# EPIGENETIC REGULATION OF EPITHELIAL-MESENCHYMAL TRANSITION IN EPITHELIAL STEM CELLS AND TRIPLE NEGATIVE BREAST CANCER

Nicole Marie Vincent Jordan

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Pharmacology.

Chapel Hill
2012

Approved by:

Gary L. Johnson, Ph.D.

Terry Magnuson, Ph.D.

Charles M. Perou, Ph.D.

Zefeng Wang, Ph.D.

Adrienne Cox, Ph.D.

## ABSTRACT

NICOLE MARIE VINCENT JORDAN: Epigenetic Regulation of Epithelial-Mesenchymal Transition in Epithelial Stem Cells and Triple Negative Breast Cancer (Under the direction of Gary L. Johnson, Ph.D.)

Epithelial-mesenchymal transition (EMT) is a cellular program responsible for the conversion of stationary epithelial cells characterized by apical-basolateral polarity to motile mesenchymal cells with front-back end polarity. EMT drives the fundamental developmental processes of implantation, gastrulation and neural crest formation. This essential developmental program becomes reactivated during cancer progression to promote metastasis. These studies examine the molecular and epigenetic mechanisms responsible for the induction of EMT in several developmental epithelial stem cell models and Basal-like and Claudin-low breast cancer subtypes.

Using trophoblast stem (TS) cells paused in EMT, we have defined the molecular and epigenetic mechanisms responsible for modulating the induction of EMT. Targeted inactivation of MAP3K4, knockdown of CBP, or overexpression of SNAI1 in TS cells induced intermediate EMT phenotypes. These TS cells exhibited epigenetic changes in the histone acetylation landscape that caused loss of epithelial maintenance while preserving self-renewal and multipotency. MAP3K4 controls the activity of the histone acetyltransferase CBP, and acetylation of histone H2B by CBP is required to maintain the epithelial phenotype of TS cells. Combined

loss of MAP3K4/CBP activity represses expression of epithelial genes and causes TS cells to undergo EMT while maintaining self-renewal and multipotency.

A similar EMT phenotype and an EMT gene signature shared with TS cells in EMT were found in Claudin-low breast cancer cells with properties of EMT and stemness. Therefore, it was hypothesized that the upregulated genes within the intersecting EMT signature define a novel set of genes critical to the induction of EMT and maintenance of the mesenchymal phenotype in breast cancer cells. We designed an RNAi screen in the Claudin-low EpCAM$^{(-)}$ population of SUM149 and SUM229 Basal-like breast cancer cells to identify genes responsible for maintaining the mesenchymal phenotype. Using this RNAi screening strategy for EMT-regulatory genes, the chromatin-remodeling factor Smarcd3/Baf60c was identified as a novel epigenetic regulator controlling the induction of EMT. Expression of Smarcd3/Baf60c in human mammary epithelial cells promoted EMT in a Wnt5a-dependent manner. These results highlight a new mechanism of epigenetic regulation of EMT, whereby Smarcd3/Baf60c promotes loss of epithelial cell polarity and dissolution of cell-cell junctions through activation of non-canonical WNT signaling.

To my “growing” family

## ACKNOWLEDGEMENTS

I would like to thank Dr. Gary Johnson for his wonderful mentorship. I have been very fortunate to be part of the Johnson lab, where if you can “design it, then you can do it”. Gary has been instrumental in helping me to think about the larger scientific question, and he has continually challenged me to design experiments to address these questions. I will forever be grateful for the scientific resources and independence that the Johnson lab has provided me.

I am also deeply indebted to Dr. Amy Abell for her outstanding co-mentorship. Amy has taught me the experimental side of science from how to carefully culture trophoblast stem cells to how to plan out a well-controlled experiment. I have benefitted tremendously from working collaboratively with Amy throughout my graduate school career.

I would also like to thank my thesis committee members Drs. Adrienne Cox, Terry Magnuson, Chuck Perou and Zefeng Wang for their scientific guidance. I am very fortunate to have had a thesis committee comprised of such a diverse and expert group of scientists. My thesis committee has provided tremendous perspective toward addressing a wide range of scientific questions.

I am also grateful to Drs. Ken Harden and Lee Graves for making the Pharmacology Department at UNC-Chapel Hill a collegial and welcoming environment. Ken and Lee made me feel at home in this department from the

moment I walked through the doors for an interview. Without them, my decision to pursue my graduate training at UNC-Chapel Hill would not have been so easy.

Lastly, I would like to thank my family for their continual love and support. My parents, Penny and Larry, have supported my scientific interests ever since my middle school science fair projects. Without their encouragement, I would not be pursuing science to this day. My husband, David, has been my sounding board. He has patiently understood when I have come home consistently late on account of lengthy experiments. I will always appreciate his easy-going attitude with the mantra that it will all work out. And it has.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

BL: Basal-like intrinsic molecular subtype of breast cancer

CBP: CREB binding protein

Cdh1: E-cadherin gene

ChIP: chromatin immunoprecipitation

ChIP-seq: chromatin immunoprecipitation coupled to high throughput sequencing

CL: Claudin-low intrinsic molecular subtype of breast cancer

CTCs: circulating tumor cells

DAPI: 4',6'-diamidino-2-phenylindole

D3-HMECs: Smarcd3/Baf60c expressing human mammary epithelial cells

EMT: epithelial-mesenchymal transition

$EpCAM^{+/high}$: Basal-like breast cancer cells with high EpCAM expression

$EpCAM^{(-)}$: Basal-like breast cancer cells lacking EpCAM expression

ER: estrogen receptor

FACS: fluorescently activated cell sorting

FLAG: polypeptide tag consisting of DYKDDDDK

FGF4: fibroblast growth factor 4

GFP: green fluorescent protein

GO: Gene Ontology

H2AK5Ac: histone H2A lysine 5 acetylation

H2BK5Ac: histone H2B lysine 5 acetylation

H2BK12Ac: histone H2B lysine 12 acetylation

H2BK15Ac: histone H2B lysine 15 acetylation

H2BK20Ac: histone H2B lysine 20 acetylation

H3K9Ac: histone H3 lysine 9 acetylation

H4K8Ac: histone H4 lysine 8 acetylation

H2A/H2BAc: histones H2A and H2B acetylation

H3K4me3: histone H3 lysine 4 trimethylation

H3K9me3: histone H3 lysine 9 trimethylation

H3K27me3: histone H3 lysine 27 trimethylation

HA: hemagglutinin polypeptide tag

HAT: histone acetyltransferase

HDAC: histone deacetylase

HGFR: hepatocyte growth factor receptor

HMEC: human mammary epithelial cells

iPS: induced pluripotent stem cells

JNK: c-Jun N-terminal kinase

KI4: MAP3K4 kinase-inactive mouse knock-in model

MAPK: mitogen activated protein kinase

MAP3K4: mitogen activated protein kinase kinase kinase 4

MaSC: mammary stem cell

MET: mesenchymal-epithelial transition

miRNA: microRNA

Myc-tag: polypeptide tag derived from the c-Myc gene product

PDGFR: platelet-derived growth factor receptor

PR: progesterone receptor

PS: penicillin and streptomycin

qRT-PCR: quantitative real-time polymerase chain reaction

RNA: ribonucleic acid

RNAi: RNAi

shRNA: short hairpin RNA

siRNA: short interfering RNA

Slug-HMECs: Slug expressing human mammary epithelial cells

Snail-HMECs: Snail expressing human mammary epithelial cells

TGFβ: transforming growth factor beta

TIC: tumor initiating cells

TNBC: triple negative breast cancer

$T^{Diff}$: trophoblasts differentiated for four days

$T^{INV}$: invasive trophoblasts isolated from Matrigel-coated transwells following four days of differentiation

TS: trophoblast stem cells

$TS^{KI4}$: MAP3K4 kinase-inactive trophoblast stem cells

$TS^{shCBP}$: trophoblast stem cells lacking CBP expression

$TS^{Snail}$: trophoblast stem cells expressing Snail

$TS^{WT}$: wild-type trophoblast stem cells

TSS: transcription start site

# Chapter I

# Introduction

## Epithelial-Mesenchymal Transition

Epithelial-mesenchymal transition (EMT) is a morphogenic cellular program, whereby stationary epithelial cells convert to a motile mesenchymal morphology. The initiation and subsequent completion of EMT occurs through the precise coordination of numerous molecular events, including activation of EMT-inducing transcription factors, altered expression of cell-surface proteins, reorganization of the actin cytoskeleton and enhanced invasive properties (Thiery *et al*, 2009). Epithelial cells exhibit an organized apical-basolateral polarity maintained by the precise arrangement of actin filaments and adhesive structures such as tight junctions, adherens junctions and desmosomes. Specialized adhesive molecules, such as cadherins, integrins and other cell-surface proteins, are essential for maintenance of the epithelial phenotype by stabilizing cell-cell contacts (Thiery & Sleeman, 2006). Inhibition of these adhesive structures, most important of which is E-cadherin, induces a morphological switch to the mesenchymal phenotype. Conversely, mesenchymal cells are characterized by a unique spindle morphology defined by a front-back end polarity and enhanced invasive potential (Hay, 2005; Kalluri & Weinberg, 2009). In addition to promoting cellular migration and invasion, the transient phenotypic changes associated with formation of the mesenchymal state

during EMT have been associated with the acquisition of stem-like properties (Mani *et al*, 2008). The intermediate stage of EMT, coined the metastable phenotype, describes the simultaneous existence of both epithelial and mesenchymal characteristics and is of great importance for understanding the cellular changes associated with progression of the EMT program and maintenance of stem cell-like properties (Lee, 2006). Due to the tremendous difficulty in capturing cells in the intermediate states of EMT, most studies have focused on the initiation or completion of EMT. In chapter three, I describe the development of multiple trophoblast stem (TS) cell models to examine the molecular and epigenetic mechanisms underlying the intermediate stages of EMT and the connections between EMT and stemness.

**Classifications of EMT**

EMT is essential for the proper formation of the body plan and differentiation of many tissues and organs during embryonic development (Yang & Weinberg, 2008). During organ development, epithelia convert between epithelia and mesenchyme through multiple rounds of EMT and the reversible process mesenchymal-epithelial transition (MET) (Acloque *et al*, 2009). In adult tissue, the EMT program is reactivated during wound healing, organ fibrosis and tumor progression. EMTs are classified into three main categories based upon involvement in different biological processes (**Figure 1.1**) (Kalluri & Weinberg, 2009; Zeisberg & Neilson, 2009). Developmental EMTs constitute the first EMT subtype and include implantation, gastrulation and neural crest formation. Developmental EMT events

occur in a temporally controlled setting and contribute to embryonic morphogenesis and tissue remodeling during development. The second EMT subtype is associated with wound healing and tissue regeneration in adult tissue and occurs in response to inflammation. Persistent inflammatory signals produce continuous activation of EMT resulting in organ fibrosis. Organ fibrosis is commonly associated with inflammation of epithelia from the kidney, liver, lung and intestine and can ultimately lead to failure of these organs. The third EMT subtype occurs during cancer progression, whereby EMT in epithelial cancer cells produces cells with increased invasive capacity and promotes intravasation and dissemination of malignant cancer cells into the lumina of blood vessels.

Although the three subtypes of EMT have seemingly diverse outcomes, a common set of molecular features underlies each EMT event. Critical developmental signaling pathways (i.e. WNT, Notch, Hedgehog and TGFβ) with demonstrated importance in promoting developmental EMT are reactivated in the generation of EMT-associated pathologies, including organ fibrosis and the metastatic spread of cancer (**Figure 1.1**) (Thiery & Sleeman, 2006; Kalluri, 2009; Timmerman, 2004; Zavadil *et al*, 2004; Zavadil & Böttinger, 2005). Due to the considerable overlap between signaling pathways that regulate both developmental EMT and EMT-associated pathologies, it is important to examine developmental and cancer EMT in parallel. These parallel studies may increase our understanding of universal EMT regulatory mechanisms.

**Trophoblasts and the First Developmental EMT Event**

Implantation is the first developmental EMT event and is essential for proper placental development necessary to support a developing fetus. Failure of implantation prevents all successive EMT processes, such as gastrulation and neural crest formation (Rossant & Cross, 2001). During implantation, the epithelial trophoectoderm of the blastocyst undergoes EMT with changes in cell polarity, motility and adhesiveness (**Figure 1.1**). These changes in cell polarity are accompanied by extension of protrusive structures and loss of expression of the master-regulator of the epithelial phenotype E-cadherin, which are necessary for trophoblast invasion of the uterine epithelium. After implantation, the epithelial polar trophoectoderm overlying the inner cell mass of the blastocyst differentiates into specialized trophoblast lineages to form a functional placenta. In mouse, trophoblast stem cells in the polar trophoectoderm differentiate into three regions of the placenta including the labyrinth, which is comprised of syncytiotrophoblasts, the spongiotrophoblasts and the trophoblast giant cell layer. In humans, the epithelial trophoectoderm differentiates into an inner layer of proliferative cytotrophoblasts (i.e. trophoblast stem cells) and an outer layer of differentiated, multinucleated syncytiotrophoblasts. Invading trophoblasts generated through differentiation penetrate and remodel the maternal blood vessels, forming the chorionic villi. Considered the functional unit of the placenta, the chorionic villi provide a large surface area for nutrient exchange between mother and developing fetus. Formation of the chorionic villi, requiring extensive blood vessel remodeling, constitutes the second developmental EMT event (Gonzales *et al*, 1996; Pafilis *et al*, 2007;

Sutherland, 2003; Zhou *et al*, 1997b). Due to the functional importance of trophoblast stem (TS) cells to the first developmental EMT events, in chapter three, we examine TS cells during the progression of EMT.

***Pathologies Associated with Dysregulated Trophoblast EMT***

Precise regulation of trophoblast EMT is essential for successful implantation and placentation. Under normal conditions, the trophoblast invasive potential is restricted both temporally and spatially according to trophoblast differentiation signals. Several disease states are resultant from dysregulated trophoblast EMT. Two notable pathologies, preeclampsia and accreta, manifest as clinical outcomes directly linked to aberrant trophoblast invasion (Goldman-Wohl & Yagel, 2002). Preeclamsia is classically defined as an increase in blood pressure during pregnancy due to abnormally low blood flow into the placenta. At the molecular level, preeclamsia is the result of decreased or absent trophoblast invasion and remodeling of the maternal arteries (McMaster *et al*, 2004; Zhou *et al*, 1997a). The loss of trophoblast invasive potential has been connected to lack of adhesive molecules such as integrins and glycoproteins and to the increased production of extracellular matrix proteins such as collagens. As the disease state progresses, the invading trophoblasts are incapable of penetrating and remodeling the maternal spiral arteries, resulting in an immunological response against the invading trophoblasts and ultimately leading to damage and occlusion of the maternal arteries.

Conversely, placenta accreta is defined as abnormal adherence of the placenta to the uterine epithelium resulting from excessive invasion of hyperinvasive

trophoblasts into the uterine wall. Trophoblasts at the maternal-placental interface have demonstrated increased invasive capacity. Rather than being more differentiated and syncytiotrophoblastic in nature, these trophoblasts resemble self-renewing and highly proliferative cytotrophoblasts (i.e. trophoblast stem cells) in the early stages of implantation and placenta development (Khong, 2008; Perry *et al*, 2009). Due to the similarity between invasive trophoblasts and malignant cancer cells, it is likely that similar regulatory paradigms persist between these EMT subtypes (Ferretti *et al*, 2006).

**Metastatic Progression of Breast Cancer and EMT**

Cancer metastasis is responsible for the majority of cancer related deaths, and EMT is the biological program linked to the cellular traits associated with the metastatic progression of cancer (Nguyen *et al*, 2009; Berx *et al*, 2007; Yilmaz & Christofori, 2009; Huber *et al*, 2005; Christiansen, 2006). During metastasis, epithelial cancer cells utilize the EMT program to detach from the primary tumor, invade the surrounding stroma, intravasate into blood vessels and disseminate to distant tissues. The reverse process MET is required for the metastatic colonization and proliferation of cancer cells at distant tissues, resulting in the formation of macroscopic metastases (Valastyan & Weinberg, 2011). By understanding the underlying molecular mechanisms contributing to EMT and the reverse process MET, new strategies can potentially be developed for the prevention and treatment of malignant breast cancer.

Breast cancer is classified into five intrinsic subtypes on the basis of gene expression profiling. These intrinsic subtypes, including Luminal A, Luminal B, HER2-enriched, Basal-like and Claudin-low, are predictive of treatment outcome (Perou *et al*, 2000; Sørlie *et al*, 2003; 2001; Prat & Perou, 2011). Clinically, Basal-like and Claudin-low breast cancers are poor prognostic breast cancer subtypes, largely consisting of triple negative ER-, PR-negative and HER2-low invasive ductal carcinomas (TNBC). Patient prognosis with TNBC is poor due to the increased rates of recurrence, brain and lung metastasis and EMT-marker expression (Fulford *et al*, 2007; Gaedcke *et al*, 2007; Banerjee, 2006; Rakha *et al*, 2008a). At the molecular level, Basal-like and Claudin-low breast cancer share similar expression patterns, such as low expression of luminal- and HER2-related genes. Whereas the Basal-like subtype is characterized by high expression of proliferation and basal/myoepithelial cell-specific genes, including cytokeratins 5/14/17, the Claudin-low subtype is characterized by high expression of mesenchymal- and stem cell-related genes and low expression of epithelial-specific and tight junctions genes, including claudins 3/4/7 (Prat & Perou, 2011; Prat *et al*, 2010; Hennessy *et al*, 2009; Rakha *et al*, 2008b; Taube *et al*, 2010). Functionally, Basal-like and Claudin-low breast cancer demonstrate increased rates of cellular migration and invasion. The aggressive metastatic properties of the Basal-like and Claudin-low intrinsic molecular subtypes of breast cancer make these subtypes important to study with respect to the EMT. In chapter four, I describe Basal-like breast cancer cell lines with Claudin-low features and examine the molecular and epigenetic mechanisms regulating EMT with respect to breast cancer.

## Molecular Mechanisms of EMT

Multiple reports demonstrate that the very same transcription factors (i.e. SNAI1, SNAI2, TWIST1), adhesive structures (i.e. aν-β-integrin family) and signaling pathways (i.e. Tgfβ, WNT, MAPK, integrin) controlling defined developmental EMT programs are also responsible for regulating tumor progression (Thiery & Sleeman, 2006; Grünert *et al*, 2003). Loss of E-cadherin expression and gain of cellular invasiveness is central to the acquisition of EMT properties. Transcriptional and epigenetic repression of E-cadherin results in loss of E-cadherin at junctional complexes, destabilization of epithelial junctions and loss of epithelial cell apical-basolateral polarity (Larue *et al*, 1994; Derksen *et al*, 2006). Subsequent to the loss of E-cadherin at cellular junctions, β-catenin relocalizes from the cytoplasm to the nucleus and activates WNT signaling pathways. In addition to WNT signaling, significant crosstalk exists between transforming growth factor β (TGFβ), mitogen activated protein kinase (MAPK) and β1-integrin signaling pathways to promote the morphological changes associated with EMT.

### *Molecular Mechanisms of Trophoblast EMT*

Trophoblast stem (TS) cells can be isolated from the trophoectoderm of the developing blastocyst and the extraembryonic ectoderm of the E6.5 conceptus (Rossant & Cross, 2001). Self-renewal in TS cells, defined as cell division with the maintenance of multipotency, is maintained in culture conditions by fibroblast growth factor 4 (FGF4) (Erlebacher *et al*, 2004; Tanaka, 1998). E-cadherin and cortical actin localize around the periphery of TS cell epithelial colonies and maintain the integrity of the epithelial colonies (Larue *et al*, 1994; Shih *et al*, 2002). Removal of FGF4

induces TS cell differentiation into different trophoblast lineages. Differentiating trophoblasts undergo an EMT involving downregulation of E-cadherin expression, loss of cortical actin with concomitant gain of filamentous actin and the acquisition of invasiveness (Abell *et al*, 2009). Since TS cell differentiation pathways are essential to EMT processes of implantation and placentation, the, *in vitro*, TS cell differentiation model induced by removal of FGF4 is highly useful for interrogating the connections between EMT and TS cell differentiation.

The genetically engineered mouse model with targeted inactivation of MAP3K4 (KI4), a kinase regulating c-Jun N-terminal kinase (JNK) and p38 MAPK pathways, exhibits defective decidualization, fetal growth restriction and implantation defects attributable to trophoblast hyperinvasion and dysregulated EMT (Abell *et al*, 2005). When cultured in non-differentiating conditions in the presence of FGF4, $TS^{KI4}$ cells demonstrate a 20-fold increase in invasion compared to wild-type trophoblast stem cells ($TS^{WT}$ cells). Furthermore, the acquisition of the mesenchymal phenotype in $TS^{KI4}$ cells is demonstrated by several phenotypic and molecular changes, including loss of epithelial apical-basolateral polarity and a gain of mesenchymal front-back polarity with reduced levels of peripheral E-cadherin. Importantly, $TS^{KI4}$ cells maintain self-renewal properties as demonstrated by expression of trophoblast stem cell markers CDX2 and ESRRβ at levels similar to $TS^{WT}$ cells (Abell *et al*, 2009; Niwa *et al*, 2005; Ralston & Rossant, 2008; Strumpf, 2005). The simultaneous existence of stemness and EMT properties indicates that $TS^{KI4}$ cells represent a metastable EMT state. Metastable EMT is defined as an intermediate stage in the progression of the EMT program, whereby cells express

attributes of both epithelial and mesenchymal phenotypes while simultaneously expressing stem cell markers (Lee, 2006). Due to the transient nature of the metastable phenotype, it is difficult to observe this intermediate stage during the progression of EMT. Using $TS^{KI4}$ cells, we can examine the simultaneous maintenance of stemness and the acquisition of EMT properties that promote an unrestricted trophoblast invasive potential. During both development and the progression of cancer, hyperinvasion resulting from uncontrolled EMT signaling events can be detrimental to survival. Therefore, the $TS^{KI4}$ cell model system in metastable EMT is a unique tool to define novel transcriptional and epigenetic reprogramming events and signaling pathways important for progression of the EMT program.

***Molecular Mechanisms of Breast Cancer EMT***

In recent years, evidence for the role of EMT in the metastatic progression of breast cancer has continued to build (Huber *et al*, 2005; Christiansen, 2006). For instance, loss of E-cadherin expression is a hallmark of metastatic carcinoma, and expression of the transcriptional repressor of E-cadherin SNAI1 correlates with poor survival rates. Additionally, SNAI1 expression levels are elevated at the invading front of tumors (Derksen *et al*, 2006; Moody *et al*, 2005; Yook *et al*, 2006). The controversy surrounding the role of EMT in cancer metastasis stems from the difficulty of directly following the progression of EMT during tumor formation. In support of the connection between EMT and cancer metastasis, both proteomic and multiphoton microscopy analysis demonstrate the conversion of epithelial breast tumor cells to circulating mammary tumor cells with a mesenchymal phenotype

(Bonnomet *et al*, 2010; Willipinski-Stapelfeldt, 2005; Wyckoff, 2004). Furthermore, the genetically-engineered pancreatic cancer mouse model with YFP-labeled pancreatic epithelial cells reveals that EMT occurs prior to tumor cell dissemination and that circulating mouse pancreatic tumor cells have a mesenchymal phenotype (Rhim *et al*, 2012). The phenotypic similarity between the primary mammary tumor and the secondary tumor metastasis further supports a role for both EMT and MET in the metastatic progression of cancer.

Similar to EMT in development, reduced E-cadherin expression is a critical feature of EMT in breast cancer, causing loss of the epithelial phenotype and increased cellular motility and invasion. Conditional deletion of E-cadherin in the mammary glands of mice results in invasive lobular carcinomas (Derksen *et al*, 2006; Moll *et al*, 1993). These tumors are highly invasive while demonstrating partial EMT properties, indicative of an aggressive metastable EMT phenotype similar to that described for $TS^{KI4}$ epithelial stem cells in EMT. Based upon live-cell microscopy experiments, it is hypothesized that collective cell migration is the predominant mode of tissue invasion in epithelial cancers (Friedl & Wolf, 2003; Brabletz *et al*, 2005). Furthermore, receptor tyrosine kinases including the TGFβ-receptor, hepatocyte growth factor receptor (HGFR) and platelet-derived growth factor receptor (PDGFR) induce EMT in breast cancer with TGFβ being the dominant signaling pathway for activation of EMT (Valcourt *et al*, 2005; Zavadil *et al*, 2001; Jechlinger, 2006; Huber, 2004; Dong *et al*, 2007). In H-Ras transformed mammary epithelial cells, TGFβ promotes a mesenchymal cellular morphology and the expression of EMT-inducing transcription factors SNAI1, SNAI2 and TWIST1,

while PDGFR signaling is stimulated by TGFβ to promote a TGFβ-PDGFR autocrine signaling loop.

## Transcriptional Regulation of EMT

Expression of the adhesive glycoprotein and master-regulator of the epithelial phenotype, E-cadherin, is transcriptionally repressed by several transcription factors of the Snail and Zeb superfamilies (Nieto, 2002). In multiple EMT systems, SNAI1, SNAI2, ZEB1 and ZEB2 are coordinately upregulated during the initiation and progression of EMT. SNAI1 and SNAI2 are well-characterized transcriptional repressors of E-cadherin and function by binding the canonical E-box consensus sequence within the E-cadherin gene promoter. Ectopic expression of SNAI1 can induce EMT in several epithelial cell models, including human mammary epithelial cells (Mani *et al*, 2008).

### *Transcriptional Regulation of EMT in Trophoblasts*

As TS cells are differentiated, *in vitro*, via the withdrawal of FGF4, SNAI1 expression gradually increases. The expression of SNAI1 has also been demonstrated in the invading extravillous trophoblasts in human placenta development, signifying the importance of SNAI1 to the induction of trophoblast EMT (Zhou *et al*, 1997b; Perry *et al*, 2009). $TS^{KI4}$ cells with properties of stemness and EMT express elevated levels of the EMT-inducing transcription factors SNAI2 and TWIST1, but not SNAI1 (Abell *et al*, 2009). This expression pattern contrasts with EMT induced by differentiation where SNAI1 is induced, but not SNAI2 or TWIST1, suggesting that different sets of transcription factors may be used to induce EMT in

the context of stemness. These findings suggest that maintenance of stemness is important for regulating the trophoblast invasive potential during the induction of EMT.

***Transcriptional Regulation of EMT in Breast Cancer***

In patients with breast carcinoma, the EMT-inducing transcriptional factors SNAI1, SNAI2 and TWIST1 correlate with disease relapse and poor patient survival. The transcription factors SNAI1, SNAI2, ZEB1, ZEB2, KLF8 and E47 directly repress E-cadherin expression, resulting in the induction of EMT (Hajra *et al*, 2002; Peinado *et al*, 2007; Vandewalle, 2005; Vandewalle *et al*, 2008; Moody *et al*, 2005). Furthermore, expression of SNAI1, TWIST1 and ZEB1 induces EMT in human mammary epithelial cells (Mani *et al*, 2008). In addition to these direct transcriptional repressors, TWIST1, GSC (Goosecoid) and FOXC2 promote EMT by repressing E-cadherin through indirect mechanisms. TWIST1 expression is upregulated in metastatic breast cancer and facilitates the formation of invadopodia to promote cancer cell invasion (Yang *et al*, 2004; Eckert *et al*, 2011). GSC is overexpressed in the majority of human breast tumors, and its expression in human mammary epithelial cells enhances cell motility, a functional indicator of EMT (Hartwell *et al*, 2006). FOXC2 expression is associated with the Basal-like subtype of invasive ductal breast cancers, and expression of FOXC2 in MDCK cells induces the expression of mesenchymal markers and the serine proteases MMP2 and MMP9 with concomitant increases in cellular invasion (Mani *et al*, 2007). Collectively, multiple EMT-inducing transcription factors result in decreased E-cadherin

expression, suggesting the importance of precise control of this gene for the progression of the EMT-traits associated with breast cancer.

## Epigenetic Regulation of EMT

Recent studies have demonstrated an increasingly significant role for epigenetic mechanisms in the regulation of critical developmental programs important to the maintenance of cellular plasticity and to the establishment of cell fate commitment (Bernstein *et al*, 2006; Ng *et al*, 2008). EMT is a cellular reprogramming event that is reactivated during the metastatic progression of cancer, suggesting epigenetic control of the EMT program. The main mechanisms of epigenetic regulation include DNA methylation, histone methylation, histone acetylation and non-coding RNAs (Kouzarides, 2007; Jenuwein, 2001). These epigenetic mechanisms regulate how chromatin is organized and how genes are expressed without directly altering the DNA sequence.

### *Epigenetic Regulation of EMT in Trophoblasts*

The first cell fate decision during development generates two primary lineages of the blastocyst, the trophoectoderm that differentiates into the trophoblast subtypes of the placenta and the inner cell mass that forms the embryo (Sutherland, 2003). In contrast to pluripotent embryonic stem cells responsible for the formation of the embryo, DNA of the trophoblast epigenome is hypomethylated. This hypomethylation is critical for implantation and trophoblast differentiation. As trophoblasts differentiate into a more invasive subtype, occurring simultaneously with the progression of EMT, the trophoblast epigenome is continually demethylated.

Reduced DNA methylation promotes activation of genetic information critical to trophoblast lineage commitment (Hemberger, 2007; 2010; Rossant *et al*, 1986; Ng *et al*, 2008; Santos *et al*, 2002). Despite global hypomethylation of the trophoblast lineage, DNA methylation is indispensable for extraembryonic development. For example, the DNA methyltransferase regulatory factor DNMT3L is highly expressed in the epithelial trophoectoderm, and DNMT3L-deficient mice exhibit multiple trophoblast defects, including failure to form syncytiotrophoblasts of the placenta (Arima *et al*, 2006; Bourc'his, 2001; Hata *et al*, 2002). Furthermore, recent genome-wide sequencing studies have demonstrated that DNA hypomethylation is located primarily in the intergenic regions, which raises the question of the importance of intergenic DNA methylation for the control of genetic information.

Histone modifications comprise the second layer of trophoblast epigenetic regulation. Gene expression levels are controlled by the degree of chromatin compaction, which is mediated by wrapping DNA around an octamer of four core histones H2A, H2B, H3 and H4. The charged tails of histones are often modified by acetylation and methylation marks, which generally contribute to the respective activation or repression of genetic information. The repressive histone modification H3K27 tri-methylation (H3K27me3) conferred by the multi-subunit Polycomb complex has been extensively studied for control of gene expression during differentiation of embryonic lineages. Immunohistochemistry of the mouse blastocyst reveals reduced levels of H3K27me3 in the trophectoderm compared to the inner cell mass (Hemberger, 2007; O'Carroll *et al*, 2001; Pasini *et al*, 2004; Wang *et al*, 2002). Furthermore, genome-wide sequencing studies demonstrate that few

promoters in TS cells are marked by this modification (Rugg-Gunn *et al*, 2010). These findings contrast with post-implantation stage extraembryonic tissues, where high gene-specific levels of H3K27me3 are detected. These results suggest that an alternative repressive modification is dominant in TS cells and that the repressive H3K27me3 modification becomes important only after implantation, likely driving later lineage commitment decisions critical to the formation of the placenta. This could explain why mutants of the Polycomb complex, such as EZH2, SUZ12 and EED, demonstrate defects later during the development of extraembryonic tissues while not directly affecting TS cells (O'Carroll *et al*, 2001; Pasini *et al*, 2004). Specifically, EED mutants exhibit trophoblast differentiation defects demonstrated by failure to produce secondary invasive trophoblast giant cells, suggesting that H3K27me3 modifications might be important during activation of a specific transcriptional program (Mager *et al*, 2003; Kalantry *et al*, 2006). For example, invading trophoblast giant cells play important roles during the EMT events of placentation. Furthermore, the activating H3K4me3, the repressive H3K9me3 and the activating histone acetylation modifications demonstrate importance in TS cell maintenance, but the biological connections to EMT are unclear (Hemberger, 2007). Further studies with these histone modifications could provide insight into the importance of these marks in specific EMT programs. In chapter three, I examine the role of histone acetylation on maintenance of the epithelial phenotype of TS cells.

***Epigenetic Regulation of EMT in Breast Cancer***

Increasingly, epigenetic changes have been linked to the metastatic process. These epigenetic changes are often directed to a common gene target and provide a

selective advantage to multiple metastatic cancer cell types. Most attempts to identify the epigenetic contributions to metastasis have focused on the repressive modification DNA methylation. In primary breast tumors, hypermethylation of the CDH1 (E-cadherin) promoter was associated with lymph node metastasis, and CDH1 hypermethylation also correlated with TNBC (Bissell & Radisky, 2001; Perl *et al*, 1998; Caldeira *et al*, 2006). Furthermore, promoter hypermethylation has also been described for important tumor suppressor genes, while hypomethylation was observed in essential molecular pathways driving metastatic breast cancer, including the invasive pathways involving the proteases uPA, MMP2 and MMP9 (Rodríguez-Paredes & Esteller, 2011; Esteller, 2008; Tomaskovic-Crook *et al*, 2009).

In addition to DNA methylation, microRNAs (miRNAs) contribute to EMT and the subsequent progression of breast cancer. miRNAs are small non-coding RNAs that control gene expression through inhibition of translation or RNA decay. A variety of miRNAs demonstrate specificity for molecules important for cell-cell junction integrity, cell-cell adhesion, cellular migration and angiogenesis (Garzon *et al*, 2009). Notable EMT-regulatory miRNAs include the miR200 family, miR205 and miR34. These miRNAs, which repress ZEB1/2, TGFβ2, β-catenin, BMI1 and SNAI1, are upregulated in invasive ductal and metaplastic breast cancers (Gregory *et al*, 2008). Each of the aforementioned miRNA-regulated genes encodes critical EMT-inducing molecules. Additionally, Rho/Rac signaling molecules are inhibited by a second set of miRNAs, miR10b, miR21 and Let7i, thereby affecting EMT-specific functions, such as cancer cell motility and invasion. Lastly, miR9 and miR155 regulate cell-cell junction integrity, adhesion and angiogenesis through direct inhibition of E-cadherin

and RhoA (Garzon *et al*, 2009; Johnson *et al*, 2005; Yanaihara *et al*, 2006; Volinia *et al*, 2006; Chan, 2005). Collectively, the miRNAs have emerged as important regulatory mechanisms controlling feedback loops in the metastatic progression of breast cancer.

**Connections between EMT and Stemness Properties**

Several studies have reported the relationship between EMT and stemness properties. The transcription factor SNAI1 is well characterized for its ability to induce EMT in multiple epithelial cell systems. Interestingly in human mammary epithelial cells, overexpression of SNAI1 or TWIST1 not only induces EMT, but also results in the acquisition of stem cell properties (Mani *et al*, 2008; Morel *et al*, 2008; Floor *et al*, 2011). Furthermore, TWIST1 promotes EMT while bypassing cellular senescence, connecting EMT to an unrestricted proliferative capacity (Ansieau *et al*, 2008). More recently, mesenchymal-epithelial transition (MET), the reverse process of EMT, was shown to occur in parallel with the reprogramming of fibroblasts to induced pluripotent stem cells (iPS) (Li *et al*, 2010). Similarly, in TS cells, the coexistence of EMT and stemness properties is demonstrated by $TS^{KI4}$ cells. $TS^{KI4}$ cells have robust stem cell features similar to those of $TS^{WT}$ cells, including equivalent levels of expression of the trophoblast stem cell marker CDX2 (Abell *et al*, 2009). Thus, $TS^{KI4}$ cells, characterized as self-renewing, multipotent stem cells in a highly invasive intermediate, metastable EMT state, highlight the association between stemness properties and EMT. Obviously, such a phenotype has the potential to cause severe disease pathologies. During the pathophysiological

progression of placenta accreta, trophoblasts at the maternal-placental interface with increased invasive capacity resemble self-renewing and highly proliferative TS cells in the early stages of implantation and placenta development. Although the genetic basis for accreta is unknown, the phenotype highlights the ability of stem cells to display EMT properties (Perry *et al*, 2009). In addition to $TS^{KI4}$ cells, EMT metastability has been reported in colorectal cancer cells and progenitor cells of various organs undergoing EMT. During colorectal tumor remodeling and metastasis, immunohistochemistry of colon adenocarcinomas demonstrated the simultaneous retention of branched epithelial structures stabilized by E-cadherin and invading mesenchymal protrusions marked by high levels of β-catenin, SNAI1 and fibronectin expression (Brabletz *et al*, 2001). Furthermore, the Claudin-low subtype of breast cancer, which is characterized by its EMT features, also exhibits stem-like properties, as demonstrated by overlapping gene expression profiles with breast cancer stem cells identified by the $CD44^{+}/CD24^{-/low}$ antigenic phenotype and bipotent progenitor cells identified by the $EpCAM^{(-)}/CD49f^{+}$ profile (Al-Hajj *et al*, 2003; Hennessy *et al*, 2009; Prat *et al*, 2010; Sheridan *et al*, 2006; Reya *et al*, 2001). Collectively, these studies reveal a connection between the cellular plasticity associated with the induction of EMT and the maintenance of stem-like characteristics. In the context of EMT and stemness, the importance of identifying $TS^{KI4}$ cells is that they were the first example of a point mutation targeting the activity of a kinase (MAP3K4) that causes loss of epithelial cell maintenance, while allowing a self-renewing, multipotent stem cell to be in a permanent metastable EMT. The $TS^{KI4}$ cell phenotype clearly shows that dysregulation of kinase signaling networks

can induce a metastable EMT with properties arguably similar to those proposed for the controversial tumor initiating or cancer stem cell. These results suggest that characterization of these signaling networks might provide new therapeutic strategies to reverse these EMT-associated pathologies.

**Conclusions and Outlook**

Unlike cancer metastasis, trophoblast invasive potential is temporally and spatially restricted during normal development. We propose that studies aimed at understanding the regulatory mechanisms restricting EMT and invasiveness of primary epithelial stem cells can also provide insight into signaling networks responsible for management of metastatic cancer. Due to the intersecting features of developmental and malignant EMT subtypes, a complete picture of this complex cellular program, which is fundamental to development and reactivated in adult tissues during disease progression, can and should be studied from multiple biological perspectives. Significant therapeutic interest in the reversal of EMT is highlighted by the exploration of synthetic molecular compounds to restore E-cadherin expression (Stoops *et al*, 2011). Furthermore, the controlled induction of MET could have benefit for the reversal of several pathophysiological states where there is a dysregulated EMT. A greater understanding of the genes important for EMT may elucidate new signaling pathways to target for treatment of metastatic Claudin-low and Basal-like breast cancer, which are refractory to most current treatments.

## Objectives for this Project

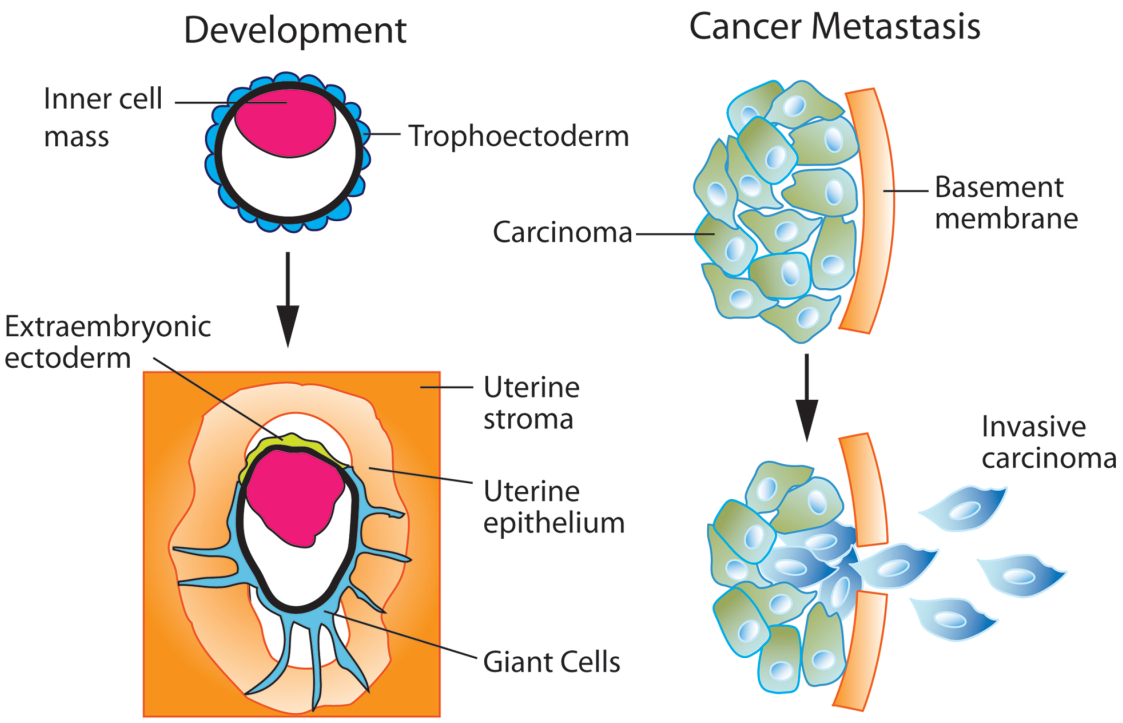
### *Define the Epigenetic Regulatory Mechanisms of EMT in Epithelial Stem Cells*

Using several epithelial trophoblast stem (TS) cells model systems, we examined the molecular and epigenetic mechanisms important for the first developmental EMT events of implantation and placenta formation. One of these TS cell models was isolated from conceptuses with a point mutation in MAP3K4, a kinase regulating JNK and p38 MAPK pathways, that renders the kinase inactive ($TS^{KI4}$ cells) and results in cells in an intermediate metastable state with both EMT and stemness properties. Due to the overlapping phenotypes between KI4 mice and mice deficient for the histone acetyltransferase CBP, it was hypothesized that trophoblast EMT programs are regulated by epigenetic mechanisms of histone acetylation. Histone acetylation is associated with active gene transcription due to its ability to decondense and open chromatin. KI4 and CBP knock-out mice display skeletal, neural tube and craniofacial defects combined with growth retardation and embryonic lethality. Studies with $TS^{KI4}$ cells, in addition to two separate models of TS cell EMT ($TS^{SNAI1}$ and $TS^{shCBP}$ cells), allowed the discovery of an epigenetic mechanism reliant on H2B acetylation important for the maintenance of the epithelial phenotype of TS cells. Due to incomplete modulation of the epigenetic landscape, these TS cell models are uniquely paused in an intermediate EMT state and provide tractable systems to examine the molecular and epigenetic mechanisms contributing to the transient, intermediate EMT states. Furthermore, an intersecting EMT gene signature was found in mesenchymal Claudin-low breast cancer cells characterized

by EMT and stemness properties, highlighting the importance of studying EMT from multiple biological perspectives.

***Define Novel Genes Important for the Induction of EMT in Triple Negative Breast Cancer***

In order to understand the similar connections between developmental and breast cancer EMT, we identified an intersecting EMT gene signature from analysis of microarray gene expression data shared by $TS^{KI4}$ cells developmentally entering EMT and the Claudin-low subtype of TNBC. Similar to human Claudin-low breast cancer cell lines and tumors, $TS^{KI4}$ cells exhibit an increase in the mesenchymal markers VIM, CDH2, SNAI2 and TWIST1 with loss of the epithelial differentiation and cell adhesion markers CD24, KRT7/8/19 and CLDN4. With a few notable exceptions such as SNAI2, TWIST1, VIM and CDH2 (N-cadherin), many of the genes in this EMT signature are uncharacterized with respect to EMT. It was hypothesized that the upregulated genes within this EMT gene signature define a novel set of genes with importance in EMT and maintenance of the mesenchymal phenotype. Using a phenotypic RNAi screening approach, we identified ten genes important for establishing the Claudin-low mesenchymal phenotype of an $EpCAM^{(-)}$ population of SUM149 and SUM229 Basal-like breast cancer cells. Of particular importance, the SWI/SNF chromatin-remodeling factor Smarcd3/Baf60c was identified and characterized for its ability to induce EMT in a Wnt5a-dependent manner in human mammary epithelial cells, highlighting a new mechanism of epigenetic regulation of EMT.




**Similarities between Development and Cancer Metastasis EMT Subtypes**

- Loss of epithelial markers and adherens junctions, especially E-CADHERIN
- Expression of mesenchymal markers (VIM, N-CAD)
- Gain of EMT-inducing transcription factors including SNAI1, SNAI2, TWIST1, ZEB1/2, β-CATENIN
- Loss of apical-basal polarity with acquisition of a mesenchymal morphology
- Increased cellular invasiveness
- Activation of TGFβR and HGFR cellular pathways

**Differences between EMT Subtypes**

| Development | Cancer Metastasis |
|---|---|
| Proceeds to designated endpoint controlled by cellular differentiation | Uncontrolled systemic dissemination |
| Defined reversal of EMT by MET | MET is experimentally unconfirmed |
| Stemness is associated with retention of the epithelial phenotype | Stem-like properties emerge during the progression of EMT |

**Figure 1.1: Comparison of EMT Subtypes in Development and Cancer Metastasis**

The first developmental EMT that occurs during implantation of the blastocyst into the uterine epithelium (left panel) and EMT associated with the metastatic progression of cancer (right panel) are shown. MET indicates mesenchymal-epithelial transition.

# Chapter II

# Materials and Methods

## Chapter 3 Methods

### *Cell Lines, Culture Conditions and Transfections*

$TS^{WT}$ and $TS^{KI4}$ cells of normal karyotype were isolated from heterozygote crosses of mice with a targeted mutation of MAP3K4 (K1361R) as previously described (Abell *et al*, 2009). TS cells were cultured without feeders in 30% TS media (RPMI 1640, 20% fetal bovine serum, 1% penicillin and streptomycin (PS), 1% L-glutamine, 1% sodium pyruvate and 100 μM β-mercaptoethanol) and 70% MEF conditioned TS cell media, supplemented with FGF4 (37.5 ng/ml) and Heparin (1 mg/ml). For differentiation, TS cells were cultured in TS media only. 293T cells were cultured in Dulbecco's modified Eagle's high glucose medium with 10% FBS and 1% PS. Transfection of 293T cells was performed in 10 cm dishes for 24 hours with LipofectaminePlus (Invitrogen) according to manufacturer's specifications. The Snail expression plasmid was a kind gift from Dr. Paul Wade. Snail was subcloned between EcoRI and XhoI restriction sites of pcDNA3.1(+). G418 was used to stably select for TS cells overexpressing Snail. Published plasmids include HA-MAP3K4, HA-MAP3K4 kinase inactive, HA-JNK, FLAG-p38, HA-CBP, HAT-inactive HA-CBP, Myc-p300 and HAT-inactive Myc-p300 (Gerwins, 1997; Abell, 2004; Bradney, 2003).

***Lentivirus Production, Infections of TS cells and shRNA Knockdown of CBP and p300***

To produce replication incompetent lentivirus, 293T cells were co-transfected with either GFP or shRNA constructs in pLKO.1 (Open Biosystems) in combination with pMD2.G and psPAX2 (Addgene) using Lipofectomine Plus reagent (Invitrogen). 24 hours later, cell media was changed. Viral supernatants were harvested at 48 and 72 hours post-transfection by ultracentrifugation. Viral pellets were resuspended in 100 μl of 30% TS media (RPMI 1640, 20% fetal bovine serum, 1% PS, 1% L-glutamine, 1% sodium pyruvate and 100 μM β-mercaptoethanol) and 70% MEF conditioned TS cell media, supplemented with FGF4 (37.5 ng/ml) and Heparin (1 mg/ml) and stored at -80° C. Infection of TS cells was performed essentially as previously described (Odiatis & Georgiades, 2010). Briefly, $TS^{WT}$ cells were plated overnight at 1 X $10^4$ cells per 24 mm well in a Nunc 4-well dish, and then infected with 100 μl of lentivirus and polybrene (5 μg/ml). The next day, cells were washed five times in PBS and TS media was replaced. Three days after infection, cells were examined for GFP positive cells as a indicator of transduction efficiency, showing a 40-50% infection efficiency. Puromycin (2 μg/ml) was used to select for transduced cells. Selection was complete after seven days of puromycin treatment. Knockdown of CBP or p300 was measured by qRT-PCR as described in **Real-time Quantitative RT-PCR.**

***Invasion Assays and Isolation of Invasive Trophoblasts ($T^{INV}$)***

Invasion assays were performed essentially as previously described (Abell *et al*, 2009). Briefly, cells were plated on growth factor-reduced Matrigel (BD Biosciences)

coated 8 μM pore transwell chambers. After 48 hours, invasion assays were terminated. Non-invading cells were removed from the top of the transwells by washing and swabbing. Invaded cells ($T^{INV}$) were isolated from the bottom of the transwell chambers by either direct lysis in RNAeasy buffer for RNA isolation or by trypsinization for immunofluorescence, ChIP, and Western blotting. For quantitation of invasion, chambers were fixed in 3% paraformaldehyde for ten minutes and stained with DAPI nuclear stain. For each transwell, five 10X fields were imaged and counted.

***Real-Time Quantitative RT-PCR (qRT-PCR)***

Total RNA was isolated using RNeasy Plus minikit (Qiagen). cDNA was synthesized from 3 μg RNA using High-Capacity reverse transcription kit (Applied Biosystems). Applied Biosystems 7500 RT-PCR system with inventoried TaqMan probes was used to quantify gene expression levels normalized to β-actin expression. PCR for TS cell marker Cdx2 was carried out at 61$^{o}$C for 35 cycles to generate a 284 bp product. The following primers were used for semi-quantitative RT-PCR: sense 5'-CATCCCCGCCTCTACAGCTTA-3' and 5'-ATACATGCTCACGTCCTTGTCC-3'.

ChIP assays were quantified by real-time PCR using Absolute Blue SYBR green PCR mix (Thermoscientific) and the Applied Biosystems Fast 7500 Real-Time PCR System. Fold enrichment was determined by the $2^{-\Delta CT}$ method. PCR primers were designed to amplify approximately 75-100 bp fragments from genomic DNA using Primer Express 3.0 (Applied Biosystems). Primers used for ChIP-qRT-PCR are described in **Table 2.1**.

***Agilent Gene Expression Arrays***

For gene arrays, cDNA was competitively hybridized to a 4 x 24,000 Agilent gene array. Raw gene expression data was analyzed using Agilent GeneSpringGX software to produce a ranked gene list based upon the $\log_2$ ratio of Agilent probe intensity appearing in at least three gene arrays with a Benjamini-Hochberg p-value < 0.05. Heat maps were prepared by hierarchical clustering in Cluster 3.0 and JavaTreeView. Published breast cancer cell line data from Neve et al. was clustered according to the methods described above (Neve *et al*, 2006).

***Comparative Expression and GO Pathway Analysis***

Comparative gene expression analysis was used to determine the $T^{INV}$ gene signature by filtering for genes enriched in the $T^{INV}$ versus $T^{DIFF}$ statistically analyzed gene sets (Benjamini-Hochberg p-value < 0.05 and a $\log_2$ ratio $\geq$ +/- 1). Enrichment was quantified by the absolute value of the $\log_2$ ratios between $T^{INV}$ and $T^{DIFF}$ gene sets using a secondary cutoff $\log_2$ ratio $\geq$ 0.5. The $T^{INV}$ gene signature was assessed for overlap with the $TS^{KI4}$ statistically analyzed gene set to produce the minimal trophoblast EMT signature. The NIH DAVID gene ontology bioinformatics database was used to cluster genes into canonical KEGG signaling pathways from both the ChIP-seq and Agilent array datasets of statistically significant genes with a p-value < 0.05 and a $\log_2$ ratio $\leq$ -1.

***Microarray Analyses of Human Breast Tumors and Cell Lines***

Two publicly available gene expression microarray datasets (UNC337 [GSE18229] for tumors and Neve et al. for cell lines) were analyzed as previously described (Neve *et al*, 2006; Prat *et al*, 2010). For human tumors, we obtained the Claudin-low

human tumor expression signature of up- and down-regulated genes from Supplemental Data.xls in Prat et al. (http://breast-cancer-research.com/content/supplementary/bcr2635-s2.xls). For human cell lines, we performed a Significance Analysis of Microarrays (SAM) by comparing the previously described CL cell lines versus the rest (false discovery rate <5%). The resulting up- and down-regulated gene sets defined the CL cell line signature analyzed. Finally, overlap between two signatures was estimated using the exact hypergeometric probability implemented in R package (http://cran.r-project.org). In each overlap analysis, only genes present in both platforms were considered. The accession number for the microarray data is GSE27883.

***Generation and Analysis of Chimeras***

GFP expressing $TS^{WT}$ and $TS^{KI4}$ cells generated as described above in ***Lentivirus Production and Infections of TS Cells*** were used to generate chimeric embryos. The Animal Models Core at the University of North Carolina at Chapel Hill performed blastocyst injections using standard procedures. Conceptuses were dissected at E6.5 and E9.5, where date of blastocyst injection was considered E3.5. Samples were fixed for one hour in PBS containing 4% paraformaldehyde and 1% sucrose, washed, and put through a sucrose gradient. After embedding and freezing in OCT, samples were cryosectioned at 18 μm sections. All animal and embryo work was performed according to both university and federal guidelines for the use of animals.

***Immunofluorescence***

Immunostaining was performed as previously described (Abell *et al*, 2009). Briefly, TS cells cultured on glass coverslips for two days were fixed in 3%

paraformaldehyde, 1% sucrose, 1% PBS for 10 min. Coverslips were washed with PBS, permeabilized for 5 min with 0.2% Triton and blocked for one hour with 10% goat serum. Coverslips were incubated for one hour at room temperature with anti-E-cadherin antibody (1:1000), washed for 30 min and then incubated at room temperature for one hour with Alexa 555 anti-mouse secondary antibody (Invitrogen) and DAPI nuclear stain.

***Western Blotting of Whole-cell, Nuclear and Histone Lysates***

Whole cell and nuclear lysates were isolated as previously described (Abell *et al*, 2009). For histone lysates, cells were lysed on ice in buffer containing PBS, 1% Triton-X and 1 mM PMSF. Pellets were spun at 2000 rpm for 10 min at $4^{o}$ C and extracted overnight in 0.2 N HCl with shaking at $4^{o}$ C. Western blots were performed with the antibodies specified in **Table 2.2**.

***In vitro Histone Acetyltransferase (HAT), Immunoprecipitation, and Kinase Assays***

HAT assays were performed using Millipore (cat. # 17-289) HAT assay kit. Briefly, HA-CBP and/or Myc-p300 were immunoprecipitated in PBS with 5 μg anti-HA or anti-myc antibodies specified below in **Table 2.2**. Following immunoprecipitation, CBP/p300-bound protein G sepharose beads were incubated for 30 min at 30° C in a reaction mixture containing 100 μM acetyl-CoA, 1.4 μCi [$^{3}$H]-acetyl-CoA (specific activity 200 mCi/mmol), 10 μg purified core histones and HAT buffer provided by the manufacturer (Millipore). The reaction mixture was spotted onto P18 filter paper, washed with sodium phosphate buffer (pH 9) and quantitated with a Beckman liquid scintillation counter. Total HAT activity was measured from nuclear lysates

according to the same method. Kinase assays were performed as previously described (Abell et al., 2007). Briefly, HA-CBP or Myc-p300 immunoprecipitates were incubated for 20 min at 30° C with HA-MAP3K4 and HA-JNK or FLAG-p38 in kinase buffer containing 0.5 mM ATP and 10 μCi γ- $^{32}$P[ATP]. Proteins were separated by SDS-PAGE and quantitated by autoradiography using a Phosphorimager.

***Chromatin Immunoprecipitation Coupled to High Throughput Sequencing (ChIP-seq)***

Cells were fixed for 10 min in 1% formaldehyde, sonicated (VCX130 Ultrasonicator) and immunoprecipitated with 5 μg anti-H2BK5Ac and Protein A dynabeads (Invitrogen) (Wang *et al*, 2008). Crosslinking was reversed by overnight incubation at 65$^{o}$C. DNA was purified with the MinElute PCR purification kit (Qiagen). Library preparation for Illumina ChIP-seq was performed according to manufacturer's instructions (Illumina). Illumina Solexa GA II was used to produce ~36 bp sequence reads, which were mapped to the mouse genome using Mapping and Alignment with Quality (MAQ) software in conjunction with EpiCenter for comparative analysis as described under ***Epicenter ChIP-seq Analysis Tools***. PCR conditions and primers used for ChIP-seq validation are described in **Table 2.2**.

***Epicenter ChIP-seq Analysis Tools***

For both ChIP-seq and mRNA-seq data analysis, we used MAQ (Li et al., 2008) to map read tags to the mouse reference genome (NCBI build 37.1/mm9), and used EpiCenter to identify significant changes between TS$^{WT}$ and TS$^{KI4}$ cells (Huang *et al*, 2011). EpiCenter is a versatile tool for comparative analysis of both ChIP-seq and

mRNA-seq data to identify differential changes. EpiCenter is freely available at http://www.niehs.nih.gov/research/resources/software/epicenter.

**Read mapping and quality filtering:** For the analysis of H2BK5Ac ChIP-seq data, read tags were mapped directly to the mouse reference genome. Reads with MAQ alignment quality scores <10 (i.e. number of mismatches > 2) were filtered out to remove poorly aligned reads or non-uniquely mapped reads. After the quality filter, about 38.7 (87.8%) and 24.1 (88.9%) million reads were uniquely mapped to the mouse reference genome in $TS^{WT}$ and $TS^{KI4}$ cells, respectively. For the analysis of gene expression mRNA-seq data, read tags were directly mapped to cDNA sequences of all mouse RefSeq genes. After filtering out reads with alignment quality scores <10, about 41.6 and 33.6 million reads were uniquely mapped to the RefSeq cDNA sequences. We also mapped the reads to the mouse reference genome, but we had about 20% fewer mapable reads because reads from exon junctions could not be mapped back to the reference genome. The results reported here were based on cDNA sequence mapping data.

**Data normalization:** Since H2BK5Ac ChIP-seq data showed broad change between $TS^{WT}$ and $TS^{KI4}$ samples, we chose the ratio of the numbers of mapped reads to adjust the difference of read coverage depths. We also used EpiCenter's linear least squares regression method to estimate the read coverage ratio for normalization. The 300 most significant genes were very similar between the two approaches, although fewer genes were identified with the regression-based approach. For mRNA-seq data analysis, we used EpiCenter's "parsimony" normalization procedure to adjust read coverage difference. The "parsimony"

automatically adjusted normalization ratio to minimize the number of significant genes by the EpiCenter's exact rate ratio test.

**Statistical tests of significance:** EpiCenter first tested whether the read tag count in each region or gene was significantly larger than that expected under cell-type specific background noise. Regions/genes whose read-tag counts did not significantly exceed the background noise in both $TS^{WT}$ and $TS^{KI4}$ samples were filtered out. For those regions/genes retained, we then used EpiCenter exact rate ratio test to find whether the difference in read counts was significantly larger than that expected from Poisson variation. In mRNA-seq data analysis, we also used EpiCenter's Z-test to determine if the difference in read counts exceeded normal variation. We used Benjamini-Hochberg correction to adjust p-value for multiple testing in analysis of both ChIP-seq and mRNA-seq data. The technical details of EpiCenter were published in Huang et al. (Huang *et al*, 2011).

## Chapter 4 Methods

### *Cell lines, Culture Conditions, Constructs and Transfections*

The primary human mammary epithelial cell line (HMEC), immortalized using the retrovirus pBabe-hygro-hTERT, was cultured as previously described (Troester, 2004). Basal-like breast cancer lines SUM149 and SUM229 cells were cultured in HuMEC medium, 5% FBS, 1% penicillin and streptomycin (PS) plus bovine pituitary extract and HuMEC supplement or Ham's F12 medium, 5% FBS, 1% PS plus 5 μg/mL insulin and 1 μg/mL hydrocortisone, respectively. 293T cells were cultured in Dulbecco's modified Eagle's high glucose medium with 10% FBS and 1% PS.

Transfection of 293T cells was performed in 15 cm dishes for 24 hours with Lipofectamine Plus (Invitrogen) according to manufacturer's specifications. The Smarcd3/Snail/Slug Gateway entry plasmid was subcloned into a lentiviral FLAG-Gateway destination vector. The lentiviral FLAG-Gateway destination vector was a kind gift from Dr. Ben Major. The EpCAM reporter was subcloned between XbaI(2708) and NotI(4101) restriction sites and replaced the CMV-IE-Promoter-Enhancer/Turbo-GFP tag in the pGIPZ empty vector. The EpCAM promoter expression plasmid was a kind gift from Dr. Marianne Rots.

***Lentivirus Production and Infections of HMEC Cells***

To produce replication incompetent lentivirus, 293T cells were co-transfected with either pGIPZ empty vector, FLAG-Smarcd3/Snail/Slug lentiviral Gateway constructs or the lentiviral EpCAM reporter construct in combination with pMD2.G and psPAX2 (Addgene) using Lipofectomine Plus reagent (Invitrogen). 24 hours later, cell media was changed. Viral supernatants were harvested at 48 hours post-transfection by ultracentrifugation, and viral pellets were resuspended in 300 μL base media. HMECs were infected overnight with 100 μL lentivirus in 6 μg/mL polybrene. Puromycin (3 μg/mL) was used to select transduced cells. Selection was completed 7 days after infection. Expression of Smarcd3/Snail/Slug was measured by qRT-PCR.

***Fluorescently Activated Cell Sorting (FACS)***

Cells were trypsinized and filtered into single cell suspensions, resuspended in Hank's balanced salt solution containing 2% FBS (HF media), and incubated with antibodies for 30 min at 4$^{o}$C. Cells were stained with the following antibodies

EpCAM-FITC (Stem Cell Technologies 10109), EpCAM-APC (BD 347200), Cd49f-PE-Cy5 (BD 551129), Cd44-APC (BD 559942) or Cd24-FITC (BD 555427). Unbound antibodies were washed from cells using HF media. For analytical flow, cells were fixed with 3% paraformaldehyde and analyzed using Beckman-Coulter Cyan Instrument. For sterile live-cell flow cytometry, cells were sorted using a Sony iCyt/Reflection Instrument.

***RNAi Screening Conditions, Immunofluorescence and High Content Imaging***

For the RNAi screen, $EpCAM^{(-)}$ SUM149 and SUM229 cells were reverse transfected in fibronectin-coated 96-well plates (Greiner) with 25 nM siRNA smart pools (Dharmacon) containing the combination of 4 different siRNA oligonucleotides for each target or control gene (siUBB for siRNA knockdown efficiency; siGAPDH for a negative control; siSnail/Slug for a positive control). Target genes were plated in duplicate on triplicate plates, and control genes were plated 2X for UBB, 6X for GAPDH and 8X for siSnail/Slug on triplicate plates. After 6-days in culture, cells were fixed with 3% paraformaldehyde, permeabilized with 0.1% Triton X and stained with nuclear DAPI stain, EpCAM (Stem Cell Technologes 01420) and E-cadherin (Cell Signaling 3195S) antibodies. EpCAM (conjugated to secondary Alexa 555) and E-cadherin (conjugated to secondary Alexa 488) fluorescence was measured using the high-content BD Pathway 855 microscope system. Changes in EpCAM and E-cadherin expression were quantitatively measured on a single-cell basis as a function of the percent change in the cellular fluorescence using the CellProfiler image analysis software. Results were the compilation of two independent screens performed in triplicate.

***Chromatin Immunoprecipitation Assays (ChIP)***

Cells were fixed for 10 min at 37$^{o}$C in 1% formaldehyde, sonicated (VCX130 Ultrasonicator) and immunoprecipitated with 5 μg anti-H3K4me3 (Abcam 8580), anti-Smarcd3 or anti-FLAG (Sigma F7425) and Protein A dynabeads (Invitrogen). Crosslinking was reversed by overnight incubation at 65$^{o}$C. DNA was purified with the MinElute PCR purification kit (Qiagen). ChIP assays were quantified by qRT-PCR using Absolute Blue SYBR green PCR mix (Thermoscientific) and the Applied Biosystems Fast 7500 Real-Time PCR System. Fold enrichment was determined by the $2^{-\Delta CT}$ method. PCR primers were designed to amplify approximately 75-100 bp fragments from genomic DNA using Primer Express 3.0 (Applied Biosystems). ChIP-qRT-PCR primers include: Cdh1 sense 5'-ACCCCCTCTCAGTGGCGT-3', and 5'-GGAGCGGGCTGGAGTCTG; Cd44 sense 5'-AGTGGATGGACAGGAGGATG-3', and 5'-TTATGTCCTTCTGGGCTCT; Cldn4 sense 5'-TCAGCCTTCCA GGTCCTCAA-3', and 5'-CCCCATGGAGGCCATTG-3'; Wnt5a sense 5'-CCTATTTTGCTCCCCGTT-3', and 5'-AAGAGTCAGCCCCAAATT-3'.

***Matrigel Invasion Assays***

Cells were plated on growth factor-reduced Matrigel (BD Biosciences) coated 8 μM pore transwell chambers. After 24 hours, invasion assays were terminated. Non-invading cells were removed from the top of the transwells by washing and swabbing. Invasive cells were quantified by fixing chambers in 3% paraformaldehyde for ten minutes and staining with DAPI. For each transwell, three 10X fields were imaged and counted.

***Western Blotting of Whole-cell and Nuclear Lysates***

Whole cell and nuclear lysates were isolated as previously described (Abell *et al*, 2009). Western blots were performed with the following specified antibodies: EpCAM (Abcam 71916), E-cadherin (BD 610181), Fibronectin (Calbiochem CP70), N-cadherin (Abcam 76057), Vimentin (Cell Signaling 5741S), γ-tubulin (Sigma T6557), FLAG (Sigma M5 F4042/F7425), Snail (Cell Signaling 3895S), Slug (Cell Signaling 9585S), Brg1 (Santa Cruz sc-17796X), Brm (Abcam 15597), Phospho-PKCβ (Cell Signaling 9371S), Phospho-JNK (Cell Signaling 9251S), β-catenin (Sigma C7207) and Smarcd3/Baf60c. Smarcd3/Baf60c antibody was a kind gift from Dr. Lorenzo Puri.

***Real-time Quantitative RT-PCR (qRT-PCR)***

Total RNA was isolated using RNeasy Plus minikit (Qiagen). cDNA was synthesized from 3 μg RNA using High-Capacity reverse transcription kit (Applied Biosystems). Applied Biosystems 7500 RT-PCR system with inventoried TaqMan probes was used to quantify gene expression levels normalized to β-actin expression.

***Gene Expression Microarrays***

Transfected and control HMEC cell lines were profiled as described previously using 44K human oligo microarrays (Agilent Technologies, Santa Clara, CA, USA) (Hu et al, 2006; Prat et al, 2010). The probes or genes for all analyses were filtered by requiring the lowest normalized intensity values in both sample and control to be > 10. The normalized $\log_2$ ratios (Cy5 sample/Cy3 control) of probes mapping to the same gene (Entrez ID as defined by the manufacturer) were averaged to generate independent expression estimates. At least three biological replicates were used for

each transfected gene (i.e. each Cy5-sample). For Cy3-controls, we used the empty vector cell line, which was the same for all samples. All microarray data is available in the University of North Carolina (UNC) Microarray Database (https://genome.unc.edu/) and has been deposited in the Gene Expression Omnibus under the accession number GSE40145.

***Gene Signatures***

The previously published Claudin-low predictor and the Differentiation Score were evaluated as previously described (Prat et al, 2010). For each HMEC infected cell line (i.e. Smarcd3, Slug and Snail), we derived a gene signature by performing a one-class Significance Analysis of Microarrays (SAM) with a false discovery rate of 0%. An enrichment score for each individual signature (up- and down-regulated genes combined) was evaluated in the previously published UNC337 dataset (GSE18229) of breast cancer samples representing all the intrinsic molecular subtypes. To do so, we calculated the inner product of each signature (gene ratio) and the gene expression value of each breast tumor sample. Subtype calls in the UNC337 dataset were used as provided in GEO (Prat et al, 2010).

***Statistical Analysis***

Gene overlap between two signatures was estimated using exact hypergeometric probabilities implemented in R package (http://cran.r-project.org). All microarray cluster analyses were displayed using Java Treeview version 1.1.4r2. Average-linkage hierarchical clustering was performed using Cluster v3.0 (Eisen et al, 1998).

| Gene Name | Forward Primer | Reverse Primer |
|---|---|---|
| *Mouse* | | |
| Acsl6 | ACCCCCGCTGCCATTT | GACCGGCCTCCAAGGTTT |
| Itgav | AAAGTCCCGCCGAGTATGC | AGAAGTCCACGGCGAATCC |
| Lamb2 | TTCCCTTGCTTTCCAACTTTACTC | GCAGTCACTCCTCCACAATCTG |
| Trim54 | TAGCCGGGATCCCAAGAAT | TGCCTGTGCACCCTTCCT |
| Cdx2 | TGGCTCCGCAGAACTTTGTC | CCGCCACGTGGTAACCAA |
| Eomes | GCCTTCCACCTTTGATGTATCC | AAAGCTTTGGCGCCTTCTCT |
| Aim1 | CCAGCTTTGCACTTGTTGCA | GCACGCTTGGCTCTCAGTTT |
| Arfgef1 | GGCAAGAAGACGAAGAAGAACATGTT | ACTTCCTTGTCGGCCAGAATC |
| Galnt3 | CCTTAAGCGACTAGTAAAATTGCACAT | TGACTGCACCCAGCTTCCA |
| Irx5 | AACCAGACGGTGTTGAATCGT | GGCTCCGCAACATTTTCG |
| Krt7 | CCGCATGGAGGAATAAAAGG | GAGGAGCAGCAGCGCGAACTC |
| Krt19 | AGTCCGCGGTGGAAGTTTTA | CCAGACAGCAGCCCATCAG |
| Rab25 | TCTCCACCCGCACTGTAATG | GTCCCAGATCTGTGCCTTGAC |
| Scyl3 | AAGGCTTCGGTGCCAGAGT | TGGAAACTGGCCTCAGTCAAG |
| Cobll1 | TCAGACACCTGACACTGAACTTTTT | AAAAGGTTCCTATGTGTCAGGATGA |
| Ctnnd1 | CGCCGCTCCTCTTCCA | CTACGACGAGCTGGGCAATT |
| Dbndd2 | CCCTGGAGCGCCAACA | GGCTGTAAAATGTCCTCAAAGAACT |
| Def6 | GCACTTCCGGGATGACGAT | TGAGGTAGGGCATGTAACCTTGA |
| F11r | AGGACCGATTGGTTGGAAAAG | GCACCCTGAGGAAAACTGTCA |
| Ipp | GACAAACATGCCCAGCTCATC | TCACAGAAATGCTGTCCACTTCTC |
| Isg20 | CGCTGCAGCATTGTGAACA | CGGGTCGGATGTACTTGTCA |
| Plxnb1 | TGCAGAGGAAGGCCGTGTAC | ATACTCGCCCATCCCATGAG |
| Pou3f1 | GCCCACCGCAGGATCTC | TTCTGGAGGTCCCTGCTGTAG |
| Wwc1 | GACGGCAAGGTCTACTACATAGATCA | CCGCGGGTCGATCCA |
| Yme1l1 | GTGTTACCCTGCGTCAGAAATG | CTATAACAGCATGTTCATATTCACAAAATC |
| *Human* | | |
| Aim1 | AACGCGGCCAGCGATA | CCGCGGAGATTTCACTTTCT |
| Arfgef1 | GGCCGACAAGGAAGTGAAGA | CCACCTCGCAAGCTTTGC |
| Galnt3 | GCTTGGTGCAGTAATTTTTTTCTTT | TTTCCATCCTTGATTCCTCTTTG |
| Irx5 | CCATTGCCTTGGTCAAAAATG | GAGGCGGACGGCTGGTA |
| Krt7 | CAGATCAAGACCCTCAACAACAAG | GTCCAGTCCCGCTCACCTT |
| Krt19 | AAGCTAACCATGCAGAACCTCAA | CAGGGCGCGCACCTT |

| | | |
|---|---|---|
| Rab25 | TCTCCACCCGCACTGTGAT | TGTCCCAGATCTGAGCCTTGA |
| Scyl3 | ACACTGAGAGAACCACCATTTACCT | TCTTGCAGTACAGCGGGATAAAC |
| Cobll1 | CCAAAGCTCTGACCCAGAACA | TCTCCCGAACGGATTGCA |
| Ctnnd1 | CGACCGCCAGCATCTTG | CCCGGGTCAGCTTCTCAAA |
| Dbndd2 | GGCAAAAATTCTTCGAGGACAT | GATGCAGATGGGACAGAGGAA |
| Def6 | GCCCTGCGCAAGGAACT | ACGTCCAGCGCGGTAAAG |
| F11r | CCTGGCATTGGGCAGTGT | TTCTCAGGAATTCTGACTTCAGGTT |
| Ipp | AGGCTGCTGATAGTCCTTTTTCA | CATCTTATTGATTTGGGCCAAGA |
| Isg20 | TCGTTGCAGCCTCGTGAA | GGCCGGATGAACTTGTCGTA |
| Plxnb1 | CCCCTTCCACCAACTGCAT | TTGCCAGGTGCTGCAGATAC |
| Pou3f1 | CCGAATTTCTGACCCATCTCTATT | ACAAACAGAAGAAGAAACGGAGAAG |
| Wwc1 | TCGACGGCAAGGTCTACTACATAG | CCGCGGGTCGATCCA |
| Yme1l1 | CGGTGCAACCCCAGGTTA | TGGTGTATGGAAGGCATTGATG |

**Table 2.1 ChIP-PCR Primers**
Table lists each primer used for ChIP-qRT-PCR to confirm the results of the H2BK5Ac ChIP-seq experiments that measure H2B promoter acetylation levels described in chapter three.

| Antibody specificity | Company | Catalogue number |
|---|---|---|
| vimentin (rabbit mAb) | Cell Signaling | 5741 |
| vinculin (mouse mAb) | Sigma | V-9131 |
| E-cadherin (mouse mAb) | BD Transduction | 610181 |
| actin (mouse mAb) | Sigma | A-3853 |
| fibronectin (rabbit Ab) | EMD Calbiochem | CP70-100UG |
| N-cadherin (rabbit Ab) | Abcam | AB76057 |
| gamma tubulin | Sigma | T6557 |
| p300 (mouse mAb) | Millipore | 05-257 |
| CBP (rabbit pAb) | Abcam | ab2832 |
| H2AK5Ac (rabbit pAb) | Cell Signaling | 2576 |
| H2BK5Ac (rabbit pAb) | Cell Signaling | 2574 |
| H2BK5Ac (rabbit pAb, ChIP) | Abcam | ab40886 |
| H2BK12Ac (rabbit pAb) | Epitomics | 1755-1 |
| H2BK15Ac (rabbit pAb) | Epitomics | 2170-1 |
| H2BK20Ac (rabbit pAb) | Cell Signaling | 2571S |
| H3K9Ac (rabbit pAb) | Cell Signaling | 9671 |
| H4K8Ac (rabbit pAb) | Cell Signaling | 2594 |
| Total H2A (rabbit pAb) | Cell Signaling | 2572 |
| Total H3 (rabbit pAb) | Cell Signaling | 9715 |
| H3K9me3 (rabbit pAb) | Abcam | ab8898 |
| H3K4me3 (rabbit pAb) | Abcam | ab8580 |
| Anti-HA (rabbit pAb) | Santa Cruz | 12CA5 |
| Anti-Myc (mouse mAb) | Santa Cruz | sc-40 |

**Table 2.2 Antibodies and Sources**
Table describes the type and source of each antibody used in the Western blotting, ChIP and immunoprecipitation experiments performed in chapter three

# Chapter III

## MAP3K4/CBP Regulated H2B Acetylation Controls Epithelial-Mesenchymal Transition in Trophoblast Stem Cells[1]

### Introduction

The transition of epithelial cells to motile mesenchymal cells occurs through a process known as epithelial-mesenchymal transition (EMT), in which epithelial cells lose cell-cell contacts and apical-basal polarity concomitantly with the acquisition of a mesenchymal morphology and invasive properties. Several molecular events are coordinated for the initiation and completion of EMT including loss of the adhesive cell-surface protein E-cadherin, activation of EMT-inducing transcription factors and reorganization of the actin cytoskeleton (Yang & Weinberg, 2008). During development, EMT is responsible for proper formation of the body plan and for differentiation of many tissues and organs. Examples of EMT in mammalian development include implantation, gastrulation and neural crest formation (Thiery *et al*, 2009). Initiation of placenta formation regulated by trophoectoderm differentiation is the first, and yet most poorly defined developmental EMT.

The commitment of stem cells to specialized cell types requires extensive reprogramming of gene expression, in part, involving epigenetic control of

[1] The data presented in this chapter was performed in collaboration with Dr. Amy Abell and was published as a co-first author manuscript in *Cell Stem Cell* (Abell *et al*, 2011).

transcription. The first cell fate decision is the formation of the trophoectoderm and the inner cell mass of the blastocyst. Trophoblast stem (TS) cells within the trophoectoderm maintain a self-renewing state in the presence of FGF4 (Tanaka, 1998). For TS cells and most other tissue stem cells, self-renewal is defined as cell division with the maintenance of multipotency . Diminished exposure to FGF4 induces TS cells to give rise to multiple differentiated trophoblast lineages, each required for establishment of the placenta. For implantation to occur, TS cells must undergo morphological changes to a more invasive phenotype, thereby exhibiting the functional hallmarks of EMT.

An emerging topic in the EMT field is the intersection between EMT and stemness. We have previously characterized the developmental defects of a genetically engineered mouse with the targeted inactivation of MAP3K4, a serine-threonine kinase important for JNK and p38 activation in response to FGF4 (Abell *et al*, 2009). In addition to embryonic defects, the MAP3K4 kinase-inactive mouse (KI4) has trophoblast defects resulting in hyperinvasion, defective decidualization, fetal growth restriction and implantation defects (Abell *et al*, 2005; 2009). TS cells isolated from the conceptuses of KI4 mice ($TS^{KI4}$ cells) exhibit the hallmarks of EMT, while maintaining TS cell multipotency and are a unique developmental stem cell model to examine parallels between EMT and stemness.

Recently, EMT has been linked to the metastatic progression of cancer and to the acquisition of stem cell properties (Mani *et al*, 2008). The Claudin-low (CL) intrinsic molecular subtype of breast cancer is characterized by its mesenchymal and stem cell-like properties. In concordance with the stem cell-like $CD44^{+}/CD24^{-/low}$

and CD49f$^{+}$/EpCAM$^{-}$ antigenic phenotypes of breast tumor initiating cells (TICs) and mammary stem cells, gene expression profiling demonstrated that CL tumors have reduced expression of several epithelial differentiation markers, while exhibiting increased expression of certain stemness and mesenchymal markers (Lim *et al*, 2009; Prat *et al*, 2010).

Herein, we define an epigenetic mechanism important for the initiation of the first EMT event during development. Using TS$^{KI4}$ cells uniquely trapped in EMT prior to initiation of the trophoblast differentiation program, we capture the genetic and epigenetic profile of the intersection between the properties of EMT and stemness. Importantly, we identify a molecular mechanism reliant on the histone acetyltransferase CBP that is responsible for controlling epigenetic remodeling during EMT in TS$^{KI4}$ cells. TS cells lacking CBP (TS$^{shCBP}$) expression exhibit an EMT similar to TS$^{KI4}$ cells, which is mediated by the selective loss of H2A and H2B acetylation. Comparison across developmental and cancer EMT models exhibiting stem-like properties demonstrated a highly significant intersection between the gene expression profiles of TS$^{KI4}$ cells and CL human tumors and cell lines. Repressed genes from the EMT gene signature demonstrated loss of H2BK5Ac in TS$^{KI4}$, TS$^{shCBP}$ and CL cells. These results highlight the importance of MAP3K4/CBP-mediated acetylation of H2BK5 for maintenance of the TS cell epithelial phenotype.

## Results

### *$TS^{KI4}$ Cells Exhibit Properties of Both Stemness and EMT*

In the presence of FGF4, TS cells maintain self-renewal as defined by the maintenance of cell division with multipotency (Niwa *et al*, 2005; Tanaka, 1998). $TS^{KI4}$ cells lack MAP3K4 activity but are self-renewing and multipotent. When cultured in the presence of FGF4, $TS^{KI4}$ cells expressed the TS cell markers Cdx2, Eomes, Esrrβ and FGFR2 at levels similar to $TS^{WT}$ cells (**Figures 3.1**A-D). Removal of FGF4 promoted the differentiation of $TS^{KI4}$ cells into all trophoblast lineages similar to $TS^{WT}$ cells with the loss of expression of stem cell markers (**Figures 3.1**A-D) and the gain of expression of trophoblast lineage markers including the giant cell marker PLI, the spongiotrophoblast marker Tpbpα, and the syncytiotrophoblast marker Gcm1 (**Figures 3.1**E-G). Developmental potency of these stem cells was established through the injection of GFP-labeled $TS^{WT}$ or $TS^{KI4}$ cells into wild-type blastocysts. Both $TS^{WT}$ and $TS^{KI4}$ cells produced chimeric conceptuses with similar frequencies (**Figure 3.1**H) and contributed selectively to the trophoblast lineages including the extraembryonic ectoderm, ectoplacental cone, and giant cells (**Figures 3.1**I-N and **3.2**A-H). These cells were not found in the embryo or in any ICM-derived extraembryonic tissue (**Figures 3.1**I-N).

$TS^{KI4}$ cells also exhibit the molecular and cellular characteristics of EMT. $TS^{WT}$ cells grew with an epithelial morphology (**Figure 3.1**O), with strong peripheral E-cadherin (**Figure 3.1**P) and cortical actin expression at cell-cell contacts (**Figures 3.1**Q-R). In contrast, $TS^{KI4}$ cells exhibited a polarized front-back end, mesenchymal morphology (**Figure 3.1**S). E-cadherin was significantly reduced in $TS^{KI4}$ cells

(**Figures 3.1**T). Also, filamentous actin was observed at the leading edge of $TS^{KI4}$ cells (**Figures 3.1**U-V). Reduction of total E-cadherin in $TS^{KI4}$ cells relative to $TS^{WT}$ cells was further validated by Western blotting (**Figure 3.1**W). Real-time quantitative RT-PCR (qRT-PCR) showed a modest increase in the expression of the mesenchymal markers vimentin and N-cadherin in $TS^{KI4}$ cells, and Western blotting revealed a 2.9-fold and 2.2-fold increase in vimentin and N-cadherin protein expression respectively (**Figures 3.2**I and **3.2**J)**.** Differentiation of both $TS^{WT}$ and $TS^{KI4}$ cells also increased vimentin and N-cadherin protein expression (**Figure 3.2**J). $TS^{KI4}$ cells and 4-day differentiated $TS^{WT}$ cells exhibited a fourteen to sixteen-fold increase in invasiveness compared to $TS^{WT}$ cells (**Figure 3.1**X). Hyperinvasiveness of the $TS^{KI4}$ cells was also seen, *in vivo*. Defective decidualization is induced by excessive trophoblast invasion (Norwitz *et al*, 2001). Compared to injection of $TS^{WT}$ cells (**Figure 3.1**Y), injection of $TS^{KI4}$ cells into wild-type blastocysts resulted in defective decidualization consistent with the hyperinvasiveness of $TS^{KI4}$ cells (**Figures 3.1**Z-BB). Cumulatively, these data demonstrate that $TS^{KI4}$ cells are self-renewing stem cells with properties of EMT including loss of E-cadherin and gain of invasiveness.

***Developmental EMT in $TS^{WT}$ Cells***

Removal of FGF4 promotes the differentiation of TS cells and increased invasiveness of trophoblasts through Matrigel-coated transwells (**Figure 3.3**A). The invasive population was largest at four days post FGF4 withdrawal ($T^{INV}$) (**Figure 3.3**A). Morphologically, $TS^{WT}$ cells cultured in FGF4 grew in tight epithelial sheets with actin localized around the cell periphery (**Figures 3.3**B-C). In contrast, $T^{INV}$ cells

isolated from the bottom of Matrigel-coated transwells exhibited mesenchymal cell characteristics with prominent actin stress fibers, filamentous actin and loss of E-cadherin (**Figures 3.3**D-G). E-cadherin was more significantly reduced in $T^{INV}$ cells compared to four day differentiated trophoblasts ($T^{DIFF}$) not selected for invasiveness (**Figures 3.3**F-G). Expression of the mesenchymal marker fibronectin was significantly increased and vimentin was also expressed in $T^{INV}$ cells (**Figure 3.4**A). These changes in invasiveness, morphology, and E-cadherin expression indicated that the $T^{INV}$ cells have undergone EMT.

***Induction of EMT in $TS^{WT}$ Cells by Ectopic Expression of Snail***

To induce EMT in TS cells, we ectopically expressed a defined inducer of EMT, the transcription factor Snail, in $TS^{WT}$ cells ($TS^{Snail}$) (Cano *et al*, 2000). $TS^{Snail}$ cells exhibited a polarized, mesenchymal morphology with increased filamentous actin and mesenchymal markers vimentin and N-cadherin (**Figures 3.3**H-J and **3.4**B). A ten-fold increase in invasiveness compared to $TS^{WT}$ cells and loss of E-cadherin showed Snail overexpression induced TS cell EMT (**Figures 3.3**K-L). $TS^{Snail}$ cells expressed TS stem cell markers and were able to differentiate into trophoblast lineages upon removal of FGF4 (**Figures 3.4**C-D). To identify transcription factors responsible for inducing EMT in TS cells, we compared the expression levels of eight transcription factors known to regulate EMT in other systems. FGF4 withdrawal from $TS^{WT}$ cells induced Snail message at four days of differentiation and in $T^{INV}$ cells suggesting that Snail is important for TS cell EMT (**Figures 3.3**M and **3.4**E). The EMT-inducing transcription factors Slug and Twist were elevated in undifferentiated $TS^{KI4}$ cells, but were not induced with differentiation

or invasion of $TS^{WT}$ cells (**Figures 3.3**N-O and **3.4**E). Upregulation of Lef1 and Ets1 were also detected in $TS^{KI4}$ and $TS^{Snail}$ cells (**Figures 3.3**P-Q). Similar to Snail, Lef1 was induced at four days of $TS^{WT}$ cell differentiation and in $T^{INV}$ cells (**Figure 3.3**P and **3.4**E). Other known transcriptional regulators of EMT like Zeb1, Foxc2, and Gsc were not induced with EMT of $T^{INV}$ or $TS^{KI4}$ cells (**Figures 3.4**F-H).

***Gene Expression Changes with the Acquisition of Trophoblast Invasiveness***

To identify genes related to the acquisition of invasiveness in trophoblast EMT, we measured gene expression changes in $T^{INV}$ cells compared to $TS^{WT}$ cells using genome-wide Agilent microarrays. With a Benjamini-Hochberg p-value <0.05, 6641 genes exhibited a two-fold change in expression in $T^{INV}$ cells (**Table S1**, http://www.sciencedirect.com/science/article/pii/S1934590911001184). Since $T^{INV}$ cells were differentiated for four days, a component of the 6641 changed genes reflected differentiation-specific changes unrelated to the acquisition of invasiveness. When gene expression changes of $T^{DIFF}$ cells compared to $TS^{WT}$ cells were measured, 5706 genes exhibited altered expression by a minimum of two-fold (**Table S1**, http://www.sciencedirect.com/science/article/pii/S1934590911001184). Direct comparison of gene expression changes in $T^{INV}$ and $T^{DIFF}$ cells revealed that 2359 genes had significantly different expression measured as a minimum 1.5-fold change with 80% of these genes being changed only in $T^{INV}$ cells (**Table S2**, http://www.sciencedirect.com/science/article/pii/S1934590911001184). Gene ontology (GO) analysis of the 2359 genes showed significant enrichment in canonical KEGG pathways regulating focal adhesions, actin cytoskeleton, and

adherens junctions (**Figure 3.5**A). These findings define an invasive gene signature for trophoblast EMT.

### *Gene Expression Changes in $TS^{KI4}$ Cells*

Unlike $T^{INV}$ cells that have undergone EMT but have lost stemness, $TS^{KI4}$ cells are self-renewing stem cells in EMT. $TS^{KI4}$ cells serve as a useful model to distinguish between genes that mediate invasiveness and EMT. Gene expression profiling defined changes in $TS^{KI4}$ cells compared to $TS^{WT}$ cells, identifying 1083 significantly upregulated genes and 702 significantly downregulated genes by more than two-fold (Benjamini-Hochberg p-value <0.05) (**Table S1**, http://www.sciencedirect.com/science/article/pii/S1934590911001184). GO analysis of downregulated genes showed a significant enrichment in nine canonical signaling pathways (**Figure 3.5**B). The top three enriched signaling pathways included focal adhesions, ECM-interactions, and regulation of the actin cytoskeleton, the same pathways that were identified in the $T^{INV}$ signature.

### *Comparison of $TS^{KI4}$ and $T^{INV}$ Signatures Identify Genes Important for EMT*

Unsupervised hierarchical clustering of $TS^{KI4}$, $T^{INV}$ and $T^{DIFF}$ cells showed a common gene set for $T^{INV}$ and $T^{DIFF}$ cells indicated by a node correlation coefficient of 0.82 (**Figure 3.6**A). This strong correlation reflects the significant changes in gene expression due to FGF4 removal and loss of stemness in $T^{INV}$ and $T^{DIFF}$ cells, as compared to $TS^{KI4}$ cells, showing that $TS^{KI4}$ cells are not differentiated cells. Venn diagrams of upregulated (**Figure 3.6**B) and downregulated (**Figure 3.6**C) genes highlight the significant intersection between $T^{DIFF}$ and $T^{INV}$ cells, while demonstrating the limited intersection of 416 genes between $TS^{KI4}$ and $T^{INV}$ cells. Significantly

changed genes were categorized according to biological function, and 25% of these genes regulate cell adhesion and motility including Lamb2, Fn1 and RhoB (**Figures 3.6**D and **3.5**C). In addition, genes were identified whose importance to invasion and EMT has not been previously defined, including the enrichment for genes regulating RNA splicing, transcription, translation, and protein degradation (**Figure 3.5**C). However, these genes must be functionally tested to prove their role in invasion and EMT. Using qRT-PCR, we validated the expression changes of several genes showing similar changes in $TS^{KI4}$ and $T^{INV}$ cells (**Figures 3.5**D-E).

***Acetylation of Specific H2A and H2B Lysines Is Inhibited in $TS^{KI4}$ Cells***

Induction of differentiation by FGF4 removal from $TS^{WT}$ cells was accompanied by reduced acetylation of all core histones, specifically at H2AK5, H2BK5/K12/K15/K20, H3K9, and H4K8 (**Figure 3.7**A). Trimethylation of H3K4 and H3K9 was unchanged with differentiation (data not shown). The differentiation-induced loss of histone acetylation suggests the importance of histone acetylation in maintaining the undifferentiated epithelial state of TS cells. We therefore examined histone modifications in $TS^{KI4}$ cells. **Figure 3.7**B shows the loss of acetylation at H2AK5, H2BK5, H2BK12, and H2BK15 in $TS^{KI4}$ cells compared to $TS^{WT}$ cells. H3K9Ac and H4K8Ac were unaffected demonstrating that loss of histone acetylation was selective for H2A and H2B. Examination of histone methylation showed that trimethylation of H3K4 and H3K9 was not altered, suggesting that loss of H2A and H2B acetylation (H2A/H2BAc) in $TS^{KI4}$ cells occurs independently of changes in histone methylation (**Figure 3.7**C). Examination of $TS^{Snail}$ cells that exhibit properties of stemness and EMT similar to $TS^{KI4}$ cells revealed the selective loss of

H2A/H2BAc with no change in H3 and H4 acetylation (**Figure 3.7**D). These data show the association of loss of H2A/H2BAc with changes characteristic of EMT while maintaining stemness in $TS^{KI4}$ cells.

***Reduction of H2BK5Ac on Select Gene Promoters in $TS^{KI4}$ cells***

We used ChIP-seq to identify genes associated with the epigenetic mark H2BK5Ac during TS cell reprogramming important for the induction of EMT. After obtaining DNA samples by ChIP using a highly specific H2BK5Ac antibody (**Figures 3.8**A-B), we generated a total of 44 and 27 million Illumina sequence reads for $TS^{WT}$ and $TS^{KI4}$ cells, respectively, consistent with Western blotting data (**Figure 3.7**B). This genome-wide analysis of the read-tag distribution demonstrated that H2BK5Ac is significantly enriched near the transcription start sites (TSS) of 13625 genes in TS cells (**Figure 3.9**A and **Table S3**, http://www.sciencedirect.com/science/article/pii/S1934590911001184). This H2BK5Ac peak location profile near the TSS is consistent with published studies in $CD4^+$ T cells (Wang *et al*, 2008).

Limiting the analysis to the well-annotated RefSeq gene set (NCBI37/mm9), we compared the read tag density between $TS^{WT}$ and $TS^{KI4}$ cells within 1kb upstream and downstream of the TSS. We normalized read-tag counts based upon the ratio of genome-wide total mapped reads between $TS^{WT}$ and $TS^{KI4}$ cells. We removed background regions by filtering genes whose read tag counts did not significantly exceed cell-type specific background noise (<20 read tags per 2kb region). After removing duplicate genes, 4163 genes were significantly different in H2BK5Ac between $TS^{WT}$ and $TS^{KI4}$ cells, as determined by the exact rate ratio test

with Benjamini-Hochberg adjusted p-value $\leq$ 0.05. From the 4163 genes with a significant change in H2BK5Ac, 3515 genes had a significant loss of H2BK5Ac, while 648 genes had a significant gain of H2BK5Ac in $TS^{KI4}$ compared to $TS^{WT}$ cells (**Figures 3.9**B-C).

Changes in H2BK5Ac were visualized by normalized read tag density plots. We demonstrated the dramatic loss of H2BK5Ac in $TS^{KI4}$ cells for a select set of genes including Acsl6, Dbndd2, Itgav, Krt19 and Trim54 (**Figures 3.9**D and **3.8**C). These are examples of genes with a highly significant loss of H2BK5Ac (i.e. Benjamini-Hochberg p-values $<10^{-18}$) and occur in the top 3% of affected genes. Loss of H2BK5Ac in $TS^{KI4}$ compared to $TS^{WT}$ cells was confirmed by ChIP-qRT-PCR (**Figure 3.9**E) and correlated with the loss of gene expression (**Figure 3.5**E). Furthermore, Acsl6, Itgav, Lamb2 and Trim54 demonstrated a similar loss of H2BK5Ac in $T^{INV}$ cells by ChIP-qRT-PCR, suggesting the importance of these genes in regulating the EMT program that occurs during trophoblast differentiation (**Figure 3.9**E). Although the majority of genes had a loss of H2BK5Ac in $TS^{KI4}$ compared to $TS^{WT}$ cells, 648 genes had an increase in H2BK5Ac density in $TS^{KI4}$ cells. Normalized density plots of Dkk3 and Mycn highlight two genes with a significant increase in H2BK5Ac (i.e. Benjamini-Hochberg p-values $< 10^{-12}$) and a coordinate increase in gene expression in $TS^{KI4}$ cells (**Figure 3.9**F and **Tables S1** and **S3**, http://www.sciencedirect.com/science/article/pii/S1934590911001184).

Consistent with the maintenance of TS cell multipotency, enrichment of H2BK5Ac occurs at the promoters of the TS cell markers Cdx2, Eomes, and Fgfr2 (**Figure 3.9**G). As indicated from normalized promoter density plots, Eomes and

Esrrβ demonstrated unchanged levels of H2BK5Ac between $TS^{WT}$ and $TS^{KI4}$ cells, while Cdx2 and Fgfr2 demonstrate a 50% decrease in H2BK5Ac density (**Figure 3.9**G). H2BK5Ac levels were validated by H2BK5Ac ChIP-qRT-PCR for the Eomes and Cdx2 promoters (**Figure 3.8**D).

***Loss of H2BK5Ac Correlates with Repression of Genes Critical to Maintenance of the TS Cell Epithelial Phenotype***

Using mRNA-seq to quantitatively compare the changes in H2BK5Ac with the changes in gene expression for $TS^{WT}$ and $TS^{KI4}$ cells, there was a modest positive correlation between changes in gene expression and H2BK5Ac, as determined by a Pearson correlation coefficient of 0.62 (p-value $<10^{-16}$) (**Figure 3.9**H). This finding was further supported by GO analysis of genes both significantly downregulated and hypoacetylated showing shared gene changes important for maintenance of focal adhesions, the actin cytoskeleton, and extracellular matrix interactions (**Figure 3.5**B and **Figure 3.8**E). Collectively, these results highlight the importance of H2BK5Ac in regulating the active gene transcription program of TS cells, whereby loss of H2BK5Ac results in repression of genes critical to maintenance of the epithelial phenotype.

***$TS^{KI4}$ Cells and Claudin-low Breast Cancer Share EMT Properties***

Recent studies have shown that the CL subtype of triple negative breast cancer exhibits both mesenchymal and stem-like properties (Prat *et al*, 2010; Hennessy *et al*, 2009). Compared to the four other breast tumor subtypes (i.e. Luminal A, Luminal B, HER2-enriched and Basal-like), CL tumors have the lowest expression of epithelial differentiation markers CD24, EpCAM, KRT7/19 and the cell

adhesion proteins CLDN3/4/7 and CDH1, while exhibiting highest expression of the mesenchymal markers VIM, N-cadherin, SNAI2 and TWIST1 (Prat & Perou, 2011; Prat *et al*, 2010). Hierarchical clustering of 22 genes characteristic of EMT and stemness from gene expression data of $TS^{KI4}$ cells and 52 breast cancer cells lines reported by Neve et al. revealed that $TS^{KI4}$ cells clustered most closely with the CL breast cancer subtype (**Figure 3.10**A) (Neve *et al*, 2006). Similar to the CL cell lines, $TS^{KI4}$ cells exhibited an increase in the mesenchymal markers VIM, CDH2, SNAI2 and TWIST1 with loss of the epithelial differentiation and cell adhesion markers CD24, KRT7/8/19 and CLDN4 (**Figure 3.10**A). Next, we examined the intersection between genes with unique Entrez identifiers from gene array data of CL cell lines compared to gene array data of $TS^{KI4}$ cells. The intersection between upregulated and downregulated genes in the CL cell lines compared to $TS^{KI4}$ cells was determined to be significant with 62 upregulated (p-value < 0.005) and 31 downregulated (p-value < 0.01) overlapping genes (**Figure 3.10**B). This overlapping gene set was plotted on the basis of $\log_2$ ratio values from $TS^{KI4}$ cells to demonstrate gene expression changes of the intersecting $TS^{KI4}$/CL cell EMT gene signature (**Figure 3.10**C). Genes important for the induction of the mesenchymal phenotype, such as CDH2, DKK1, MET, PDGFR$\beta$, SNAI2, TIMP2, THY1, TWIST1 and VIM were significantly upregulated, while genes important for maintenance of the epithelial phenotype and cell adhesion, such as AIM1, BCAM, KRT7, KRT19 and RAB25, were repressed (**Figure 3.10**C). In addition to these known regulators of EMT, this significant genetic intersection between two distinct EMT models with stem cell characteristics highlights a gene set important for both EMT and stemness.

### *Epigenetic Repression of $TS^{KI4}$/CL EMT Genes by Reduction of H2BK5Ac*

TS cell EMT models, $TS^{KI4}$ and $TS^{Snail}$, demonstrated selective loss of histone H2A/H2BAc (**Figures 3.7**D). By H2BK5Ac ChIP-qRT-PCR, we examined the levels of H2BK5Ac on 32 downregulated genes that have overlapping gene expression profiles between CL SUM159 and $TS^{KI4}$ cells and are known to have a significant loss of H2BK5Ac in $TS^{KI4}$ cells by ChIP-seq (**Figures 3.10**C and **Table S3**, http://www.sciencedirect.com/science/article/pii/S1934590911001184). Of the 32 genes tested by H2BK5Ac ChIP-qRT-PCR, 75 percent of these genes were validated to have a loss of H2BK5Ac and a coordinate loss of gene expression in $TS^{KI4}$ compared to $TS^{WT}$ cells (**Figure 3.10**D and **Figures 3.11**A-C). Furthermore, 81 percent of these genes had a similar loss of H2BK5Ac in $TS^{Snail}$ cells (**Figure 3.10**D and **3.11**B-C). Due to the significant genetic intersection between CL cell lines and the $TS^{KI4}$ EMT model, we determined the levels of H2BK5Ac on the promoters of the same 32 overlapping genes in the CL SUM159 cell line. Importantly, 80 percent of these genes had a loss of H2BK5Ac in SUM159 cells compared to human mammary epithelial cells (HMECs), as determined by ChIP-qRT-PCR (**Figure 3.10**E and **3.11**D-E). These results suggest that loss of H2BK5Ac represses genes with an important role in maintenance of the epithelial phenotype, thereby contributing to the progression of two distinct EMT programs.

The pathological significance of the $TS^{KI4}$/CL association was further emphasized by comparing the gene expression profile of $TS^{KI4}$ cells to the five intrinsic molecular subtypes of breast tumors cataloged in the UNC337 data set (Prat *et al*, 2010). Tumors from the CL breast cancer subtype showed highest

expression of the $TS^{KI4}$ gene expression signature compared to the Basal-like, HER2-enriched, Luminal A and Luminal B breast tumors (**Figure 3.10**F). Further analysis of the overlapping gene expression profiles in CL human tumors and $TS^{KI4}$ cells demonstrated a highly significant intersection between the gene array profiles of $TS^{KI4}$ cells and CL human tumors (p-value < 0.0001 for upregulated genes) (**Figure 3.10**G). Although the intersection between the downregulated genes from $TS^{KI4}$ cells and CL human tumors consists of only 13 genes, approximately 50 percent of these genes AIM1, IRX5, KRT7, KRT19, RAB25 and SCYL3 were present in the intersecting $TS^{KI4}$/CL EMT gene signature; these same genes also exhibited a coordinate loss of H2BK5Ac in $TS^{KI4}$ and CL cells (**Figure 3.10**D). These findings highlight the importance of H2BK5Ac on genes whose repression is important for EMT in both developmental stem cell and metastatic human tumor models of EMT with stem cell properties.

***MAP3K4 regulates CBP acetylation of H2A and H2B***

Previously, we showed the requirement of MAP3K4 kinase activity for neural tube, skeletal, and placental development (Abell *et al*, 2005; 2009). We systematically examined the MAP3K4 signaling network for genes whose targeted disruption resulted in phenotypes similar to that of KI4 mice (**Table 3.1**). Strikingly, the genes overlapping most closely with the developmental defects observed with loss of MAP3K4 kinase activity were the histone acetyltransferase (HAT) CBP and its closely related family member p300. This phenotypic overlap suggested that the loss of H2A/H2BAc in $TS^{KI4}$ cells may be related to altered CBP and/or p300 HAT activity. Nuclear extracts isolated from $TS^{KI4}$ cells have significantly diminished HAT

activity relative to $TS^{WT}$ cells (**Figure 3.12**A). Total CBP and p300 protein expression was unchanged in $TS^{KI4}$ cells (**Figure 3.13**A). MAP3K4-dependent JNK phosphorylation of CBP and p300 increased HAT activity, which was blocked by the JNK inhibitor SP600125 (**Figures 3.12**B and **3.13**B). $TS^{KI4}$ cells have a strongly diminished JNK activity (Abell *et al*, 2009), consistent with MAP3K4-dependent JNK activation regulating CBP/p300 HAT activity. Co-expression of MAP3K4 and JNK resulted in a 17.8 and 8.3-fold increase in the phosphorylation of CBP and p300, respectively (**Figures 3.12**C and **3.13**C). CBP/p300 phosphorylation was JNK dependent, as p38 activation did not significantly alter phosphorylation of CBP or p300 (**Figures 3.12**C and **3.13**C). To determine if CBP or p300 regulate endogenous TS cell functions, we infected $TS^{WT}$ cells with shRNAs to either CBP or p300. We achieved greater than 85% knockdown of CBP or p300 with three to four individual shRNAs for each (**Figures 3.12**D-E and data not shown). Unlike control virus infected cells (**Figure 3.12**F), loss of CBP resulted in a dramatic change in morphology with $TS^{shCBP}$ cells exhibiting a front-back end polarized morphology (**Figures 3.12**G-H). In contrast, cells with loss of p300 maintained a normal epithelial morphology (data not shown). Compared to control virus infected cells, stemness markers in $TS^{shCBP}$ cells were unchanged for Eomes and FGFR2 and decreased by 25% for Cdx2 and Esrrβ (**Figure 3.13**E). Further, expression of the mesenchymal marker vimentin was increased at both the level of message and protein in $TS^{shCBP}$ cells (**Figure 3.13**F and data not shown). Most importantly, $TS^{shCBP}$ cells exhibited a five to fifteen-fold increase in invasiveness as compared to control virus infected cells (**Figure 3.12**I). Changes in morphology and the expression of mesenchymal

markers combined with increases in invasiveness suggest that loss of CBP in $TS^{WT}$ cells induces EMT. Examination of histone acetylation in $TS^{shCBP}$ cells revealed the selective and specific loss of H2A/H2BAc (**Figure 3.12**J). In contrast, loss of p300 resulted in a reduction in H3 and H4 acetylation, but did not affect H2A/H2BAc (**Figure 3.12**K). These data strongly suggest that CBP is the primary HAT that regulates H2A/H2BAc in TS cells and that the loss of H2A/H2BAc is sufficient to induce EMT in TS cells.

Genes downregulated in the $TS^{KI4}$/CL EMT gene signature were similarly decreased in $TS^{shCBP}$ cells (**Figure 3.13**G). Because $TS^{shCBP}$ cells exhibit the selective loss of H2A/H2BAc acetylation similar to $TS^{KI4}$ and $TS^{Snail}$ cells, (**Figures 3.7**D and **3.12**J), we used H2BK5Ac ChIP-qRT-PCR to measure H2BK5Ac on downregulated genes from the intersecting $TS^{KI4}$/CL gene profile (**Figure 3.13**G). Of these genes, 72 percent demonstrated a loss of H2BK5Ac in $TS^{shCBP}$ cells compared to control virus-infected cells (**Figure 3.13**H). These data show the coordinate loss of H2BK5Ac and gene expression in CL, $TS^{KI4}$, $TS^{Snail}$, and $TS^{shCBP}$ cells. Together, these findings show the importance of CBP-mediated H2BK5Ac in maintaining the epigenetic landscape important for the epithelial phenotype of TS cells.

## Discussion

We have shown MAP3K4 dependent activation of JNK in response to FGF4 controls CBP activity for the maintenance of the TS cell epithelial phenotype (**Figure 3.14**). Loss of MAP3K4 kinase activity in $TS^{KI4}$ cells results in gain of EMT properties including reduced E-cadherin, and morphological changes characteristic of

mesenchymal cells and increased invasiveness. $TS^{KI4}$ cells also retain stemness defined by self-renewal with the maintenance of multipotency. These properties of $TS^{KI4}$ cells show a functional separation of FGF4 dependent control of epithelial maintenance and stemness, with MAP3K4 signaling being critical for the epithelial phenotype.

$TS^{KI4}$ cells exhibit the selective loss of H2A/H2BAc, whereas histone H3 and H4 acetylation was largely unaffected. Loss of H2BK5Ac is restricted to a select set of genes in $TS^{KI4}$ cells whose expression is significantly reduced. Epithelial maintenance was also disrupted by CBP knockdown, causing the loss of H2A/H2BAc similar to that observed with $TS^{KI4}$ cells. Loss of CBP expression induced a phenotype similar to $TS^{KI4}$, including gain of invasiveness and EMT characteristics while maintaining stemness. Consistent with the novel role of CBP and H2BK5Ac in regulation of gene expression profiles important for the epithelial phenotype, H3K4me3 and H3K9me3 are unchanged in $TS^{KI4}$ cells. Additionally, H3K27me3 has been shown as unimportant in TS cell differentiation (Rugg-Gunn *et al*, 2010). Thus, histone acetylation by CBP is a primary mechanism for maintenance of the epithelial phenotype of TS cells, whereby loss of H2BK5Ac results in EMT. This finding is consistent with the role for CBP in maintaining hematopoietic stem cell self-renewal (Rebel *et al*, 2002). In addition to direct inhibition of CBP HAT activity, it is possible that a secondary mechanism of regulation exists to target loss of H2A/H2BAc to select gene promoters, whereby changes in CBP phosphorylation controls interactions with transcriptional regulators of EMT (Shenghui *et al*, 2009).

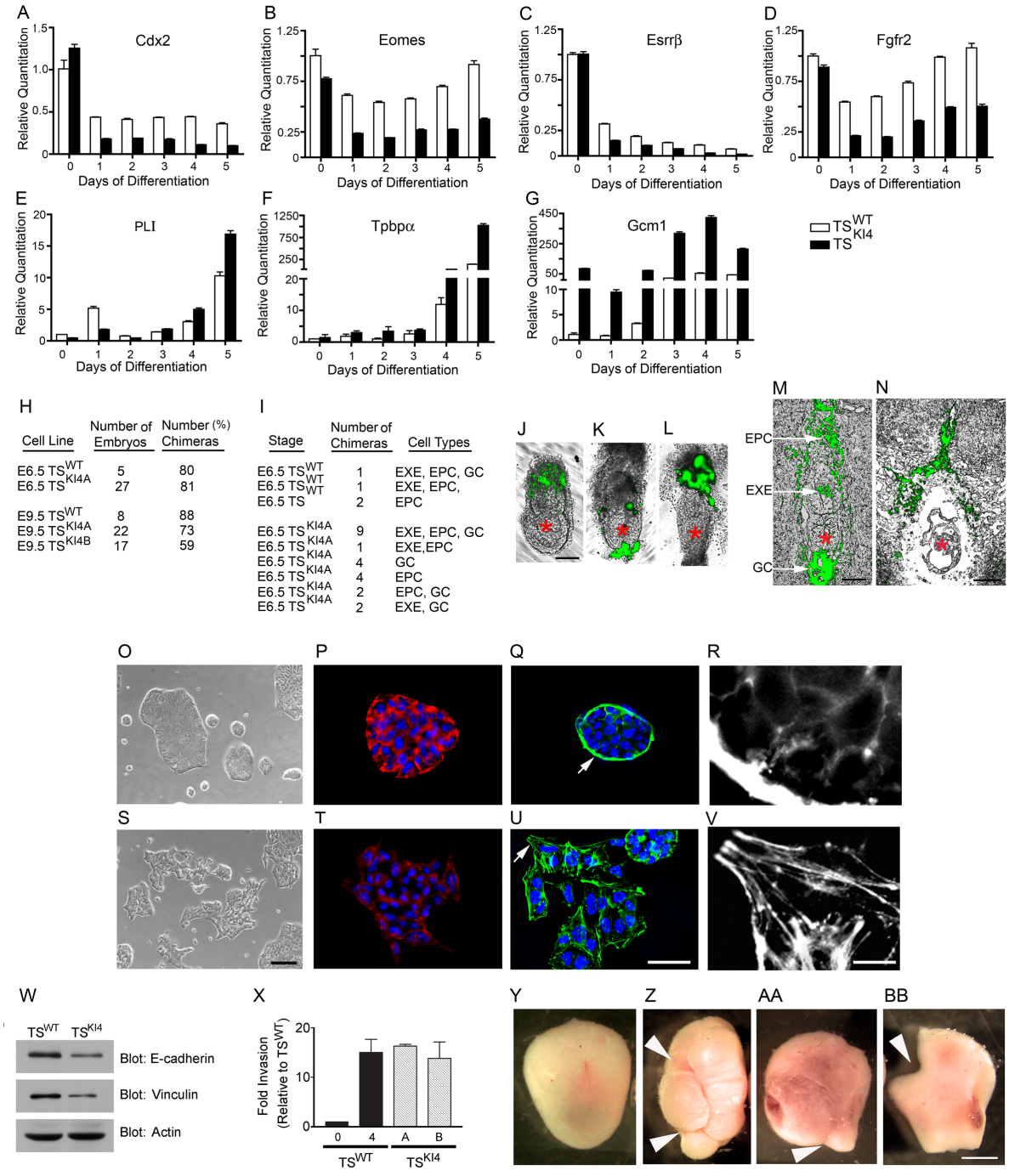
Ectopic expression of Snail has been used to induce EMT in different cell types, and overexpression of Snail in HMECs induced a mesenchymal phenotype with expression of specific stem cell markers (Mani *et al*, 2008). This phenotype is reminiscent of $TS^{KI4}$ cells, which induce EMT while maintaining stemness. Stable expression of Snail in TS cells resulted in the selective loss of H2A/H2BAc and properties of EMT and stemness, similar to $TS^{KI4}$ and $TS^{shCBP}$ cells. ChIP-qRT-PCR studies showed loss of H2BK5Ac on an overlapping set of genes for $TS^{KI4}$, $TS^{shCBP}$, and $TS^{Snail}$ cells, defining each as a unique model system for the epigenetic control of EMT in a self-renewing primary cell.

In contrast to $TS^{KI4}$, $TS^{shCBP}$, and $TS^{Snail}$ cells, $T^{INV}$ cells have completed EMT, having fully lost their epithelial morphology, as evidenced by the increase in invasiveness and gain in the mesenchymal morphology associated with filamentous actin and increased expression of the mesenchymal marker fibronectin. $T^{INV}$ cells do not self-renew and have lost acetylation of all four core histones. Since $T^{INV}$ cells have completed EMT, $T^{INV}$ gene expression profiles probably lack the EMT initiators, instead showing the expression of EMT executors (Thiery *et al*, 2009). Compared to $T^{INV}$ cells, $TS^{KI4}$, $TS^{shCBP}$, and $TS^{Snail}$ cells are in an intermediate state of EMT, where they are not fully mesenchymal but exhibit properties of both EMT and stemness. $TS^{KI4}$ cells are uniquely trapped in this intermediate EMT state prior to complete acquisition of the mesenchymal phenotype, which can still be induced by the removal of FGF4.

TS cell EMT shares many key properties with neural crest and cancer cell EMTs including loss of E-cadherin, gain of front-to-back polarity, and increased

invasiveness (Yang & Weinberg, 2008). However, there are differences in marker expression among these different EMT models indicative of cell type and stage-specific EMT. For example, mesenchymal markers such as vimentin and N-cadherin are differentially expressed in these different EMT models, with N-cadherin repression being required for neural crest EMT (Yang & Weinberg, 2008). Lamb2 is increased in hepatocyte EMT, but reduced in neural crest, CL, $T^{INV}$, $TS^{KI4}$ and $TS^{Snail}$ EMTs. Fibronectin is elevated in breast and gastric cancers and in $T^{INV}$ cells but reduced in $TS^{KI4}$ and $TS^{Snail}$ cells. The critical property for each EMT model is increased invasiveness (Kalluri, 2009).

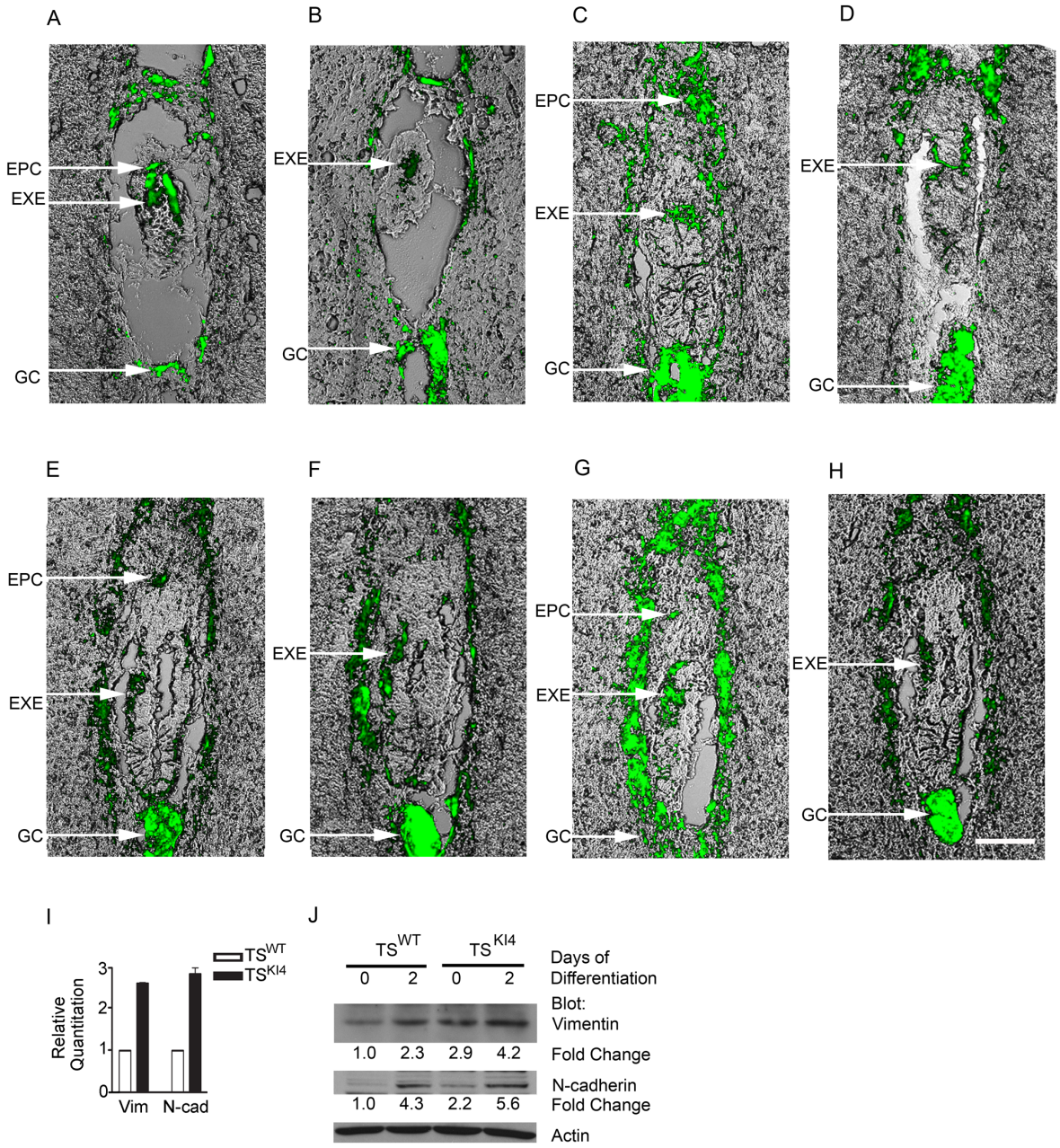
Finally, $TS^{KI4}$ and CL human breast cancer cells share properties of stemness and EMT with a common gene expression profile also found in patient CL tumors. Intersecting genes with loss of expression had a correlative loss of H2BK5Ac in both $TS^{KI4}$ and CL cells. Some of these genes have defined roles in maintenance of the epithelial phenotype such as Aim1, Rab25 and Galnt3 (Maupin *et al*, 2010; Ray *et al*, 1997), but most of the shared genes in the $TS^{KI4}$/CL intersecting list have not been characterized for their role in epithelial maintenance or EMT and should be analyzed in different EMT models. Discovery of how H2A/H2BAc controls maintenance of the epithelial TS cell phenotype provides unique insight into how signaling networks controlling tissue stem cell EMT can be used to define previously unrecognized genes contributing to cancer cell EMT. This discovery may lead to defining novel gene targets or combinations of targets whose inhibition can be used to selectively inhibit TICs.




***Note:** Mouse chimera experiments were courtesy of Dr. Amy Abell.

**Figure 3.1: $TS^{KI4}$ Cells Deficient in MAP3K4 Activity Maintain Self-renewal, Multipotency, and Developmental Potency While Exhibiting Properties of EMT.** (A-D) Expression of TS cell stemness genes in $TS^{KI4}$ cells measured by qRT-PCR. (E-G) Maintenance of multipotency in $TS^{KI4}$ cells demonstrated by induction of trophoblast differentiation markers measured by qRT-PCR. Differentiation was induced by FGF4 withdrawal for the indicated number of days. (H-N) Developmental

potency of $TS^{WT}$ and $TS^{KI4}$ cells demonstrated by injection of GFP positive $TS^{WT}$ and $TS^{KI4}$ cells into wild-type blastocysts. (H) Data shows the number of GFP positive chimeras per cell type injected expressed as a percent of total embryos examined. A and B are two independent $TS^{KI4}$ cell clones. (I) Sites of incorporation of GFP positive $TS^{WT}$ and $TS^{KI4}$ cells in E6.5 conceptuses. (J, K, and L) Three representative whole embryo E6.5 $TS^{KI4}$ chimeras and (M) 18 μm section from E6.5 $TS^{KI4}$ chimera are shown. Black bar represents 100 μm. EPC, ectoplacental cone; EXE, extraembryonic ectoderm; GC, giant cells. (N) Representative E9.5 $TS^{WT}$ cell chimera. Red asterisks indicate location of embryos. (O-V) Mesenchymal phenotype of $TS^{KI4}$ cells compared to epithelial morphology of $TS^{WT}$ cells. (O) Phase microscopy of $TS^{WT}$ cells. (P) Immunostaining with anti-E-cadherin antibody (red) shows strong staining around the cellular periphery in $TS^{WT}$ cells with nuclear DAPI stain (blue). (Q) Cortical actin staining (green) and nuclear DAPI stain (blue) in $TS^{WT}$ cells. (R) Enlarged inset of region indicated by arrow in (Q) showing peripheral cortical actin in $TS^{WT}$ cells. (S) Phase microscopy of $TS^{KI4}$ cells. Black bar represents 100 μm. (T) Relocalization and loss of E-cadherin (red) from the periphery of $TS^{KI4}$ cells with nuclear DAPI stain (blue). (U) Filamentous actin staining (green) and nuclear DAPI stain (blue) in $TS^{KI4}$ cells. White bar represents 50 μm. (V) Enlarged inset of region indicated by arrow in (U) showing filamentous actin in $TS^{KI4}$ cells. White bar represents 10 μm. White arrow indicates site for insets in (R) and (V). (W) Reduced protein expression of E-cadherin and vinculin in $TS^{KI4}$ cells shown by Western blotting. (X) Invasion through Matrigel by undifferentiated $TS^{WT}$ cells (0), $TS^{WT}$ cells differentiated for four days by FGF4 removal (4), or two independent $TS^{KI4}$ cell clones (A and B) cultured in the presence of FGF4. Fold change is relative to $TS^{WT}$ cells. (Y) Representative E9.5 decidua from blastocyst injected with $TS^{WT}$ cells. (Z-BB) Three representative E9.5 deciduas from blastocysts injected with $TS^{KI4}$ cells. Arrowheads indicate sites of defective decidualization. (A-G) and (X) represent the mean ± range of two independent experiments.




***Note:** Mouse chimera experiments were courtesy of Dr. Amy Abell.

**Figure 3.2: Developmental Potency and Mesenchymal Properties of $TS^{KI4}$ Cells**
(A-H) Developmental potency of $TS^{WT}$ and $TS^{KI4}$ cells demonstrated by injection of GFP positive $TS^{WT}$ and $TS^{KI4}$ cells into wild-type blastocysts. Representative 18 μm section from $TS^{WT}$ cell (A,B) or $TS^{KI4}$ cell (C-H) E6.0 to E6.5 chimeras. (E-H) Serial sections through an E6.5 chimera generated by injection of GFP positive $TS^{KI4}$ cells. EPC, ectoplacental cone; EXE, extraembryonic ectoderm; GC, giant cells. White bar equals 100 μm. (I) Increased expression of mesenchymal markers in $TS^{KI4}$ cells relative to $TS^{WT}$ cells as measured by qRT-PCR. Data shows the mean ± range of

two independent experiments performed in triplicate. (J) Increased expression of mesenchymal markers in $TS^{KI4}$ relative to $TS^{WT}$ cells and with trophoblast differentiation as measured by Western blotting with the indicated antibodies. Fold changes are relative to undifferentiated $TS^{WT}$ cells. Data shows representative blots from two independent experiments.

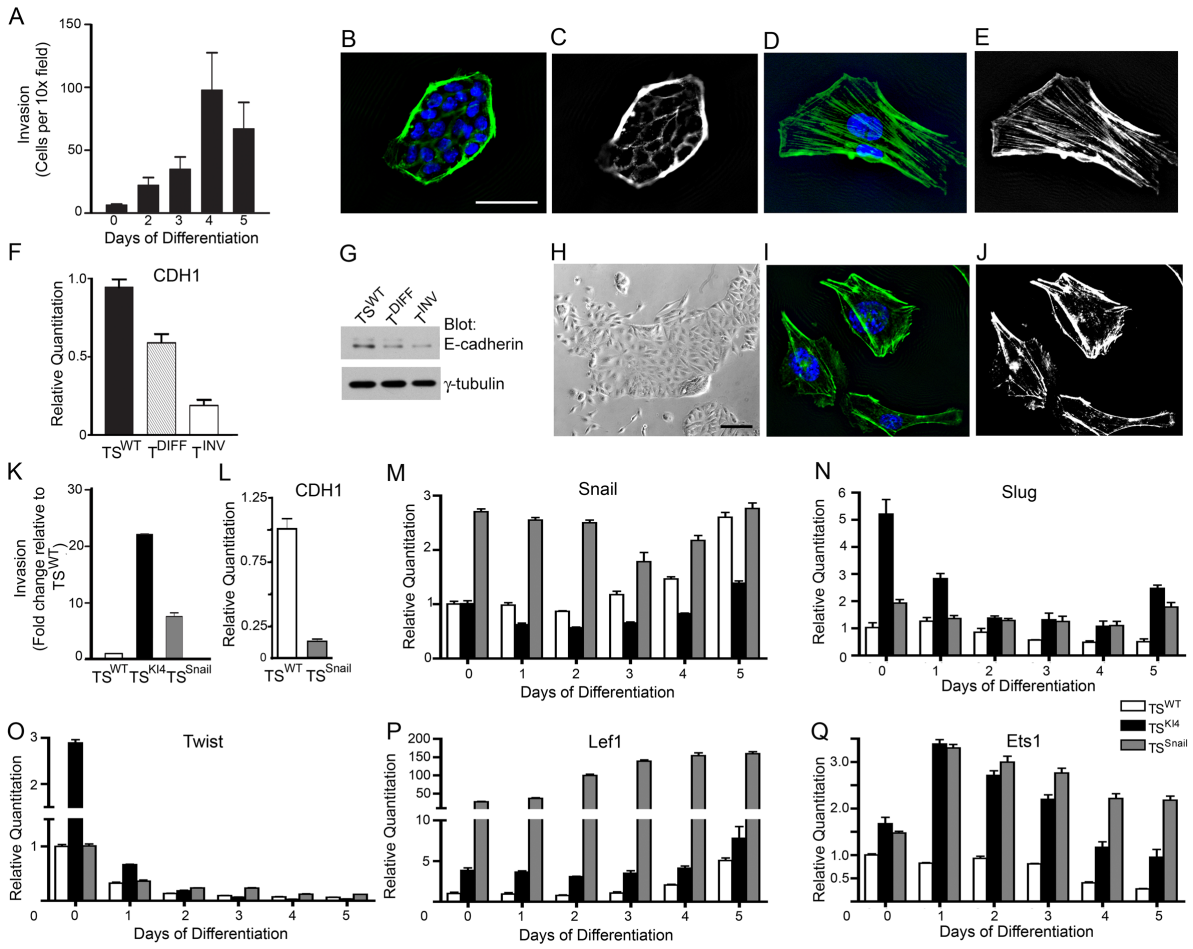



**Figure 3.3: Differentiation of TS$^{WT}$ Cells Induces EMT.**
(A) Differentiation of TS cells results in increased invasiveness. Invasion through Matrigel by undifferentiated TS cells (0) or TS cells differentiated by FGF4 withdrawal for the indicated number of days (2,3,4,5) is shown. Data are the mean ± range of two independent experiments performed in duplicate. Undifferentiated cells (B,C), or invasive cells isolated from the bottom of Matrigel-coated transwell chambers (D,E) plated on glass coverslips. (B,D) Cells were stained for actin (green) and nuclei (blue). (C,E) Actin staining alone shows peripheral cortical actin in undifferentiated cells (C) and filamentous actin in invasive trophoblasts (E). White bar represents 50 μm. (F) Reduced Cdh1 (E-cadherin) in invasive trophoblasts (T$^{INV}$) compared to undifferentiated TS cells (TS$^{WT}$) and TS cells differentiated for four days (T$^{DIFF}$). Data normalized to TS$^{WT}$ samples represent the mean ± SEM of three independent experiments. (G) Reduced E-cadherin protein in T$^{INV}$ cells. (H) Mesenchymal morphology of TS$^{Snail}$ cells. Black bar represents 100 μm. (I) Filamentous actin staining (green) and nuclear DAPI stain (blue) in TS$^{Snail}$ cells. (J) Filamentous actin staining in TS$^{Snail}$ cells. (K) Increased invasiveness of TS$^{KI4}$ and TS$^{Snail}$ cells relative to TS$^{WT}$ cells. (L) Reduced Cdh1 in TS$^{Snail}$ cells compared to TS$^{WT}$ cells. (M-Q) Expression changes of EMT-inducing genes in TS$^{WT}$, TS$^{KI4}$, and TS$^{Snail}$ cells. Data show the mean ± range of two independent experiments.

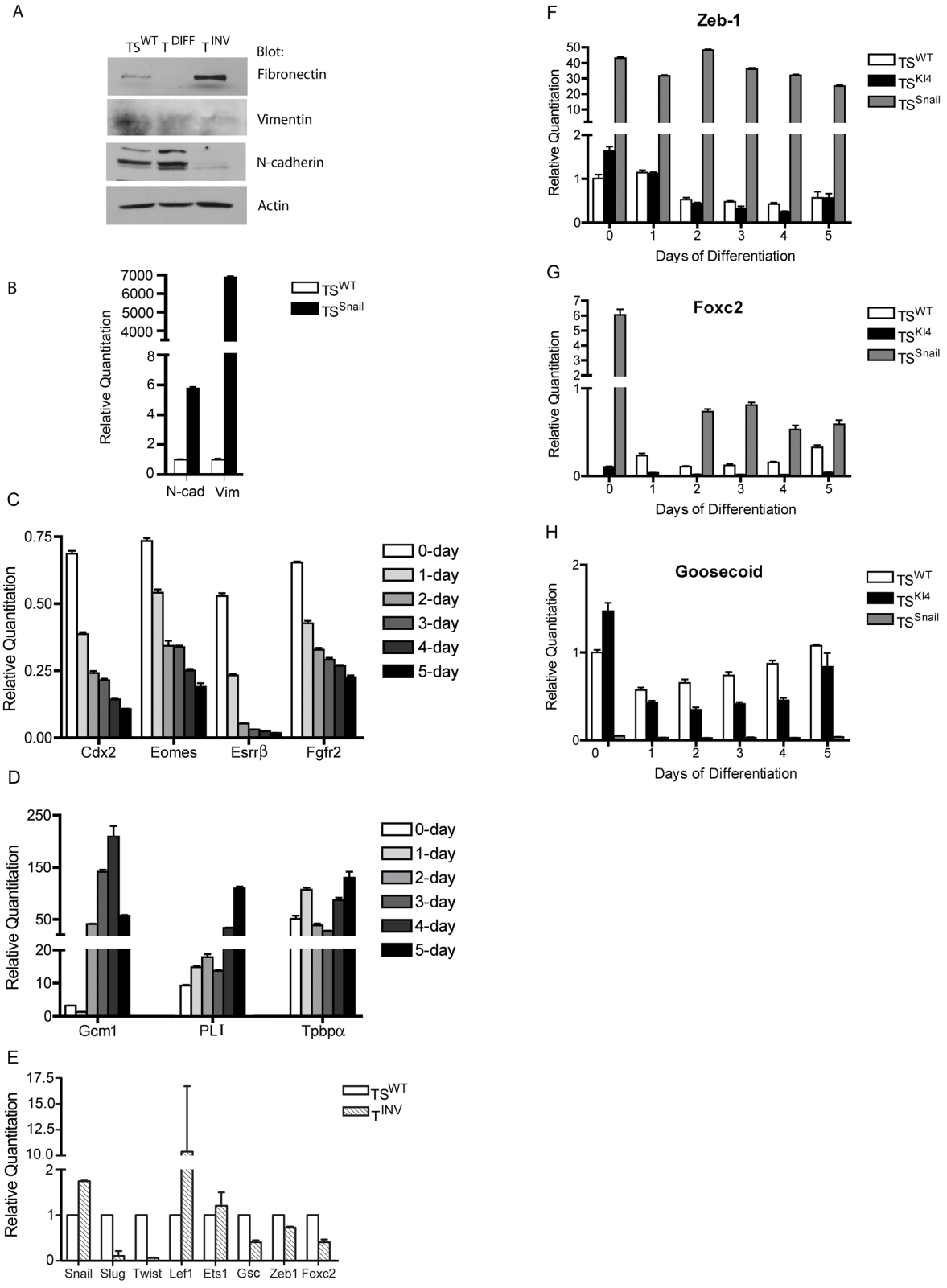



**Figure 3.4: Expression of Stem and Mesenchymal Cell Markers, and EMT-Inducing Transcription Factors in $T^{INV}$, $TS^{KI4}$, and $TS^{Snail}$ Cells**

(A) Increased expression of mesenchymal markers in $T^{INV}$ cells relative to $TS^{WT}$ cells as measured by Western blotting with the indicated antibodies. Data shows representative blots from two independent experiments. (B) Increased expression of mesenchymal markers in $TS^{Snail}$ cells relative to $TS^{WT}$ cells as measured by qRT-PCR. Data shows the mean ± range of two independent experiments performed in triplicate. (C) Expression of TS cell stemness markers in $TS^{Snail}$ cells relative to $TS^{WT}$ cells measured by qRT-PCR. Data shows the mean ± range of two independent experiments performed in triplicate. (D) Maintenance of multipotency in $TS^{Snail}$ cells demonstrated by normal induction of trophoblast differentiation markers 0 to 5 days after removal of FGF4. Changes in message relative to $TS^{WT}$ cells were measured by qRT-PCR. Data shows the mean ± range of two independent experiments performed in triplicate. (E) Expression of EMT-inducing transcription factors in $TS^{WT}$ and $T^{INV}$ cells measured by qRT-PCR. Data shows the mean ± range of two independent experiments. (F-H) Expression of indicated EMT-inducing transcription factors 0 to 5 days after differentiation of $TS^{WT}$, $TS^{KI4}$, and $TS^{Snail}$ cells measured by qRT-PCR. Data shown are the mean ± range of two independent experiments.

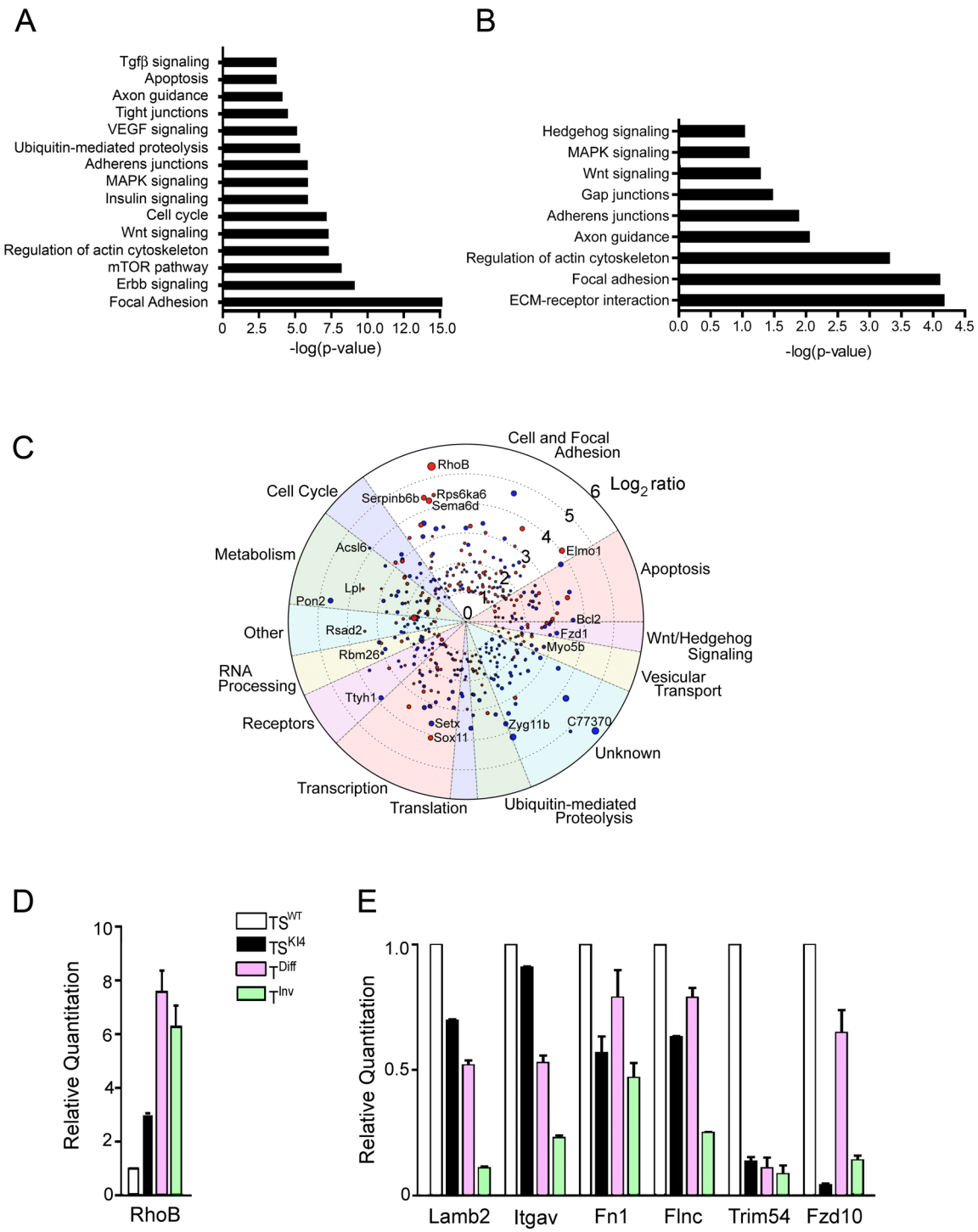



**Figure 3.5: Intersecting Gene Expression Changes in Focal Adhesion and Actin Cytoskeleton Pathways in $T^{INV}$ and $TS^{KI4}$ Cells**

(A) Enrichment for focal adhesion GO category in $T^{INV}$ cells. Canonical KEGG pathways are plotted according to p-values of pathway enrichment. (B) Enrichment for focal adhesion and actin cytoskeleton GO categories in $TS^{KI4}$ cells. The p-values

for canonical KEGG pathways are shown. (C) Diagram shows $T^{INV}$ cell $log_2$ ratios for the intersecting 416 genes with $TS^{KI4}$ cells categorized according to biological function. Red and blue indicate upregulated and downregulated genes, respectively. Circle diameter shows the relative difference in $log_2$ ratios in $T^{INV}$ cells versus $TS^{KI4}$ cells; abs($T^{INV}$-$TS^{KI4}$). (D,E) Validation of gene expression changes by qRT-PCR in $TS^{WT}$, $TS^{KI4}$, $T^{DIFF}$, and $T^{INV}$ cells. Data represent gene expression changes relative to $TS^{WT}$ cells expressed as the mean ± range for two independent experiments.

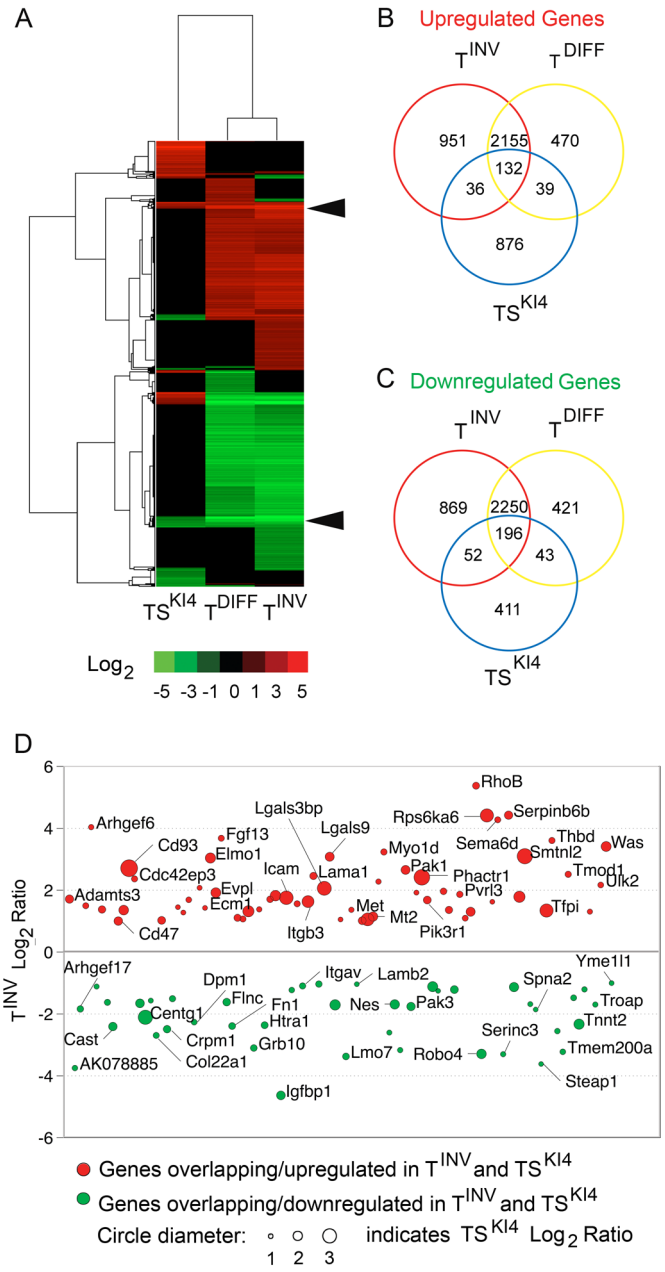



**Figure 3.6: Gene Expression Profiling of EMT in $TS^{KI4}$ Compared to $T^{INV}$ Cells**
(A) Loss of MAP3K4 activity in $TS^{KI4}$ cells results in significant changes in gene expression. Heat map compares gene expression changes in $TS^{KI4}$ cells (1785 genes), $T^{DIFF}$ cells (5706 genes), and $T^{INV}$ cells (6641 genes) relative to $TS^{WT}$ cells. Significantly changed genes have a Benjamini-Hochberg p-value < 0.05 and $\log_2$ ratio ≥ abs(1). Arrowheads indicate overlapping genes. (B,C) Venn diagrams show the overlap of upregulated (B) and downregulated (C) genes in $TS^{KI4}$ cells, $T^{DIFF}$ cells, and $T^{INV}$ cells. (D) 25% of the overlapping genes between $TS^{KI4}$ and $T^{INV}$ EMT models control cell adhesion and motility. Diagram depicts $T^{INV}$ cell $\log_2$ ratios from (A) for genes that regulate adhesion and motility having shared changes between $TS^{KI4}$ and $T^{INV}$ cells. Circle diameter depicts changes in $TS^{KI4}$ cell $\log_2$ ratios.

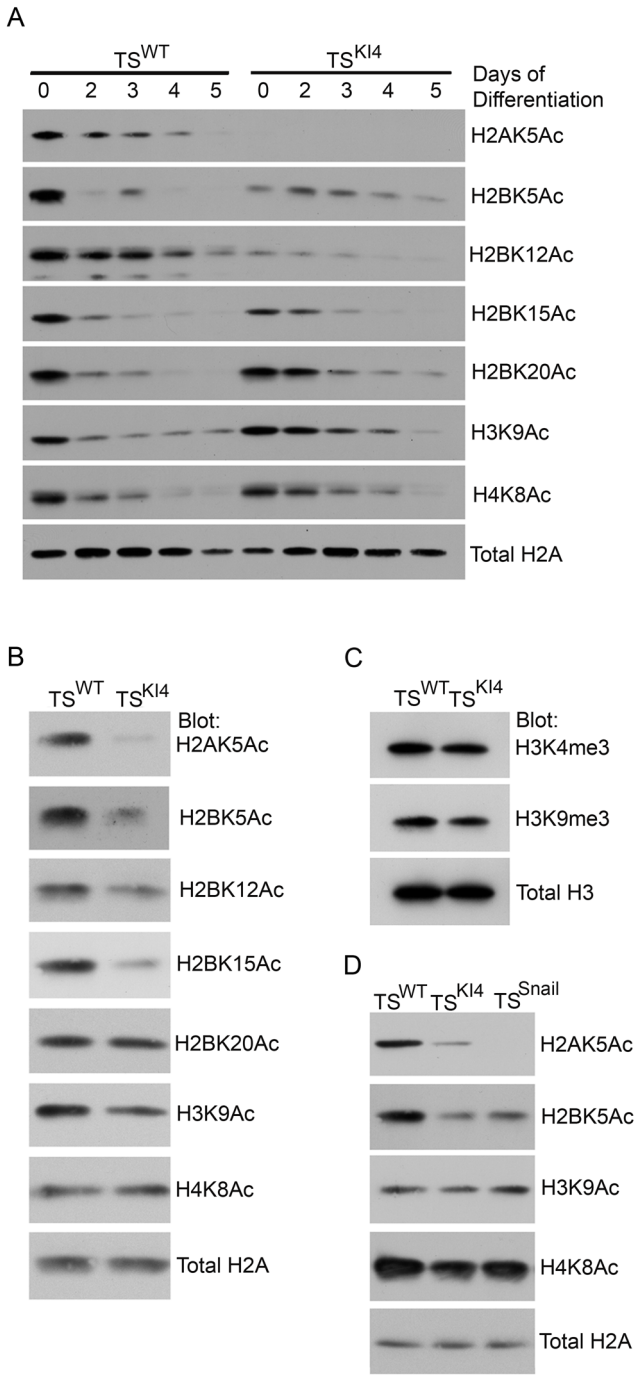



**Figure 3.7: Selective Loss of H2A and H2B Acetylation in Undifferentiated TS$^{KI4}$ Cells**

(A) Acetylation of histones H2A, H2B, H3 and H4 is decreased upon induction of TS cell differentiation. Western blotting of lysates from TS$^{WT}$ or TS$^{KI4}$ cells differentiated for the indicated days is shown. (B,C) TS$^{KI4}$ cells exhibit selective loss of acetylation on histones H2A and H2B and no change H3 methylation. (D) TS$^{Snail}$ cells exhibit selective loss of acetylation on histones H2A and H2B. (A-D) Results are representative of two to three independent experiments.

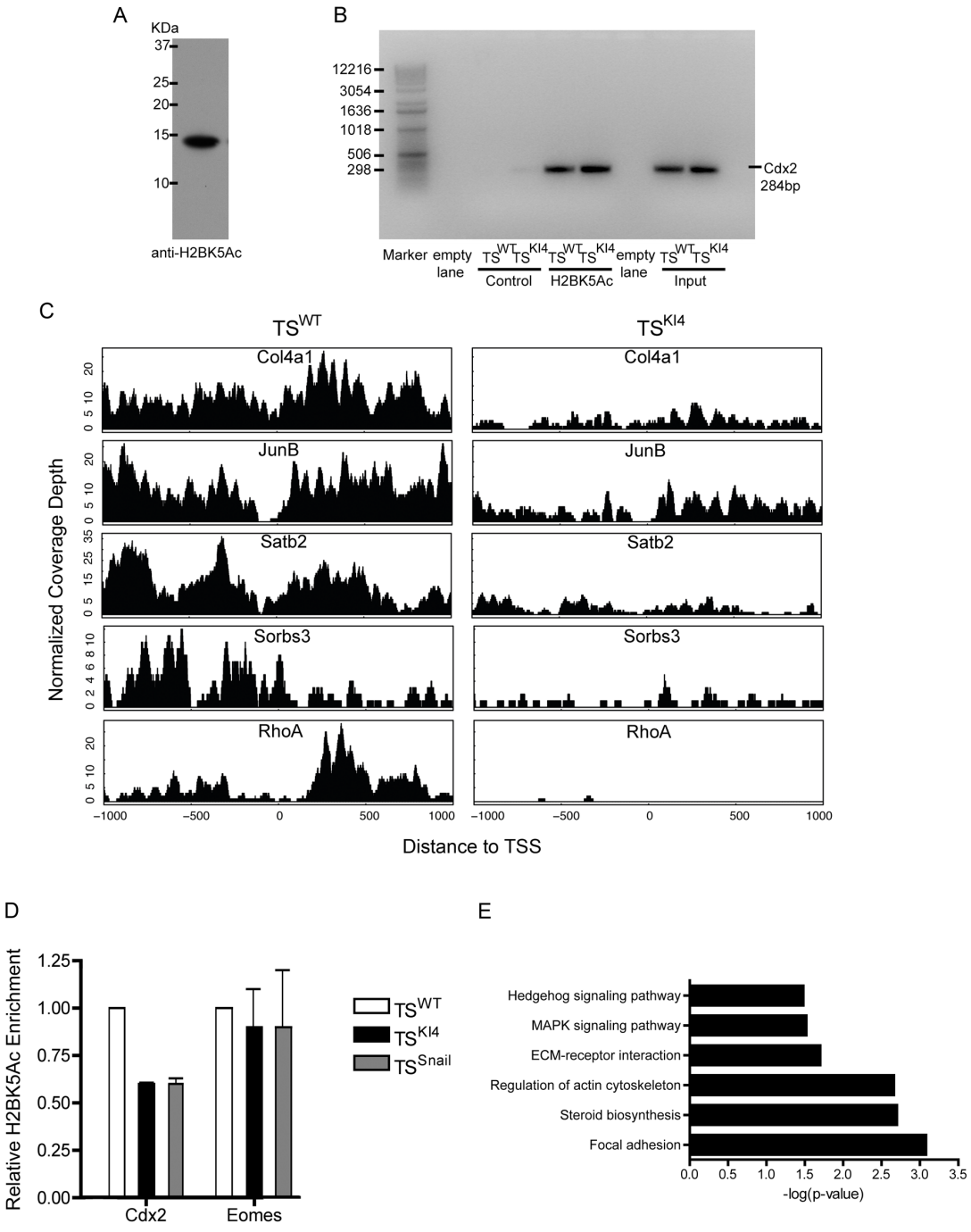



**Figure 3.8: Validation of H2BK5Ac ChIP-seq Analysis**
(A) Validation of the H2BK5Ac antibody used for ChIP-seq**.** Specificity of anti-acetylated H2BK5 antibody was measured by Western blotting of acid-extracted histone lysates from $TS^{WT}$ cells using anti-H2BK5Ac antibody. (B) Specificity of anti-acetylated H2BK5 antibody by ChIP. Semi-quantitative RT-PCR for the promoter of the TS cell marker Cdx2 from ChIP samples isolated from $TS^{WT}$ or $TS^{KI4}$ cells is shown. Results are representative of three independent experiments. (C) Promoter

acetylation density plots of additional genes with significant loss of H2BK5 acetylation. Comparison of normalized coverage depth of acetylation density within 1 kb of the transcription start site (TSS) between $TS^{WT}$ and $TS^{KI4}$ cells for indicated genes with p-values $\leq 1.4e^{-14}$. The x-axis for each subplot is the distance to the TSS, and the y-axis is the normalized coverage depth. (D) Validation of H2BK5Ac enrichment on promoters of TS cell stemness markers, Cdx2 and Eomes using H2BK5Ac ChIP-qRT-PCR. Data represent the mean $\pm$ range of two independent experiments performed in triplicate. (E) GO categories enriched for genes with altered H2BK5Ac in $TS^{KI4}$ cells. The p-values of canonical KEGG pathways are shown.

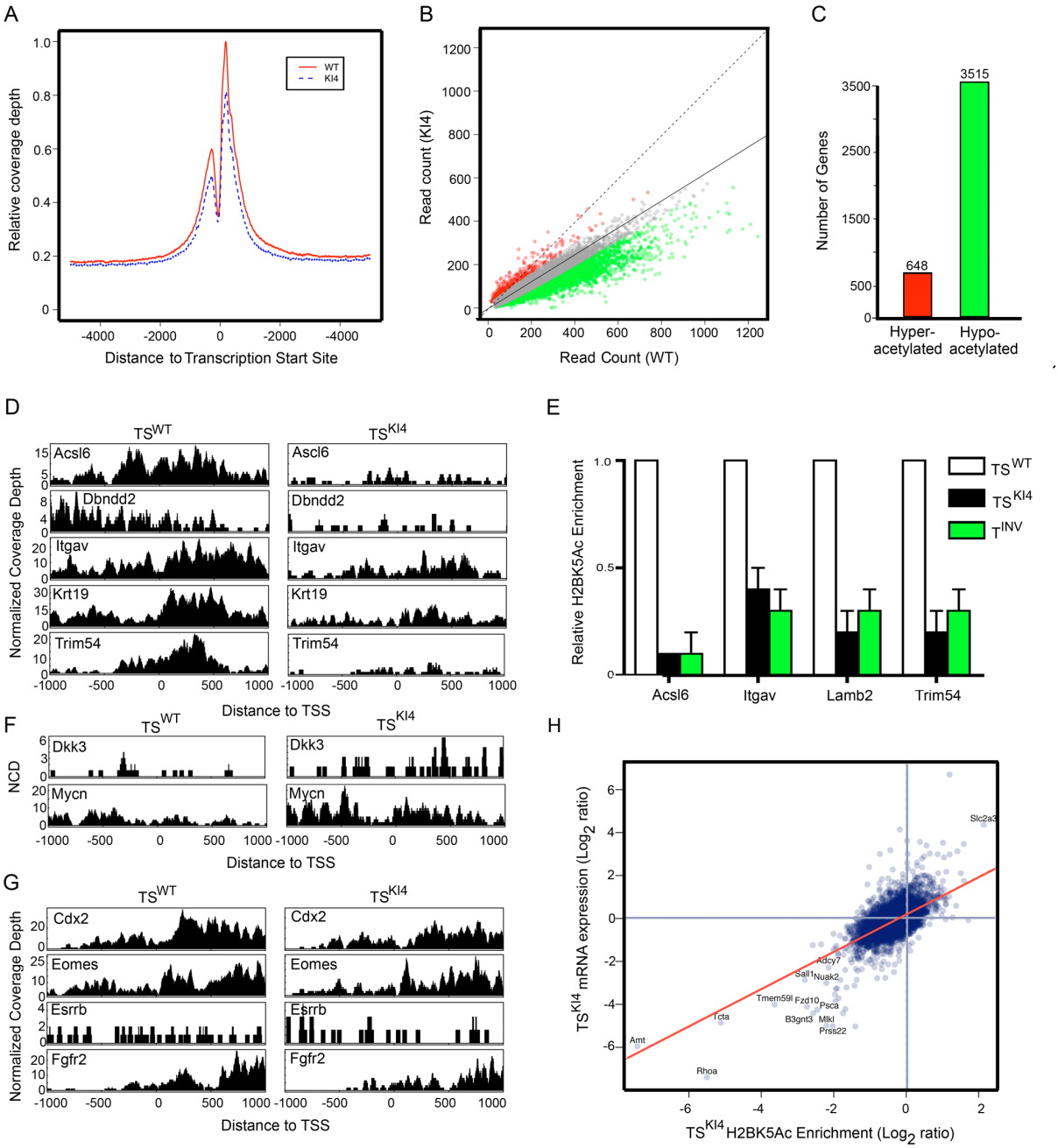



**Figure 3.9: Reduction of H2BK5Ac on Select Promoters in TS$^{KI4}$ Cells Contributes to Repression of the Epithelial Phenotype.**

(A) Peak H2BK5Ac density occurs within 1kb of the TSS. Plot depicts the distribution of read tags around the TSS of mouse RefSeq genes. Solid red line represents the read tag distribution density of all genes in TS$^{WT}$ cells, and dotted blue line represents TS$^{KI4}$ cells. Data was pooled from two independent experiments sequenced in duplicate. (B) Loss of H2BK5Ac occurs near the TSS of 3515 genes in TS$^{KI4}$ cells. Scatter plot of RefSeq genes with read tag counts above 20 reads per 2kb for genes identified from H2BK5Ac ChIP-seq of TS$^{WT}$ (x-axis) and TS$^{KI4}$ (y-axis) cells. Solid line represents the data normalization reference line for delineating

genes with significant read count differences between $TS^{WT}$ and $TS^{KI4}$ cells. Significance is based upon Benjamini-Hochberg FDR adjusted p-value $\leq$ 0.05. Green and red dots represent 3515 genes with a significant decrease and 648 genes with a significant increase in read tag density between $TS^{WT}$ and $TS^{KI4}$ cells, respectively. Gray dots represent all non-significantly changed genes. Dotted line is a $45^{o}$ reference line. (C) Bar graph summarizes all genes in (B) with significantly altered H2BK5Ac density in $TS^{KI4}$ cells. Data for (B and C) were pooled as described in (A). (D) Reduced H2BK5Ac near the TSS of specific genes in $TS^{KI4}$ cells. Plots of H2BK5Ac density compare the normalized coverage depth (NCD) of read tags around the TSS for indicated genes with loss of H2BK5Ac (p-values $\leq 1e^{-18}$) between $TS^{WT}$ and $TS^{KI4}$ cells. The x-axis for each subplot is the distance to TSS, and the y-axis is the NCD of read tags. (E) Validation of reduced H2BK5Ac on specific genes in both $TS^{KI4}$ and $T^{INV}$ cells shown by ChIP qRT-PCR of $TS^{WT}$, $TS^{KI4}$ and $T^{INV}$ cells. Data show the mean ± range of two independent experiments performed in duplicate. (F) Genes with increased H2BK5Ac near the TSS in $TS^{KI4}$ cells plotted as described in (D). (G) Plots of H2BK5Ac density comparing the NCD for stem cell markers between $TS^{WT}$ and $TS^{KI4}$ cells. (H) Changes in H2BK5Ac correlates with altered gene expression in $TS^{KI4}$ cells. Scatter plot of 6722 RefSeq genes with read tag counts above 50 reads per 1kb for genes identified by H2BK5Ac ChIP-seq (x-axis) and mRNA-seq (y-axis) in $TS^{KI4}$ cells. Read tag counts were converted to $\log_2$ ratio values for comparison. Solid red line represents the linear regression line of the data with a Pearson correlation coefficient of 0.62 (p-value < $10^{-16}$). Labeled dots are genes with a four-fold change in both mRNA abundance and H2BK5Ac enrichment.

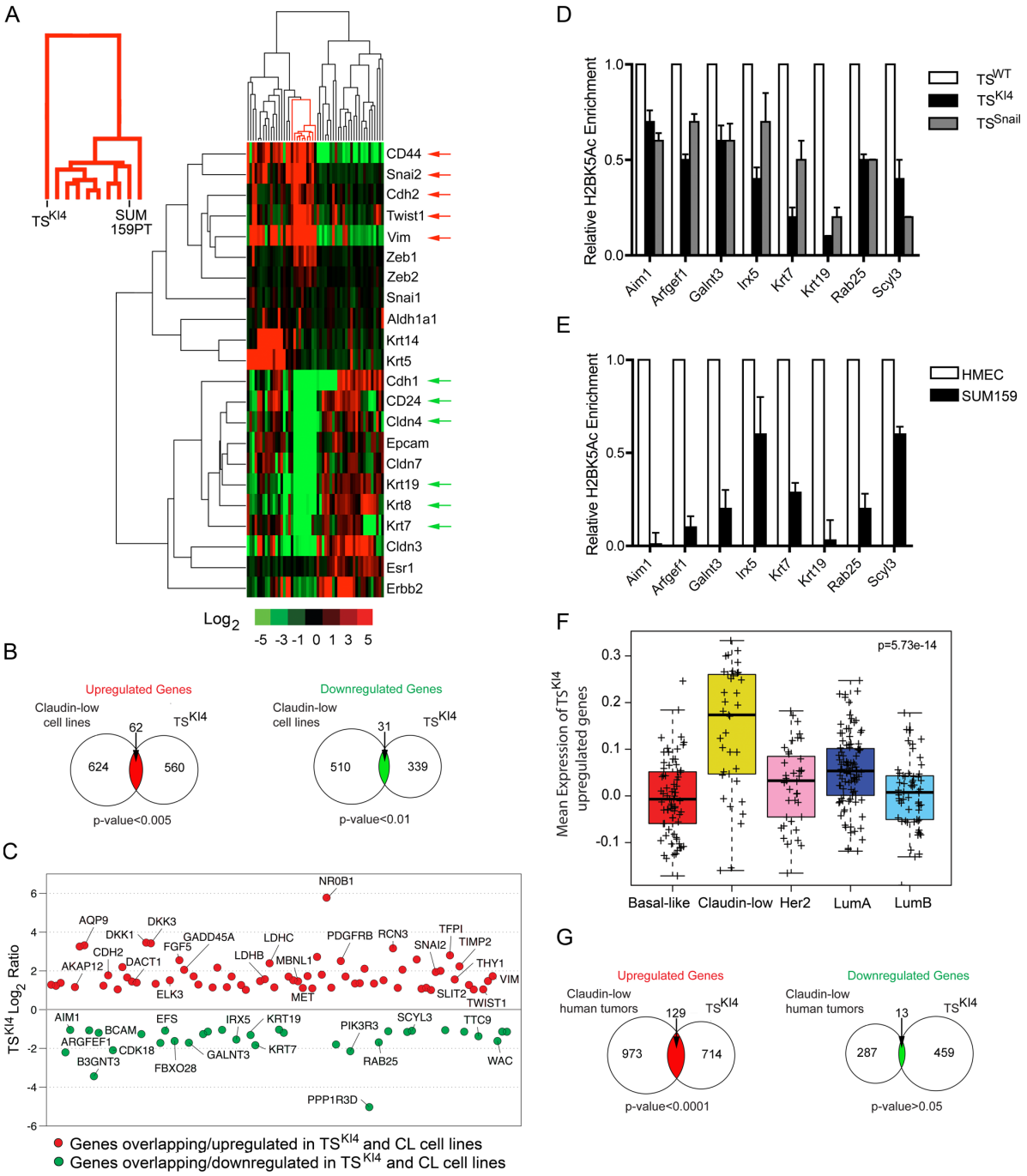



**Figure 3.10: TS$^{KI4}$ Cells and Claudin-low Breast Cancer With Properties of EMT and Stemness Show Loss of H2BK5Ac on Shared Genes.**
(A) TS$^{KI4}$ cells cluster with the CL subtype of breast cancer cells. Heat map compares expression of 22 breast cancer EMT genes in TS$^{KI4}$ cells and 52 breast cancer cell lines. TS$^{KI4}$ cell and CL breast cancer cluster is outlined in red. Red and green arrows indicate shared upregulated and downregulated genes respectively. (B) Venn diagram depicts the intersection between genes elevated and

downregulated in CL cell lines compared to $TS^{KI4}$ cells. (C) Plot shows $log_2$ ratios (y-axis) of $TS^{KI4}$ cells for intersecting genes depicted in (B). (D) Reduced H2BK5Ac on specific genes in both $TS^{KI4}$ and $TS^{Snail}$ cells from the intersecting downregulated genes in (C) measured by qRT-PCR of ChIP samples is shown. (E) Reduced H2BK5Ac on specific genes in CL breast cancer lines compared to normal HMEC breast cells as measured by ChIP-qRT-PCR. (D,E) Data are the mean ± range of two independent experiments performed in duplicate. (F) CL human tumors show the highest expression of $TS^{KI4}$ genes among breast cancer subtypes. Mean expression of $TS^{KI4}$ cell upregulated genes across the subtypes of breast cancer in the UNC337 data set. p-value was calculated by comparing gene expression means across all subtypes using an ANOVA test. Each + represents a distinct tumor sample within the data set. (G) Venn diagram of intersecting genes elevated and downregulated in CL human tumors and $TS^{KI4}$ cells.

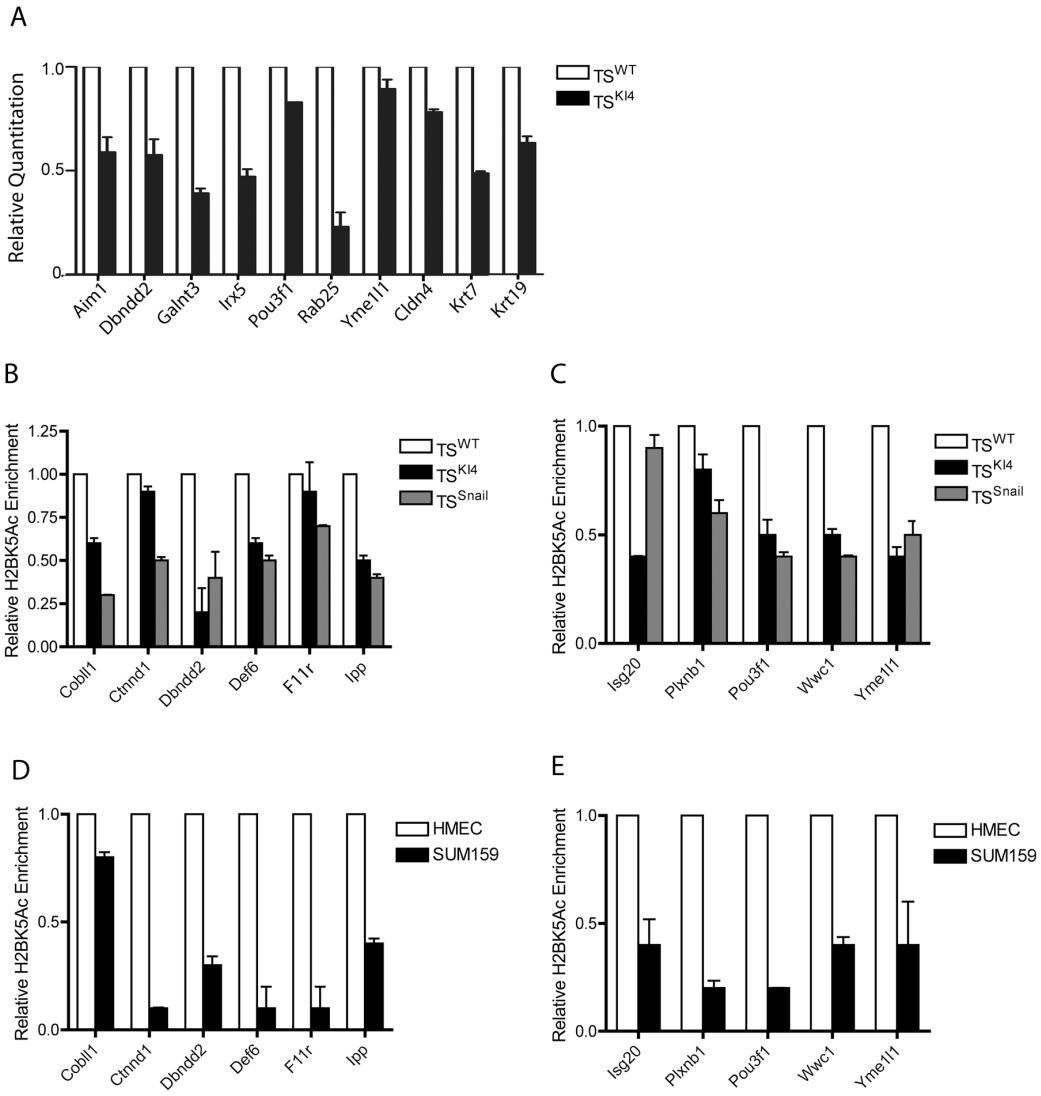



**Figure 3.11: Reduced Gene Expression and H2BK5 Acetylation in Intersecting Gene Sets between $TS^{KI4}$, $TS^{Snail}$, and CL Cells**

(A) Validation of reduced expression of $TS^{KI4}$/CL intersecting genes in $TS^{KI4}$ cells relative to $TS^{WT}$ cells measured by qRT-PCR. (B-E) Validation of loss of H2BK5Ac on promoters of $TS^{KI4}$/CL intersecting genes measured by H2BK5Ac ChIP coupled to qRT-PCR. Data shows loss of H2BK5Ac from shared set of genes between $TS^{KI4}$, $TS^{Snail}$, and CL cells. (A-E) Data represent the mean $\pm$ range of two independent experiments performed in triplicate.

| Gene | Skeletal | Neural Tube | Cranio-facial | Embryonic Lethality | Growth Retardation | Cardiac | Ref. |
|---|---|---|---|---|---|---|---|
| MAP3K4 | + | + | + | + | + | - | 1,2,4 |
| CBP | + | + | + | + | + | - | 11,12 |
| P300 | + | + | - | + | + | + | 15 |
| TRAF4 | + | + | - | + | - | - | 9 |
| GCN5 | + | + | - | + | + | - | 3,7,13,14 |
| Dvl1/2 | + | + | - | + | - | + | 5 |
| JNK1/2 | - | + | - | + | + | - | 6,10 |
| TWIST | - | + | + | + | + | - | 8 |
| P38IP | - | + | - | + | + | - | 17 |
| AXIN | + | + | - | + | + | + | 16 |
| PCAF | - | - | - | - | - | - | 14 |

**Table 3.1: Phenotypic Similarity between MAP3K4 (KI4) and CBP/p300 -/- Mice**
Yellow indicates phenotypic overlap of the MAP3K4 (KI4) mouse with the knockout mice for the genes listed. + indicates presence of the phenotype specified. - indicates absence of the phenotype specified. Table is based upon comprehensive literature review, as indicated by the references below.

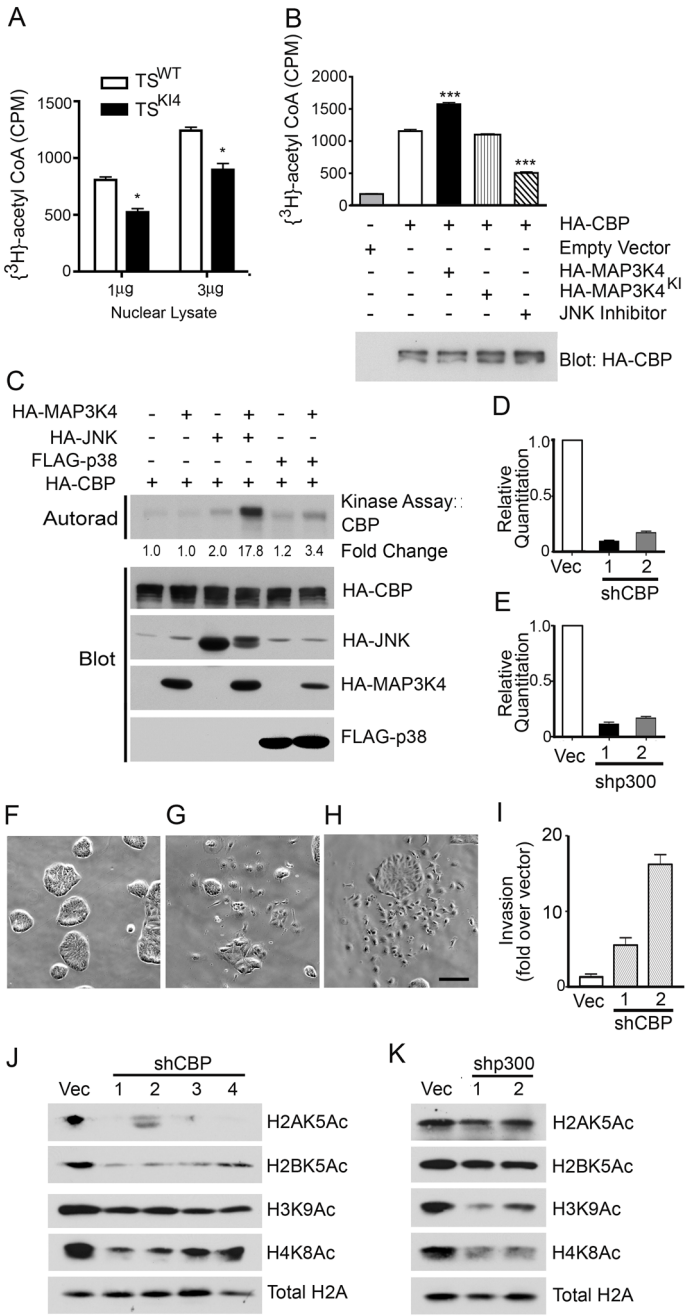



**Figure 3.12: CBP Expression Is Required for MAP3K4-dependent Regulation of the Epithelial Phenotype in TS Cells.**

(A) HAT activity is significantly decreased in $TS^{KI4}$ cells. Activity of one or three μg of nuclear lysate isolated from $TS^{WT}$ or $TS^{KI4}$ cells is measured. Significance of change in HAT activity was evaluated by an unpaired Student's T-test, *p-value < 0.05. (B) MAP3K4/JNK increases HAT activity of CBP measured as incorporation of [$^3$H]-acetyl-CoA in counts per minute (CPM) in 293 cells co-expressing CBP with the constitutively active kinase domain of MAP3K4 or kinase-inactive (KI). Significance of change in HAT activity was evaluated by an unpaired Student's T-test. ***p-value

< 0.001; *p-value < 0.05. Data represents the mean ± SEM of three independent experiments performed in triplicate. (C) MAP3K4 and JNK promote phosphorylation of CBP as measured by kinase assay. $\gamma^{32}$P-ATP labeled proteins were visualized by autoradiography and quantified by phosphorimaging. Data are representative of two independent experiments. (D,E) qRT-PCR shows reduced expression of CBP and p300 in $TS^{WT}$ cells infected with unique shRNAs for CBP (D) or p300 (E) compared to control virus-infected cells (Vec). Data represents the mean ± range of two independent experiments performed in triplicate. (F, G, and H) shRNA knockdown of CBP results in loss of the epithelial phenotype. Phase microscopy shows morphology of control virus-infected cells (F) compared to mesenchymal morphology of cells expressing CBP shRNAs (G,H). Black bar represents 100 μm. (I) Increased invasiveness of shCBP expressing cells compared to control virus-infected cells. Data shown represents the mean ± range of two independent experiments performed in triplicate. (J,K) Selective loss of H2A/H2BAc in CBP knockdown cells. Western blotting of lysates from $TS^{WT}$ cells infected with control virus, four independent CBP shRNAs (J) or two independent p300 shRNAs is shown. Results are representative of two independent experiments.

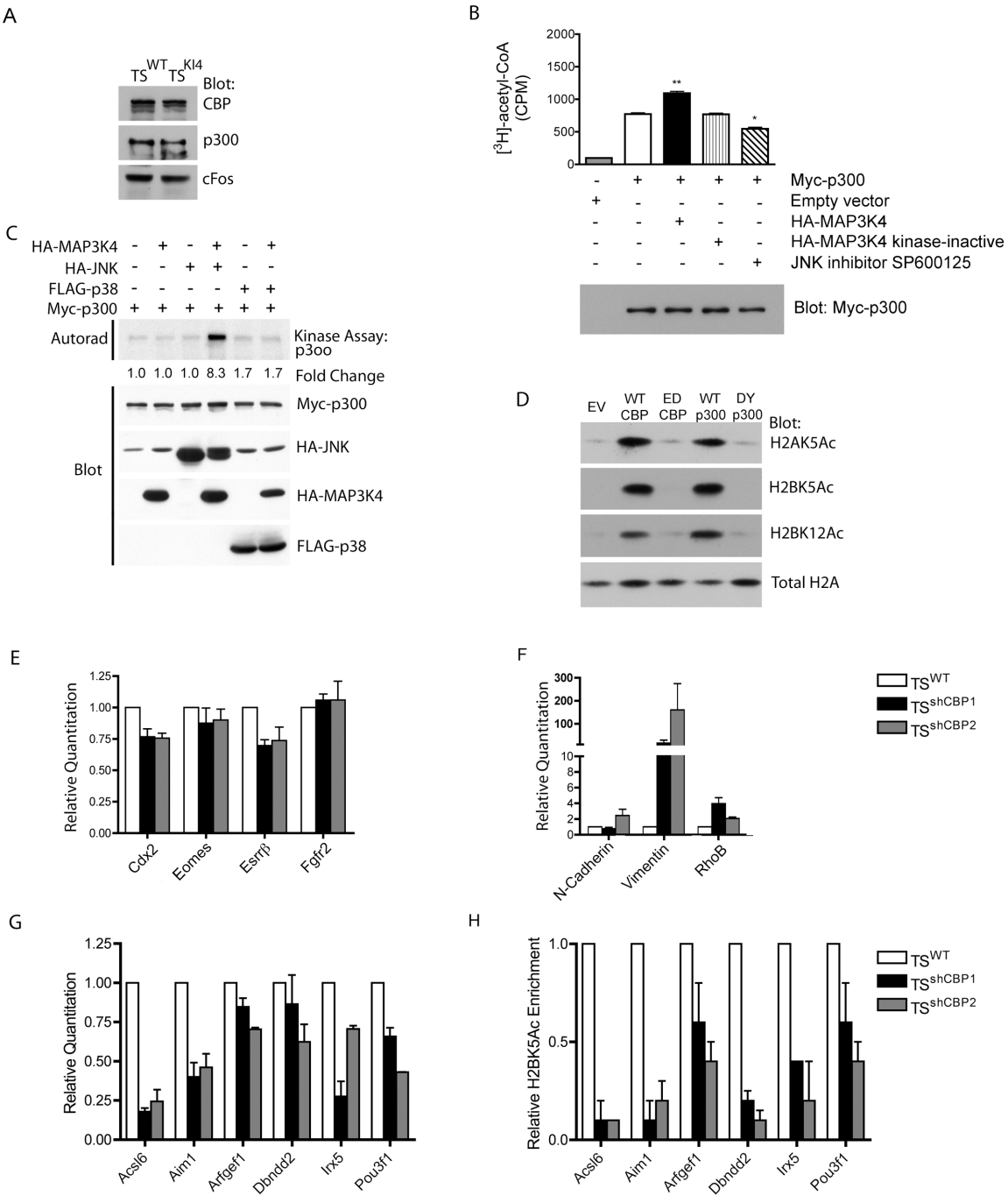



**Figure 3.13: TS$^{shCBP}$ Cells Display Properties of EMT and Stemness with Reduced Expression and H2BK5 Acetylation of TS$^{KI4}$/CL Intersecting Genes.**
(A) CBP and p300 protein expression levels are unchanged in TS$^{KI4}$ cells relative to TS$^{WT}$ cells. Western blots were performed on nuclear lysates from either TS$^{WT}$ or TS$^{KI4}$ cells using anti-CBP or anti-p300 antibodies. Results are representative of two independent experiments. (B) MAP3K4/JNK increases HAT activity of p300 measured as incorporation of [$^3$H]-acetyl-CoA in counts per minute (CPM) in 293 cells co-expressing p300 with the constitutively active kinase domain of MAP3K4 or

kinase-inactive. Significance of change in HAT activity was evaluated by an unpaired Student's T-test. **p-value < 0.01; *p-value < 0.05. Data represents the mean ± SEM of three independent experiments performed in triplicate. (C) MAP3K4 and JNK promote phosphorylation of p300 as measured by kinase assay. $\gamma^{32}$P-ATP labeled proteins were visualized by autoradiography and quantified by phosphorimaging. Protein expression levels were monitored by Western blotting with the indicated antibodies. Data are representative of two independent experiments. (D) CBP and p300 HAT activity is required to selectively acetylate histones H2A and H2B as HAT deficient CBP and p300 mutants are unable to acetylate H2A and H2B. Western blots were performed using acid-extracted histone cell lysates from 293 cells expressing wild-type (WT) or HAT-inactive CBP (ED) and p300 (DY) and probed with the indicated anti-acetylated H2A and H2B antibodies. Results are representative of two independent experiments. (E) Expression of TS cell stemness genes in $TS^{shCBP}$ cells compared to control virus (Vec) infected cells measured by qRT-PCR. (F) Increased expression of mesenchymal markers in $TS^{shCBP}$ cells relative to control virus (Vec) infected cells as measured by qRT-PCR. (G) Reduced expression in $TS^{shCBP}$ cells of genes common to $TS^{KI4}$, $TS^{Snail}$ and CL cells as measured by qRT-PCR. (H) Reduced H2BK5Ac on the promoters of the same genes in G shown by H2BK5Ac ChIP-qRT-PCR. (E-H) Data shows the mean ± range of two independent experiments performed in triplicate.

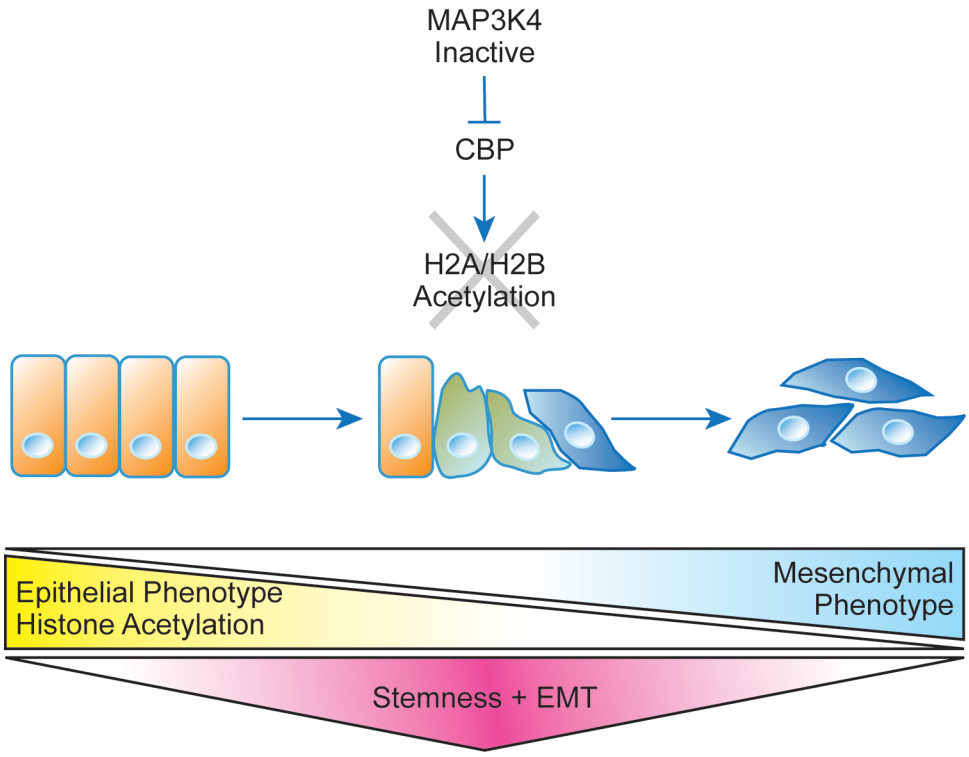



**Figure 3.14: Induction of EMT in TS cells Is Epigenetically Regulated by Loss of H2A/H2B Acetylation.**
Graphical summary demonstrating that inactivation of MAP3K4/CBP results in loss of H2A/H2B acetylation and the induction of the intermediate EMT state characterized by the simultaneous existence stemness and EMT properties.

# Chapter IV

## Chromatin Remodeler Smarcd3/Baf60c Induces EMT by Epigenetic Regulation of Non-Canonical WNT Signaling

### Introduction

Epithelial-mesenchymal transition (EMT) is a reversible developmental process, whereby epithelial cells lose cell-cell adhesion and apical-basolateral polarity, while acquiring a mesenchymal front-back end polarity and increased cellular invasiveness (Thiery & Sleeman, 2006). During metastatic tumor progression, this fundamental developmental program becomes reactivated to confer tumor cells with enhanced migratory and invasive properties, thereby promoting intravasation and dissemination of malignant cancer cells into the bloodstream. The opposing process of mesenchymal-epithelial transition (MET) reestablishes the epithelial state, as demonstrated by formation of epithelial tissues and organs in development and by metastatic colonization of epithelial tumors to distant organs (Yang & Weinberg, 2008; Valastyan & Weinberg, 2011; Thiery *et al*, 2009). By understanding the underlying molecular mechanisms contributing to EMT and the reverse process MET, new strategies may be developed for the prevention and treatment of metastatic cancer.

Studies based on global gene expression profiling have identified five breast cancer subtypes (Luminal A, Luminal B, HER2-enriched, Basal-like and Claudin-low)

(Perou *et al*, 2000; Sørlie *et al*, 2001; Prat & Perou, 2011). Among them, the Basal-like and Claudin-low subtypes largely consist of poor prognostic tumors with an estrogen receptor (ER)-negative, progesterone receptor (PR)-negative and HER2-non-overexpressing phenotype (i.e. triple-negative). These breast tumors demonstrate a high propensity to metastasize to the lung and brain parenchyma (Fulford *et al*, 2007; Gaedcke *et al*, 2007; Banerjee, 2006). At the molecular level, Basal-like and Claudin-low tumors share similar gene expression patterns such as low expression of Luminal and HER2-enriched genes. The Basal-like subtype is characterized by high expression of proliferation- and basal/myoepithelial cell-specific genes, including cytokeratins 5, 14 and 17, while the Claudin-low subtype is characterized by high expression of mesenchymal- and stem cell-related genes and low expression of epithelial features such as claudins 3, 4 and 7 (Prat *et al*, 2010; Hennessy *et al*, 2009; Rakha *et al*, 2008b). These gene expression profiles are indicative of differences in epithelial and mesenchymal cellular states between Basal-like and Claudin-low subtypes.

We previously identified a statistically significant EMT gene signature shared by epithelial stem cells developmentally entering EMT and the Claudin-low subtype of triple negative breast cancer (Abell *et al*, 2011). Based on this EMT gene signature across development and breast cancer models, we hypothesized that the upregulated genes within the signature represent genes critical to the induction of EMT and maintenance of the mesenchymal cell phenotype. To test this hypothesis, we designed an RNAi screen to identify individual genes within the EMT signature responsible for maintaining the mesenchymal phenotype. We used SUM149 and

SUM229 Basal-like breast cancer cells that exist in culture as dual populations of EpCAM$^{(-)}$ and EpCAM$^{+/high}$ cells (Prat *et al*, 2010; Gupta *et al*, 2011; Proia *et al*, 2011). The two populations have a common genetic background but different phenotypes that are distinguishable based upon EpCAM expression.

Herein, we describe an RNAi screening approach that identified a novel set of genes important for establishing the mesenchymal phenotype of EpCAM$^{(-)}$ SUM149 and SUM229 cells. Using this RNAi screening strategy, the SWI/SNF chromatin-remodeling factor Smarcd3/Baf60c was identified a novel regulator of EMT. While siRNA-mediated knockdown of Smarcd3/Baf60c induced MET in mesenchymal EpCAM$^{(-)}$ SUM149 and SUM229 cells, expression of Smarcd3/Baf60c in human mammary epithelial cells (HMECs) promoted the reverse process of EMT. This EMT program was activated by Smarcd3/Baf60c-mediated epigenetic induction of Wnt5a and non-canonical WNT signaling. These results define a new epigenetic mechanism of EMT regulation through the Smarcd3/Baf60c-mediated induction of non-canonical WNT signaling.

## Results

### *SUM149 and SUM229 Breast Cancer Cells Maintain Epithelial and Mesenchymal Populations*

SUM149 and SUM229 breast cancer cells maintain a consistent distribution of two distinct populations of EpCAM$^{(-)}$ and EpCAM$^{+/high}$ cells that profile by gene expression analysis as Claudin-low and Basal-like breast cancer subtypes, respectively (Prat *et al*, 2010). Lander and colleagues modeled the behavior of these

populations as a measure of phenotypic equilibrium in SUM149 breast cancer cells and concluded that EpCAM$^{(-)}$ cells give rise to EpCAM$^{+/high}$ cells (Gupta *et al*, 2011). Consistent with this conclusion, following separation of the two populations by fluorescence activated cell sorting (FACS), the EpCAM$^{(-)}$ population of SUM149 cells gives rise to EpCAM$^{+/high}$ cells. After 10 days in culture, the EpCAM$^{(-)}$ cells consist of two distinct populations containing 5-10% EpCAM$^{+/high}$ cells, while the EpCAM$^{+/high}$ SUM149 cells remained in the EpCAM$^{+/high}$ state (**Figure 4.1**A). Similarly, the EpCAM$^{(-)}$ population of SUM229 cells gives rise to EpCAM$^{+/high}$ cells with a frequency of 1-3 percent over 10 days, while the EpCAM$^{+/high}$ population remained in the EpCAM$^{+/high}$ state (**Figure 4.1**A). To rule out the possibility of intrinsic contamination following FACS with EpCAM and Cd49f antigens, the EpCAM$^{(-)}$ SUM149 cells were resorted 10 days following the initial sort of the parental population. After repeating FACS with the EpCAM$^{(-)}$ SUM149 cells, the EpCAM$^{(-)}$ population gave rise to EpCAM$^{+/high}$ cells with a similar frequency of 5-10 percent over 10 days (**Figure 4.1**B). Based upon the purity of the sorted populations and a doubling time of approximately 24 hours for SUM149 and SUM229 cells, the number of EpCAM$^{+/high}$ cells within the sorted EpCAM$^{(-)}$ population is consistent with EpCAM$^{(-)}$ cells giving rise to EpCAM$^{+/high}$ cells.

Further characterization revealed the mesenchymal versus epithelial characteristics of EpCAM$^{(-)}$ and EpCAM$^{+/high}$ cells, respectively. EpCAM$^{(-)}$ cells demonstrated a spindle-like mesenchymal morphology contrasting with the rounded, cobblestone epithelial morphology of the EpCAM$^{+/high}$ cells (**Figure 4.1**C). Immunofluorescence and Western blotting confirmed the absence of the epithelial

markers EpCAM and E-cadherin in EpCAM$^{(-)}$ SUM149 and SUM229 cells (**Figure 4.1**C-E). EpCAM$^{(-)}$ but not EpCAM$^{+/high}$ cells expressed the mesenchymal markers fibronectin, N-cadherin and vimentin and had filamentous actin stress fibers consistent with a front-back end polarized mesenchymal phenotype (**Figures 4.1**D-E and **4.2**A). When compared to the EpCAM$^{+/high}$ SUM149 and SUM229 cells, the EpCAM$^{(-)}$ cells had decreased mRNA expression for the epithelial differentiation markers E-cadherin and EpCAM, as measured by quantitative real-time PCR (qRT-PCR) (**Figure 4.2**B and C) and increased expression of the mesenchymal markers N-cadherin, vimentin, TGFβ, and the EMT-inducing transcription factors Snail, Slug, and Twist (**Figure 4.2**B-E). Functionally, the EpCAM$^{(-)}$ SUM149 and SUM229 cells were 3- and 8-fold more invasive through growth factor-reduced Matrigel than the EpCAM$^{+/high}$ cells, respectively (**Figure 4.1**F). Collectively, these results demonstrate that the epithelial EpCAM$^{+/high}$ and mesenchymal EpCAM$^{(-)}$ populations of SUM149 and SUM229 cells are morphologically and functionally distinct. The finding that the mesenchymal EpCAM$^{(-)}$ SUM149 and SUM229 cells give rise to epithelial EpCAM$^{+/high}$ cells was used to develop an RNAi screen to characterize genes within our EMT signature that are sufficient to induce EMT and MET.

***RNAi Screen Identifies Genes Important for EMT***

EpCAM$^{(-)}$ SUM149 and SUM229 cells were used in RNAi screens to identify genes important to EMT and maintenance of the mesenchymal phenotype (**Figure 4.3**A). Four independent siRNA oligonucleotides were used for targeting upregulated genes within the EMT gene signature shared by epithelial stem cells developmentally entering EMT and Claudin-low breast cancer (**Table 4.1**) (Abell *et*

*al*, 2011). With the exception of Slug, Twist, vimentin and N-cadherin, these genes are uncharacterized with respect to EMT (**Table 4.1**).

As proof-of-principle for the RNAi screening strategy, we silenced known EMT-inducing genes Snail and Slug and measured the increase in epithelial markers EpCAM and E-cadherin by immunofluorescence (**Figure 4.4**A and B). These experiments demonstrated that siRNA-mediated knockdown of Snail or Slug in the mesenchymal EpCAM$^{(-)}$ SUM149 cells resulted in increased EpCAM and E-cadherin expression, loss of the mesenchymal phenotype and gain of the epithelial phenotype within six days after transfection (**Figure 4.4**A-C). At six days post transfection GAPDH siRNA-treated EpCAM$^{(-)}$ cells had a background of less than 3 percent EpCAM or E-cadherin positive cells (**Figure 4.4**A and B). In contrast, dual knockdown of Snail and Slug resulted in 26 and 10 percent EpCAM or E-cadherin positive cells, respectively (**Figure 4.4**A-C). A highly significant 3-5-fold dynamic range of quantitative EpCAM and E-cadherin percent fluorescence intensity was observed when comparing individual wells of a 96-well plate with siRNA-mediated knockdown of Slug or Snail versus GAPDH (**Figure 4.4**A-C). Furthermore, FACS purity checks before each screen showed initially less than 0.5% contamination of the EpCAM$^{+/high}$ cells within the EpCAM$^{(-)}$ population, suggesting that the observed increases in EpCAM and E-cadherin were indicative of a transition from mesenchymal EpCAM$^{(-)}$ to epithelial EpCAM$^{+/high}$ cell state rather than outgrowth of a contaminating fraction (**Figures 4.3**B and **4.5**A-B). Validation of the image analysis method was achieved by titrating set percentages of EpCAM$^{+/high}$ SUM149 cells into an EpCAM$^{(-)}$ background. The percent fluorescence of EpCAM and E-cadherin

measured by high-content image analysis closely recapitulated the titrated percentages (**Figure 4.5**C). Therefore, this robust RNAi screening strategy was used for identification of genes whose knockdown induces a mesenchymal to epithelial transition (MET) in EpCAM$^{(-)}$ SUM149 and SUM229 cells.

The RNAi screen targeted 140 upregulated genes within the EMT gene signature and identified ten genes whose inhibition resulted in the morphological transition between mesenchymal EpCAM$^{(-)}$ and epithelial EpCAM$^{+/high}$ phenotypes in both SUM149 and SUM229 cells (**Figures 4.3**C, **4.6**A-D, **4.7**A-C and **Table 4.1**). Compared to GAPDH, siRNA-mediated knockdown of these ten EMT-regulatory genes demonstrated a statistically significant increase in the percent EpCAM and E-cadherin cellular fluorescence (**Figure 4.6**A and B). The ten identified EMT-regulatory genes included: Slug, Cdh2, Smarcd3/Baf60c, ABR, Rnf130, EphA4, Met, Ptprb, Fhl1 and RRAGD (**Figure 4.6**A-D; see **Table 4.2** for gene functions). In addition to upregulation of EpCAM and E-cadherin protein expression, siRNA-mediated knockdown of these EMT-regulatory genes induced a morphological change characterized by the acquisition of an epithelial, cobblestone morphology (**Figure 4.6**C and D).

Of the ten EMT-regulatory genes identified in the phenotypic RNAi screen, Slug, Cdh2, Smarcd3/Baf60c, Fhl1, Rnf130 and RRAGD were upregulated in the mesenchymal EpCAM$^{(-)}$ cells relative to the epithelial EpCAM$^{+/high}$ SUM149 and SUM229 cells (**Figure 4.8**A). Deconvolution of siRNA oligonucleotide pools showed efficient knockdown of the EMT-regulatory target genes by at least two individual oligonucleotides, suggesting that increases in EpCAM and E-cadherin expression

were the result of on-target effects (**Figure 4.8**B). Functionally, knockdown of ABR, Cdh2, Smarcd3/Baf60c or Slug expression resulted in significant decreases in cellular invasiveness through growth factor-reduced Matrigel (**Figure 4.8**C and **Table 4.2**). Furthermore, RNAi against each of the ten EMT-regulatory genes induced a decrease in mRNA expression of the EMT-inducing transcription factors Snail and Slug (**Figure 4.8**D and E and **Table 4.2**). Based upon characterization of traditional EMT markers and cellular invasiveness, this RNAi-screening strategy effectively identified genes whose individual loss was sufficient to inhibit the mesenchymal phenotype of $EpCAM^{(-)}$ SUM149 and SUM229 breast cancer cells.

***Smarcd3/Baf60c is a Novel Epigenetic Regulator of EMT***

Smarcd3/Baf60c has the potential to affect a broad network of genes through its epigenetic-regulatory function as a member of the SWI/SNF chromatin-remodeling complex (Ho & Crabtree, 2010; Lickert *et al*, 2004; Takeuchi *et al*, 2007; Forcales *et al*, 2011). FLAG-tagged Smarcd3/Baf60c was stably expressed in human mammary epithelial cells (HMECs) to determine its role in the regulation of EMT. HMECs expressing Smarcd3/Baf60c (D3-HMECs) demonstrated phenotypic changes indicative of EMT (**Figure 4.9**A). D3-HMECs lost their epithelial cell morphology and became mesenchymal-like, losing epithelial cell-cell adhesion markers EpCAM and E-cadherin and gaining expression of the mesenchymal marker vimentin (**Figure 4.9**A). Western blotting confirmed the loss of expression of the epithelial markers EpCAM and E-cadherin with gain of the mesenchymal markers fibronectin and vimentin (**Figure 4.9**B). Of the EMT-inducing transcription factors, Lef1 and Zeb2 gene expression levels were increased compared to empty

vector control HMECs (EV-HMECs) (**Figure 4.9**C). Furthermore, the epithelial cell-cell adhesion genes E-cadherin, EpCAM, Krt7, Krt19, Cldn4 and Cldn7 were downregulated (**Figure 4.9**C). Notably, mRNA expression of the tight junction genes Cldn4 and Cldn7 was reduced 90 percent compared to EV-HMECs (**Figure 4.9**C). Similar to EV-HMECs, D3-HMECs when separated by FACS, profiled as a single cellular population according to the antigenic expression levels of epithelial differentiation markers EpCAM/Cd49f and breast cancer stem cell markers Cd44/Cd24. However, D3-HMECs demonstrated a preferential reduction in the epithelial integrin Cd49f and gain in the pro-metastatic glycoprotein Cd44 with no change in the pro-differentiation glycoprotein Cd24 when compared to EV-HMECs (**Figure 4.9**D). These results support the role of Smarcd3/Baf60c as an EMT-inducing chromatin- remodeling factor that promotes the mesenchymal phenotype with loss of the epithelial phenotype.

***Smarcd3/Baf60c Induces a Claudin-low Gene Signature***

Since Smarcd3 is a SWI/SNF chromatin-remodeling factor capable of epigenetically regulating the expression of a gene network, we performed gene arrays to measure expression changes in D3-HMECS relative to EV-HMECs. Additionally, we compared gene expression changes of D3-HMECS with Snail-, and Slug expressing HMECs (Snail-HMECs and Slug-HMECs) to define gene networks important for EMT. Hierarchical cluster analysis of gene array data from D3-, Slug-, and Snail-HMECs with respect to the 9-cell line Claudin-low predictor described by Prat et al., revealed that the gene expression profiles of D3-HMECs clustered similarly to the mesenchymal Claudin-low signature of human breast cancer (**Figure**

**4.10**A and B) (Prat *et al*, 2010). Conversely, the gene expression profiles of Slug- and Snail-HMECs demonstrated substantially less overlap (**Figure 4.10**A and B). Of the D3-, Slug-, and Snail-HMECs, D3-HMECs had the most significant gene intersection (p-value < 0.0001) with the Claudin-low signature sharing 7% upregulated and 13% downregulated genes (**Figure 4.10**B). D3- and Slug-HMECs also exhibited a significant gene intersection (p-value < 0.0001) with 41% upregulated and 18% downregulated shared genes, indicating that Smarcd3/Baf60c and Slug affect an overlapping EMT gene network (**Figure 4.11**A). To further emphasize the significance of the D3-HMEC/Claudin-low association, the gene expression profile of D3-HMECs was compared to the five intrinsic molecular subtypes of human breast tumors catalogued in the UNC337 dataset. Tumors from the Claudin-low subtype showed highest expression of the D3-HMEC signature in comparison with Basal-like, HER2-enriched, Luminal A and Luminal B tumors (**Figure 4.10**C). The differentiation predictor described by Prat et al, which uses gene expression profiles to determine the cellular relatedness to the normal breast epithelial differentiation hierarchy from mammary stem cells (MaSCs) to mature luminal cells, showed that D3-HMECs had the lowest differentiation score compared to Slug- and Snail-HMECs (**Figure 4.10**D) (Prat *et al*, 2010; Lim *et al*, 2009). This finding reveals that D3-HMECs are more similar to MaSCs and more importantly to a less differentiated mesenchymal, Claudin-low phenotype than Slug- and Snail-HMECs. Collectively, these results demonstrate that Smarcd3/Baf60c induces an EMT gene expression program in HMECs similar to that of the Claudin-low breast cancer subtype.

Molecular characterization of the Smarcd3/Baf60c-induced EMT in HMECs showed a more profound phenotypic conversion than that of the Snail- or Slug-induced EMT in HMECs. Morphological changes included a loss of cell-cell adhesions in D3-HMECs, whereas Snail- and Slug-HMECs maintained epithelial cell-cell adhesions through either E-cadherin or EpCAM expression (**Figure 4.11**B). Immunofluorescence, Western blotting and qRT-PCR analysis of D3- and Slug-HMECs showed decreased E-cadherin and increased vimentin, N-cadherin and fibronectin expression (**Figure 4.11**B-D). Snail-HMECs had preferentially decreased EpCAM expression (**Figure 4.11**B-D). Both Snail- and Slug-HMECs had a 2-4-fold increase in Smarcd3/Baf60c mRNA expression (**Figure 4.11**D). Furthermore, Snail- and Slug-HMECs transcriptionally activated the EMT-inducing transcription factor Zeb2 similarly to Smarcd3/Baf60c (**Figures 4.11**E and **4.9**C). Like D3-HMECs, Slug-HMECs preferentially repressed the cell-cell adhesion tight junction genes Cldn4 and Cldn7, possibly contributing to the disruption of epithelial cell polarity (**Figure 4.11**E). Cumulatively, these findings suggest that Smarcd3/Baf60c and Slug might impact a similar EMT network.

### *SWI/SNF Complex Regulates Non-Canonical WNT Signaling through Epigenetic Activation of Wnt5a*

Wnt5a is a well-known regulator of cell polarity and cell adhesion through activation of non-canonical WNT signaling (Veeman *et al*, 2003; Zhu *et al*, 2012; Dissanayake *et al*, 2007). Expression of Wnt5a, the primary WNT ligand controlling activation of the non-canonical WNT signaling network, was upregulated in at least two independent D3-HMEC cell lines (**Figure 4.12**A). Elevated Wnt5a expression in

Slug-HMECs and EpCAM$^{(-)}$ SUM149 and SUM229 cells further validates the importance of Wnt5a to the induction of EMT in three additional EMT cell models with upregulated Smarcd3/Baf60c expression (**Figure 4.13**A and B). Activation of non-canonical WNT signaling pathways in D3-HMECs was demonstrated by increased PKCβ phosphorylation and expression of TGFβ (**Figure 4.12**A and B). Meanwhile, nuclear recruitment of β-catenin was diminished in the D3-HMECs, suggesting the possible inhibition of canonical WNT signaling (**Figure 4.12**B).

Smarcd3/Baf60c is a member of a multi-subunit SWI/SNF chromatin-remodeling complex. The ATP-dependent catalytic components of these complexes are either Brm or Brg1 (Ho & Crabtree, 2010; Bultman *et al*, 2000). In D3-HMECs, Brg1 was preferentially induced with Smarcd3/Baf60c expression, while Brm was unchanged (**Figure 4.12**C). Furthermore, Brg1 and FLAG-Smarcd3/Baf60c co-immunoprecipitated (**Figure 4.12**D), consistent with Brg1 and Smarcd3/Baf60c interacting to form a complex. Brg1 is known to regulate transcription of E-cadherin and Cd44, so we used these genes as controls for chromatin immunoprecipitation (ChIP) assays to determine if the Brg1/Smarcd3/Baf60c complex resided on the Wnt5a gene promoter (**Figure 4.12**E) (Banine, 2005). ChIP assays with antibodies against Brg1 and/or Smarcd3/Baf60c demonstrated that Smarcd3/Baf60c and Brg1 bound to the promoters of E-cadherin, Cd44 and Wnt5a genes but not the Cldn4 gene (**Figure 4.12**E and F). Simultaneous ChIP assays for the activating histone methylation mark H3K4me3 demonstrated transcriptional repression of E-cadherin and Cldn4 with activation of Cd44 and Wnt5a (**Figure 4.12**G), demonstrating that Smarcd3/Baf60c and Brg1 contribute to the epigenetic activation of Wnt5a.

### *Inhibition of Wnt5a Restores Epithelial Adherens Junctions*

To determine if Wnt5a was responsible for inducing the loss of epithelial characteristics resulting from Smarcd3/Baf60c expression, we inhibited Wnt5a expression using RNAi in D3-HMECs. siRNA-mediated knockdown of Wnt5a phenotypically converted mesenchymal-like cells lacking cell-cell adhesions to epithelial colonies with a cobblestone morphology (**Figure 4.14**A). Wnt5a knockdown resulted in the increased mRNA and protein expression of epithelial markers E-cadherin and EpCAM (**Figure 4.14**B and C). Although expression of epithelial tight junction genes Cldn4 and Cldn7 was reduced by 90 percent in D3-HMECs relative to EV-HMECs (**Figure 4.9**C), knockdown of Wnt5a resulted in a 10-fold increase in Cldn4 and Cldn7 (**Figure 4.14**B). Furthermore, the EMT-inducing transcription factor Lef1 and the mesenchymal marker vimentin, which were induced with Smarcd3/Baf60c expression (**Figure 4.9**B and C), were repressed with knockdown of Wnt5a (**Figure 4.14**B). Finally, knockdown of Wnt5a repressed PKCβ phosphorylation, consistent with inhibition of the non-canonical WNT signaling pathway (**Figure 4.14**C). Treatment of D3-HMECs with a Wnt5a blocking antibody inhibited cellular invasiveness by 70 percent compared to untreated D3-HMECs (**Figure 4.14**D), demonstrating that the phenotypic changes coincided with functional changes indicative of a more epithelial cell state. Functionally, the blocking antibody was as effective as siRNA-mediated knockdown of Wnt5a (**Figure 4.14**D). These results demonstrate that Smarcd3/Baf60c and Brg1 are members of a SWI/SNF chromatin-remodeling complex that promotes EMT through epigenetic regulation of Wnt5a expression.

## Discussion

In this report, we describe a directed phenotypic RNAi screening strategy to identify individual genes whose siRNA-mediated knockdown is sufficient to promote MET in the mesenchymal $EpCAM^{(-)}$ population of SUM149 and SUM229 breast cancer cells. Although EMT has been accepted as a reversible process during development, few studies have examined the opposite process of mesenchymal-epithelial transition (MET) during the progression of breast cancer (Kalluri & Weinberg, 2009; Valastyan & Weinberg, 2011). This phenotypic RNAi screening approach demonstrates the reversible nature of EMT and MET in SUM149 and SUM229 breast cancer cells. The RNAi screen targeted 140 genes from an EMT gene signature shared by epithelial stem cells developmentally entering EMT and Claudin-low breast cancer cells with properties of EMT and stemness (Abell *et al*, 2011). Ten of the 140 genes were identified as EMT-regulatory genes, whose individual siRNA-mediated knockdown was sufficient to promote MET. Three of the genes, Slug, Met receptor tyrosine kinase and N-cadherin, are well-defined regulators of EMT and provide further validation of the screen (Thiery & Sleeman, 2006). The other seven genes are not previously defined for having a functional role in cancer cell EMT, but the functions of many of these genes suggest importance in the phenotypic reprogramming associated with EMT. For instance, EphA4 and Ptprb/VE-Ptp represent a receptor tyrosine kinase and phosphatase, respectively, that in development have been defined to be critical regulators of angiogenesis (Holder & Klein, 1999; Pasquale, 2010; Dominguez *et al*, 2007; Mori *et al*, 2010). Fhl1 localizes to focal adhesions where it promotes cell spreading, and Rnf130 is a

ring finger protein with putative E3 ligase activity (Rafael *et al*, 2012; Maetzel *et al*, 2009; Ding *et al*, 2009; Sheikh *et al*, 2008; Guais *et al*, 2006). Consistent with its knockdown inhibiting invasion of mesenchymal EpCAM$^{(-)}$ SUM149 cells, ABR is a GEF/GAP for Rho and Cdc42 GTPases, and RRAGD is a poorly characterized GTPase with homology to Ras (Chuang *et al*, 1995; Tcherkezian & Lamarche-Vane, 2012; Sekiguchi, 2000; Sancak *et al*, 2010). Among these uncharacterized EMT-regulatory genes, Rnf130 is an ideal gene candidate for future mechanistic studies with respect to EMT. Our preliminary findings have demonstrated that combined knockdown of Rnf130, Slug and Smarcd3/Baf60c produced synergistic effects in promoting an epithelial morphology in mesenchymal Claudin-low SUM159 breast cancer cells. Our data suggest these six EMT-regulatory genes in addition to Slug, Met, N-cadherin and Smarcd3/Baf60c are likely members of a signaling network controlling EMT.

Herein, we demonstrate that Smarcd3/Baf60c is a novel epigenetic EMT-regulatory gene. Smarcd3/Baf60c was the only gene identified in the RNAi screen whose knockdown strongly increased both EpCAM and E-cadherin expression, decreased Snail and Slug expression, inhibited invasiveness and was also upregulated in EpCAM$^{(-)}$ relative to EpCAM$^{+/high}$ SUM149 and SUM229 breast cancer cells. Smarcd3/Baf60c is a member of the multi-subunit ATP-dependent SWI/SNF chromatin-remodeling complex with the ability to epigenetically modulate gene expression programs (Lickert *et al*, 2004; Ho & Crabtree, 2010). Smarcd3/Baf60c induced an EMT gene expression program in HMECs similar to that of the Claudin-low breast cancer subtype, indicating that Smarcd3/Baf60c plays an important role in

maintaining the EMT properties and mesenchymal phenotype of these cells. Consistent with their related roles as transcriptional regulators, siRNA-mediated knockdown of Slug induced a MET phenotype similar to inhibition of Smarcd3/Baf60c expression. Notably, ectopic expression of either Smarcd3/Baf60c or Slug similarly induced the expression of Wnt5a, suggesting that Smarcd3/Baf60c affect an overlapping EMT gene network. To our knowledge, this is the first report of Smarcd3/Baf60c promoting the epigenetic reprogramming events responsible for inducing EMT and for maintaining the mesenchymal phenotype in Claudin-low $EpCAM^{(-)}$ populations of two independent breast cancer cell lines.

There is an expanding awareness that SWI/SNF chromatin-remodeling complexes play critical functions in reprogramming events relevant to EMT and lineage commitment. Brg1, one of the catalytic components of SWI/SNF chromatin-remodeling complex, is important in the reprogramming of somatic cells to induced pluripotent stem (iPS) cells (Singhal *et al*, 2010; Kidder *et al*, 2009). MET was shown recently to occur simultaneously with the reprogramming of fibroblasts to iPS cells, directly connecting MET to iPS cell generation (Li *et al*, 2010). Our results indicate that studies are warranted to determine if Smarcd3/Baf60c expression is repressed in these cells as a switch to induce MET during iPS cell reprogramming. In the context of regulating EMT and MET, the non-canonical WNT signaling pathway has been shown to be an important collaborator with TGFβ for inducing EMT. Non-canonical WNT signaling, for example, contributes to the metastatic progression of pancreatic cancer and melanoma (Scheel *et al*, 2011; Yu *et al*, 2012; Dissanayake *et al*, 2007). In D3-HMECs, we showed that both Smarcd3/Baf60c and Brg1 bind the

promoters of Wnt5a and E-cadherin. Wnt5a expression is induced, stimulating non-canonical WNT signaling, while E-cadherin expression is repressed to promote EMT. Together, these findings show the importance of Smarcd3/Baf60c in controlling the commitment to the mesenchymal cellular phenotype.

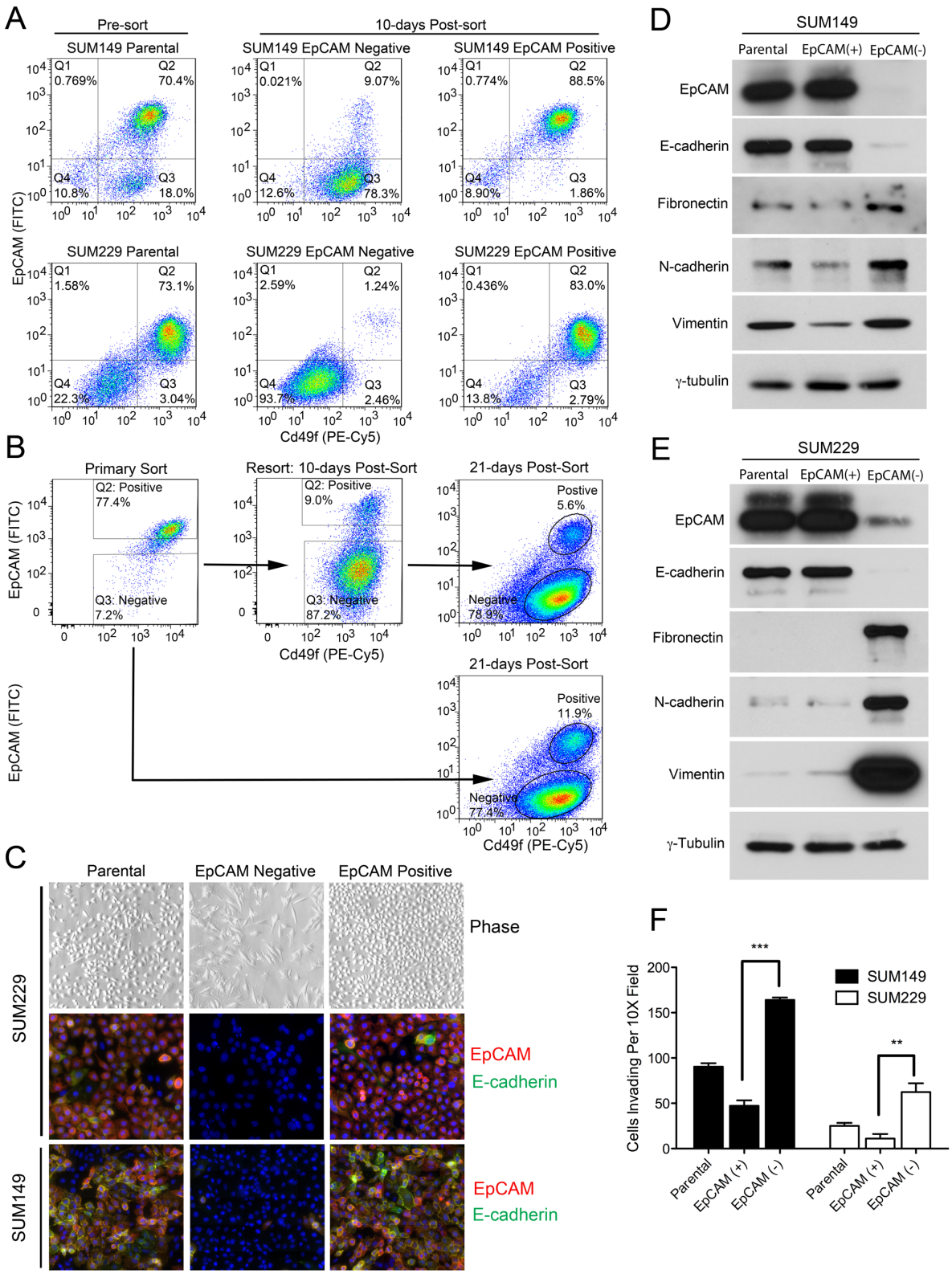



**Figure 4.1: SUM149 and SUM229 Breast Cancer Cells Maintain Epithelial and Mesenchymal Populations.**

(A) FACS analysis of SUM149 and SUM229 cells showing two $EpCAM^{(-)}$ and $EpCAM^{+/high}$ populations. (B) Serial FACS analysis of SUM149 $EpCAM^{(-)}$ cells demonstrating conversion from $EpCAM^{(-)}$ to $EpCAM^{+/high}$ cells after two repeated FACS experiments. (C) Epithelial and mesenchymal properties of $EpCAM^{(-)}$ and $EpCAM^{+/high}$ populations shown by phase microscopy and immunostaining with nuclear DAPI stain and anti-E-cadherin and anti-EpCAM antibodies. (D-E) Reduced EpCAM and E-cadherin protein expression with elevated fibronectin, N-cadherin and vimentin protein in $EpCAM^{(-)}$ cells. (A-E) Data are representative of at least two independent experiments. (F) Increased invasiveness of mesenchymal $EpCAM^{(-)}$ compared to epithelial $EpCAM^{+/high}$ cells. Statistical significance was evaluated by an unpaired Student's T-test. ***p-value < 0.001; **p-value < 0.01. Data are the mean $\pm$ range of two independent experiments performed in triplicate.

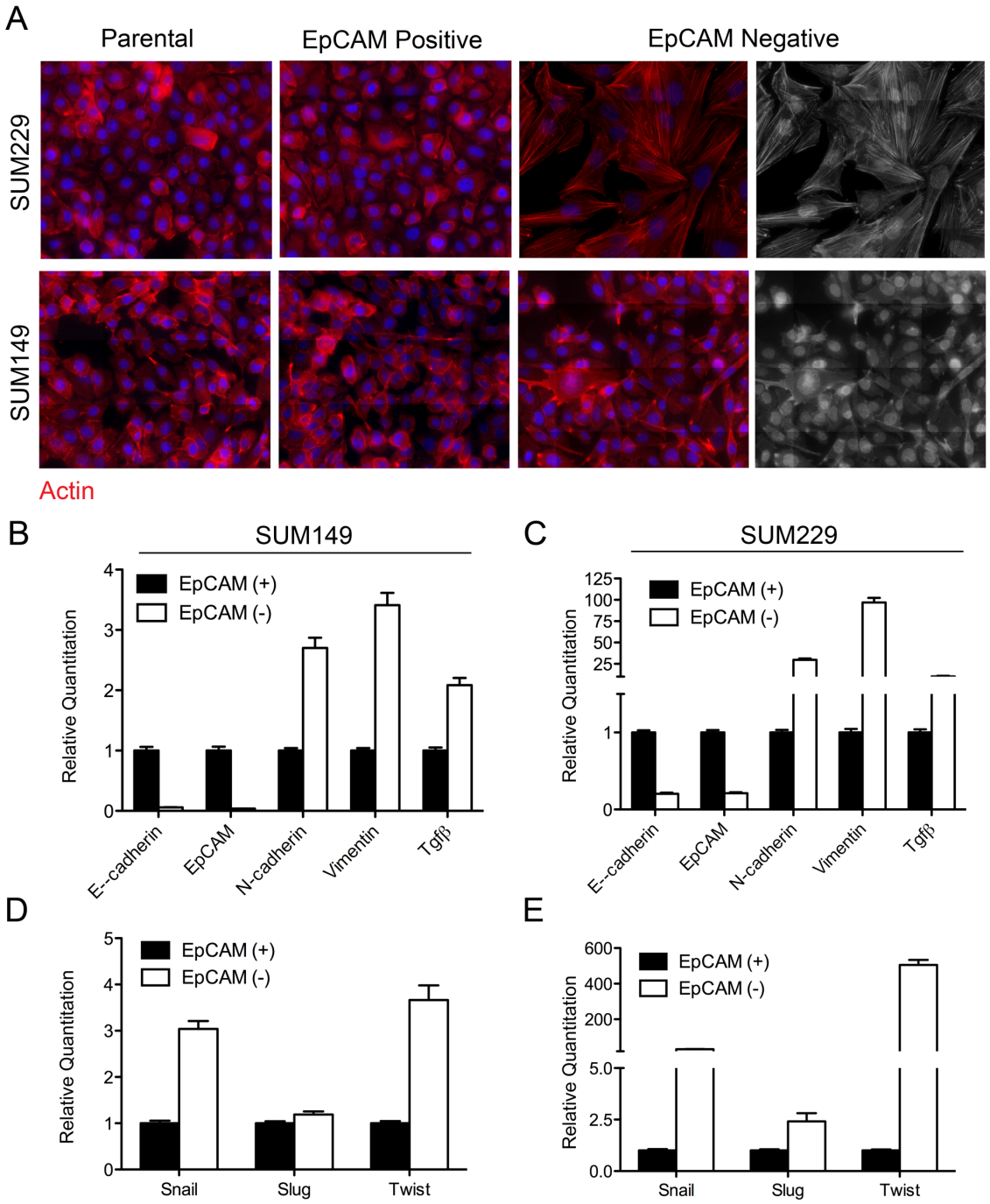



**Figure 4.2: Characterization of Epithelial and Mesenchymal Populations in SUM149 and SUM229 Cells**

(A) Filamentous actin stress fibers in EpCAM$^{(-)}$ cells compared to cortical actin in EpCAM$^{+/high}$ cells as shown by immunostaining with nuclear DAPI and rhodamine Phalloidin stains. (B-C) Reduced EpCAM and E-cadherin mRNA expression and elevated N-cadherin, vimentin and Tgfβ expression in EpCAM$^{(-)}$ compared to EpCAM$^{+/high}$ cells. (D-E) Elevated mRNA expression of EMT-inducing transcription factors Snail, Slug and Twist in EpCAM$^{(-)}$ compared to EpCAM$^{+/high}$ cells. (B-E) Data are the mean ± range of three independent experiments performed in triplicate.

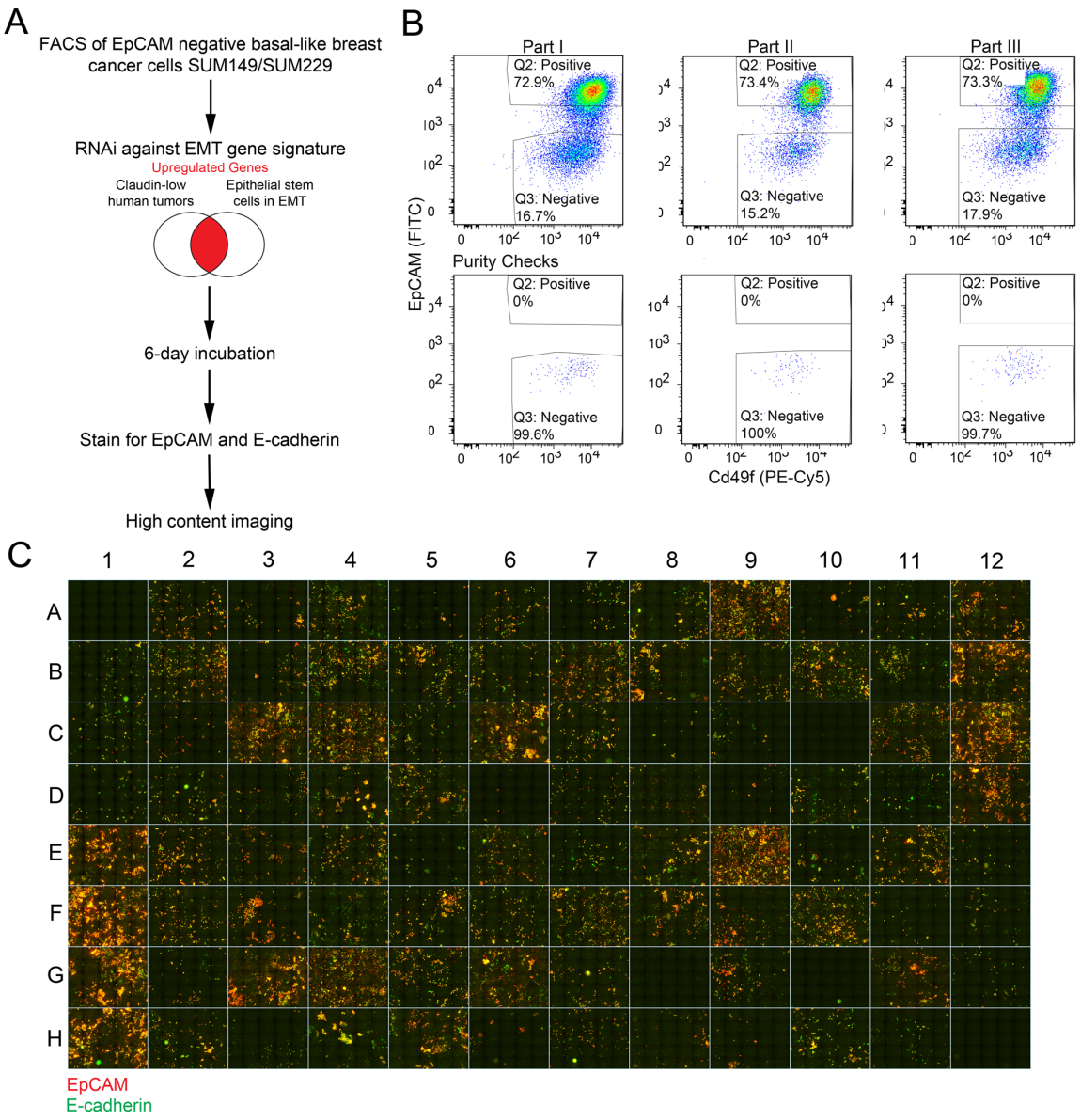



**Figure 4.3: RNAi Screening Strategy for EMT-Regulatory Genes**
(A) Schematic depicting the RNAi screening strategy. (B) Separation of EpCAM$^{+/high}$ and EpCAM$^{(-)}$ SUM149 cells by FACS shows less than 0.5 percent EpCAM$^{+/high}$ contamination in the EpCAM$^{(-)}$ cells. Data depicts FACS analysis of three individual cell sorts used for the primary RNAi screen. (C) Representative plate image from the RNAi screen shows expression of epithelial markers E-cadherin and EpCAM by immunofluorescence in EpCAM$^{(-)}$ SUM149 cells following siRNA-mediated knockdown of target genes in duplicate wells shown in columns 2-11, rows A-H; siGAPDH in columns 1 and 12, rows B-D and E-H, respectively; siSnail/Slug in columns 1 and 12, rows E-H and A-D, respectively; and siUBB in wells A1 and H12. Data are representative of duplicate RNAi screens for 140 genes performed in triplicate.

| Gene Symbol | Gene ID | Gene Accession | GI Number | Dhamacon Pool # |
|---|---|---|---|---|
| MET | 4233 | NM_000245 | 42741654 | M-003156-02 |
| PDGFRB | 5159 | NM_002609 | 68216043 | M-003163-03 |
| PTPRB | 5787 | NM_002837 | 157952212 | M-004994-03 |
| SYK | 6850 | NM_003177 | 34147655 | M-003176-03 |
| HCK | 3055 | NM_002110 | 30795228 | M-003141-03 |
| PDGFRA | 5156 | NM_006206 | 61699224 | M-003162-04 |
| AXL | 558 | NM_001699 | 157779133 | M-003104-03 |
| EPHA4 | 2043 | NM_004438 | 45439363 | M-003118-02 |
| EGFR | 1956 | NM_201283 | 41327733 | M-003114-03 |
| RET | 5979 | NM_020630 | 126273513 | M-003170-02 |
| VIM | 7431 | NM_003380 | 62414288 | M-003551-02 |
| MBNL1 | 4154 | NM_207294 | 46411167 | M-014136-01 |
| PLAT | 5327 | NM_033011 | 132626641 | M-005999-01 |
| CDH2 | 1000 | NM_001792 | 14589888 | M-011605-01 |
| FHL1 | 2273 | NM_001449 | 34147646 | M-015857-00 |
| THY1 | 7070 | NM_006288 | 19923361 | M-015337-00 |
| SERPINB1 | 1992 | NM_030666 | 20149554 | M-016240-00 |
| TIMP2 | 7077 | NM_003255 | 73858577 | M-011793-01 |
| DKK3 | 27122 | NM_001018057 | 66346688 | M-018352-01 |
| ABR | 29 | NM_001092 | 38679953 | M-008611-00 |
| SLIT2 | 9353 | NM_004787 | 4759145 | M-019853-01 |
| MCAM | 4162 | NM_006500 | 71274106 | M-003686-01 |
| DIXDC1 | 85458 | NM_033425 | 116256346 | M-007758-01 |
| ELMO1 | 9844 | NM_001039459 | 86788139 | M-012851-01 |
| KLF2 | 10365 | NM_016270 | 49574523 | M-006928-01 |
| DKK1 | 22943 | NM_012242 | 61676924 | M-003843-01 |
| SLIT3 | 6586 | NM_003062 | 11321570 | M-012600-00 |
| VAV1 | 7409 | NM_005428 | 7108366 | M-003935-01 |
| SOX17 | 64321 | NM_022454 | 145275218 | M-013028-01 |
| SOX11 | 6664 | NM_003108 | 30581115 | M-017377-02 |
| CITED1 | 4435 | NM_004143 | 20127458 | M-016022-00 |
| RAB27A | 5873 | NM_004580 | 34485707 | M-004667-00 |
| SMARCD3 | 6604 | NM_001003802 | 51477703 | M-009459-02 |
| ARHGEF6 | 9459 | NM_004840 | 22027524 | M-010231-00 |
| HEY1 | 23462 | NM_012258 | 105990527 | M-008709-02 |
| CCDC50 | 152137 | NM_174908 | 33186926 | M-017781-01 |
| HIC | 29969 | NM_199072 | 110735446 | M-025441-01 |
| METRNL-LOC284207 | 284207 | NM_001004431 | 52345386 | M-022983-01 |

| | | | | |
|---|---|---|---|---|
| PHACTR2 | 9749 | NM_001100166 | 154354971 | M-020445-02 |
| TSPAN-2 | 10100 | NM_005725 | 58331193 | M-017266-01 |
| GNG11 | 2791 | NM_004126 | 91807126 | M-016036-01 |
| LY96 | 23643 | NM_015364 | 34098964 | M-020298-00 |
| RAFTLIN | 23180 | NM_015150 | 41872576 | M-023452-01 |
| DPYSL2 | 1808 | NM_001386 | 50811906 | M-009519-01 |
| AKAP12 | 9590 | NM_005100 | 21493021 | M-010232-00 |
| CTSC | 1075 | NM_001814 | 22538438 | M-005835-03 |
| KHDRBS3 | 10656 | NM_006558 | 5730072 | M-012748-00 |
| LDHB | 3945 | NM_002300 | 142386935 | M-009779-01 |
| PTPRM | 5797 | NM_002845 | 157419149 | M-006326-01 |
| MMP9 | 4318 | NM_004994 | 74272286 | M-005970-03 |
| DPYSL3 | 1809 | NM_001387 | 4503378 | M-009821-00 |
| ITM2A | 9452 | NM_004867 | 74316000 | M-019433-01 |
| RGL1 | 23179 | NM_015149 | 93102366 | M-008387-01 |
| CTSH | 1512 | NM_148979 | 148536856 | M-005839-01 |
| ELK3 | 2004 | NM_005230 | 4885202 | M-010320-00 |
| NEXN | 91624 | NM_144573 | 148839338 | M-016402-01 |
| JAM2 | 58494 | NM_021219 | 21704284 | M-017389-00 |
| AMICA | 120425 | NM_001098526 | 148664206 | M-016407-01 |
| CORO1A | 11151 | NM_007074 | 68161542 | M-012771-01 |
| CD93 | 22918 | NM_012072 | 88758612 | M-007815-02 |
| FBLN2 | 2199 | NM_001998 | 51873054 | M-011655-01 |
| STX11 | 8676 | NM_003764 | 33667037 | M-019469-01 |
| HEM1 | 3071 | NM_005337 | 153792638 | M-019219-01 |
| PPAP2B | 8613 | NM_003713 | 29171739 | M-017312-01 |
| RNF130 | 55819 | NM_018434 | 38176162 | M-007021-00 |
| ARHGEF6 | 9459 | NM_004840 | 22027524 | M-010231-00 |
| RASSF4 | 83937 | NM_032023 | 30474868 | M-017587-00 |
| CXCL12 | 6387 | NM_000609 | 76563934 | M-007873-01 |
| ICAM1 | 3383 | NM_000201 | 4557877 | M-003502-01 |
| ACSL4 | 2182 | NM_004458 | 4758331 | M-009364-00 |
| GADD45A | 1647 | NM_001924 | 9790904 | M-003893-02 |
| NDRG1 | 10397 | NM_006096 | 37655182 | M-010563-01 |
| TBX3 | 6926 | NM_005996 | 47419904 | M-012197-00 |
| PCAF | 8850 | NM_003884 | 6382075 | M-005055-00 |
| SFRP2 | 6423 | NM_003013 | 52630413 | M-024062-02 |
| NDRG2 | 57447 | NM_016250 | 10280619 | M-013858-00 |
| DACT1 | 51339 | NM_001079520 | 118498363 | M-016440-01 |
| SH3KBP1 | 30011 | NM_031892 | 13994241 | M-014748-00 |
| WAS | 7454 | NM_000377 | | |

| | | | | |
|---|---|---|---|---|
| SMARCA4 | 6597 | NM_003072 | 21071055 | M-010431-00 |
| MT2A | 4502 | NM_005953 | 31543214 | M-018338-00 |
| PLP2 | 5355 | NM_002668 | 4505892 | M-019592-00 |
| COL5A1 | 1289 | NM_000093 | 89276750 | M-011015-01 |
| ANTXR1 | 84168 | NM_018153 | 16933552 | M-010679-00 |
| LIFR | 3977 | NM_002310 | 73486667 | M-008017-01 |
| LAMP2 | 3920 | NM_002294 | 4504956 | M-011715-00 |
| LOXL2 | 4017 | NM_002318 | 67782347 | M-008020-01 |
| GYPC | 2995 | NM_002101 | 21614502 | M-011680-00 |
| EPB41L3 | 23136 | NM_012307 | 32490571 | M-015330-00 |
| COL6A1 | 1291 | NM_001848 | 87196338 | M-011620-01 |
| HEXA | 3073 | NM_000520 | 119395761 | M-011049-01 |
| ABCC4 | 10257 | NM_001105515 | 157502202 | M-007313-01 |
| TFPI | 7035 | NM_001032281 | 98991771 | M-019922-01 |
| SLC2A3 | 6515 | NM_006931 | 5902089 | M-007516-01 |
| CLU | 1191 | NM_001831 | 42716296 | M-019513-00 |
| CSPG4 | 1464 | NM_001897 | 126091140 | M-011632-01 |
| LARP6 | 55323 | NM_197958 | 37537705 | M-021333-00 |
| CHST2 | 9435 | NM_004267 | 71067350 | M-008963-01 |
| RCN3 | 57333 | NM_020650 | 28626509 | M-017962-01 |
| PVRL3 | 25945 | NM_015480 | 11386198 | M-013952-01 |
| DHRS8 | 51170 | NM_016245 | 142976728 | M-008859-01 |
| PMP22 | 5376 | NM_153322 | 24430164 | M-010616-02 |
| RIS1 | 25907 | NM_015444 | 116805333 | M-013960-01 |
| THBD | 7056 | NM_000361 | 40288292 | M-012551-02 |
| PAG | 55824 | NM_018440 | 63054863 | M-012956-02 |
| CYBA | 1535 | NM_000101 | 68509913 | M-011020-01 |
| NT5E | 4907 | NM_002526 | 4505466 | M-008217-00 |
| CLGN | 1047 | NM_004362 | 4758003 | M-011830-00 |
| C1S | 716 | NM_001734 | 4502494 | M-005796-01 |
| ITGA11 | 22801 | NM_001004439 | 52485852 | M-008000-02 |
| SMPD1 | 6609 | NM_001007593 | 56117841 | M-006676-01 |
| ST3GAL6 | 10402 | NM_006100 | 31377787 | M-018009-01 |
| PDLIM3 | 27295 | NM_014476 | 7656882 | M-020229-00 |
| CHST7 | 56548 | NM_019886 | 23238227 | M-008443-00 |
| MYO5A | 4644 | NM_000259 | 115511013 | M-019321-01 |
| RRAGD | 58528 | NM_021244 | 142369198 | M-016120-01 |
| RGS2 | 5997 | NM_002923 | 142365756 | M-009887-02 |
| MOXD1 | 26002 | NM_015529 | 118421086 | M-009446-02 |
| LIN7B | 64130 | NM_022165 | 56676320 | M-009653-01 |
| AQP9 | 366 | NM_020980 | 157266306 | M-007768-02 |

| | | | | |
|---|---|---|---|---|
| RGS17 | 26575 | NM_012419 | 56790938 | M-009513-02 |
| MMP11 | 4320 | NM_005940 | 58331147 | M-005953-01 |
| FGF13 | 2258 | NM_033642 | 16306542 | M-011859-02 |
| GAS1 | 2619 | NM_002048 | 4503918 | M-011665-00 |
| PCDH18 | 54510 | NM_019035 | 156630996 | M-013372-01 |
| CMKOR1 | 57007 | NM_020311 | 114155149 | M-013212-03 |
| CCL2 | 6347 | NM_002982 | 56119169 | M-007831-01 |
| LDHC | 3948 | NM_017448 | 9257227 | M-008759-02 |
| GDF6 | 392255 | NM_001001557 | 48475061 | M-033055-01 |
| EPDR1 | 54749 | NM_017549 | 24475585 | M-013378-01 |
| CDW52 | 1043 | NM_001803 | 68342029 | M-019509-01 |
| PRDM5 | 11107 | NM_018699 | 41349475 | M-020332-00 |
| NR0B1 | 190 | NM_000475 | 5016089 | M-003409-00 |
| ECM1 | 1893 | NM_004425 | 12707564 | M-011849-00 |
| SLC9A6 | 10479 | NM_006359 | 110227627 | M-007626-01 |
| SLC7A8 | 23428 | NM_182728 | 33286429 | M-007618-01 |
| LDHA | 3939 | NM_005566 | 5031856 | M-008201-01 |
| TMEM88 | 92162 | NM_203411 | 44662812 | M-031924-00 |
| NPL | 80896 | NM_030769 | 13540532 | M-008230-01 |

**Table 4.1: EMT Gene Signature Shared between Epithelial Stem Cells Developmentally Entering EMT and Claudin-low Breast Cancer.**
EMT gene signature identified from the combined analysis of intersecting gene array and RNA-seq data from epithelial stem cells described in Abell et al., *Cell Stem Cell*, 2011, and the Claudin-low breast cancer subtype (Abell *et al*, 2011).

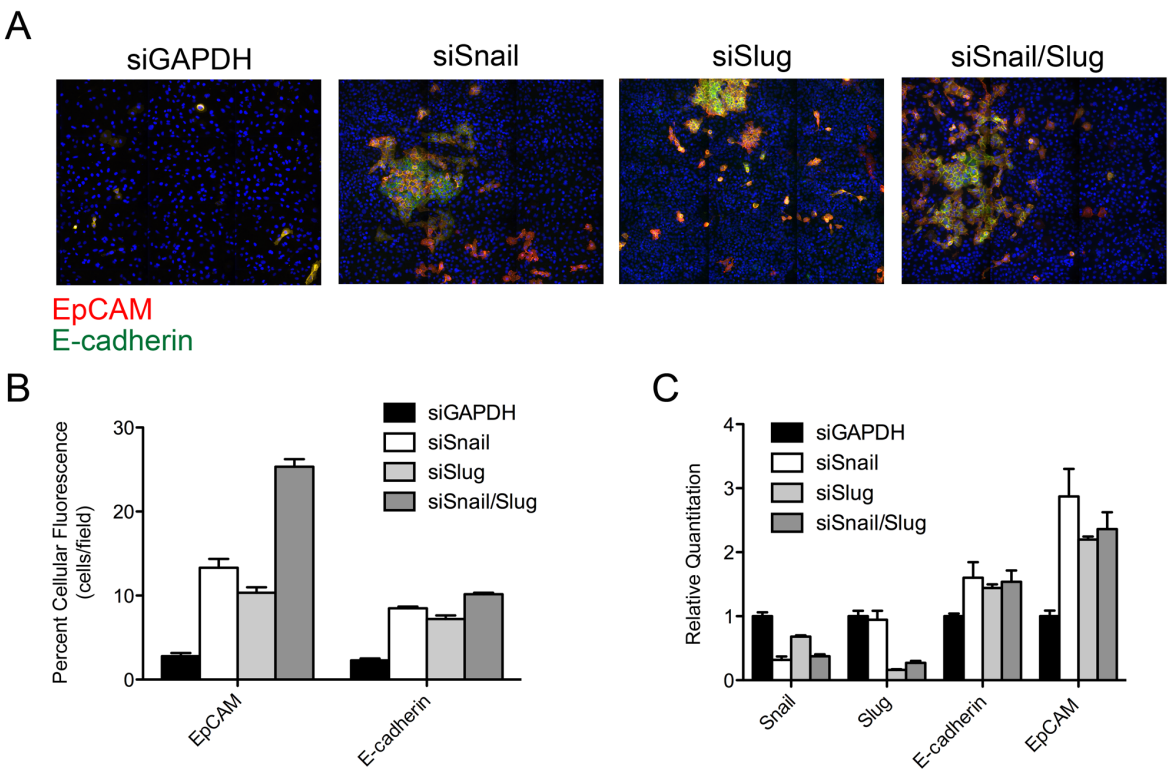



**Figure 4.4: RNAi Screen Development**
(A) Acquisition of the epithelial phenotype by mesenchymal EpCAM$^{(-)}$ cells following siRNA-mediated knockdown of Snail, Slug or Snail/Slug genes as shown by immunostaining with nuclear DAPI stain, anti-E-cadherin and anti-EpCAM. (B) Elevated EpCAM and E-cadherin percent cellular fluorescence quantitated on a single cell basis following siRNA knockdown of indicated genes. Data represent the mean $\pm$ range of 8 independent wells. (C) Elevated mRNA expression of epithelial markers EpCAM and E-cadherin following siRNA knockdown of Snail and Slug. Data represent the mean $\pm$ the range of three independent experiments performed in triplicate.

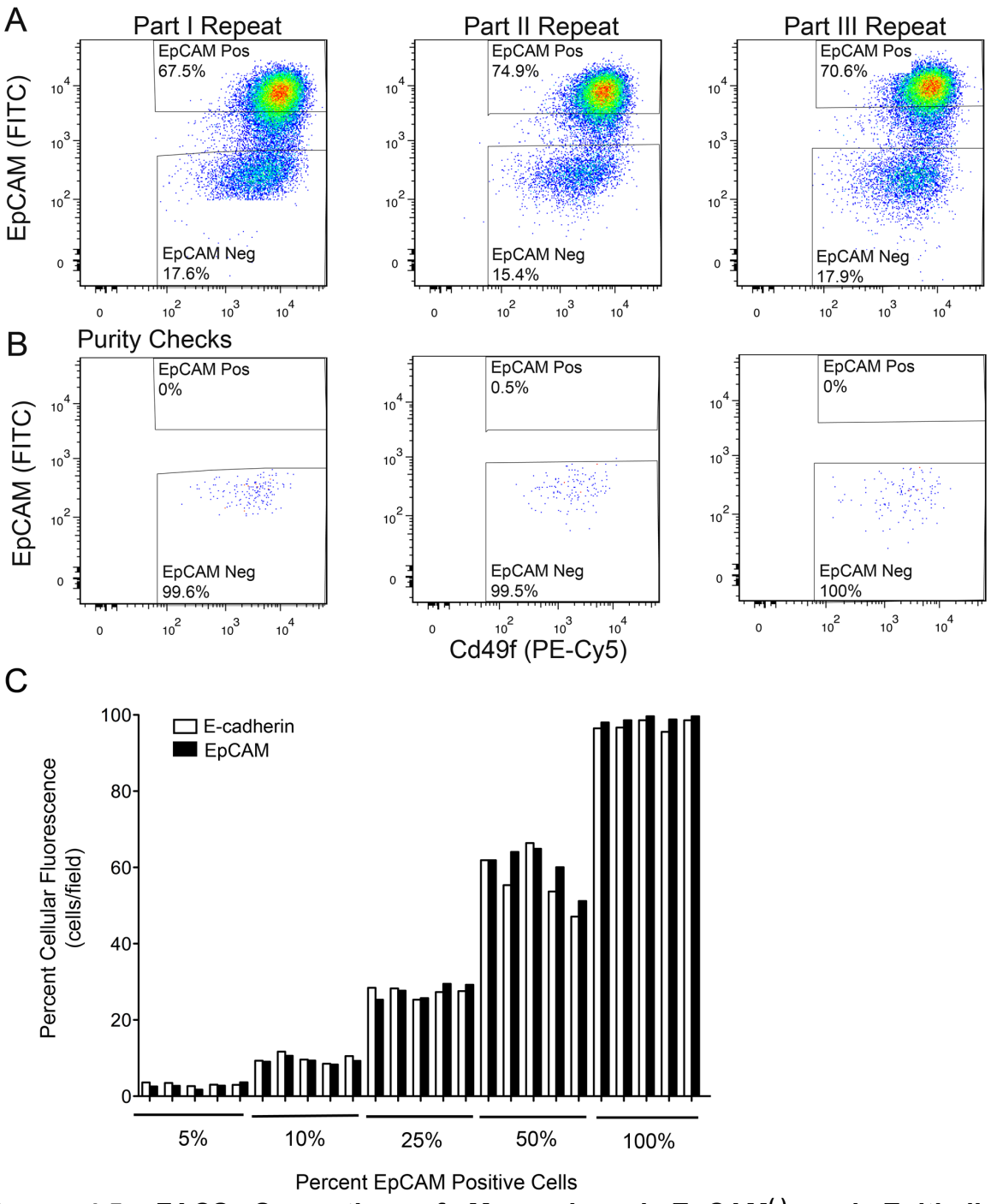



**Figure 4.5: FACS Separation of Mesenchymal EpCAM$^{(-)}$ and Epithelial EpCAM$^{+/high}$ Populations**

(A-B) Separation of EpCAM$^{+/high}$ and SUM149 EpCAM$^{(-)}$ cells by FACS shows less than 0.5 percent EpCAM$^{+/high}$ contamination of the EpCAM$^{(-)}$ cells. Data represents FACS analysis of three individual cell sorts used for confirming the RNAi screen. (C) Titration of known percentages of EpCAM$^{+/high}$ SUM149 cells into an EpCAM$^{(-)}$ background confirms accurate quantitation of the percent EpCAM and E-cadherin fluorescence on a single cell basis.

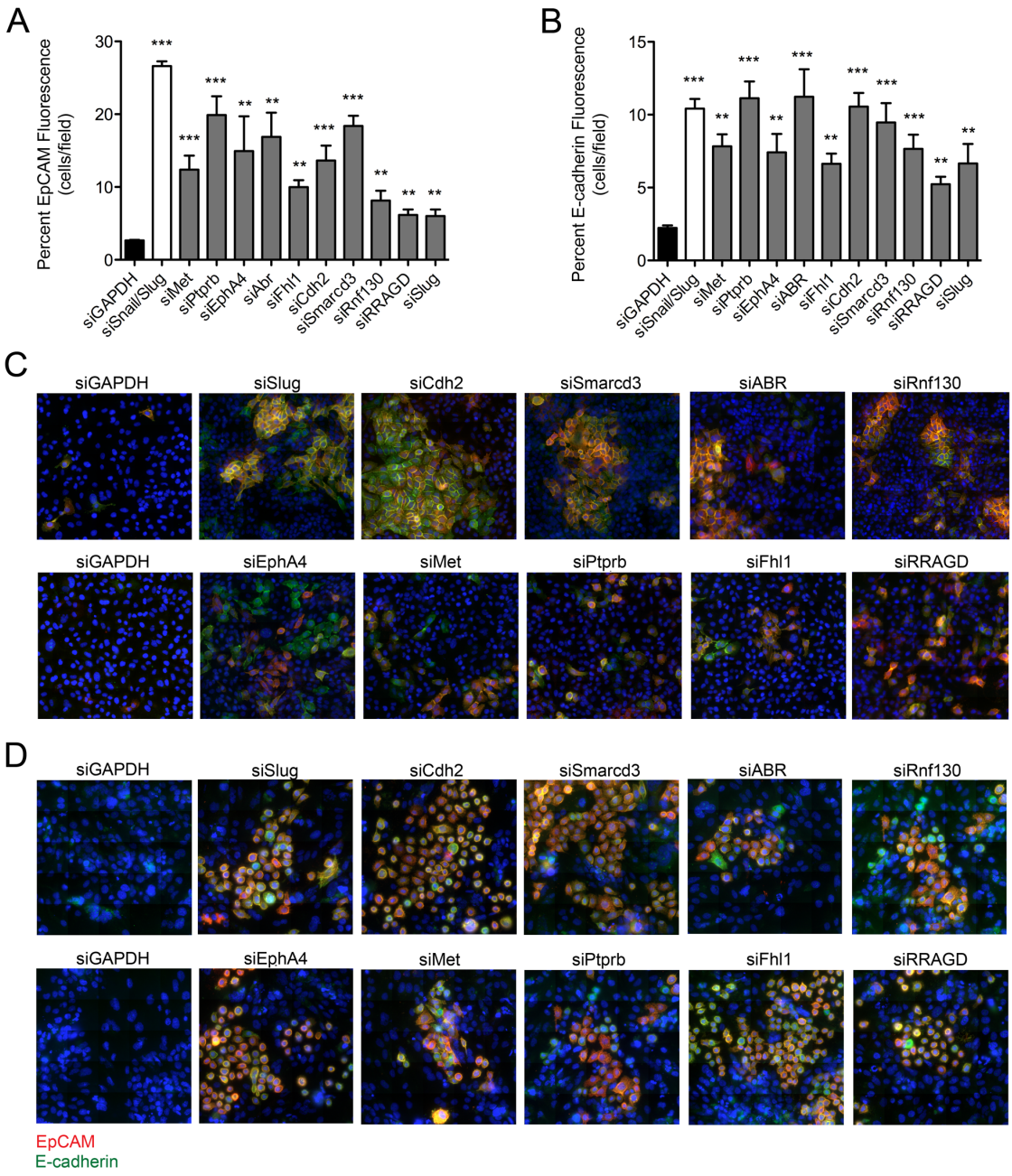



**Figure 4.6: RNAi screen Identifies Genes Important for EMT.**
(A-B) Elevated EpCAM and E-cadherin percent cellular fluorescence quantitated on a single cell basis following siRNA-mediated knockdown of indicated genes. Data represent the mean ± range of 6 independent wells from the primary screen. Statistical significance was evaluated by an unpaired Student's T-test. ***p-value < 0.001; **p-value < 0.01. (C-D) Acquisition of the epithelial phenotype by mesenchymal EpCAM$^{(-)}$ cells following siRNA-mediated knockdown of indicated genes is shown by mmunostaining of (C) SUM149 and (D) SUM229 EpCAM$^{(-)}$ cells with nuclear DAPI stain, anti-E-cadherin and anti-EpCAM antibodies.

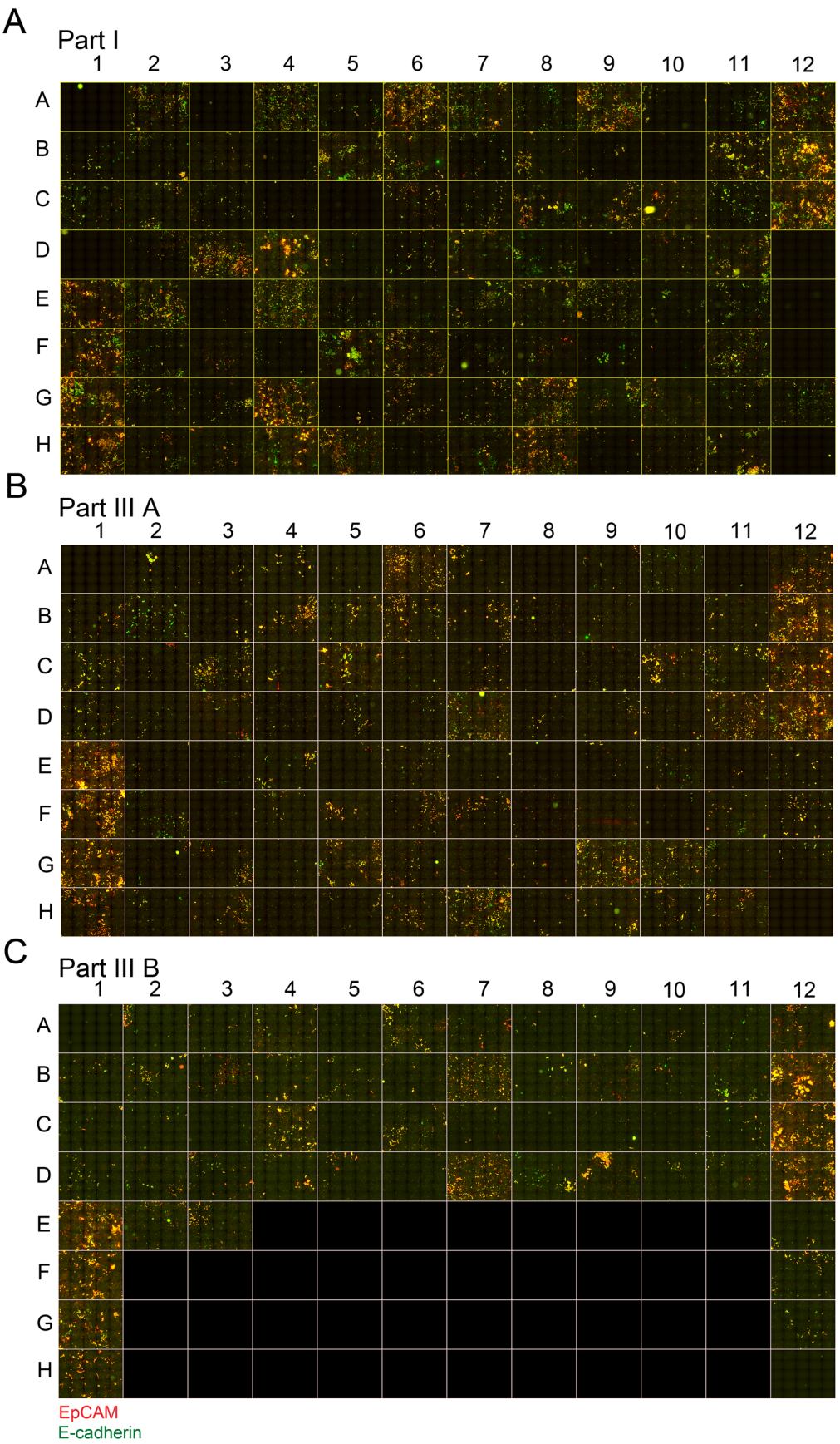

A
Part I
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
B
Part III A
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
C
Part III B
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
EpCAM
E-cadherin

**Figure 4.7: RNAi Screen Raw Images**
(A-C) Representative plate images from the RNAi screen showing expression of epithelial markers E-cadherin and EpCAM by immunofluorescence in EpCAM(-) SUM149 cells following siRNA-mediated knockdown of target genes in duplicate wells shown in columns 2-11, rows A-H; siGAPDH in columns 1 and 12, rows B-D and E-H, respectively; siSnail/Slug in columns 1 and 12, rows E-H and A-D, respectively; and siUBB in wells A1 and H12. Data is representative of duplicate RNAi screens for 140 genes performed in triplicate.

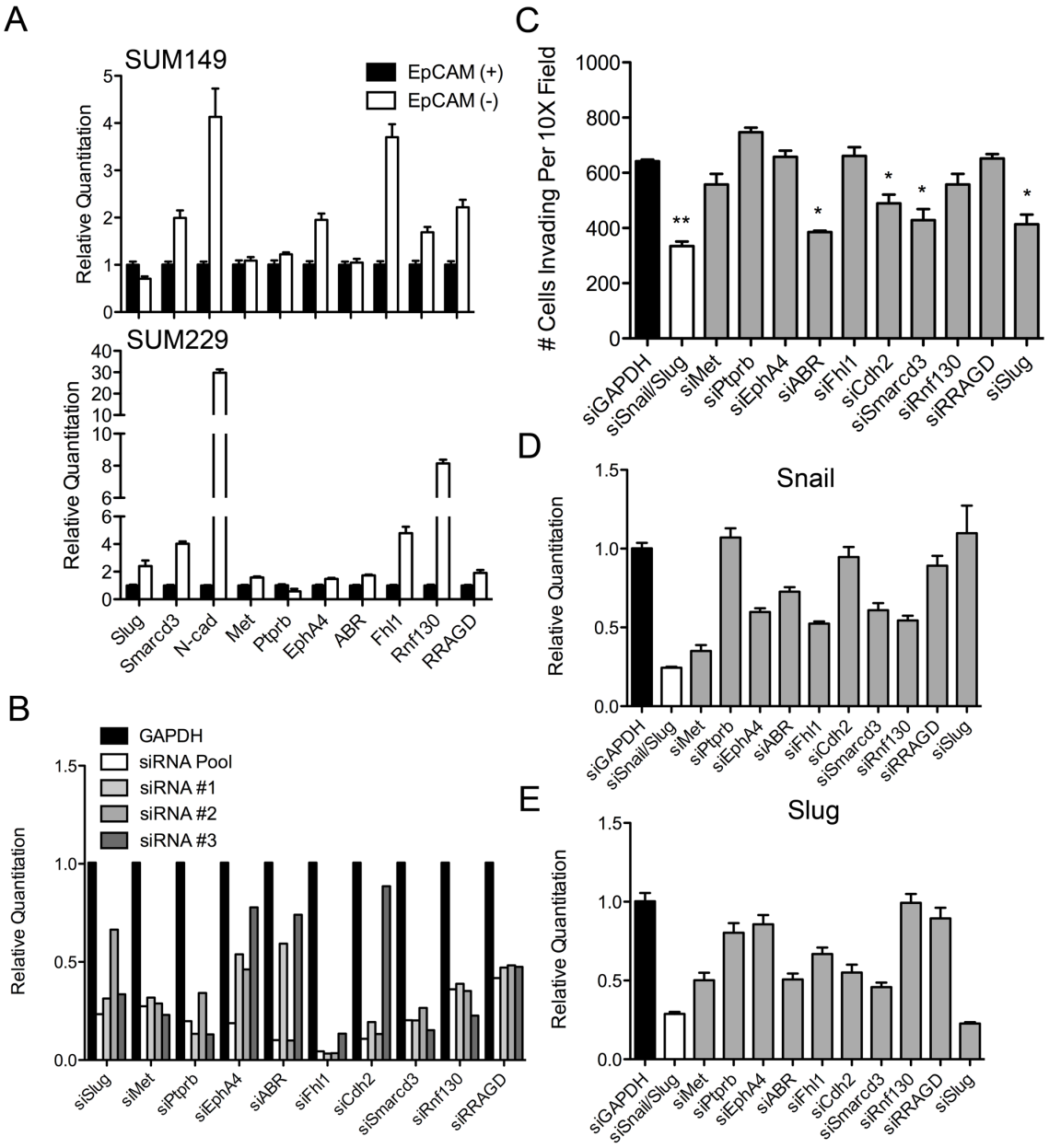



**Figure 4.8: EMT-Regulatory Genes Affect Cellular Invasiveness and Expression of EMT-Inducing Transcription Factors.**

(A) Expression levels of EMT-regulatory genes in EpCAM$^{+/high}$ compared to EpCAM$^{(-)}$ cells shown by qRT-PCR. (B) Deconvolution of siRNA smart pools demonstrates knockdown of EMT-regulatory genes with at least two individual siRNA oligonucleotides per gene in SUM149 EpCAM$^{(-)}$ cells as shown by qRT-PCR. (C) Cellular invasiveness of SUM149 EpCAM$^{(-)}$ cells following siRNA-mediated knockdown of indicated EMT-regulatory genes. **p-value < 0.01; *p-value < 0.05. Data are the mean $\pm$ range of two independent experiments performed in triplicate. (D-E) Expression changes of EMT-inducing transcription factors in SUM149

EpCAM$^{(-)}$ cells following siRNA knockdown of the indicated EMT-regulatory genes. Data are the mean $\pm$ the range of at least two independent experiments performed in triplicate.

| Gene Name | Function | Pathway/Interacting Proteins | Physiology |
|---|---|---|---|
| Smarcd3/ Baf60c | SWI/SNF chromatin remodeling factor | Binds transcription factors (GATA4, Smad4) and nuclear receptors (RAR, PPAR, ER, PR, ROR) | Defects in heart morphogenesis, outflow tract remodeling and abnormal cell differentiation in mice |
| Ptprb/VE-Ptp | Receptor tyrosine phosphatase | Regulates adherens junctions. Binds VE-cadherin and Tie2 | Failure to remodel vascular scaffold and defects in higher order branching in mice |
| Fhl1 | Zinc finger protein | Localized to focal adhesions to promote cell spreading. Binds Smad proteins and regulates TGF-beta responsive transcription | Human mutations induce cardio/skeletal myopathies. Expression observed strongly in developing heart outflow tract |
| EphA4 | Receptor tyrosine kinase | Regulates adherens junction signaling | Important in angiogenesis |
| Met | Receptor tyrosine kinase | Regulates adherens junction signaling via crosstalk with EGFR to activate PI3K/AKT pathways | Important in angiogenesis and cell spreading |
| N-cadherin | Cell-cell adhesion glycoprotein | Regulates adherens junction signaling by interacting with Fn/FGFR1 to activate MAPK/ERK pathways | Regulates gastrulation and neural crest migration |
| Slug | Transcription factor | Directly represses the epithelial marker E-cadherin | Functions during gastrulation and heart outflow tract remodeling |
| ABR | GEF/GAP | Activates Rac/Cdc42 GTPases | Important role in vestibular morphogenesis |
| Rnf130 | E3 ubiquitin ligase | Unknown | Unknown |
| RRAGD | GTPase with Ras homology | Unknown | Unknown |

**Table 4.2: Functional Description of EMT-Regulatory Genes Identified in the RNAi screen.**
Table describes the known gene function of the ten EMT-regulatory genes identified in the RNAi screen.

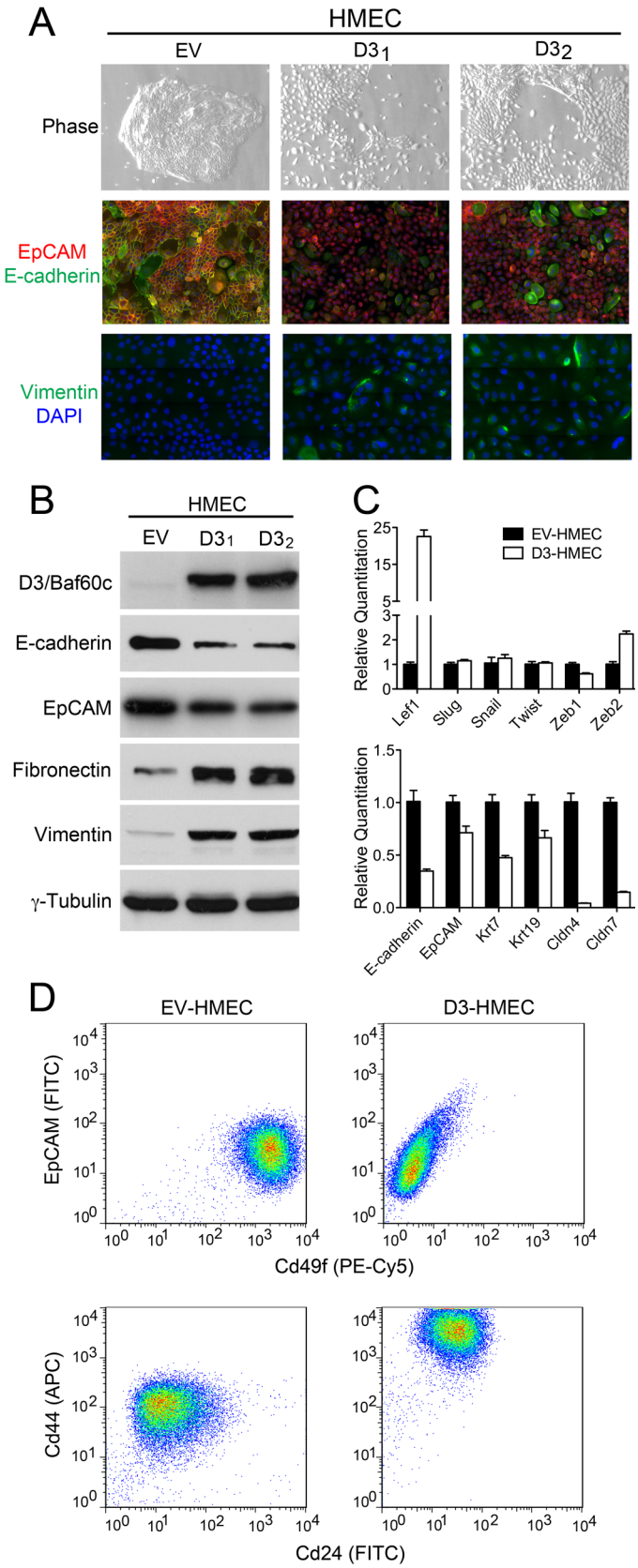

A
HMEC
EV
D3 1
D3 2
Phase
EpCAM
E-cadherin
Vimentin
DAPI
B
HMEC
EV
D3 1
D3 2
D3/Baf60c
E-cadherin
EpCAM
Fibronectin
Vimentin
γ-Tubulin
C
EV-HMEC
D3-HMEC
Relative Quantitation
Lef1
Slug
Snail
Twist
Zeb1
Zeb2
E-cadherin
EpCAM
Krt7
Krt19
Cldn4
Cldn7
D
EV-HMEC
D3-HMEC
EpCAM (FITC)
Cd49f (PE-Cy5)
Cd44 (APC)
Cd24 (FITC)

**Figure 4.9: Smarcd3/Baf60c is a Novel Regulator of EMT.**
(A) Smarcd3/Baf60c-expressing HMECs (D3-HMECs) gain a mesenchymal phenotype with loss of epithelial properties as shown by phase microscopy and immunostaining of EV- and D3-HMECs with nuclear DAPI stain, anti-E-cadherin anti-EpCAM, or anti-vimentin antibodies. (B) Loss of epithelial markers EpCAM and E-cadherin and gain of mesenchymal markers fibronectin and vimentin in D3-HMEC cells are shown by Western blot. (C) Elevated mRNA expression of EMT-inducing transcription factors Lef1 and Zeb2 and reduced expression of epithelial and cell adhesion markers in D3-HMECs. Data represent the mean $\pm$ range of three independent experiments. (D) Different antigenic profiles in EV- versus D3-HMECs demonstrated by FACS analysis with epithelial differentiation markers anti-EpCAM (FITC) and anti-Cd49f (PE-Cy5) or cancer stem-like markers anti-Cd44 (APC) and anti-Cd24 (FITC). (A-B and D) Data are representative of at least two independent experiments.

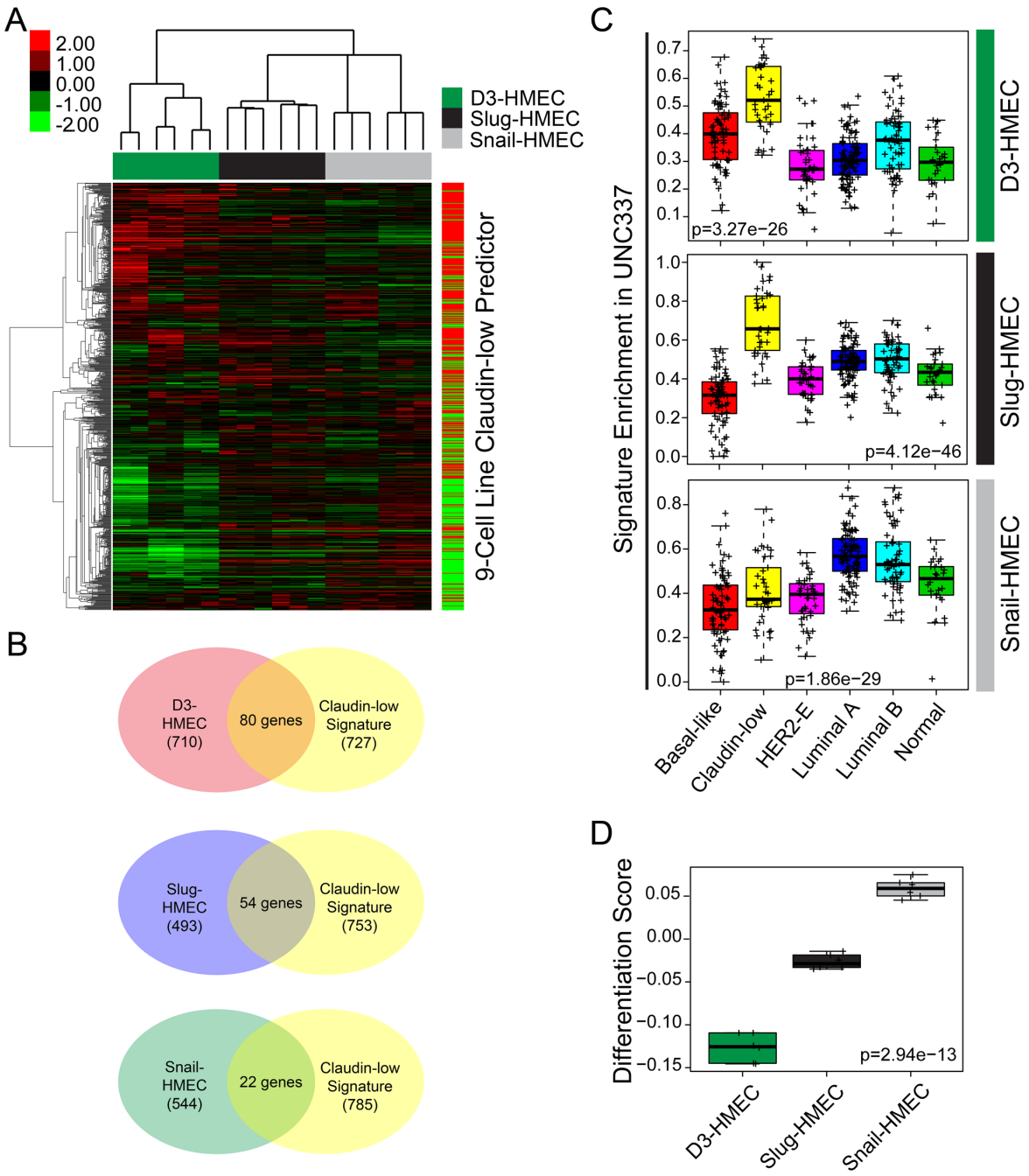



**Figure 4.10: Smarcd3/Baf60c Induces the Claudin-low Gene Signature.**
(A) Gene expression profiles of D3-HMECs cluster with the Claudin-low gene signature. Heat map compares expression of D3-HMECs, Slug-HMECs and Snail-HMECs to the 9-cell line Claudin-low predictor. Red and green indicates upregulated and downregulated genes, respectively. (B) Venn diagram depicts shared genes between D3-, Slug- and Snail-HMECs compared to the Claudin-low predictor. (C) Claudin-low human tumors show the highest expression of D3-HMEC genes among breast cancer subtypes, as demonstrated by the mean expression of D3-, Slug- and

Snail-HMEC genes across the subtypes of breast cancer in the UNC337 data set. p-values were calculated by comparing gene expression means across all subtypes using an ANOVA test. Each + symbol represents a distinct tumor sample within the data set. (D) Differentiation scores compare D3-, Slug- and Snail-HMECs relative to EV-HMECs showing the lowest differentiation propensity of D3-HMECs. p-value was calculated by comparing gene expression means across all breast epithelial cell lineages.

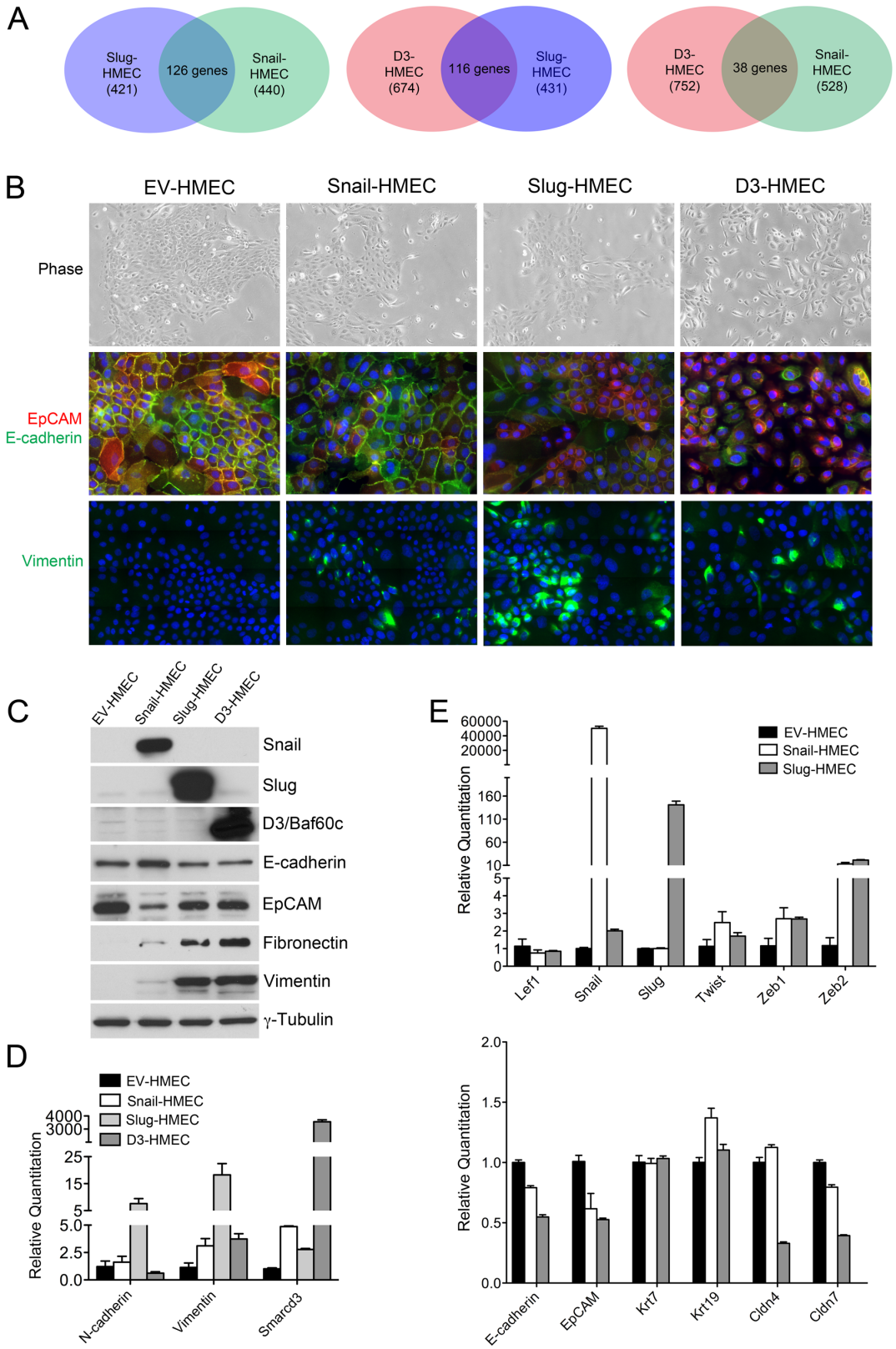



**Figure 4.11: Phenotypic Comparison of Snail and Slug to Smarcd3-Induced EMT**

(A) Venn diagram depicts shared genes between D3-, Slug- and/or Snail-HMEC cells. (B) Loss of the epithelial phenotype with gain of mesenchymal properties in D3-, Slug- and/or Snail-HMEC cells as shown by phase microscopy and immunostaining with nuclear DAPI stain, anti-E-cadherin and anti-EpCAM. (C) Reduced protein expression of epithelial markers EpCAM or E-cadherin and elevated protein expression of mesenchymal markers fibronectin and vimentin in Snail-, Slug- and D3-HMEC cells. (D) Elevated gene expression of mesenchymal markers N-cadherin, vimentin and Smarcd3/Baf60c in Snail-, Slug- and D3-HMECs. (E) Elevated gene expression of EMT-inducing transcription factors and reduced expression of epithelial cell-cell adhesion markers in Snail- and Slug-HMECs. (B-C) Data represent at least two independent experiments. (D-E) Data represent the mean $\pm$ range of three independent experiments performed in triplicate.

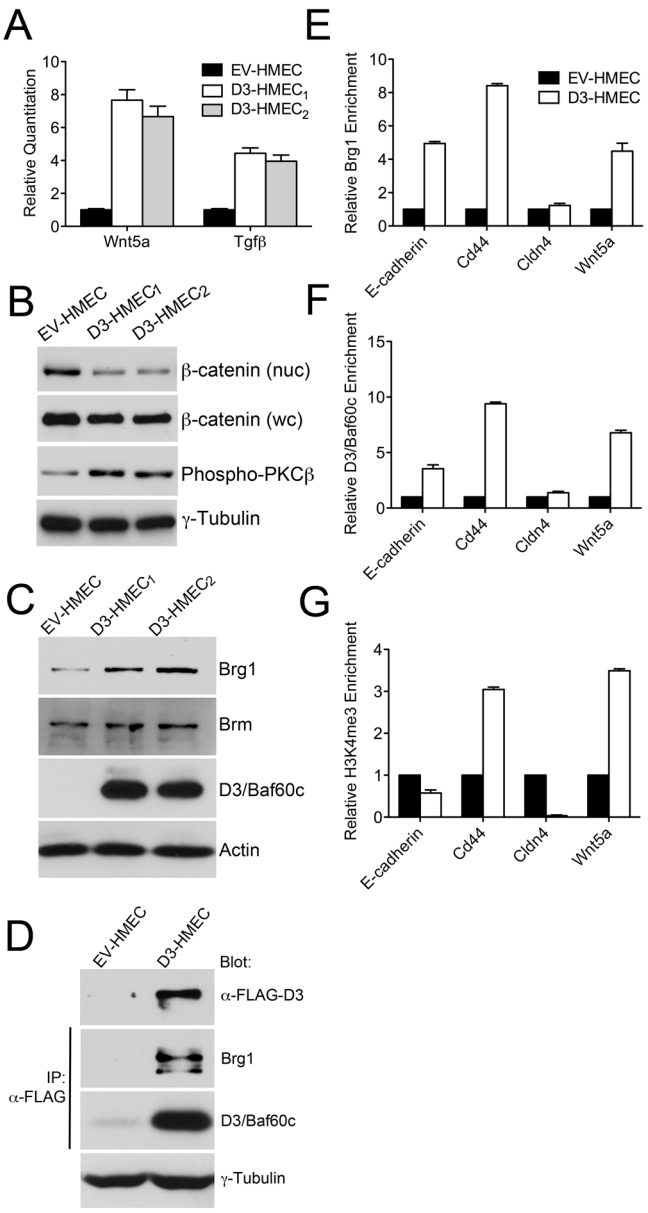



**Figure 4.12: SWI/SNF Complex Regulates Non-Canonical WNT Signaling through Epigenetic Activation of Wnt5a.**

(A) Elevated expression of Wnt5a and Tgfβ in D3-HMECs shown by qRT-PCR. Data are the mean ± the range of two independent experiments performed in triplicate. (B) Activation of non-canonical WNT signaling with inhibition of canonical WNT signaling is shown by elevated phosphorylation of PKCβ and reduced nuclear (nuc) but not whole cell (wc) β-catenin protein expression. (C) Induced Brg1 protein expression with the expression of Smarcd3/Baf60c in HMECs. (D) Brg1 co-immunoprecipitates with FLAG-Smarcd3/Baf60c in HMECs. (E-G) Elevated Smarcd3/Baf60c, Brg1 and H3K4me3 at the promoters of indicated genes in D3-HMECs compared to EV-HMEC cells as measured by ChIP-qRT-PCR. Data are the mean ± the range of three independent experiments performed in duplicate.

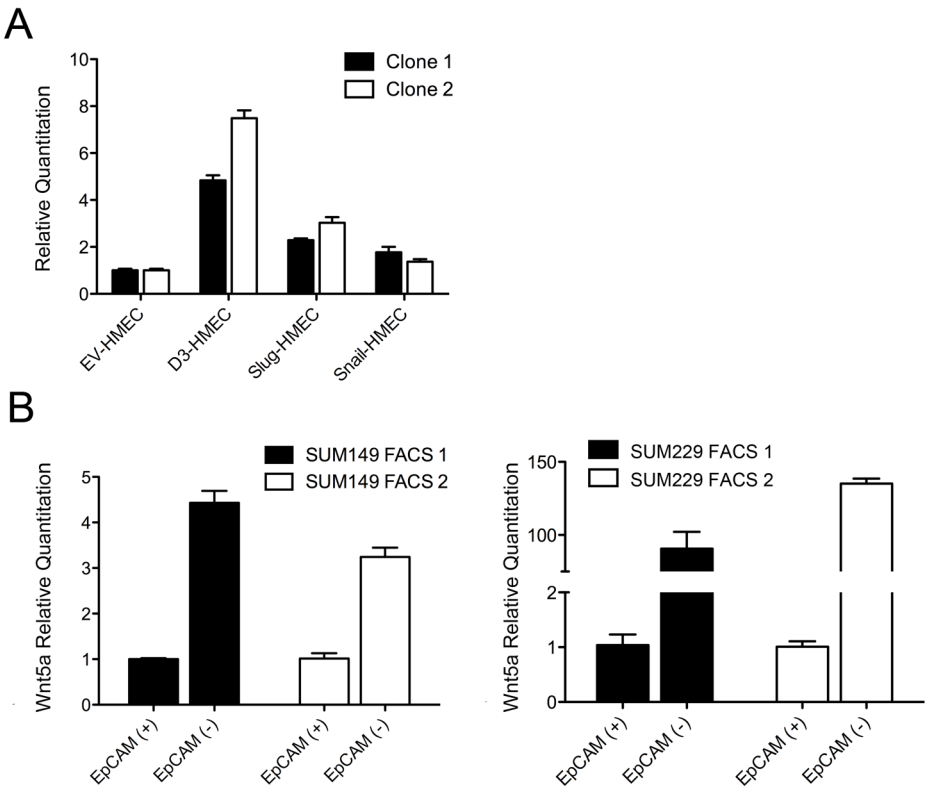



**Figure 4.13: Wnt5a Expression Correlates with the Mesenchymal Phenotype of EpCAM$^{(-)}$ Cells.**

(A) Elevated gene expression of Wnt5a in Slug-HMECs similar to D3-HMECs as shown by qRT-PCR. (B) Elevated gene expression of Wnt5a in EpCAM$^{(-)}$ compared to EpCAM$^{+/high\ cells}$. (A-B) Data represent the mean $\pm$ range of two independent experiments performed in triplicate.

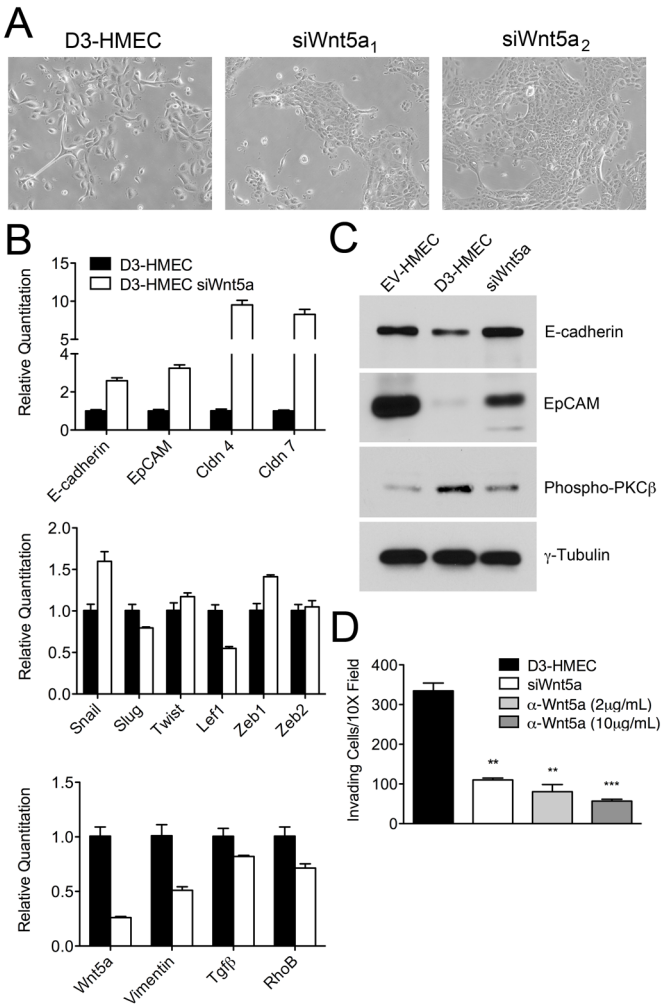



**Figure 4.14: Inhibition of Wnt5a Restores Epithelial Adherens Junctions.**
(A) siRNA-mediated knockdown of Wnt5a partially restores the epithelial phenotype of D3-HMECs as shown by phase microscopy. (B) siRNA-mediated knockdown of Wnt5a increases mRNA expression of epithelial and cell-cell adhesion markers, whereas expression of the EMT-inducing transcription factors and mesenchymal markers is unchanged. Data are the mean $\pm$ the range of two independent experiments performed in triplicate. (C) Increased protein expression of EpCAM and E-cadherin with loss of PKCβ phosphorylation in D3-HMECs transfected with siWnt5a. (D) Reduced invasiveness through growth-factor reduced Matrigel of D3-HMECs treated with siWnt5a or the Wnt5a blocking antibody compared to untreated D3-HMECs. Significance was evaluated by an unpaired Student's T-test. ***p-value < 0.001; **p-value < 0.01. Data are the mean $\pm$ the range of two independent experiments performed in triplicate.

# Chapter V

# Conclusion

## Summary

The studies described herein define the molecular and epigenetic mechanisms responsible for the induction of EMT in epithelial stem cells and Basal-like and Claudin-low subtypes of triple negative breast cancer (TNBC). Trophoblast stem (TS) cells paused in the intermediate stages of EMT allowed the discovery of epigenetic mechanisms regulating EMT in trophoblasts and highlighted an EMT gene signature overlapping with Claudin-low TNBC. Using $TS^{KI4}$ cells, we have shown that MAP3K4 regulates the activity of CBP and that acetylation of histones H2A/H2B is required to maintain the epithelial phenotype of TS cells. Loss of MAP3K4/CBP activity results in the acquisition of an intermediate EMT phenotype associated with simultaneous maintenance of stemness properties and an unrestricted trophoblast invasive potential. During both development and the progression of cancer, hyperinvasion resulting from uncontrolled EMT signaling events can be detrimental to survival. TS cells in the intermediate EMT state have proven to be a unique tool to define novel epigenetic reprogramming events and signaling pathways important for induction of the EMT program.

To understand the parallel mechanisms regulating EMT and the hyperinvasion associated with cancer progression, we identified an intersecting EMT gene signature from analysis of microarray data shared by $TS^{KI4}$ cells developmentally entering EMT and the Claudin-low subtype of TNBC. We designed a phenotypic RNAi screening strategy to define individual genes sufficient to promote the epithelial $EpCAM^{+/high}$ phenotype in a mesenchymal Claudin-low $EpCAM^{(-)}$ subpopulation of SUM149 and SUM229 breast cancer cells. Using this RNAi screening approach, the SWI/SNF chromatin-remodeling factor Smarcd3/Baf60c was identified as a novel epigenetic EMT-regulatory gene. RNAi silencing of Smarcd3/Baf60c in mesenchymal $EpCAM^{(-)}$ breast cancer cells promoted an epithelial phenotype, whereas expression of Smarcd3/Baf60c in HMECs was responsible for the Wnt5a-dependent induction of EMT. These results demonstrate that Smarcd3/Baf60c induces an EMT gene expression program in HMECs similar to that of the Claudin-low breast cancer subtype, suggesting that Smarcd3/Baf60c plays an important role in maintaining the EMT properties and mesenchymal morphology of these breast cancer cells. Due to the intersecting features of developmental and cancer EMT subtypes, a complete picture of this complex cellular program, which is fundamental to development and reactivated in adult tissues during disease progression, should be studied from multiple biological perspectives.

## Significance and Future Directions

### *Defined New Epigenetic Reprogramming Mechanisms for the Induction of EMT*

EMT is a cellular reprogramming event responsible for the morphological and functional conversion of epithelial cells to the mesenchymal phenotype (Thiery *et al*, 2009). Cells undergoing EMT maintain the same genomic background in both mesenchymal and epithelial states. However, during the progression of EMT, the gene expression profile significantly changes, as demonstrated by the repression of epithelial-specific genes and the activation of mesenchymal-specific genes (Nieto, 2011). Without altering the primary DNA sequence, epigenetic mechanisms regulate gene expression levels by physically modifying DNA and/or its associated histone proteins. Epigenetic mechanisms including the acquisition of promoter specific DNA-methylation, expression of miRNAs and the transcription factors Snail, Slug and Twist have been shown to induce EMT (Kouzarides, 2007; Nieto, 2002). *Herein, we have defined epigenetic reprogramming factors controlling histone acetylation and chromatin remodeling as important for the induction of EMT.* Specifically, we have demonstrated that the loss of acetylation on histones H2A/H2B and the increase in SWI/SNF chromatin-remodeling factors Smarcd3/Baf60c and Brg1 constitute two new epigenetic mechanisms important for the regulation of EMT. Rather than affecting the expression of a single gene, these epigenetic changes in H2A/H2B acetylation and SWI/SNF chromatin remodeling factors Smarcd3/Baf60c and Brg1 affect a broad network of genes, which suppresses the epithelial phenotype and induces the mesenchymal phenotype.

The role of histone acetylation and SWI/SNF chromatin-remodeling complexes in reprogramming events relevant to lineage commitment and the reversible processes of EMT and MET are not without precedent. Weinberg and colleagues have shown that the induction of EMT endows HMECs with stem cell-like properties, possibly suggesting a connection between the acquisition of epithelial plasticity and stemness properties (Mani *et al*, 2008; Morel *et al*, 2008). More recently, MET was shown to occur simultaneously with the reprogramming of fibroblasts to iPS cells, connecting the reprogramming events of the MET program to iPS cell generation and establishing MET is an intermediate event essential to generating iPS cells (Li *et al*, 2010). Signaling pathways such as TGFβ that are known to suppress MET reduce the efficiency of the iPS reprogramming process (Li *et al*, 2010; Plath & Lowry, 2011). Furthermore, each of the Yamanaka reprogramming factors makes a specific contribution toward the induction of MET. For example, OCT4 and SOX2 suppress Snail, KLF4 binds and activates E-cadherin, and MYC reduces TGFβ signaling (Li *et al*, 2010). These studies demonstrate a connection between the reprogramming of somatic cells to iPS cells and the induction of MET and EMT programs.

Epigenetic factors controlling histone acetylation and chromatin remodeling provide an additional link to the cellular reprogramming events relevant to iPS production. Proteomics studies reveal that SWI/SNF chromatin-remodeling BAF complex members Brg1 and Baf155 can enhance the reprogramming efficiency of somatic cells to induced pluripotent stem (iPS) cells by functionally replacing the transcription factor c-Myc (Singhal *et al*, 2010). Furthermore, in lineage commitment

events related to neural development, Smarcd3/Baf60c was identified with the acquisition of neural progenitor status (Lessard *et al*, 2007; Lessard & Crabtree, 2010; Ho & Crabtree, 2010). During mouse development, loss of Smarcd3/Baf60c expression produces heart outflow tract defects, resulting from dysregulated EMT (Lickert *et al*, 2004). Our results demonstrated that Smarcd3/Baf60c induces an EMT gene expression program in HMECs similar to that of the Claudin-low breast cancer subtype, suggesting that Smarcd3/Baf60c plays an important role in maintaining the EMT properties and mesenchymal morphology of these cells. Additionally, we showed that Smarcd3/Baf60c is upregulated in a mesenchymal $EpCAM^{(-)}$ subpopulation of Basal-like breast cancer cells when compared to the epithelial $EpCAM^{+/high}$ population, indicating a negative role for the Smarcd/Baf60 family in the progression to a more epithelial cell state. In support of this hypothesis, the SWI/SNF factors from the Smarcd/Baf60 family are silenced at high frequency in Luminal breast cancers, demonstrating the relevance of the Baf60 family to the inhibition of more differentiated epithelial breast cancers (Stephens *et al*, 2012). Conversely, studies linking the importance of histone acetylation to reprogramming demonstrate the requirement for precise temporal and spatial modulation of histone acetylation levels for improved iPS reprogramming efficiency. Treatment of fibroblasts with a narrow range of concentrations of histone deacetylase (HDAC) inhibitors Trichostatin A, sodium butyrate or valproic acid during reprogramming can enhance iPS reprogramming efficiency (Mali *et al*, 2010; Huangfu *et al*, 2008; Liang *et al*, 2010). Furthermore, genetic ablation of class I or II HDACs during mouse development adversely affects differentiation and lineage commitment. Notably, the

knockout mouse for the class I HDAC3 is embryonic lethal due to gastrulation defects, suggesting that this critical developmental EMT event requires functional HDACs to proceed toward completion (Haberland *et al*, 2009). Collectively, these studies demonstrate the importance of precise modulation of histone acetylation and SWI/SNF chromatin-remodeling factors for the control of reprogramming events related to EMT and MET.

Future studies should focus on defining: 1) Additional epigenetic mechanisms contributing to changes in histone acetylation associated with the induction of EMT; and 2) Connections between the changes in histone acetylation and SWI/SNF chromatin-remodeling factors. Unpublished data from Dr. Amy Abell revealed that HDAC6 coordinately regulates the loss of acetylation on many of the same genes affected by the inhibition of MAP3K4/CBP activity. Based upon these results, we hypothesize that the histone acetylation landscape is modulated by multiple epigenetic factors and that the interplay of these factors is essential for establishing optimal histone acetylation levels for maintenance of the epithelial phenotype. An active area of research in the Johnson lab is to use ChIP-seq to identify genes coordinately regulated by CBP and HDAC6 in TS cells and to define how these genes regulate the epithelial phenotype. Furthermore, multiple reports describe crosstalk between histone acetylation marks and chromatin-remodeling complexes (Kouzarides, 2007; Jenuwein, 2001). For example, HATs and HDACs can recruit chromatin-remodeling machinery either directly through physical interactions or indirectly through modulation of the epigenetic landscape. Importantly, histone H2A/H2B dimer exchange is an important step in ATP-dependent SWI/SNF

chromatin-remodeling that can be mediated via histone acetylation (Ito *et al*, 2000). Alternatively, histone acetylation can also negatively impact the ATPase activity of the chromatin-remodeling machinery. Since EMT was induced by selective loss of H2A/H2B acetylation and the increase in expression of chromatin-remodeling factors Smarcd3/Baf60c and Brg1, we hypothesize that H2A/H2B acetylation and Smarcd3-dependent SWI/SNF chromatin-remodeling occur sequentially, whereby changes in histone acetylation can promote chromatin-remodeling events. Future studies might examine the connection between loss of H2A/H2B acetylation and increased ATP-dependent chromatin-remodeling factors Smarcd3/Baf60c and Brg1. Using TS cell models, $TS^{KI4}$, $TS^{shCBP}$ and $TS^{shHDAC6}$, we can examine the effect of selective loss of H2A/H2B acetylation on chromatin-remodeling by micrococcal nuclease digestion coupled to high throughput sequencing, enabling the study of nucleosome occupancy on a genome-wide scale. These studies could provide insight into the crosstalk between the epigenetic modulation of histone acetylation and SWI/SNF chromatin-remodeling events that induce EMT.

***Defined the Metastable Phenotype in Developmental EMT***

The intermediate stage of EMT, coined the metastable phenotype, describes the simultaneous existence of both epithelial and mesenchymal characteristics and is of great importance for understanding the cellular changes associated with progression of the EMT program and maintenance of stem cell-like properties (Lee, 2006). The metastable EMT phenotype is defined by enhanced cellular invasiveness with the coexistence of stem cell properties, making this intermediate stage of EMT particularly important to study with respect to cancer metastasis (Jordan *et al*, 2011).

Recent studies reveal that in the malignant progression of cancer, there exists a partial rather than a complete cellular morphology transition with the invading front of most metastatic tumors moving as epithelial sheets rather than as individual cells, suggesting the importance of the intermediate EMT state to cancer metastasis (Friedl & Wolf, 2003; Brabletz *et al*, 2005). Due to the tremendous difficulty in capturing cells in the intermediate states of EMT, most studies focus on the initiation or completion of EMT. *To our knowledge, this is the first report that characterizes the intermediate stages of developmental EMT, using multiple TS cell models to track the metastable EMT continuum.* Furthermore, the EpCAM$^{(-)}$ Basal-like SUM149 and SUM229 breast cancer cells capable of converting between mesenchymal and epithelial states might also represent a metastable EMT phenotype, since these cells are also characterized by enhanced cellular plasticity associated with the intermediate stage of EMT (Gupta *et al*, 2011; Proia *et al*, 2011). The EpCAM$^{(-)}$ SUM149 and SUM229 breast cancer cells have significant mesenchymal Claudin-low features, but these cells do not profile as completely Claudin-low by gene array analysis, suggesting that they might be in an intermediate stage of EMT rather than fully mesenchymal.

The coexistence of EMT and stem cell properties makes the intermediate stage of EMT important to fully understand as a state distinctly different from the beginning epithelial or ending mesenchymal states typically studied. It is hypothesized that the intermediate stage of EMT is the only fully reversible point on the EMT continuum and therefore represents the point with greatest cellular plasticity and cellular invasiveness, attributing to the acquisition of enhanced stem-

like and metastatic properties. Future studies should focus on further molecular characterization of the metastable EMT phenotype in both development and cancer. The following broad questions are important to address: 1) What are the molecular and epigenetic signaling pathways important for inducing the progression of the EMT continuum from epithelial to intermediate to mesenchymal stages; and 2) Do invasive cancer cells resemble the metastable phenotype characteristic of the intermediate stage of EMT? Using TS cell models, including $TS^{KI4}$, $TS^{shCBP}$, $TS^{Snail}$ cells, to represent the developmental EMT program and Basal-like breast cancer cells simultaneously expressing mesenchymal $EpCAM^{(-)}$ and epithelial $EpCAM^{+/high}$ states to represent the breast cancer EMT program, we can examine the signaling pathways controlling the progression of EMT. Since the TS cell models indicate the importance of kinase signaling in promoting the epigenetic changes associated with the metastable EMT phenotype, we would initially examine the connections between broader kinase signaling networks and the intermediate EMT state using the multiplexed inhibitor bead/mass spectrometry (MIB/MS) technology developed in the Johnson lab (Duncan *et al*, 2012). Following, *in vitro*, studies with our cell culture systems, we could extend these experiments to human-in-mouse tumors exhibiting morphologically distinct features similar to Basal-like SUM149 and SUM229 breast cancer cells. Collectively, these studies might provide insight into the connections between the metastable EMT phenotype and the metastatic progression of cancer. It is our hypothesis that cancer cells become metastatic upon reaching the intermediate stage of EMT.

***Identified Genes Sufficient to Induce EMT in Development and Breast Cancer***

EMT has largely been studied from the singular perspective of individual model systems, including chick and frog models of neural crest development, fibrotic kidney cell lines, metastatic cancer cell lines, fibroblasts or epithelial cell lines transduced with factors such as TGFβ, Snail or Twist (Yang & Weinberg, 2008; Thiery *et al*, 2009). *To our knowledge, the studies described herein are the first to examine EMT from two distinct biological perspectives and EMT subtypes.* From the analysis of gene array data, we identified a statistically significant intersecting EMT gene signature shared by epithelial stem cells developmentally entering EMT and the Claudin-low subtype of TNBC (Abell *et al*, 2011). Using SUM149 and SUM229 Basal-like breast cancer cells, we designed a phenotypic RNAi screening strategy to identify individual genes sufficient to promote the $EpCAM^{+/high}$ epithelial phenotype in the mesenchymal Claudin-low $EpCAM^{(-)}$ population of breast cancer cells. These results demonstrate that parallel mechanisms control EMT during development and breast cancer.

In addition to Smarcd3/Baf60c, the phenotypic RNAi screen identified nine additional genes sufficient to promote the $EpCAM^{+/high}$ epithelial phenotype in the mesenchymal Claudin-low $EpCAM^{(-)}$ population of breast cancer cells. With the exception of Slug, N-cadherin, and the receptor tyrosine kinase Met, the remaining genes are uncharacterized with respect to EMT. Therefore, future studies should examine the underlying mechanisms contributing to EMT by molecular characterization of the undefined EMT-regulatory genes identified in the RNAi screen. Among these uncharacterized EMT-regulatory genes, Rnf130 and Fhl1 are

interesting genes to initially study. Rnf130 is a putative E3 ubiquitin ligase preferentially expressed in the mesoderm during development (Guais *et al*, 2006). It is hypothesized that Rnf130 may regulate apoptosis in hematopoietic cells, but its function is largely unknown. Preliminary findings from our group have demonstrated that combined knockdown of Rnf130, Slug and Smarcd3/Baf60c produced subtle synergistic effects in promoting an epithelial morphology in fully mesenchymal Claudin-low SUM159 breast cancer cells. These results indicate that Rnf130, Slug and Smarcd3 might be part of a common EMT-regulatory signaling network. Additionally, Fhl1 is another gene identified in the EMT-regulatory RNAi screen. Fhl1 is a zinc finger protein that localizes to focal adhesions to promote cell spreading and binds Smad proteins to promote TGFβ-responsive transcription (Rafael *et al*, 2012; Sheikh *et al*, 2008; Ding *et al*, 2009). Furthermore, the closely-related family member Fhl2 has been directly connected to EpCAM expression through direct interaction with the intercellular domain of EpCAM (Maetzel *et al*, 2009). Lastly, in $TS^{KI4}$ cells, we have shown knockdown of Fhl1 to be capable of reversing the mesenchymal phenotype. Therefore, Rnf130 and Fhl1 are ideal gene candidates to mechanistically study with respect to EMT.

***Demonstrated the Reversible Nature of EMT in Breast Cancer***

While it has long been accepted that EMT is a reversible process during development, few studies have examined the opposite process of mesenchymal-epithelial transition (MET) in the progression of breast cancer. *In addition to identifying ten genes important for regulating the phenotypic transitions between mesenchymal and epithelial cellular morphologies in breast cancer, the phenotypic*

*RNAi screening approach described herein highlights the reversible nature of EMT and MET in SUM149 and SUM229 breast cancer cells.* To date, EMT is accepted as a morphological program controlling tumor cell invasion into the surrounding stroma and intravasation into the bloodstream (Valastyan & Weinberg, 2011). Although it is hypothesized that MET controls metastatic colonization of invasive tumor cells to distant organs, the process of MET has not been definitively connected with cancer metastasis. In support of this hypothesis, unpublished work from Dr. Aleix Prat demonstrates that injection of mesenchymal $EpCAM^{(-)}$ SUM149 and SUM229 cells into the mammary fat pads of immunocompromised mice resulted in the formation of $EpCAM^{+/high}$ epithelial tumors, possibly suggesting that the mesenchymal $EpCAM^{(-)}$ population converts to an epithelial $EpCAM^{+/high}$ population for tumor formation (unpublished data from A. Prat).

Therefore, future studies should focus on understanding the role of MET in the metastatic progression of cancer. Currently, it is hypothesized that the primary tumor and secondary metastatic tumor are fully epithelial, while the tumor cells invading into the surrounding stroma and the circulating tumor cells (CTCs) are mesenchymal (Rhim *et al*, 2012; Yu *et al*, 2012). Using xenograft mouse models and metastatic breast cancer cells transduced with epithelial- and mesenchymal-specific fluorescent reporter genes, we could track the processes of EMT and MET during the, *in vivo*, progression of breast cancer. With this model system, we could examine the primary tumor, CTCs and the secondary metastasis for the presence of epithelial or mesenchymal markers in conjunction with fluorescent reporter expression. Alternatively, we could use human-in-mouse tumor models to study the epithelial

and/or mesenchymal nature of the primary tumor, CTCs and the secondary metastasis by marker analysis. By understanding the role of EMT and MET in the progression of cancer, we can begin to develop treatment strategies for cancer metastasis using our knowledge of the cellular signaling pathways controlling these programs.

## References


Abell AN, Granger DA, Johnson NL, Vincent-Jordan N, Dibble CF & Johnson GL (2009) Trophoblast Stem Cell Maintenance by Fibroblast Growth Factor 4 Requires MEKK4 Activation of Jun N-Terminal Kinase. *Mol. Cell. Biol.* **29:** 2748–2761

Abell AN, Jordan NV, Huang W, Prat A, Midland AA, Johnson NL, Granger DA, Mieczkowski PA, Perou CM, Gomez SM, Li L & Johnson GL (2011) MAP3K4/CBP-Regulated H2B Acetylation Controls Epithelial-Mesenchymal Transition in Trophoblast Stem Cells. *Cell Stem Cell* **8:** 525–537

Abell AN, Rivera-Perez JA, Cuevas BD, Uhlik MT, Sather S, Johnson NL, Minton SK, Lauder JM, Winter-Vann AM, Nakamura K, Magnuson T, Vaillancourt RR, Heasley LE & Johnson GL (2005) Ablation of MEKK4 Kinase Activity Causes Neurulation and Skeletal Patterning Defects in the Mouse Embryo. *Mol. Cell. Biol.* **25:** 8948–8959

Acloque H, Adams MS, Fishwick K, Bronner-Fraser M & Nieto MA (2009) Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *Journal of Clinical Investigation* **119:** 1438–1449

Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ & Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 3983–3988

Ansieau S, Bastid J, Doreau A, Morel A-P, Bouchet BP, Thomas C, Fauvet F, Puisieux I, Doglioni C, Piccinin S, Maestro R, Voeltzel T, Selmi A, Valsesia-Wittmann S, Caron de Fromentel C & Puisieux A (2008) Induction of EMT by Twist Proteins as a Collateral Effect of Tumor-Promoting Inactivation of Premature Senescence. *Cancer Cell* **14:** 79–89

Arima T, Hata K, Tanaka S, Kusumi M, Li E, Kato K, Shiota K, Sasaki H & Wake N (2006) Loss of the maternal imprint in Dnmt3Lmat−/− mice leads to a differentiation defect in the extraembryonic tissue. *Developmental Biology* **297:** 361–373

Banerjee S (2006) Basal-like breast carcinomas: clinical outcome and response to chemotherapy. *Journal of Clinical Pathology* **59:** 729–735

Banine F (2005) SWI/SNF Chromatin-Remodeling Factors Induce Changes in DNA Methylation to Promote Transcriptional Activation. *Cancer Research* **65:** 3542–3547

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL & Lander ES (2006) A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125:** 315–326

Berx G, Raspé E, Christofori G, Thiery JP & Sleeman JP (2007) Pre-EMTing metastasis? Recapitulation of morphogenetic processes in cancer. *Clin Exp Metastasis* **24:** 587–597

Bissell MJ & Radisky D (2001) Putting tumours in context. *Nat Rev Cancer* **1:** 46–54

Bonnomet A, Brysse A, Tachsidis A, Waltham M, Thompson EW, Polette M & Gilles C (2010) Epithelial-to-Mesenchymal Transitions and Circulating Tumor Cells. *J Mammary Gland Biol Neoplasia* **15:** 261–273

Bourc'his D (2001) Dnmt3L and the Establishment of Maternal Genomic Imprints. *Science* **294:** 2536–2539

Brabletz T, Jung A, Reu S, Porzner M, Hlubek F, Kunz-Schughart LA, Knuechel R & Kirchner T (2001) Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proc. Natl. Acad. Sci. U.S.A.* **98:** 10356–10361

Brabletz T, Jung A, Spaderna S, Hlubek F & Kirchner T (2005) Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nat Rev Cancer* **5:** 744–749

Bultman S, Gebuhr T, Yee D, La Mantia C, Nicholson J, Gilliam A, Randazzo F, Metzger D, Chambon P, Crabtree G & Magnuson T (2000) A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Molecular Cell* **6:** 1287–1295

Caldeira J, Prando ÉC, Quevedo FC, Neto F, Rainho CA & Rogatto SR (2006) CDH1 promoter hypermethylation and E-cadherin protein expression in infiltrating breast cancer. *BMC Cancer* **6:** 48--57

Cano A, Perez-Moreno MA, Rodrigo I, Locascio A, Blanco MJ, del Barrio MG, Portillo F & Nieto MA (2000) The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* **2:** 76–83

Chan JA (2005) MicroRNA-21 Is an Antiapoptotic Factor in Human Glioblastoma Cells. *Cancer Research* **65:** 6029–6033

Christiansen JJ (2006) Reassessing Epithelial to Mesenchymal Transition as a Prerequisite for Carcinoma Invasion and Metastasis. *Cancer Research* **66:** 8319–8326

Chuang TH, Xu X, Kaartinen V, Heisterkamp N, Groffen J & Bokoch GM (1995) Abr and Bcr are multifunctional regulators of the Rho GTP-binding protein family. *Proc. Natl. Acad. Sci. U.S.A.* **92:** 10282–10286

Derksen PWB, Liu X, Saridin F, van der Gulden H, Zevenhoven J, Evers B, van Beijnum JR, Griffioen AW, Vink J, Krimpenfort P, Peterse JL, Cardiff RD, Berns A & Jonkers J (2006) Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell* **10:** 437–449

Ding L, Wang Z, Yan J, Yang X, Liu A, Qiu W, Zhu J, Han J, Zhang H, Lin J, Cheng L, Qin X, Niu C, Yuan B, Wang X, Zhu C, Zhou Y, Li J, Song H, Huang C, et al (2009) Human four-and-a-half LIM family members suppress tumor cell growth through a TGF-β–like signaling pathway. *Journal of Clinical Investigation* **119**: 349--361

Dissanayake SKS, Wade MM, Johnson CEC, O'Connell MPM, Leotlela PDP, French ADA, Shah KVK, Hewitt KJK, Rosenthal DTD, Indig FEF, Jiang YY, Nickoloff BJB, Taub DDD, Trent JMJ, Moon RTR, Bittner MM & Weeraratna ATA (2007) The Wnt5A/protein kinase C pathway mediates motility in melanoma cells via the inhibition of metastasis suppressors and initiation of an epithelial to mesenchymal transition. *Journal of Biological Chemistry* **282:** 17259–17271

Dominguez MG, Hughes VC, Pan L, Simmons M, Daly C, Anderson K, Noguera-Troise I, Murphy AJ, Valenzuela DM, Davis S, Thurston G, Yancopoulos GD & Gale NW (2007) Vascular endothelial tyrosine phosphatase (VE-PTP)-null mice undergo vasculogenesis but die embryonically because of defects in angiogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **104:** 3243–3248

Dong M, How T, Kirkbride KC, Gordon KJ, Lee JD, Hempel N, Kelly P, Moeller BJ, Marks JR & Blobe GC (2007) The type III TGF-β receptor suppresses breast cancer progression. *Journal of Clinical Investigation* **117:** 206–217

Duncan JS, Whittle MC, Nakamura K, Abell AN, Midland AA, Zawistowski JS, Johnson NL, Granger DA, Jordan NV, Darr DB, Usary J, Kuan P-F, Smalley DM, Ben Major, He X, Hoadley KA, Zhou B, Sharpless NE, Perou CM, Kim WY, et al (2012) Dynamic Reprogramming of the Kinome in Response to Targeted MEK Inhibition in Triple-Negative Breast Cancer. *Cell* **149:** 307–321

Eckert MA, Lwin TM, Chang AT, Kim J, Danis E, Ohno-Machado L & Yang J (2011) Twist1-Induced Invadopodia Formation Promotes Tumor Metastasis. *Cancer Cell* **19:** 372–386

Eisen MB, Spellman PT, Brown PO & Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U.S.A.* **95:** 14863–14868

Erlebacher A, Price KA & Glimcher LH (2004) Maintenance of mouse trophoblast stem cell proliferation by TGF-β/activin. *Developmental Biology* **275:** 158–169

Esteller M (2008) Epigenetics in cancer. *N. Engl. J. Med.* **358:** 1148–1159

Ferretti C, Bruni L, Dangles-Marie V, Pecking AP & Bellet D (2006) Molecular circuits shared by placental and cancer cells, and their implications in the proliferative, invasive and migratory capacities of trophoblasts. *Human Reproduction Update* **13:** 121–141

Floor S, van Staveren WCG, Larsimont D, Dumont JE & Maenhaut C (2011) Cancer cells in epithelial-to-mesenchymal transition and tumor-propagating–cancer stem cells: distinct, overlapping or same populations. **30:** 4609–4621

Forcales SV, Albini S, Giordani L, Malecova B, Cignolo L, Chernov A, Coutinho P, Saccone V, Consalvi S, Williams R, Wang K, Wu Z, Baranovskaya S, Miller A, Dilworth FJ & Puri PL (2011) Signal-dependent incorporation of MyoD–BAF60c into Brg1-based SWI/SNF chromatin-remodelling complex. *EMBO J.* **2**: 301–316

Friedl P & Wolf K (2003) Tumour-cell invasion and migration: diversity and escape mechanisms. *Nat Rev Cancer* **3:** 362–374

Fulford LG, Reis-Filho JS, Ryder K, Jones C, Gillett CE, Hanby A, Easton DF & Lakhani SR (2007) Basal-like grade III invasive ductal carcinoma of the breast: patterns of metastasis and long term survival. *Breast Cancer Res* **9:** R4

Gaedcke J, Traub F, Milde S, Wilkens L, Stan A, Ostertag H, Christgen M, Wasielewski von R & Kreipe HH (2007) Predominance of the basal type and HER-2/neu type in brain metastasis from breast cancer. *Mod Pathol* **20:** 864–870

Garzon R, Calin GA & Croce CM (2009) MicroRNAs in Cancer. *Annu. Rev. Med.* **60:** 167–179

Goldman-Wohl D & Yagel S (2002) Regulation of trophoblast invasion: from normal implantation to pre-eclampsia. *Molecular and Cellular Endocrinology* **187:** 233–238

Gonzales DS, Jones JM, Pinyopummintr T, Carnevale EM, Ginther OJ, Shapiro SS & Bavister BD (1996) Trophectoderm projections: a potential means for locomotion, attachment and implantation of bovine, equine and human blastocysts. *Hum. Reprod.* **11:** 2739–2745

Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y & Goodall GJ (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **10:** 593–601

Grünert S, Jechlinger M & Beug H (2003) Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. *Nat. Rev. Mol. Cell Biol.* **4:** 657–665

Guais A, Siegrist S, Solhonne B, Jouault H, Guellaën G & Bulle F (2006) h-Goliath, paralog of GRAIL, is a new E3 ligase protein, expressed in human leukocytes. *Gene* **374:** 112–120

Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C & Lander ES (2011) Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells. *Cell* **146:** 633–644

Haberland M, Montgomery RL & Olson EN (2009) The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet* **10:** 32–42

Hajra KM, Chen DY-S & Fearon ER (2002) The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Research* **62:** 1613–1618

Hartwell KA, Muir B, Reinhardt F, Carpenter AE, Sgroi DC & Weinberg RA (2006) The Spemann organizer gene, Goosecoid, promotes tumor metastasis. *Proc. Natl. Acad. Sci. U.S.A.* **103:** 18969–18974

Hata K, Okano M, Lei H & Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129:** 1983–1993

Hay ED (2005) The mesenchymal cell, its role in the embryo, and the remarkable signaling mechanisms that create it. *Dev. Dyn.* **233:** 706–720

Hemberger M (2007) Epigenetic landscape required for placental development. *Cell. Mol. Life Sci.* **64:** 2422–2436

Hemberger M (2010) Genetic-epigenetic intersection in trophoblast differentiation: implications for extraembryonic tissue function. *Epigenetics* **5:** 24–29

Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, Fridlyand J, Sahin A, Agarwal R, Joy C, Liu W, Stivers D, Baggerly K, Carey M, Lluch A, Monteagudo C, He X, Weigman V, Fan C, Palazzo J, et al (2009) Characterization of a Naturally Occurring Breast Cancer Subset Enriched in Epithelial-to-Mesenchymal Transition and Stem Cell Characteristics. *Cancer Research* **69:** 4116–4124

Ho L & Crabtree GR (2010) Chromatin remodelling during development. *Nature* **463:** 474–484

Holder N & Klein R (1999) Eph receptors and ephrins: effectors of morphogenesis. *Development* **126:** 2033–2044

Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Orrico A, Dreher D, et al (2006) The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7:** 96--108

Huang W, Umbach DM, Vincent-Jordan N, Abell AN, Johnson GL & Li L (2011) Efficiently identifying genome-wide changes with next-generation sequencing data. *Nucleic Acids Research* **39:** e130–e130

Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE & Melton DA (2008) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol* **26:** 795–797

Huber MA (2004) NF- B is essential for epithelial-mesenchymal transition and metastasis in a model of breast cancer progression. *Journal of Clinical Investigation* **114:** 569–581

Huber MA, Kraut N & Beug H (2005) Molecular requirements for epithelial–mesenchymal transition during tumor progression. *Current Opinion in Cell Biology* **17:** 548–558

Ito T, Ikehara T, Nakagawa T, Kraus WL & Muramatsu M (2000) p300-mediated acetylation facilitates the transfer of histone H2A–H2B dimers from nucleosomes to a histone chaperone. *Genes & Development* **14:** 1899–1907

Jechlinger M (2006) Autocrine PDGFR signaling promotes mammary cancer metastasis. *Journal of Clinical Investigation* **116:** 1561–1570

Jenuwein T (2001) Translating the Histone Code. *Science* **293:** 1074–1080

Johnson SM, Großhans H, Shingara J, Byrom M, Jarvis R, Cheng A, Labourier E, Reinert KL, Brown D & Slack FJ (2005) RAS Is Regulated by the let-7 MicroRNA Family. *Cell* **120:** 635–647

Jordan NV, Johnson GL & Abell AN (2011) Tracking the intermediate stages of epithelial-mesenchymal transition in epithelial stem cells and cancer. *Cell Cycle* **10:** 2865–2873

Kalantry S, Mills KC, Yee D, Otte AP, Panning B & Magnuson T (2006) The Polycomb group protein Eed protects the inactive X-chromosome from differentiation-induced reactivation. *Nat. Cell Biol.* **8:** 195–202

Kalluri R (2009) EMT: When epithelial cells decide to become mesenchymal-like cells. *Journal of Clinical Investigation* **119:** 1417–1419

Kalluri R & Weinberg RA (2009) The basics of epithelial-mesenchymal transition. *Journal of Clinical Investigation* **119:** 1420–1428

Khong TY (2008) The pathology of placenta accreta, a worldwide epidemic. *Journal of Clinical Pathology* **61:** 1243–1246

Kidder BL, Palmer S & Knott JG (2009) SWI/SNF-Brg1 Regulates Self-Renewal and Occupies Core Pluripotency-Related Genes in Embryonic Stem Cells. *Stem Cells* **27:** 317–328

Kouzarides T (2007) Chromatin Modifications and Their Function. *Cell* **128:** 693–705

Larue L, Ohsugi M, Hirchenhain J & Kemler R (1994) E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proc. Natl. Acad. Sci. U.S.A.* **91:** 8263–8267

Lee JM (2006) The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *The Journal of Cell Biology* **172:** 973–981

Lessard J, Wu JI, Ranish JA, Wan M, Winslow MM, Staahl BT, Wu H, Aebersold R, Graef IA & Crabtree GR (2007) An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development. *Neuron* **55:** 201–215

Lessard JA & Crabtree GR (2010) Chromatin Regulatory Mechanisms in Pluripotency. *Annu. Rev. Cell Dev. Biol.* **26:** 503–532

Li R, Liang J, Ni S, Zhou T, Qing X, Li H, He W, Chen J, Li F, Zhuang Q, Qin B, Xu J, Li W, Yang J, Gan Y, Qin D, Feng S, Song H, Yang D, Zhang B, et al (2010) A Mesenchymal-to-Epithelial Transition Initiates and Is Required for the Nuclear Reprogramming of Mouse Fibroblasts. *Stem Cell* **7:** 51–63

Liang G, Taranova O, Xia K & Zhang Y (2010) Butyrate Promotes Induced Pluripotent Stem Cell Generation. *Journal of Biological Chemistry* **285:** 25516–25521

Lickert H, Takeuchi JK, Both Von I, Walls JR, McAuliffe F, Adamson SL, Henkelman RM, Wrana JL, Rossant J & Bruneau BG (2004) Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* **432:** 107–112

Lim E, Vaillant F, Di Wu, Forrest NC, Pal B, Hart AH, Asselin-Labat M-L, Gyorki DE, Ward T, Partanen A, Feleppa F, Huschtscha LI, Thorne HJ, kConFab, Fox SB, Yan M, French JD, Brown MA, Smyth GK, Visvader JE, et al (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15**: 907–915

Maetzel D, Denzel S, Mack B, Canis M, Went P, Benk M, Kieu C, Papior P, Baeuerle PA, Munz M & Gires O (2009) Nuclear signalling by tumour-associated antigen EpCAM. *Nat. Cell Biol.* **11:** 162–171

Mager J, Montgomery ND, de Villena FP-M & Magnuson T (2003) Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nat Genet* **33:** 502–507

Mali P, Chou B-K, Yen J, Ye Z, Zou J, Dowey S, Brodsky RA, Ohm JE, Yu W, Baylin SB, Yusa K, Bradley A, Meyers DJ, Mukherjee C, Cole PA & Cheng L (2010) Butyrate Greatly Enhances Derivation of Human Induced Pluripotent Stem Cells by Promoting Epigenetic Remodeling and the Expression of Pluripotency-Associated Genes. *Stem Cells* **28:** 713–720

Mani SA, Guo W, Liao M-J, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M, Campbell LL, Polyak K, Brisken C, Yang J & Weinberg RA (2008) The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. *Cell* **133:** 704–715

Mani SA, Yang J, Brooks M, Schwaninger G, Zhou A, Miura N, Kutok JL, Hartwell K, Richardson AL & Weinberg RA (2007) Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers. *Proc. Natl. Acad. Sci. U.S.A.* **104:** 10069–10074

Maupin KA, Sinha A, Eugster E, Miller J, Ross J, Paulino V, Keshamouni VG, Tran N, Berens M, Webb C & Haab BB (2010) Glycogene Expression Alterations Associated with Pancreatic Cancer Epithelial-Mesenchymal Transition in Complementary Model Systems. *PLoS ONE* **5:** e13002

McMaster MT, Zhou Y & Fisher SJ (2004) Abnormal placentation and the syndrome of preeclampsia. *Seminars in Nephrology* **24:** 540–547

Moll R, Mitze M, Frixen UH & Birchmeier W (1993) Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. *The American Journal of Pathology* **143:** 1731–1742

Moody SE, Perez D, Pan T-C, Sarkisian CJ, Portocarrero CP, Sterner CJ, Notorfrancesco KL, Cardiff RD & Chodosh LA (2005) The transcriptional repressor Snail promotes mammary tumor recurrence. *Cancer Cell* **8:** 197–209

Morel A-P, Lièvre M, Thomas C, Hinkal G, Ansieau S & Puisieux A (2008) Generation of Breast Cancer Stem Cells through Epithelial-Mesenchymal Transition. *PLoS ONE* **3:** e2888

Mori M, Murata Y, Kotani T, Kusakari S, Ohnishi H, Saito Y, Okazawa H, Ishizuka T, Mori M & Matozaki T (2010) Promotion of cell spreading and migration by vascular endothelial-protein tyrosine phosphatase (VE-PTP) in cooperation with integrins. *J. Cell. Physiol.* **224**: 195--204

Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe J-P, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo W-L, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, et al (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10:** 515–527

Ng RK, Dean W, Dawson C, Lucifero D, Madeja Z, Reik W & Hemberger M (2008) Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nat. Cell Biol.* **10:** 1280–1290

Nguyen DX, Bos PD & Massagué J (2009) Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer* **9:** 274–284

Nieto MA (2002) THE SNAIL SUPERFAMILY OF ZINC-FINGER TRANSCRIPTION FACTORS. *Nat. Rev. Mol. Cell Biol.* **3:** 155–166

Nieto MA (2011) The Ins and Outs of the Epithelial to Mesenchymal Transition in Health and Disease. *Annu. Rev. Cell Dev. Biol.* **27:** 347–376

Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R & Rossant J (2005) Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. *Cell* **123:** 917–929

Norwitz ER, Schust DJ & Fisher SJ (2001) Implantation and the survival of early pregnancy. *N. Engl. J. Med.* **345:** 1400–1408

O'Carroll D, Erhardt S, Pagani M, Barton SC, Surani MA & Jenuwein T (2001) The Polycomb-Group Gene Ezh2 Is Required for Early Mouse Development. *Mol. Cell. Biol.* **21:** 4330–4336

Odiatis C & Georgiades P (2010) New Insights for Ets2 Function in Trophoblast using Lentivirus-Mediated Gene Knockdown in Trophoblast Stem Cells. *Placenta* **31:** 630–640

Pafilis J, Batistatou A, Iliopoulou A, Tsanou E, Bakogiannis A, Dassopoulos G & Charalabopoulos K (2007) Expression of adhesion molecules during normal pregnancy. *Cell Tissue Res* **329:** 1–11

Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E & Helin K (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* **23:** 4061–4071

Pasquale EB (2010) Eph receptors and ephrins in cancer: bidirectional signalling and beyond. Nature Rev Cancer **10**: 165–180

Peinado H, Olmeda D & Cano A (2007) Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* **7:** 415–428

Perl AK, Wilgenbus P, Dahl U, Semb H & Christofori G (1998) A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature* **392:** 190–193

Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO & Botstein D (2000) Molecular portraits of human breast tumours. *Nature* **406:** 747–752

Perry JK, Lins RJ, Lobie PE & Mitchell MD (2009) Regulation of invasive growth: similar epigenetic mechanisms underpin tumour progression and implantation in human pregnancy. *Clin. Sci.* **118:** 451–457

Plath K & Lowry WE (2011) Progress in understanding reprogramming to the induced pluripotent state. *Nat Rev Genet* **12:** 253–265

Prat A & Perou CM (2011) Deconstructing the molecular portraits of breast cancer. *Molecular Oncology* **5:** 5–23

Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X & Perou CM (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12:** R68

Proia TA, Keller PJ, Gupta PB, Klebba I, Jones AD, Sedic M, Gilmore H, Tung N, Naber SP, Schnitt S, Lander ES & Kuperwasser C (2011) Genetic Predisposition Directs Breast Cancer Phenotype by Dictating Progenitor Cell Fate. *Stem Cell* **8:** 149–163

Rafael MS, Laizé V, Florindo C, Ferraresso S, Bargelloni L & Cancela ML (2012) Overexpression of four and a half LIM domains protein 2 promotes epithelial-mesenchymal transition-like phenotype in fish pre-osteoblasts. *Biochimie* **94:** 1128–1134

Rakha E, Ellis I & Reis-Filho J (2008a) Are Triple-Negative and Basal-Like Breast Cancer Synonymous? *Clinical Cancer Research* **14:** 618–618

Rakha EA, Reis-Filho JS & Ellis IO (2008b) Basal-Like Breast Cancer: A Critical Review. *Journal of Clinical Oncology* **26:** 2568–2581

Ralston A & Rossant J (2008) Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. *Developmental Biology* **313:** 614–629

Ray ME, Wistow G, Su YA, Meltzer PS & Trent JM (1997) AIM1, a novel non-lens member of the betagamma-crystallin superfamily, is associated with the control of tumorigenicity in human malignant melanoma. *Proc. Natl. Acad. Sci. U.S.A.* **94:** 3229–3234

Rebel VI, Kung AL, Tanner EA, Yang H, Bronson RT & Livingston DM (2002) Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal. *Proc. Natl. Acad. Sci. U.S.A.* **99:** 14789–14794

Reya T, Morrison SJ, Clarke MF & Weissman IL (2001) Stem cells, cancer, and cancer stem cells. *Nature* **414:** 105–111

Rhim AD, Mirek ET, Aiello NM, Maitra A, Bailey JM, McAllister F, Reichert M, Beatty GL, Rustgi AK, Vonderheide RH, Leach SD & Stanger BZ (2012) EMT and Dissemination Precede Pancreatic Tumor Formation. *Cell* **148:** 349–361

Rodríguez-Paredes M & Esteller M (2011) Cancer epigenetics reaches mainstream oncology. *Nat Med* **17**: 330–339

Rossant J & Cross JC (2001) Placental development: lessons from mouse mutants. *Nat Rev Genet* **2:** 538–548

Rossant J, Sanford JP, Chapman VM & Andrews GK (1986) Undermethylation of structural gene sequences in extraembryonic lineages of the mouse. *Developmental Biology* **117:** 567–573

Rugg-Gunn PJ, Cox BJ, Ralston A & Rossant J (2010) Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proceedings of the National Academy of Sciences* **107:** 10783–10790

Sancak Y, Bar-Peled L, Zoncu R, Markhard AL, Nada S & Sabatini DM (2010) Ragulator-Rag Complex Targets mTORC1 to the Lysosomal Surface and Is Necessary for Its Activation by Amino Acids. *Cell* **141:** 290–303

Santos F, Hendrich B, Reik W & Dean W (2002) Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo. *Developmental Biology* **241:** 172–182

Scheel C, Eaton EN, Li SH-J, Chaffer CL, Reinhardt F, Kah K-J, Bell G, Guo W, Rubin J, Richardson AL & Weinberg RA (2011) Paracrine and Autocrine Signals Induce and Maintain Mesenchymal and Stem Cell States in the Breast. *Cell* **145:** 926–940

Sekiguchi T (2000) Novel G Proteins, Rag C and Rag D, Interact with GTP-binding Proteins, Rag A and Rag B. *Journal of Biological Chemistry* **276:** 7246–7257

Sheikh F, Raskin A, Chu P-H, Lange S, Domenighetti AA, Zheng M, Liang X, Zhang T, Yajima T, Gu Y, Dalton ND, Mahata SK, Dorn GW II, Heller-Brown J, Peterson KL, Omens JH, McCulloch AD & Chen J (2008) An FHL1-containing complex within the cardiomyocyte sarcomere mediates hypertrophic biomechanical stress responses in mice. *Journal of Clinical Investigation* **118:** 3870–3880

Shenghui H, Nakada D & Morrison SJ (2009) Mechanisms of Stem Cell Self-Renewal. *Annu. Rev. Cell Dev. Biol.* **25:** 377–406

Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, Turner CH, Goulet R, Badve S & Nakshatri H (2006) Cd44+/Cd24- breast cancer cells exhibit enhanced invasive properites: an early stem necessary for metastasis. *Breast Cancer Res* **8:** R59

Shih I-M, Hsu M-Y, Oldt RJ III, Herlyn M, Gearhart JD & Kurman RJ (2002) The Role of E-cadherin in the Motility and Invasion of Implantation Site Intermediate Trophoblast. *Placenta* **23:** 706–715

Stephens PS, Tarpey PS, Davies H, VanLoo P, Greenman C, Wedge DC, Nik-Zainal S, Martin S, Varela I, Bignell R, Yates LR, Papaemmanvil E, Beare D, Butler A, Cheverton A, Gamble J, Hinton J, Jia M, Jayakumar A, Jones D, et al (2012) The Landscape of Cancer Genes and Mutational Processes in Breast Cancer. *Nature* **486**: 400--406

Singhal N, Graumann J, Wu G, Araúzo-Bravo MJ, Han DW, Greber B, Gentile L, Mann M & Schöler HR (2010) Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. *Cell* **141:** 943–955

Stoops SL, Pearson AS, Weaver C, Waterson AG, Days E, Farmer C, Brady S, Weaver CD, Beauchamp RD & Lindsley CW (2011) Identification and Optimization of Small Molecules That Restore E-Cadherin Expression and Reduce Invasion in Colorectal Carcinoma Cells. *ACS Chem. Biol.* **6:** 452–465

Strumpf D (2005) Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132:** 2093–2102

Sutherland A (2003) Mechanisms of implantation in the mouse: differentiation and functional importance of trophoblast giant cell behavior. *Developmental Biology* **258:** 241–251

Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lønning PE & Børresen-Dale AL (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U.S.A.* **98:** 10869–10874

Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lønning PE, Brown PO, Børresen-Dale A-L & Botstein D (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci. U.S.A.* **100:** 8418–8423

Takeuchi JK, Lickert H, Bisgrove BW, Sun X, Yamamoto M, Chawengsaksophak K, Hamada H, Yost HJ, Rossant J & Bruneau BG (2007) Baf60c is a nuclear Notch signaling component required for the establishment of left-right asymmetry. *Proc. Natl. Acad. Sci. U.S.A.* **104:** 846–851

Tanaka S (1998) Promotion of Trophoblast Stem Cell Proliferation by FGF4. *Science* **282:** 2072–2075

Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, Hartwell K, Onder TT, Gupta PB & Evans KW (2010) Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl. Acad. Sci. U.S.A.* **107:** 15449–15454

Tcherkezian J & Lamarche-Vane N (2012) Current knowledge of the large RhoGAP family of proteins. *Biology of the Cell* **99:** 67–86

Thiery JP & Sleeman JP (2006) Complex networks orchestrate epithelial–mesenchymal transitions. *Nat. Rev. Mol. Cell Biol.* **7:** 131–142

Thiery JP, Acloque H, Huang RYJ & Nieto MA (2009) Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **139:** 871–890

Timmerman LA (2004) Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes & Development* **18:** 99–115

Tomaskovic-Crook E, Thompson EW & Thiery JP (2009) Epithelial to mesenchymal transition and breast cancer. *Breast Cancer Res* **11:** 213

Troester MA (2004) Cell-Type-Specific Responses to Chemotherapeutics in Breast Cancer. *Cancer Research* **64:** 4218–4226

Valastyan S & Weinberg RA (2011) Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* **147:** 275–292

Valcourt U, Kowanetz M, Niimi H, Heldin C-H & Moustakas A (2005) TGF-beta and the Smad signaling pathway support transcriptomic reprogramming during epithelial-mesenchymal cell transition. *Mol. Biol. Cell* **16:** 1987–2002

Vandewalle C (2005) SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic Acids Research* **33:** 6566–6578

Vandewalle C, Roy F & Berx G (2008) The role of the ZEB family of transcription factors in development and disease. *Cell. Mol. Life Sci.* **66:** 773–787

Veeman MT, Axelrod JD & Moon RT (2003) A second canon. Functions and mechanisms of beta-catenin-independent Wnt signaling. *Developmental Cell* **5:** 367–377

Volinia S, Calin GA, Liu C-G, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yanaihara N, Lanza G, Scarpa A, Vecchione A, Negrini M, Harris CC & Croce CM (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U.S.A.* **103:** 2257–2261

Wang J, Mager J, Schnedier E & Magnuson T (2002) The mouse PcG gene eed is required for Hox gene repression and extraembryonic development. *Mammalian Genome* **13:** 493–503

Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh T-Y, Peng W, Zhang MQ & Zhao K (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40:** 897–903

Willipinski-Stapelfeldt B (2005) Changes in Cytoskeletal Protein Composition Indicative of an Epithelial-Mesenchymal Transition in Human Micrometastatic and Primary Breast Carcinoma Cells. *Clinical Cancer Research* **11:** 8006–8014

Wyckoff J (2004) A Paracrine Loop between Tumor Cells and Macrophages Is Required for Tumor Cell Migration in Mammary Tumors. *Cancer Research* **64:** 7022–7029

Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens RM, Okamoto A, Yokota J & Tanaka T (2006) Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* **9:** 189–198

Yang J & Weinberg RA (2008) Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis. *Developmental Cell* **14:** 818–829

Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C, Savagner P, Gitelman I, Richardson A & Weinberg RA (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117:** 927–939

Yilmaz M & Christofori G (2009) EMT, the cytoskeleton, and cancer cell invasion. *Cancer Metastasis Rev* **28:** 15–33

Yook JI, Li X-Y, Ota I, Hu C, Kim HS, Kim NH, Cha SY, Ryu JK, Choi YJ, Kim J, Fearon ER & Weiss SJ (2006) A Wnt–Axin2–GSK3β cascade regulates Snail1 activity in breast cancer cells. *Nat. Cell Biol.* **8:** 1398–1406

Yu M, Ting DT, Stott SL, Wittner BS, Ozsolak F, Paul S, Ciciliano JC, Smas ME, Winokur D, Gilman AJ, Ulman MJ, Xega K, Contino G, Alagesan B, Brannigan BW, Milos PM, Ryan DP, Sequist LV, Bardeesy N, Ramaswamy S, et al (2012) RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis_Supplemental. *Nature* **487:** 510–513

Zavadil J & Böttinger EP (2005) TGF-β and epithelial-to-mesenchymal transitions. *Oncogene* **24:** 5764–5774

Zavadil J, Bitzer M, Liang D, Yang YC, Massimi A, Kneitz S, Piek E & Bottinger EP (2001) Genetic programs of epithelial cell plasticity directed by transforming growth factor-beta. *Proc. Natl. Acad. Sci. U.S.A.* **98:** 6686–6691

Zavadil J, Cermak L, Soto-Nieves N & Böttinger EP (2004) Integration of TGF-beta/Smad and Jagged1/Notch signalling in epithelial-to-mesenchymal transition. *EMBO J.* **23:** 1155–1165

Zeisberg M & Neilson EG (2009) Biomarkers for epithelial-mesenchymal transitions. *Journal of Clinical Investigation* **119:** 1429–1437

Zhou Y, Damsky CH & Fisher SJ (1997a) Preeclampsia is associated with failure of human cytotrophoblasts to mimic a vascular adhesion phenotype. One cause of defective endovascular invasion in this syndrome? *Journal of Clinical Investigation* **99:** 2152–2164

Zhou Y, Fisher SJ, Janatpour M, Genbacev O, Dejana E, Wheelock M & Damsky CH (1997b) Human cytotrophoblasts adopt a vascular phenotype as they differentiate. A strategy for successful endovascular invasion? *Journal of Clinical Investigation* **99:** 2139–2151

Zhu Y, Tian Y, Du J, Hu Z, Yang L, Liu J & Gu L (2012) Dvl2-Dependent Activation of Daam1 and RhoA Regulates Wnt5a-Induced Breast Cancer Cell Migration. *PLoS ONE* **7:** e37823