryotic translation elongation factor 1 delta pseudogene 3

Table 3. Summary of RNA-seq of genes downregulated in $\beta 4$ knockout cells.

Rank	gene id	fold change sg3 / EV	fold change sg4 / EV	genename
1	TSTD1	0.04350126	0.069911745	thiosulfate sulfurtransferase (rhodanese)-like domain containing 1
2	SPANXA2	0.051349569	0.051349569	SPANX family, member A2
3	C16orf45	0.07342529	0.064272018	chromosome 16 open reading frame 45
4	FAM133DP	0.073443954	0.827558717	family with sequence similarity 133, member A pseudogene
5	CXCR4	0.073926139	0.057124818	chemokine (C-X-C motif) receptor 4
6	SPINK4	0.075410232	0.191924771	serine peptidase inhibitor, Kazal type 4
7	NGFR	0.086140919	0.079623097	nerve growth factor receptor
8	SPANXA2	0.092089193	0.092089014	SPANX family, member A2
9	EPCAM	0.099203079	0.120440375	epithelial cell adhesion molecule
10	CCBE1	0.116851125	0.111226787	collagen and calcium binding EGF domains 1
11	KISS1	0.121346758	0.257111317	KiSS-1 metastasis-suppressor
12	SCARNA9	0.131348741	0.131348741	small Cajal body-specific RNA 9
13	HIST1H4A	0.142589889	0.574498084	histone cluster 1, H4a
14	OR2A20P	0.145343836	0.302905078	olfactory receptor, family 2, subfamily A, member 20 pseudogene
15	C10orf32- AS3MT	0.151792632	0.151764504	
16	LOC100129 216	0.153774235	0.153774235	beta-defensin 131-like
17	ZNF521	0.154848462	0.234063645	zinc finger protein 521
18	OR2A9P	0.159888532	0.159876545	olfactory receptor, family 2, subfamily A, member 9 pseudogene
19	SCARNA27	0.168141171	0.670204372	small Cajal body-specific RNA 27
20	NAP1L3	0.171950121	0.31324291	nucleosome assembly protein 1-like 3
21	RBPMS2	0.173485324	0.202433517	RNA binding protein with multiple splicing 2
22	TNFSF12- TNFSF13	0.174179488	1.623951512	TNFSF12-TNFSF13 readthrough
23	DCN	0.17713751	0.05923473	decorin
24	IGFBP5	0.179743337	0.1656084	insulin-like growth factor binding protein 5
25	INHBB	0.185578391	0.23093968	inhibin, beta B
26	ZNF727	0.187631342	0.867948814	zinc finger protein 727
27	C2orf82	0.193909685	0.394747374	chromosome 2 open reading frame 82
28	SORCS2	0.19484684	0.433642666	sortilin-related VPS10 domain containing receptor 2
29	HIST1H3A	0.19598313	0.396238692	histone cluster 1, H3a
30	SNORA58	0.197098707	0.197098707	small nucleolar RNA, H/ACA box 58
31	SNORA55	0.204217229	1.013776963	small nucleolar RNA, H/ACA box 55
32	LGALS12	0.208078264	0.249252292	lectin, galactoside-binding, soluble, 12
33	FRMD3	0.210432574	1.261265508	FERM domain containing 3
34	MGMT	0.213281099	0.272516171	O-6-methylguanine-DNA methyltransferase

35	LOC100130 298	0.21811697	0.412259395	hCG1816373-like
36	SCGB1B2P	0.218850945	0.682335227	secretoglobin, family 1B, member 2, pseudogene
37	C19orf73	0.221246658	0.450753757	chromosome 19 open reading frame 73
38	LOC400927	0.222905072	0.774445985	TPTE and PTEN homologous inositol lipid phosphatase pseudogene
39	SHANK2	0.226461863	0.287905246	SH3 and multiple ankyrin repeat domains 2
40	DPYD-AS1	0.227782338	0.802344154	DPYD antisense RNA 1
41	RPS10- NUDT3	0.236315084	0.948134529	RPS10-NUDT3 readthrough
42	LOC339535	0.239718369	0.195576834	
43	C11orf94	0.245990115	0.245990115	chromosome 11 open reading frame 94
44	SCARNA6	0.250293093	1.028429291	small Cajal body-specific RNA 6
45	CRYAB	0.250299439	0.342318396	crystallin, alpha B
46	MCAM	0.252200635	0.128516192	melanoma cell adhesion molecule
47	ITGB4	0.255371167	0.275924966	integrin, beta 4
48	F2RL2	0.256479997	0.261992419	coagulation factor II (thrombin) receptor-like 2
49	GYPC	0.257663691	0.266597922	glycophorin C (Gerbich blood group)
50	CBWD5	0.261927415	1.013365068	COBW domain containing 5
51	PCSK1N	0.262831201	0.413415047	proprotein convertase subtilisin/kexin type 1 inhibitor
52	PDCL3P4	0.264976447	1.200811281	phosducin-like 3 pseudogene 4
53	GYG2	0.268174461	0.546234848	glycogenin 2
54	ASTN1	0.269359997	0.289423557	astrotactin 1
55	RPS4Y1	0.27005168	0.18999848	ribosomal protein S4, Y-linked 1
56	RPPH1	0.273742397	0.584716071	ribonuclease P RNA component H1
57	HLA-DRB4	0.275207359	0.124839581	major histocompatibility complex, class II, DR beta 4
58	PTGES3L	0.277866041	0.361699879	prostaglandin E synthase 3 (cytosolic)- like
59	PSG1	0.279738817	3.076952994	pregnancy specific beta-1-glycoprotein 1
60	C4orf22	0.280510889	0.516061492	chromosome 4 open reading frame 22
61	MT1JP	0.280572907	0.641128172	metallothionein 1J, pseudogene
62	CKLF- CMTM1	0.282259749	0.9087251	CKLF-CMTM1 readthrough
63	IDH1-AS1	0.288881305	0.289558529	IDH1 antisense RNA 1
64	NCOR1P1	0.289551924	0.289551924	nuclear receptor corepressor 1 pseudogene 1
65	RAB3IL1	0.290038762	1.120807546	RAB3A interacting protein (rabin3)-like 1
66	IGLL3P	0.29010229	0.640255754	immunoglobulin lambda-like polypeptide 3, pseudogene
67	LOC100270 746	0.294879562	0.612377396	uncharacterized LOC100270746
68	LRP3	0.29535187	0.956468679	low density lipoprotein receptor-related protein 3

69	HIGD1C	0.295529234	0.295529234	HIG1 hypoxia inducible domain family, member 1C
70	DIP2A-IT1	0.295945896	1.000113643	DIP2A intronic transcript 1
71	PTGES3L-AARSD1	0.296339847	0.47612212	PTGES3L-AARSD1 readthrough
72	ZG16B	0.298211566	0.298211566	zymogen granule protein 16B
73	HIST1H2BB	0.299100664	1.35017629	histone cluster 1, H2bb
74	LOC100505806	0.30304665	0.415919052	
75	SPANXE	0.30687487	0.306874152	
76	HTRA3	0.311659042	0.311219353	HtrA serine peptidase 3
77	THY1	0.313556509	0.23963804	Thy-1 cell surface antigen
78	TERC	0.314465804	1.682530418	telomerase RNA component
79	TP73-AS1	0.315250251	0.397491492	TP73 antisense RNA 1
80	SYNGR1	0.316266361	0.555500139	synaptogyrin 1
81	DAPK1	0.31656718	0.316370303	death-associated protein kinase 1
82	LINC00173	0.320532503	0.57799114	long intergenic non-protein coding RNA 173
83	DACT3-AS1	0.321700638	0.321700638	DACT3 antisense RNA 1
84	COL3A1	0.322228195	0.11756305	collagen, type III, alpha 1
85	MIR186	0.324137158	0.000175196	microRNA 186
86	SCN5A	0.327542966	0.531201419	sodium channel, voltage gated, type V alpha subunit
87	LINC00707	0.329058609	0.921487632	long intergenic non-protein coding RNA 707
88	SNORD23	0.329967959	1.476360074	small nucleolar RNA, C/D box 23
89	MMP2	0.332096761	0.082891432	matrix metalloproteinase 2
90	SNORA76	0.334330762	1.183125653	
91	NPM2	0.335076625	0.335076625	nucleophosmin/nucleoplasmin 2
92	TAS2R14	0.335832374	1.248389684	taste receptor, type 2, member 14
93	ENG	0.336152969	0.439583302	endoglin
94	SNORA66	0.337839813	0.434350617	small nucleolar RNA, H/ACA box 66
95	ARHGAP30	0.338493546	0.463041349	Rho GTPase activating protein 30
96	SOST	0.339875277	0.609284823	sclerostin
97	HIST1H4C	0.34467218	0.348488932	histone cluster 1, H4c
98	COL1A2	0.346016848	0.066684619	collagen, type I, alpha 2
99	LINC00516	0.349443687	0.713864264	
100	LOC101101776	0.349443687	0.531330429	

Table 4. Sequences of the primers used in this study.

qPCR	Forward	Reverse
BMP4	5'-AAAGTCGCCGAGATTCAGGG-3'	5'-GACGGCACTCTTGCTAGGC-3'
CD24	5'-CTCCTACCCACGCAGATTTATTC-3'	5'-AGAGTGAGACCACGAAGAGAC-3'
E-cad	5'-TTCCAGGAACCTCTGTGATG-3'	5'-TCTTGGCTGAGGATGGTGTA-3'
ETS1	5'-TACACAGGCAGTGGACCAATC-3'	5'-CCCCGCTGTCTTGTGGATG-3'
GAPDH	5'-CTGGGCTACACTGAGCACC-3'	5'-AAGTGGTCGTTGAGGGCAATG-3'
HEY1	5'-ATCTGCTAAGCTAGAAAAAGCCG-3'	5'-GTGCGCGTCAAAGTAACCT-3'
ITGB4	5'-CTCCACCGAGTCAGCCTTC-3'	5'-CGGGTAGTCCTGTGTCCTGTA-3'
JAG1	5'-GGGGCAACACCTTCAACCTC-3'	5'-CCAGGCGAAACTGAAAGGC-3'
KRT5	5'-ACATGAACAAGGTGGAGCTG-3'	5'-AGCTCCGCATCAAAGAACAT-3'
KRT14	5'-AACGAGATGCGTGACCAGTA-3'	5'-TCAGCTCCTCTGTCTTGGTG-3'
NOTCH3	5'-CGTGGCTTCTTTCTACTGTGC-3'	5'-CGTTCACCGGATTTGTGTAC-3'
OCT4	5'-GAGAAGGATGTGGTCCGAGT-3'	5'-GTGCATAGTCGCTGCTTGAT-3'
P-cad	5'-TGGAGATCCTTGATGCCAATGA-3'	5'-GCGTCCAGATCAGTGACCG-3'
PTH1R	5'-AGTGCGAAAAACGGCTCAAG-3'	5'-GATGCCTTATCTTTCCTGGGC-3'
PTHrP	5-AAGGTGGAGACGTACAAAGAGC-3'	5'-CAGAGCGAGTTCGCCGTTT-3'
TP63	5'-GGACCAGCAGATTCAGAACGG-3'	5'-AGGACACGTCGAAACTGTGC-3'
VEGF-A	5'-AGGGCAGAATCATCACGAAGT-3'	5'-AGGGTCTCGATTGGATGGCA-3'
CRISPR	Oligo 1	Oligo 2
sg3	CACCGTTGTCCAGATCATCGGACA	AAACTGTCCGATGATCTGGACAAC
sg4	CACCGAAATCCAATAGTGTAGTCGC	AAACGCGACTACACTATTGGATTC

Chapter III

Generate $\beta 4$ Reporter Mouse Using CRISPR

Abstract

To better study the role of integrin $\beta 4$ in tumor formation and track the $\beta 4^+$ population during tumor progression, I made a $\beta 4$ reporter mouse using CRISPR. I inserted a p2A-mCherry cassette before the stop codon of the ITGB4 gene, and first tested the design strategy in vitro in NMuMG cells.

Introduction

To better study the role of the $\beta 4^+$ population in breast tumors, it will be helpful if we can track the $\beta 4^+$ population during tumor progression. Previous models suggest that tumor cells need to undergo EMT to drive tumor invasion and break away from the primary tumor site. Cells will then undergo MET at the secondary site and form metastasis (Thiery et al., 2009). However, there is no clear evidence showing that EMT and MET are real and necessary during metastasis. Studies from the Ewald group show that the leader cells that drive tumor invasion are a basal epithelial population, not the cells that undergo EMT (Cheung et al., 2013; Shamir et al., 2014). Recently, two back-to-back studies showed that EMT is not required for metastasis in lung and pancreatic cancer (Fischer et al., 2015; Zheng et al., 2015). These studies show that there is a basal epithelial population that drive tumor invasion, and epithelial tumor cells can intravasate into the blood

system and form metastasis at the secondary site without undergoing EMT and MET. To show how $\beta 4$ is involved in this process and localize the $\beta 4^+$ population, I wanted to make a $\beta 4$ reporter mouse and cross it with breast tumor model so that I can track the $\beta 4^+$ population during tumor progression.

Homozygous $\beta 4$ deletion mice die shortly after birth (Dowling et al., 1996; van der Neut et al., 1996; 1999b). Our lab showed that $\beta 4$ is important for mammary gland development, and even heterozygotes have defects in the mammary gland (Li et al., 2015). Thus, this reporter mouse design strategy has to keep the ITGB4 gene intact after inserting mCherry. Moreover, to avoid blocking the ligand-binding site of $\beta 4$, it is better to add the reporter gene at the C-terminus of $\beta 4$. However, our lab showed that fusing GFP to the C-terminus of $\beta 4$ drives $\beta 4$ into nucleus. Thus, I added a p2A self-cleaved peptide between ITGB4 and mCherry. After translation, p2A will be cleaved automatically so that $\beta 4$ and mCherry will be separated with $\beta 4$ going to the cell surface while mCherry stays in the cytoplasm (Kim et al., 2011b). Thus, my design is to insert the p2A-mCherry cassette right before the TGA stop codon of ITGB4 without affecting the sequence of ITGB4.

Results and Discussion

Initially I designed 4 sgRNAs near the TGA stop codon of the ITGB4 gene. To test the efficiency of the 4 sgRNAs, I cloned the 4 sgRNAs into PX458 vector and

transfected them into NMuMG cells. The transfection efficiency was about 10%. I sorted GFP⁺ cells and did a SURVEYOR Assay (Fig 5B). I chose the sgRNA #2 that cut 3bp after the stop codon of ITGB4 since it would give me higher homology recombination (HR) efficiency. To make the plasmid donor for HR, I PCR amplified ITGB4 2kb left homology arm and ITGB4 3kb right homology arm from the genomic DNA of NMuMG cells, and cloned them flanking the p2A-mCherry cassette (Fig 5A).

I transfected PX330-sgRNA and plasmid donor together into NMuMG cells, and used PX330 vector alone with plasmid donor as negative control. I got 0.3% mCherry-positive cells without any selection 5 days post-transfection (Fig 5C). To confirm the correct insertion of the p2A-mCherry cassette, I designed a pair of primers with the Forward primer outside of the 2kb ITGB4 left homology arm and the Reverse primer inside of the mCherry sequence (Fig 5A). I did PCR with genomic DNA extracted from post-transfected NMuMG cells before mCherry sorting. I saw the correct size band of about 2.5kb in the sgRNA cells but not in the negative control cells (Fig 5D). I cut the 2.5kb band and did sequencing to make sure p2A-mCherry was inserted before TGA stop codon correctly (Fig 5D).

Western blot showed that the size and expression of integrin β 4 is not affected by mCherry insertion in sorted mCherry-positive cells, and mCherry mRNA is also detected (Fig 5E). I also tested my design strategy and HR frequency in E14

cells which is a normal mouse embryonic stem cell line, and saw the similar results (data not shown). After I confirmed the CRISPR design successfully in vitro with cell lines, I made the mouse in BL6 background at Jackson Lab, collaborating with Dr. Steve Murray. There is one potentially correct mouse, and we'll do further examination to confirm it.

Materials and Methods

Generating sgRNAs Expressing Vector and plasmid donor. Bicistronic expression vector (PX458 or PX330) expressing Cas9 and sgRNA (Cong et al., 2013) were digested with BbsI and treated with Antarctic Phosphatase, and the linearized vector was gel purified. A pair of oligos for each targeting site (Table 5) was annealed, phosphorylated, and ligated to linearized vector. For plasmid donor, Nanog-2A-mCherry plasmid (Addgene) was cut at BamHI/NheI on the left and PacI/NotI on the right. ITGB4 2kb left homologous arm and ITGB4 3kb right homologous arm was PCR amplified from the genomic DNA of NMuMG cells and cloned into digested Nanog-2A-mCherry on each side, respectively.

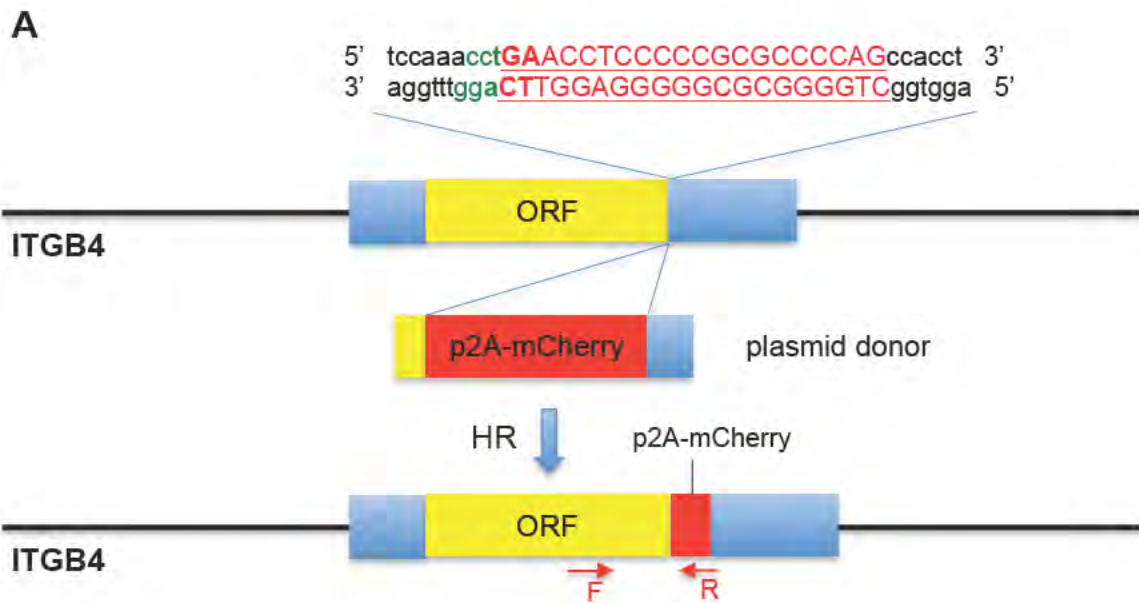
Cell Culture and Transfection. NMuMG cells were purchased from ATCC and maintained as described (Chang et al., 2013). PX458 vector (Addgene) cloned with indicated sgRNAs were transfected into NMuMG cells to examine to efficiency of sgRNAs. PX330 vector (Addgene) cloned with indicated sgRNAs

and plasmid donor were transfected in NMuMG cells to test the homology recombination in vitro.

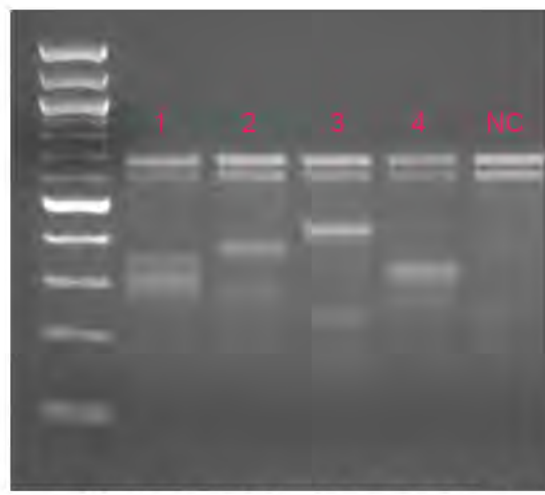
Surveyor Assay. Genomic DNA from NMuMG cells transfected with indicated sgRNAs in the PX458 vector were extracted and surveyor assay was performed as described by (Guschin et al., 2010). Sequences of the primers used for surveyor assay were provided in Table 5.

Immunoblotting and qPCR. Cells were extracted in RIPA (50mM Tris-HCL [pH 7.4], 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1mM PMSF, and protease inhibitors). The proteins were separated by SDS-PAGE and immunoblotted using antibodies for actin (Sigma-Aldrich) and β 4 (505) (Rabinovitz et al., 1999). For qPCR, RNA isolation was done with the NucleoSpin RNA kit (Macherey-Nagel) and 1ug of total RNA was used to produce cDNAs using Roche First strand synthesis kit (Roche). qPCR was performed using a SYBR green (Applied Biosystems) master mix as described by the manufacturer. Sequences for qPCR primers used are provided in Table 5.

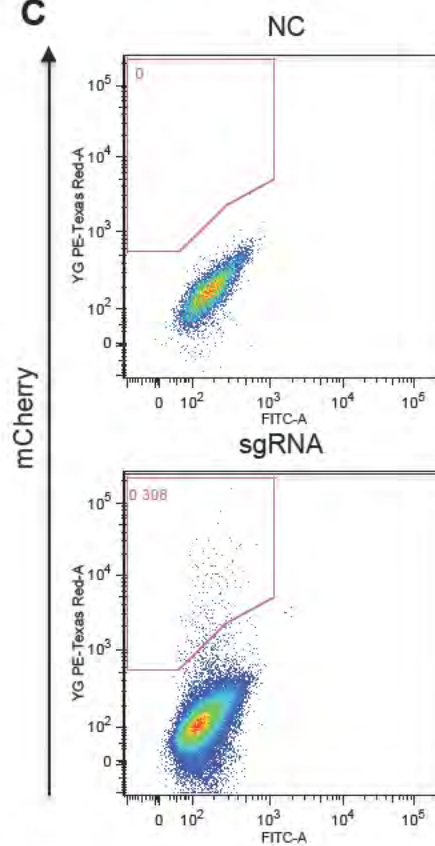
Figures



B



C



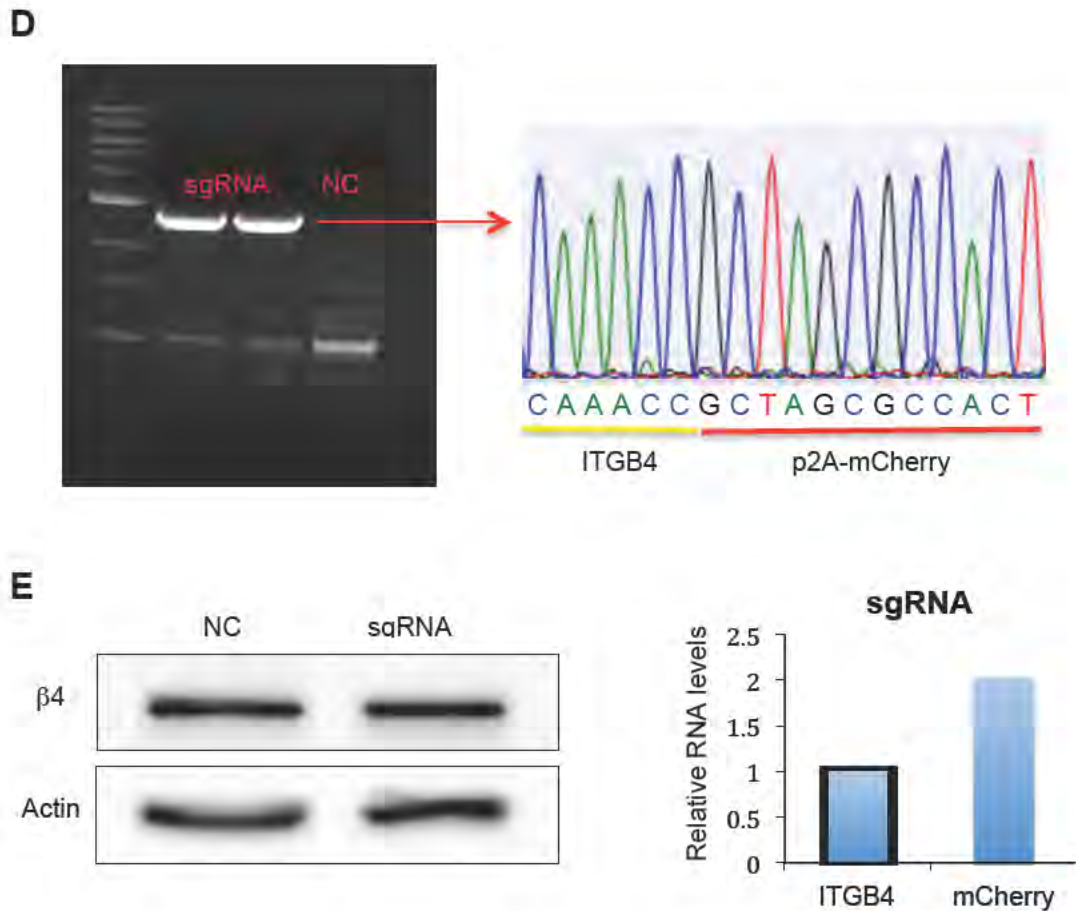


Figure 5. Generate integrin $\beta 4$ reporter mouse using CRISPR.

(A) Schematic overview of strategy to generate a ITGB4-p2A-mCherry knockin allele. The sgRNA coding sequence is underlined, capitalized, and labeled in red. The protospacer-adjacent motif (PAM) sequence is labeled in green. The stop codon of ITGB4 is in bold. The 2kb ITGB4 left homologous arm and 3kb ITGB4 right homologous arm are indicated in yellow and blue, respectively flanking the p2A-mCherry cassette in red in plasmid donor. (B) Surveyor assay of 4 sgRNAs cloned in PX458 vectors transfected to NMuMG cells. Indel percentages are

shown at the bottom. (C) mCherry expression in NMuMG cells transfected with PX330-sgRNA2 and plasmid donor. (D) PCR checking the correct insertion of p2A-mCherry. Primers are shown in (A). Sequencing result of the 2.5kb band is shown on the right. (E) Western blot and qPCR results checking β 4 and mCherry expression.

Table 5. Sequences of the primers used in this study.

qPCR	Forward	Reverse
GAPDH	5'-CTGGGCTACACTGAGCACC-3'	5'-AAGTGGTCGTTGAGGGCAATG-3'
ITGB4	5'-CTCCACCGAGTCAGCCTTC-3'	5'-CGGGTAGTCCTGTGCCTGTA-3'
mCherry	5'-AACGGCCACGAGTTCGAGAT-3'	5'-GTAGGCCTTGAGCCGTACA-3'
	Forward	Reverse
Surveyor	5'-ATCCTGACTGACGGATCTGCC-3'	5'-AAGTGCAGAACAAAAGGCTGGG-3'
mCherry Seq	5'-TAGGCCTCCCTCCTATCTGGGAAG-3'	5'-TCCTCGAAGTTCATCACGCGCT-3'
	Oligo 1	Oligo 2
CRISPR	Oligo 1	Oligo 2
sgRNA1	5'-CACCGTGACCCAGGAATTCGTGACC-3'	5'-AAACGGTCACGAATTCCTGGGTAC-3'
sgRNA2	5'-CACCGCTGGGGCGCGGGGAGGTTC-3'	5'-CACCGTGACCCAGGAATTCGTGACC-3'
sgRNA3	5'-CACCGAGGAAGAAGGCGCTAGGAG-3'	5'-AAACCTCCTAGCGCCTTCTTCCTC-3'
sgRNA4	5'-CACCGAGAGAGCCACTGGCCGTTA-3'	5'-AAACTAACGGCCAGTGGCTCTCTC-3'
	Sequence	
Cloning BamHI- Fw	5'-CGCGG TCCCATGGCTGTCACGCTCCATTT-3'	
NheI-Rv	5'-CTAGCTAGCGGTTTGAAGAAGTGGTCCAT-3'	
PacI-Fw	5'-TTAATTAATTAATGAACCTCCCCGCGCCCCAG-3'	
NotI-Rv	5'-AAGGAAAAAAGCGGCCGCAGTGGTCCTTGAGCCTGGAATCCCACAG-3'	

Chapter IV

General Discussion

EMT and CSC

A CD24⁻CD44⁺ population has been identified as CSCs in breast cancer by showing that this population alone is sufficient to form new tumors in vivo (Al-Hajj et al., 2003). EMT generates cells with stem-cell properties that express mesenchymal markers and are CD24⁻CD44⁺ (Mani et al., 2008). In the past decade, this EMT model has received lots of interest in the field. At the primary tumor site, normal epithelial cells undergo EMT and become mesenchymal-like with CSC properties, which is important for tumor initiation. According to this model, cells that undergo EMT can intravasate into the blood system and extravasate to arrive at the secondary site. Upon arrival, these cells need to undergo MET to form metastasis (Thiery et al., 2009). However, no clear evidence shows that EMT and MET processes are real and necessary for tumor progression and metastasis.

There are two CSC models during tumor metastasis. Both models propose that cells need to undergo EMT for invasion and MET to form colonization at the secondary site. In the first model, CSCs are motile and circulate through blood system (Tsai et al., 2012). In the second model, EMT factors need to be repressed for cells to gain CSC properties. It proposes that CSCs are stationary and exist in both primary and secondary site all the time (Ocaña et al., 2012).

However, recently there are two back-to-back papers indicating that EMT is not required for metastasis in lung cancer and pancreatic cancer, respectively (Fischer et al., 2015; Zheng et al., 2015). Studies also show that it is a basal epithelial population that drives tumor invasion, which contradicts the idea that cells need to undergo EMT to break away from the primary site (Thiery et al., 2009).

Intratumor heterogeneity and cell cooperation

Cancer genome sequencing studies indicate that there are multiple subpopulations that have distinct molecular properties within a single breast cancer (Ding et al., 2010; Navin et al., 2011; Nik-Zainal et al., 2012; Shah et al., 2009). It is known that tumor cells need to interact with the microenvironment and communicate with the surrounding cells of various types, such as endothelial cells, mesenchymal stem cells, cancer-associated fibroblasts and immune cells (Ungefroren et al., 2011).

However, within the tumor cells, different subpopulations also need to interact with each other. In the view of clonal evolution model, subpopulations are considered as competitors to others that the certain subpopulation will gain growth advantage over others after certain mutations. At some times, tumor cells also need to cooperate. By injecting EMT cells and non-EMT cells separately or

together, Tsuji et al. show that both EMT and non-EMT breast tumor cells are needed for lung metastasis and they work together (Tsuji et al., 2009). Cleary et al. show that in Wnt-driven mammary cancers, both basal Hras mutant and luminal Hras wild-type subpopulations are required for tumor formation (Cleary et al., 2014). Yamamoto et al. also showed that CSCs and non-CSCs interact with each other through paracrine signaling (Yamamoto et al., 2013). In this study, they found that non-CSCs produce JAG1 while CSCs express Notch3 receptor so that non-CSCs can regulate CSCs through paracrine Notch signaling by providing JAG1 ligand.

Integrin β 4 in breast tumor metastasis

In breast cancer, β 4 has been shown to switch from linking cytokeratin to actin and redistribute to cell protrusions (Lipscomb and Mercurio, 2005; Rabinovitz and Mercurio, 1997; Rabinovitz et al., 1999). In addition LM332, the ligand of β 4, is enriched at the invading front of several carcinomas (Kim et al., 2011a; Pyke et al., 1995; Xue et al., 2011). Given the important role that β 4 plays for tumor invasion, I think it is very likely that β 4⁺ cells associated with a basal epithelial program in tumor are the cells that at the invading front to drive tumor invasion. These cells don't undergo EMT or MET, and they stay epithelial through tumor metastasis. To examine this hypothesis, I made a β 4 reporter mouse so that I can cross it with breast tumor models and track β 4⁺ cells during tumor progression and metastasis.

Expression of integrin $\beta 4$ has been found to correlate with highly metastatic cells {Dedhar:1993ww, Cimino:1997uv}. Periostin, an integrin $\beta 4$ ligand, is an extracellular matrix protein induced in the lung stroma and provides a niche to support metastatic colonization (Baril et al., 2006; Malanchi et al., 2012). This could be the way that $\beta 4^+$ tumor cells migrate to the metastatic site and form metastasis. Interestingly, a recent study linked integrin $\beta 4$ to the exosomes in tumor metastasis (Hoshino et al., 2015). Exosomes are released from tumor cells into the microenvironment of the primary site and have been found to play roles in tumor invasion and metastasis (Lowry et al., 2015). This study shows that exosomes have distinct integrin expression patterns, in which the integrin $\alpha 6\beta 4$ and $\alpha 6\beta 1$ are associated with lung metastasis (Hoshino et al., 2015). The data suggests that exosomes expressing integrins will first travel to secondary site and secrete various proteins they bring to create a metastasis niche. Tumor cells will then migrate to that secondary site to form metastasis. These studies suggest a role for integrin $\beta 4$ in creating the metastatic niche at the secondary site and supporting metastatic colonization.

Integrin $\beta 4$ and PTHrP

The role of PTHrP as a mediator of osteolytic bone metastasis in breast cancer has been extensively studied in bone metastasis. PTHrP is a paracrine factor that is expressed by tumor cells and is secreted to the bone microenvironment,

which together with cytokines and growth factors play a negative impact on osteoblast function and lead to the destruction of bone. PTHrP produced by breast cancer cells enhances osteoclastic activity via the expression of osteoblastic receptor activator of nuclear factor kappa B ligand (RANKL), vascular endothelial growth factor (VEGF) and matrix metalloproteinase-13 (MMP13) (Chen et al., 2010; Ibaragi et al., 2010; Isowa et al., 2010; Zhang et al., 2010). PTHrP is also important for mammary tumor formation and invasion (Lowry et al., 2015). This thesis study also discussed the possibility that $\beta 4$ regulation of PTHrP is important for tumor formation by regulating CSCs in a non-cell-autonomous manner. We showed in a couple of different systems that PTHrP is expressed by $\beta 4^+CD24^+CD44^+$ non-CSCs that have basal epithelial markers while its receptor PTH1R is associated with $\beta 4^-CD24^-CD44^+$ CSCs. These two populations interact with each other through PTH1R paracrine signaling. This work also shows that integrin $\beta 4$ can regulate PTHrP expression which can be directly affected by the ligation of integrin $\alpha 6\beta 4$ when culturing cells on laminins.

Conclusion Remarks and Perspectives

This study investigates the role of integrin $\beta 4$ in breast tumor formation and the mechanism involved. It shows that $\beta 4$ is heterogeneously expressed in breast tumors and it is important for tumor formation though it is not directly expressed by CSCs. I characterized $\beta 4^+$ and $\beta 4^-$ subpopulations in breast tumors and

investigated how they may cooperate with each other during tumor progression. My data shows that $\beta 4$ is expressed by a basal epithelial $CD24^+CD44^+$ population which are likely to be the leader cells at the invading front to drive tumor invasion. $\beta 4^+$ non-CSCs can regulate $\beta 4^-$ CSC functions through paracrine signaling, such as Notch signaling and PTH1R signaling. Further experiments needed to be done to demonstrate the interactions between $\beta 4^+$ and $\beta 4^-$ populations and show that these interactions are essential for stem cell properties. $\beta 4^+$ and $\beta 4^-$ populations can be isolated and labeled before co-culture to examine the stem cell properties as well as signalings and downstream genes in $\beta 4^-$ population. $\beta 4^+$ and $\beta 4^-$ populations can also be cultured in transwell plates to see if the physical interactions between the two populations are needed to maintain the CSC properties of $\beta 4^-$ population.

I also investigated the possibility of PTH1R signaling involved in $\beta 4$ regulation on CSCs. More experiments need to be done to confirm the regulation of integrin $\beta 4$ on PTHrP and study the mechanism involved. It is also important to show that PTH1R signaling is critical for CSC functions. This can be studied by knocking down PTH1R in breast cancer cell lines and examine CSC properties either by mammosphere formation assay or tumor initiation assay in vivo. I also showed that $\beta 4$ deletion in SUM-159 cells affected the CSC function of $CD24^-CD44^+$ population to give rise to $CD24^+CD44^+$ population. Further examination is needed to check the changes in $CD24^-CD44^+$ population after $\beta 4$ deletion at the

molecular level. Finally, to better study the role of integrin $\beta 4$ in tumor progression and metastasis, the $\beta 4$ reporter mouse I'm making will be a very useful tool. The $\beta 4$ reporter mouse can be crossed with breast tumor model so that the role of $\beta 4^+$ population and the interactions between $\beta 4^+$ and $\beta 4^-$ populations can be better studied. Moreover, $\beta 4^+$ population can also be tracked during tumor progression from the primary tumor site to metastatic site to study the role of $\beta 4$ in tumor metastasis. One important question is that whether $\beta 4^+$ population travel to the metastatic site or the $\beta 4^-$ tumor cells from primary tumor site undergo MET and gain $\beta 4$ expression at the metastatic site.

BIBLIOGRAPHY

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. *Pnas* 100, 3983–3988.

Allred, D.C. (2010). Ductal carcinoma in situ: terminology, classification, and natural history. *J. Natl. Cancer Inst. Monographs* 2010, 134–138.

Baril, P., Gangeswaran, R., Mahon, P.C., Caulee, K., Kocher, H.M., Harada, T., Zhu, M., Kalthoff, H., Crnogorac-Jurcevic, T., and Lemoine, N.R. (2006). Periostin promotes invasiveness and resistance of pancreatic cancer cells to hypoxia-induced cell death: role of the $\beta 4$ integrin and the PI3k pathway. *Oncogene* 26, 2082–2094.

Bertotti, A., Comoglio, P.M., and Trusolino, L. (2006). B4 integrin activates Shp-2-Src signaling pathway that sustains HGF-induced anchorage-independent growth. *J Cell Biol* 175 (6), 993-1003.

Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 3, 730–737.

Borradori, L., and Sonnenberg, A. (1999). Structure and function of hemidesmosomes: more than simple adhesion complexes. *J Investig Dermatol* 112, 411–418.

Bouras, T., Pal, B., Vaillant, F., Harburg, G., Asselin-Labat, M.-L., Oakes, S.R., Lindeman, G.J., and Visvader, J.E. (2008). Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. *Cell Stem Cell* 3, 429–441.

Campbell, I.D., and Humphries, M.J. (2011). Integrin Structure, Activation, and Interactions. *Cold Spring Harb Perspect Biol* 3, a004994–a004994.

Campbell, L.L., and Polyak, K. (2014). Breast Tumor Heterogeneity: Cancer Stem Cells or Clonal Evolution? *Cell Cycle* 6, 2332–2338.

Chacón, R.D., and Costanzo, M.V. (2010). Triple-negative breast cancer. *Breast Cancer Res.* 12 *Suppl* 2, S3.

Chang, C., Yang, X., Pursell, B., and Mercurio, A.M. (2013). Id2 complexes with the SNAG domain of Snai1 inhibiting Snai1-mediated repression of integrin $\beta 4$. *Mol. Cell. Biol.* 33, 3795–3804.

- Chapman, H.A., Li, X., Alexander, J.P., Brumwell, A., Lorizio, W., Tan, K., Sonnenberg, A., Wei, Y., and Vu, T.H. (2011). Integrin $\alpha 6\beta 4$ identifies an adult distal lung epithelial population with regenerative potential in mice. *J. Clin. Invest.* *121*, 2855–2862.
- Chen, Y.-C., Sosnoski, D.M., and Mastro, A.M. (2010). Breast cancer metastasis to the bone: mechanisms of bone loss. *Breast Cancer Res.* *12*, 215.
- Cheung, K.J., Gabrielson, E., Werb, Z., and Ewald, A.J. (2013). Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell* *155*, 1639–1651.
- Chung, J., Bachelder, R.E., Lipscomb, E.A., Shaw, L.M., and Mercurio, A.M. (2002). Integrin (alpha 6 beta 4) regulation of eIF-4E activity and VEGF translation: a survival mechanism for carcinoma cells. *J Cell Biol* *158*, 165–174.
- Chung, J., Yoon, S.-O., Lipscomb, E.A., and Mercurio, A.M. (2004). The Met receptor and alpha 6 beta 4 integrin can function independently to promote carcinoma invasion. *J. Biol. Chem.* *279*, 32287–32293.
- Cimino, L., Perrotti, D., D'Agostino, G., Falcioni, R., and Sacchi, A. (1997). Beta 4 integrin subunit expression is downregulated in low metastatic carcinoma variants. *Cancer Detect. Prev.* *21*, 158–166.
- Cleary, A.S., Leonard, T.L., Gestl, S.A., and Gunther, E.J. (2014). Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* *508*, 113–117.
- Collins, A.T., Berry, P.A., Hyde, C., Stower, M.J., and Maitland, N.J. (2005). Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res* *65*, 10946–10951.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.
- Dedhar, S., Saulnier, R., Nagle, R., and Overall, C.M. (1993). Specific alterations in the expression of alpha 3 beta 1 and alpha 6 beta 4 integrins in highly invasive and metastatic variants of human prostate carcinoma cells selected by in vitro invasion through reconstituted basement membrane. *Clin. Exp. Metastasis* *11*, 391–400.
- Dent, R., Trudeau, M., Pritchard, K.I., Hanna, W.M., Kahn, H.K., Sawka, C.A., Lickley, L.A., Rawlinson, E., Sun, P., and Narod, S.A. (2007). Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin. Cancer Res.* *13*,

4429–4434.

Ding, L., Ellis, M.J., Li, S., Larson, D.E., Chen, K., Wallis, J.W., Harris, C.C., McLellan, M.D., Fulton, R.S., Fulton, L.L., et al. (2010). Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* *464*, 999–1005.

Dittmer, A., Vetter, M., Schunke, D., Span, P.N., Sweep, F., Thomssen, C., and Dittmer, J. (2006). Parathyroid hormone-related protein regulates tumor-relevant genes in breast cancer cells. *J. Biol. Chem.* *281*, 14563–14572.

Dontu, G., Jackson, K.W., McNicholas, E., Kawamura, M.J., Abdallah, W.M., and Wicha, M.S. (2004). Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. *Breast Cancer Res.* *6*, R605–R615.

Dowling, J., Yu, Q.C., and Fuchs, E. (1996). Beta4 integrin is required for hemidesmosome formation, cell adhesion and cell survival. *J Cell Biol* *134*, 559–572.

Eramo, A., Lotti, F., Sette, G., Piloizzi, E., Biffoni, M., Di Virgilio, A., Conticello, C., Ruco, L., Peschle, C., and De Maria, R. (2008). Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ.* *15*, 504–514.

Fischer, K.R., Durrans, A., Lee, S., Sheng, J., Li, F., Wong, S.T.C., Choi, H., Rayes, El, T., Ryu, S., Troeger, J., et al. (2015). Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* *527*, 472–476.

Folgiero, V., Bachelder, R.E., Bon, G., Sacchi, A., Falcioni, R., and Mercurio, A.M. (2007). The alpha6beta4 integrin can regulate ErbB-3 expression: implications for alpha6beta4 signaling and function. *Cancer Res* *67*, 1645–1652.

Gerson, K.D., Shearstone, J.R., Maddula, V.S.R.K., Seligmann, B.E., and Mercurio, A.M. (2012). Integrin β 4 regulates SPARC protein to promote invasion. *J. Biol. Chem.* *287*, 9835–9844.

Gerson, K.D., Maddula, V.S.R.K., Seligmann, B.E., Shearstone, J.R., Khan, A., and Mercurio, A.M. (2012). Effects of β 4 integrin expression on microRNA patterns in breast cancer. *Biol Open* *1*, 658–666.

Gjorevski, N., and Nelson, C.M. (2011). Integrated morphodynamic signalling of the mammary gland. *Nat. Rev. Mol. Cell Biol.* *12*, 581–593.

Glukhova, M.A., and Streuli, C.H. (2013). How integrins control breast biology. *Current Opinion in Cell Biology* *25*, 633–641.

- Goel, H.L., Pursell, B., Chang, C., Shaw, L.M., Mao, J., Simin, K., Kumar, P., Vander Kooi, C.W., Shultz, L.D., Greiner, D.L., et al. (2013). GLI1 regulates a novel neuropilin-2/ $\alpha 6\beta 1$ integrin based autocrine pathway that contributes to breast cancer initiation. *EMBO Mol Med* 5, 488–508.
- Goel, H.L., Gritsko, T., Pursell, B., Chang, C., Shultz, L.D., Greiner, D.L., Norum, J.H., Toftgard, R., Shaw, L.M., and Mercurio, A.M. (2014). Regulated splicing of the $\alpha 6$ integrin cytoplasmic domain determines the fate of breast cancer stem cells. *Cell Rep* 7, 747–761.
- Green, K.J., and Jones, J.C. (1996). Desmosomes and hemidesmosomes: structure and function of molecular components. *Faseb J* 10, 871–881.
- Guo, W., and Giancotti, F.G. (2004). Integrin signalling during tumour progression. *Nat. Rev. Mol. Cell Biol.* 5, 816–826.
- Guo, W., Pylayeva, Y., Pepe, A., Yoshioka, T., Muller, W.J., Inghirami, G., and Giancotti, F.G. (2006). Beta 4 integrin amplifies ErbB2 signaling to promote mammary tumorigenesis. *Cell* 126, 489–502.
- Guschin, D.Y., Waite, A.J., Katibah, G.E., Miller, J.C., Holmes, M.C., and Rebar, E.J. (2010). A rapid and general assay for monitoring endogenous gene modification. *Methods Mol. Biol.* 649, 247–256.
- Harrison, H., Farnie, G., Howell, S.J., Rock, R.E., Stylianou, S., Brennan, K.R., Bundred, N.J., and Clarke, R.B. (2010). Regulation of breast cancer stem cell activity by signaling through the Notch4 receptor. *Cancer Res* 70, 709–718.
- Hermann, P.C., Huber, S.L., Herrler, T., Aicher, A., Ellwart, J.W., Guba, M., Bruns, C.J., and Heeschen, C. (2007). Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* 1, 313–323.
- Hoshino, A., Costa-Silva, B., Shen, T.-L., Rodrigues, G., Hashimoto, A., Tesic Mark, M., Molina, H., Kohsaka, S., Di Giannatale, A., Ceder, S., et al. (2015). Tumour exosome integrins determine organotropic metastasis. *Nature* 527, 329–335.
- Hu, M., Yao, J., Carroll, D.K., Weremowicz, S., Chen, H., Carrasco, D., Richardson, A., Violette, S., Nikolskaya, T., Nikolsky, Y., et al. (2008). Regulation of in situ to invasive breast carcinoma transition. *Cancer Cell* 13, 394–406.
- Hynes, R.O. (2002). Integrins: bidirectional, allosteric signaling machines. *Cell* 110, 673–687.

Ibaragi, S., Shimo, T., Iwamoto, M., Hassan, N.M.M., Kodama, S., Isowa, S., and Sasaki, A. (2010). Parathyroid hormone-related peptide regulates matrix metalloproteinase-13 gene expression in bone metastatic breast cancer cells. *Anticancer Res.* 30, 5029–5036.

Isowa, S., Shimo, T., Ibaragi, S., Kurio, N., Okui, T., Matsubara, K., Hassan, N.M.M., Kishimoto, K., and Sasaki, A. (2010). PTHrP regulates angiogenesis and bone resorption via VEGF expression. *Anticancer Res.* 30, 2755–2767.

Jauliac, S., López-Rodriguez, C., Shaw, L.M., Brown, L.F., Rao, A., and Toker, A. (2002). The role of NFAT transcription factors in integrin-mediated carcinoma invasion. *Nat. Cell Biol.* 4, 540–544.

Kassam, F., Enright, K., Dent, R., Dranitsaris, G., Myers, J., Flynn, C., Fralick, M., Kumar, R., and Clemons, M. (2009). Survival outcomes for patients with metastatic triple-negative breast cancer: implications for clinical practice and trial design. *Clin. Breast Cancer* 9, 29–33.

Kennecke, H., Yerushalmi, R., Woods, R., Cheang, M.C.U., Voduc, D., Speers, C.H., Nielsen, T.O., and Gelmon, K. (2010). Metastatic behavior of breast cancer subtypes. *J. Clin. Oncol.* 28, 3271–3277.

Kim, B.G., An, H.J., Kang, S., Choi, Y.P., Gao, M.-Q., Park, H., and Cho, N.H. (2011a). Laminin-332-rich tumor microenvironment for tumor invasion in the interface zone of breast cancer. *Am. J. Pathol.* 178, 373–381.

Kim, J.H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K.Y., Kim, M.-K., Shin, B.A., and Choi, S.-Y. (2011b). High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS ONE* 6, e18556.

Klinowska, T.C., Alexander, C.M., Georges-Labouesse, E., van der Neut, R., Kreidberg, J.A., Jones, C.J., Sonnenberg, A., and Streuli, C.H. (2001). Epithelial development and differentiation in the mammary gland is not dependent on alpha 3 or alpha 6 integrin subunits. *Developmental Biology* 233, 449–467.

Kreso, A., and Dick, J.E. (2014). Evolution of the cancer stem cell model. *Cell Stem Cell* 14, 275–291.

Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367, 645–648.

Lee, J.L., and Streuli, C.H. (2014). Integrins and epithelial cell polarity. *J. Cell.*

Sci. 127, 3217–3225.

Levental, K.R., Yu, H., Kass, L., Lakins, J.N., Egeblad, M., Erler, J.T., Fong, S.F.T., Csiszar, K., Giaccia, A., Weninger, W., et al. (2009). Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* 139, 891–906.

Li, C., Heidt, D.G., Dalerba, P., Burant, C.F., Zhang, L., Adsay, V., Wicha, M., Clarke, M.F., and Simeone, D.M. (2007). Identification of pancreatic cancer stem cells. *Cancer Res* 67, 1030–1037.

Li, C.-Y., Wang, P., Zhang, S., Liu, Y., and Zhang, J. (2013). [Clinicopathological features and prognosis of triple-negative breast cancer]. *Zhonghua Zhong Liu Za Zhi* 35, 463–467.

Li, J., Karaplis, A.C., Huang, D.C., Siegel, P.M., Camirand, A., Yang, X.F., Muller, W.J., and Kremer, R. (2011). PTHrP drives breast tumor initiation, progression, and metastasis in mice and is a potential therapy target. *J. Clin. Invest.* 121, 4655–4669.

Li, J., Sun, H., Feltri, M.L., and Mercurio, A.M. (2015). Integrin $\beta 4$ regulation of PTHrP underlies its contribution to mammary gland development. *Developmental Biology* 407, 313–320.

Liedtke, C., Mazouni, C., Hess, K.R., André, F., Tordai, A., Mejia, J.A., Symmans, W.F., Gonzalez-Angulo, A.M., Hennessy, B., Green, M., et al. (2008). Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* 26, 1275–1281.

Lipscomb, E.A., and Mercurio, A.M. (2005). Mobilization and activation of a signaling competent $\alpha 6\beta 4$ integrin underlies its contribution to carcinoma progression. *Cancer Metastasis Rev.* 24, 413–423.

Lipscomb, E.A., Simpson, K.J., Lyle, S.R., Ring, J.E., Dugan, A.S., and Mercurio, A.M. (2005). The $\alpha 6\beta 4$ integrin maintains the survival of human breast carcinoma cells in vivo. *Cancer Res* 65, 10970–10976.

Lowry, M.C., Gallagher, W.M., and O'Driscoll, L. (2015). The Role of Exosomes in Breast Cancer. *Clin. Chem.* 61, 1457–1465.

Lu, S., Simin, K., Khan, A., and Mercurio, A.M. (2008). Analysis of integrin $\beta 4$ expression in human breast cancer: association with basal-like tumors and prognostic significance. *Clin Cancer Res* 14, 1050–1058.

Malanchi, I., Santamaria-Martínez, A., Susanto, E., Peng, H., Lehr, H.-A., Delaloye, J.-F., and Huelsken, J. (2012). Interactions between cancer stem cells

and their niche govern metastatic colonization. *Nature* 481, 85–89.

Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715.

Mariotti, A., Kedeshian, P.A., Dans, M., Curatola, A.M., Gagnoux-Palacios, L., and Giancotti, F.G. (2001). EGF-R signaling through Fyn kinase disrupts the function of integrin alpha6beta4 at hemidesmosomes: role in epithelial cell migration and carcinoma invasion. *J Cell Biol* 155, 447–458.

Meacham, C.E., and Morrison, S.J. (2013). Tumour heterogeneity and cancer cell plasticity. *Nature* 501, 328–337.

Mercurio, A.M. (1995). Laminin receptors: achieving specificity through cooperation. *Trends Cell Biol.* 5, 419–423.

Mercurio, A.M. (2003). Invasive skin carcinoma--Ras and alpha6beta4 integrin lead the way. *Cancer Cell* 3, 201–202.

Mukai, H. (2010). Treatment strategy for HER2-positive breast cancer. *Int. J. Clin. Oncol.* 15, 335–340.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94.

Nik-Zainal, S., Van Loo, P., Wedge, D.C., Alexandrov, L.B., Greenman, C.D., Lau, K.W., Raine, K., Jones, D., Marshall, J., Ramakrishna, M., et al. (2012). The life history of 21 breast cancers. *Cell* 149, 994–1007.

Nisticò, P., Di Modugno, F., Spada, S., and Bissell, M.J. (2014). β 1 and β 4 integrins: from breast development to clinical practice. *Breast Cancer Res.* 16, 459.

O'Brien, C.A., Pollett, A., Gallinger, S., and Dick, J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* 445, 106–110.

Ocaña, O.H., Córcoles, R., Fabra, A., Moreno-Bueno, G., Acloque, H., Vega, S., Barrallo-Gimeno, A., Cano, A., and Nieto, M.A. (2012). Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell* 22, 709–724.

Patrawala, L., Calhoun, T., Schneider-Broussard, R., Li, H., Bhatia, B., Tang, S., Reilly, J.G., Chandra, D., Zhou, J., Claypool, K., et al. (2006). Highly purified CD44+ prostate cancer cells from xenograft human tumors are enriched in tumorigenic and metastatic progenitor cells. *Oncogene* 25, 1696–1708.

Perou, C.M., Sørlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. *Nature* 406, 747–752.

Pontier, S.M., and Muller, W.J. (2009). Integrins in mammary-stem-cell biology and breast-cancer progression--a role in cancer stem cells? *J. Cell. Sci.* 122, 207–214.

Prince, M.E., Sivanandan, R., Kaczorowski, A., Wolf, G.T., Kaplan, M.J., Dalerba, P., Weissman, I.L., Clarke, M.F., and Ailles, L.E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Pnas* 104, 973–978.

Pyke, C., Salo, S., Ralfkiaer, E., Rømer, J., Danø, K., and Tryggvason, K. (1995). Laminin-5 is a marker of invading cancer cells in some human carcinomas and is coexpressed with the receptor for urokinase plasminogen activator in budding cancer cells in colon adenocarcinomas. *Cancer Res* 55, 4132–4139.

Rabinovitz, I., and Mercurio, A.M. (1997). The integrin alpha6beta4 functions in carcinoma cell migration on laminin-1 by mediating the formation and stabilization of actin-containing motility structures. *J Cell Biol* 139, 1873–1884.

Rabinovitz, I., Toker, A., and Mercurio, A.M. (1999). Protein kinase C-dependent mobilization of the alpha6beta4 integrin from hemidesmosomes and its association with actin-rich cell protrusions drive the chemotactic migration of carcinoma cells. *J Cell Biol* 146, 1147–1160.

Rakha, E.A., Reis-Filho, J.S., and Ellis, I.O. (2008). Basal-like breast cancer: a critical review. *J. Clin. Oncol.* 26, 2568–2581.

Ricci-Vitiani, L., Lombardi, D.G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C., and De Maria, R. (2007). Identification and expansion of human colon-cancer-initiating cells. *Nature* 445, 111–115.

Rimawi, M.F., Schiff, R., and Osborne, C.K. (2015). Targeting HER2 for the treatment of breast cancer. *Annu. Rev. Med.* 66, 111–128.

Rivenbark, A.G., O'Connor, S.M., and Coleman, W.B. (2013). Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine. *Am. J. Pathol.* 183, 1113–1124.

Shah, S.P., Morin, R.D., Khattra, J., Prentice, L., Pugh, T., Burleigh, A., Delaney, A., Gelmon, K., Guliany, R., Senz, J., et al. (2009). Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* *461*, 809–813.

Shamir, E.R., Pappalardo, E., Jorgens, D.M., Coutinho, K., Tsai, W.-T., Aziz, K., Auer, M., Tran, P.T., Bader, J.S., and Ewald, A.J. (2014). Twist1-induced dissemination preserves epithelial identity and requires E-cadherin. *J Cell Biol* *204*, 839–856.

Shaw, L.M., Rabinovitz, I., Wang, H.H.F., Toker, A., and Mercurio, A.M. (1997). Activation of Phosphoinositide 3-OH Kinase by the $\alpha 6 \beta 4$ Integrin Promotes Carcinoma Invasion. *Cell* *91*, 949–960.

Siegel, R.L., Miller, K.D., and Jemal, A. (2015). Cancer statistics, 2015. *CA Cancer J Clin* *65*, 5–29.

Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. *Nature* *432*, 396–401.

Skibinski, A., and Kuperwasser, C. (2015). The origin of breast tumor heterogeneity. *Oncogene*.

Sorlie, T. (2004). Molecular portraits of breast cancer: tumour subtypes as distinct disease entities. *Eur. J. Cancer* *40*, 2667–2675.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., et al. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Pnas* *100*, 8418–8423.

Soung, Y.H., Gil, H.J., Clifford, J.L., and Chung, J. (2011). Role of $\alpha 6 \beta 4$ integrin in cell motility, invasion and metastasis of mammary tumors. *Cpps* *12*, 23–29.

Srichai, M.B., and Zent, R. (2010). Integrin Structure and Function. In *Cell-Extracellular Matrix Interactions in Cancer*, (New York, NY: Springer New York), pp. 19–41.

Takada, Y., Ye, X., and Simon, S. (2007). The integrins. *Genome Biol*.

Thiery, J.P., Acloque, H., Huang, R.Y.J., and Nieto, M.A. (2009). Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* *139*, 871–890.

Trusolino, L., Bertotti, A., and Comoglio, P.M. (2001). A signaling adapter

function for $\alpha 6 \beta 4$ integrin in the control of HGF-dependent invasive growth. *Cell* 107, 643–654.

Tsai, J.H., Donaher, J.L., Murphy, D.A., Chau, S., and Yang, J. (2012). Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* 22, 725–736.

Tsuji, T., Ibaragi, S., and Hu, G.-F. (2009). Epithelial-mesenchymal transition and cell cooperativity in metastasis. *Cancer Res* 69, 7135–7139.

Ungefroren, H., Sebens, S., Seidl, D., Lehnert, H., and Hass, R. (2011). Interaction of tumor cells with the microenvironment. *Cell Commun. Signal* 9, 18.

van der Neut, R., Cachaço, A.S., Thorsteinsdóttir, S., Janssen, H., Prins, D., Bulthuis, J., van der Valk, M., Calafat, J., and Sonnenberg, A. (1999). Partial rescue of epithelial phenotype in integrin $\beta 4$ null mice by a keratin-5 promoter driven human integrin $\beta 4$ transgene. *J. Cell. Sci.* 112 (Pt 22), 3911–3922.

van der Neut, R., Krimpenfort, P., Calafat, J., Niessen, C.M., and Sonnenberg, A. (1996). Epithelial detachment due to absence of hemidesmosomes in integrin $\beta 4$ null mice. *Nat. Genet.* 13, 366–369.

Vieira, A.F., Ricardo, S., Ablett, M.P., Dionísio, M.R., Mendes, N., Albergaria, A., Farnie, G., Gerhard, R., Cameselle-Teijeiro, J.F., Seruca, R., et al. (2012). P-cadherin is coexpressed with CD44 and CD49f and mediates stem cell properties in basal-like breast cancer. *Stem Cells* 30, 854–864.

Vieira, A.F., Ribeiro, A.S., Dionísio, M.R., Sousa, B., Nobre, A.R., Albergaria, A., Santiago-Gómez, A., Mendes, N., Gerhard, R., Schmitt, F., et al. (2014). P-cadherin signals through the laminin receptor $\alpha 6 \beta 4$ integrin to induce stem cell and invasive properties in basal-like breast cancer cells. *Oncotarget* 5, 679–692.

Ward, E.M., DeSantis, C.E., Lin, C.C., Kramer, J.L., Jemal, A., Kohler, B., Brawley, O.W., and Gansler, T. (2015). Cancer statistics: Breast cancer in situ. *CA Cancer J Clin* 65, 481–495.

Wu, C., Wei, Q., Utomo, V., Nadesan, P., Whetstone, H., Kandel, R., Wunder, J.S., and Alman, B.A. (2007). Side population cells isolated from mesenchymal neoplasms have tumor initiating potential. *Cancer Res* 67, 8216–8222.

Xue, L.-Y., Zou, S.-M., Zheng, S., Liu, X.-Y., Wen, P., Yuan, Y.-L., Lin, D.-M., and Lu, N. (2011). Expressions of the $\gamma 2$ chain of laminin-5 and secreted protein acidic and rich in cysteine in esophageal squamous cell carcinoma and their relation to prognosis. *Chin J Cancer* 30, 69–78.

Yamamoto, M., Taguchi, Y., Ito-Kureha, T., Semba, K., Yamaguchi, N., and Inoue, J.-I. (2013). NF- κ B non-cell-autonomously regulates cancer stem cell populations in the basal-like breast cancer subtype. *Nat Comms* 4, 2299.

Yang, X., Pursell, B., Lu, S., Chang, T.-K., and Mercurio, A.M. (2009). Regulation of beta 4-integrin expression by epigenetic modifications in the mammary gland and during the epithelial-to-mesenchymal transition. *J. Cell. Sci.* 122, 2473–2480.

Yoon, S.-O., Shin, S., and Lipscomb, E.A. (2006). A novel mechanism for integrin-mediated ras activation in breast carcinoma cells: the alpha6beta4 integrin regulates ErbB2 translation and transactivates epidermal growth factor receptor/ErbB2 signaling. *Cancer Res* 66, 2732–2739.

Yoshioka, T., Otero, J., Chen, Y., Kim, Y.-M., Koutcher, J.A., Satagopan, J., Reuter, V., Carver, B., de Stanchina, E., Enomoto, K., et al. (2013). β 4 Integrin signaling induces expansion of prostate tumor progenitors. *J. Clin. Invest.* 123, 682–699.

Zahir, N., Lakins, J.N., Russell, A., Ming, W., Chatterjee, C., Rozenberg, G.I., Marinkovich, M.P., and Weaver, V.M. (2003). Autocrine laminin-5 ligates alpha6beta4 integrin and activates RAC and NFkappaB to mediate anchorage-independent survival of mammary tumors. *J Cell Biol* 163, 1397–1407.

Zhang, Y., Ma, B., and Fan, Q. (2010). Mechanisms of breast cancer bone metastasis. *Cancer Letters* 292, 1–7.

Zheng, X., Carstens, J.L., Kim, J., Scheible, M., Kaye, J., Sugimoto, H., Wu, C.-C., LeBleu, V.S., and Kalluri, R. (2015). Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 527, 525–530.

Zheng, Y., la Cruz, de, C.C., Sayles, L.C., Alleyne-Chin, C., Vaka, D., Knaak, T.D., Bigos, M., Xu, Y., Hoang, C.D., Shrager, J.B., et al. (2013). A Rare Population of CD24+ITGB4+Notch1 Cells Drives Tumor Propagation in NSCLC and Requires Notch3 for Self-Renewal. *Cancer Cell* 24, 59–74.