2019

GALNT3 MAINTAINS THE EPITHELIAL STATE BY PROMOTING O-GalNAC MODIFICATION OF E-CADHERIN IN TROPHOBLAST STEM CELLS AND HUMAN MAMMARY EPITHELIAL CELLS

Deepthi Raghu

GALNT3 MAINTAINS THE EPITHELIAL STATE BY PROMOTING O-GalNAc MODIFICATION OF E-CADHERIN IN TROPHOBLAST STEM CELLS AND HUMAN MAMMARY EPITHELIAL CELLS

by

Deepthi Raghu

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Biology

The University of Memphis

August 2019

Acknowledgements

I would like to thank my major professor, Dr. Amy Abell, for her support, training, mentorship, and guidance. I am grateful for having a mentor like her, as she has always believed in me and constantly helped me grow as a scientist. I am indebted to Amy for putting in the time and effort in making me a strong and independent scientist. I thank her for enabling me to be an effective scientific communicator, teaching me to generate new ideas, and directing me to ask the right questions. She has played a major role in contributing to my success.

I thank my committee members, Dr. Judith Cole for being an excellent mentor and support, Dr. Ramin Homayouni for his valuable insights on my research, Dr. Omar Skalli for mentoring and all the scientific discussions, and Dr. Thomas Sutter for his significant research ideas. Together, I thank my committee members for their valuable time and for holding me to a higher standard.

I am also thankful for my fellow lab members, Noha Shendy, Chad Perry, Dylan Hammrich, and Adiba Safi, as well as past lab members, R Jackson Mobley, Kayley Abell-Hart, Lauren Duke, and Kimber Lewis for contributing to the mentoring, supportive, and positive environment of our lab. This has been my family for the past five years, and I could not have asked for a better support system.

I would especially like to thank Dr. David Freeman for being a supportive and encouraging department chair; Dr. Beck, Dr. Mathew Parris, and Dr. Omar Skalli for being wonderful graduate coordinators; Dr. Bernie Daigle, Dr. Shawn Brown, and Dr. Keith Bowers for their constant encouragement and support, and the department administrative staff.

Lastly but most importantly, I would like to thank my husband Harshad, colleagues and friends including Gayatri, Kris, Elyan, and Malle and my family back in India for always believing, supporting, and encouraging me while I pursue my passion and career for science.

Preface

This dissertation is presented as three chapters including research from one published research article. Chapter I provides background information and rationale for this study, as well as a description of the three aims that accomplish the goals of the study. Chapters II and III explain the research performed in completing the three aims of the study. The data presented in chapters II and III and appendices A-E were published in March of 2019 in Cell Reports (PMC6501849), entitled "GALNT3 Maintains the Epithelial State in Trophoblast Stem Cells". Chapter II contains the data supporting aim one, entitled "GALNT3 is Important for Epithelial State Maintenance in Trophoblast Stem Cells, Blastocyst Trophectoderm, and Human Mammary Epithelial Cells". Chapter III contains data supporting aims two and three, entitled "GALNT3 Mediated O-GalNAc Glycosylation of E-cadherin Promotes the Epithelial State in Trophoblast Stem Cells and Human Mammary Epithelial Cells". Chapter IV includes conclusion for Chapters II and III and also future directions of this work.

Abstract

Deepthi Raghu. Ph.D. The University of Memphis. August 2019. GALNT3 Maintains the Epithelial State by Promoting O-GalNAc Modification of E-cadherin in Trophoblast Stem Cells and Human Mammary Epithelial Cells. Major Professor: Dr. Amy. N Abell, Ph.D.

O-GalNAc glycosylation is initiated in the Golgi by glycosyltransferases called GALNTs. Proteomic screens identified >600 O-GalNAc modified proteins, but biological relevance of these modifications has been difficult to determine. We have discovered a conserved function for GALNT3 in trophoblast stem (TS) cells, blastocyst trophectoderm, and human mammary epithelial cells (HMECs). Loss of GALNT3 expression in these systems reduces O-GalNAc glycosylation and induces epithelial-mesenchymal transition. Further, *Galnt3* expression is reduced in aggressive, mesenchymal Claudin-low breast cancer cells. We show that GALNT3 expression controls the O-GalNAc glycosylation of multiple proteins including E-cadherin in both TS cells and HMECs. Loss of GALNT3 results in the intracellular retention of E-cadherin in the Golgi. Significantly, re-expression of GALNT3 in TS cells increases O-GalNAc glycosylation and restores the epithelial state. Together, these data demonstrate the critical biological role of GALNT3 O-GalNAc glycosylation to promote the epithelial phenotype in TS cells, blastocyst trophectoderm, and HMECs.

Table of Contents

List of Figures

List of Figures (Continued)

Abbreviations

Term	Abbreviation
Adherens Junction	(AJ)
Chromatin Immunoprecipitation	(ChIP)
Chromatin Immunoprecipitation Coupled to Real-time qPCR	(ChIP-PCR)
cAMP-response element binding protein	(CREB)
CREB Binding Protein	(CBP)
CDC42 and Rac Interactive Binding	(CRIB)
Claudin-low Breast Cancer Cells	(CL)
Endoplasmic Reticulum	(ER)
Epithelial to Mesenchymal Transition	(EMT)
GalNAc Transferase 3	(GALNT3)
Histone 2B Lysine 5 Acetylation	(H2BK5Ac)
Histone Deacetylase	(HDAC)
Human Mammary Epithelial Cells	(HMEC)
Human Mammary Epithelial Cells Expressing GALNT3 shRNA	($HMEC^{Gsh}$)
Kinase-inactive MAP3K4	(KI4)
Mitogen Activated Protein Kinase	(MAPK)
MAP3K4 Kinase-inactive Trophoblast Stem Cells	(TS^{KI4})
MAP3K4 Kinase-inactive Trophoblast Stem Cells Re-expressing Human GALNT3	(TS^{KI4Gab})
Mesenchymal to Epithelial Transition	(MET)
Mice Expressing Kinase-inactive MAP3K4	(KI4 Mice)

Abbreviations (Continued)

Term	Abbreviation
N-acetylgalactosamine	(GalNAc)
N-acetylgalactosaminyl Transferase	(GALNT)
Short-Hairpin RNA	(shRNA)
Real-time quantitative PCR	(qPCR)
Tight Junction	(TJ)
Tumoral Calcinosis	(TC)
Trophectoderm	(TE)
Vicia Villosa Lectin	(VVL)
Wild-type Trophoblast Stem Cells	(TS^{WT})
Wild-type Trophoblast Stem Cells Expressing GALNT3 shRNA	(TS^{WTGsh})

Chapter 1

Introduction

Specific Aims

Epithelial to mesenchymal transition (EMT) is a cellular program in which non-motile epithelial cells convert to motile mesenchymal cells. In this morphogenetic cellular process, epithelial cells lose cell-to-cell contact and apical-basal polarity with the gain of mesenchymal morphology and invasive properties. EMT is critical in numerous developmental processes and is reactivated in the adult tissues during organ fibrosis and cancer metastasis. Murine trophoblast stem (TS) cells have been used to define the molecular and epigenetic mechanisms regulating developmental EMT. Wild type TS cells (TS^{WT}) are epithelial with tight cell-cell adhesion, whereas TS cells lacking MAP3K4 activity (TS^{KI4}) express a mesenchymal-like phenotype. *Galnt3* (UDP-GalNAc transferase 3) gene expression is reduced in both TS^{KI4} cells and mesenchymal, Claudin-low (CL) breast cancer cells. This O-GalNAc glycosyltransferase, GALNT3, initiates O-linked glycosylation in the Golgi by transferring N-acetylgalactosamine (GalNAc) to the hydroxyl groups of serine or threonine residues. Importantly, the function of GALNT3 is to promote folding and localization of either membrane or secreted proteins. Aberrantly glycosylated and misfolded proteins are either trapped in the Golgi or mislocalized to the endoplasmic reticulum. With the exception of FGF23, substrates of GALNT3 are unknown and undefined. Loss of function mutations in GALNT3 cause misfolding of FGF23, a hormone required for phosphate homeostasis. Aberrant glycosylation of FGF23 causes a severe autosomal disorder, Hyperphosphatemic Familial Tumoral Calcinosis. This disorder is characterized by increased phosphate levels and subcutaneous calcium

deposits. Our preliminary data suggest that *Galnt3* is reduced in all mesenchymal, CL breast cancer cells relative to human mammary epithelial cells (HMEC). Further, we observe decreased expression of *Galnt3* in the mesenchymal TS^{KI4} cells relative to epithelial TS^{WT} cells. Based on our preliminary data, we predict that GALNT3 may be responsible for the maintenance of the epithelial phenotype, and that loss of this gene is associated with an invasive, mesenchymal phenotype in both TS cells and breast cells.

Our hypothesis is that GALNT3 regulates the folding and localization of proteins that are important in the maintenance of the epithelial phenotype. The goal of this project is to define the impact of GALNT3 dependent O-GalNAc glycosylation on developmental EMT that is reactivated in breast cancer metastasis. We have addressed this goal by examining the role of GALNT3 in the regulation of cellular phenotype in TS cells, blastocyst trophectoderm, and HMECs. Alterations in the levels of GALNT3 in both TS cells and breast cells impacts EMT/MET processes. In addition to the cultured TS cells, we also examined the impact of loss of GALNT3 on EMT in the trophectoderm of preimplantation blastocysts. Importantly, we observed that GALNT3 mediates the O-GalNAc glycosylation of multiple proteins, including E-cadherin. For the first time, we have defined the critical role of GALNT3 dependent O-GalNAc glycosylation in promoting the epithelial state in both TS cells and HMECs.

Background and Significance

Epithelial to Mesenchymal Transition

Epithelial to mesenchymal transition (EMT) is a morphogenetic cellular process in which non-motile epithelial cells lose cell-to-cell contact and apical-basal polarity with the gain of motile mesenchymal cells with front-back polarity (Kalluri and Weinberg, 2009). EMT is critical for numerous developmental processes and wound healing (Thiery et al., 2009). Importantly, reactivation of EMT in the adult tissues occurs during organ fibrosis and cancer metastasis (Chapman, 2011; De Craene and Berx, 2013). EMT is reversible through mesenchymal-epithelial transition (MET), where motile, individual mesenchymal cells convert to non-motile, tightly adherent epithelial cells (Esteban et al., 2012; Mikkelsen et al., 2008). Sequential rounds of this reversible EMT process promote the differentiation of specialized cell types and result in the formation of three-dimensional structures of internal organs (Thiery et al., 2009). Developmental EMT processes can be further classified as primary, secondary and tertiary EMT. In early embryonic development, primary EMT occurs during gastrulation, where the cells from the epiblast undergo EMT to form the mesoderm, endoderm, and ectoderm (Acloque et al., 2009). In addition, epithelial neural crest cells also undergo primary EMT to form neural tubes leading to the development of the central nervous system (Acloque et al., 2009; Cheung and Briscoe, 2003; Cheung et al., 2005). Different subtypes of early mesodermal cells are condensed to form transient epithelial structures including somites, notochord, somatopleures, and splanchnopleures. These transient structures undergo secondary EMT and differentiate to form specific cell types such as hematopoietic stem cells, connective tissues, and secretary tissues (Buckingham et al., 2003). Tertiary EMT

is poorly understood and the best studied example is formation of cardiac valves and the endocardium of the heart (Nakajima et al., 2000).

Several molecular events occur during the initiation and completion of EMT. These events include the loss of cell-cell adhesion proteins such as tight-junction proteins, ZO-1 and Claudin-6, and adherens junction proteins including E-cadherin, α-catenin, β-catenin, and δ-catenin. Importantly, cells in EMT undergo reorganization of actin cytoskeleton from epithelial cells with cortical actin to mesenchymal cells with filamentous actin. In addition to morphological changes, cells lose the expression of epithelial markers like E-cadherin (De Craene and Berx, 2013; Peinado et al., 2004; Peinado et al., 2007; Yilmaz and Christofori, 2009). Cells gain mesenchymal marker expression like Vimentin and N-cadherin and increase expression of EMT-inducing transcription factors like TWIST1 and SNAI2 (Batlle et al., 2000; Cano et al., 2000; Dong et al., 2013; Herranz et al., 2008; Peinado et al., 2004; Tong et al., 2012; Wheelock et al., 2008; Yilmaz and Christofori, 2009). Most importantly, mesenchymal cells acquire increased motility and/or invasiveness (De Craene and Berx, 2013; Kalluri and Weinberg, 2009).

In addition to having an essential role during development, EMT is also reactivated during tissue or organ fibrosis, and cancer metastasis. During fibrosis, uncontrolled and constitutively active EMT results in continuous transition of epithelial cells to myofibroblasts (Iwano et al., 2002; Mansour et al., 2017; Okada et al., 1996; Sato et al., 2003; Tan et al., 2006; Zeisberg et al., 2007; Zeisberg et al., 2003). This sustained EMT process results in increased deposition of extracellular matrix proteins leading to high immune infiltration, persistent inflammation, and fibrosis (Iwano et al., 2002;

Mansour et al., 2017; Okada et al., 1996; Sato et al., 2003; Tan et al., 2006; Zeisberg et al., 2007; Zeisberg et al., 2003). During cancer metastasis, epithelial cancer cells undergo EMT and convert to motile, invasive cells (Boutet et al., 2007; Byler et al., 2014; Byler and Sarkar, 2014; Heerboth et al., 2014; Housman et al., 2014; Roche, 2018; Sarkar et al., 2013). During cancer EMT, a subpopulation of metastatic cancer cells expresses both epithelial and mesenchymal markers, suggesting that these cancer cells are in a partial EMT state. The cells in partial EMT have increased expression of EMT inducing transcription factors such as *Snail*, *Twist*, and *Zeb* family members (Saitoh, 2018). These cancer cells are highly invasive and migrate either in clusters or as individual cells leading to metastasis. Prat et al. (2010) examined microarray data from 337 patient breast tumors and from 52 breast cancer cell lines (Hennessy et al., 2009; Herschkowitz et al., 2007; Neve et al., 2006; Parker et al., 2009; Prat et al., 2010). Using these microarray data from breast tumors and cell lines, they clustered different breast cancer subtypes based on the gene expression (Prat et al., 2010; Prat and Perou, 2011). Breast cancer subtypes were defined as Luminal A and B, Basal-like, HER-2 enriched, and Claudin-low (CL) breast cancers (Prat et al., 2010; Prat and Perou, 2011). Among these, CL breast cancers have increased mesenchymal marker expression and show EMT characteristics with stemness properties (Prat et al., 2010). The CL breast cancers are triple-negative breast cancers that lack expression of estrogen, progesterone, and HER2 receptors. In addition, CL breast cancers have complete absence of the tight junction proteins, Claudin 3/4/7, and the adherens junction proteins, E-cadherin and EpCAM. Importantly, CL breast cancers are enriched in tumor-initiating factors such as CD44 and CD24. The presence of these tumor-initiating factors enables cells to form tumors, a key feature of

cancer stem cells. In addition, lack of estrogen receptors or HER2 receptors in the CL breast cancer cells makes them unresponsive to targeted therapies such as tamoxifen and lapatinib (Dias et al., 2017). Together, these data suggest that cancer cells that have undergone EMT respond poorly to therapy, and thus these patients have low survival rates. Therefore, defining the signaling mechanisms underlying EMT processes is important for identifying new therapeutic strategies to treat EMT related diseases.

Glycosylation

Protein glycosylation plays a key role in both developmental EMT and EMT that occurs in disease states. The majority of cell surface and secreted proteins undergo this complex post-translational modification. Glycosylation occurs in either the endoplasmic reticulum (ER) or the Golgi complex. During glycosylation, multiple glycoenzymes mediate the attachment of carbohydrates to hydroxyl groups or amide groups of specific proteins. This protein-oligosaccharide bonding promotes proper protein folding leading to protein structural and functional stability (Shental-Bechor and Levy, 2008). Genetic defects in the glycoenzymes or alterations in the glycosylation pathways result in aberrant protein glycosylation (Yarema and Bertozzi, 2001). There are two major types of glycosylation, N-linked and O-linked. These types of glycosylation are initiated in different organelles and have distinct impacts on protein structure and function.

In the ER, the attachment of N-glycans to asparagine (Asn) residues on specific proteins marks the initiation of N-glycosylation (Figure 1) (Stanley et al., 2015). The consensus sequence begins with Asn followed by any amino acid except proline and then serine or threonine residues (Asn-X-Ser/Thr) (Stanley et al., 2015). The biosynthesis of N-glycosylation occurs primarily in the lumen of the ER membrane. The first step of

synthesis includes a lipid-like molecule termed dolichol phosphate, which is followed by "en bloc" transfer of 14 glycan sugars to proteins with the consensus sequence Asn-X-Ser/Thr (Figure 1). The specific glycans in the "en bloc" depend on the cell type and physiological status of the cell. The most common N-glycan attachment to Asn is N-acetylglucosamine. Disorders associated with genetic defects in the N-glycans are termed congenital disorders of glycosylation (CDG) (Freeze et al., 2015; Sparks and Krasnewich, 1993). These disorders are classified into two types, I and II. Type I patients have incomplete N-glycan chains due to unoccupied sugars at the consensus sequence of the protein (Freeze et al., 2015; Sparks and Krasnewich, 1993). In type II patients, N-glycan structures are altered due to abnormal processing of glycoenzymes (Freeze et al., 2015; Sparks and Krasnewich, 1993). Importantly, N-glycan biosynthesis and types of sugars that are added to the protein depend on the cell type and the physiological status of the cell such as development, differentiation or disease states.

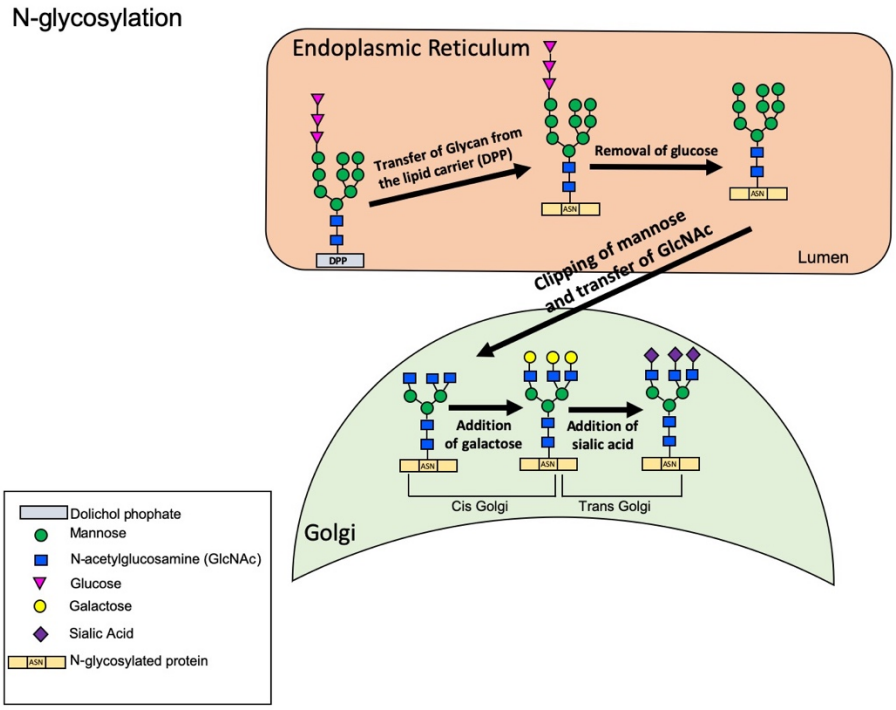

Figure 1: N-linked glycosylation occurs in both the ER and Golgi. Initiation of N-glycosylation occurs in the lumen of the ER and involves the transfer of glycan sugars from the dolichol phosphate (DPP) onto Asparagine (Asn) residues of specific proteins. Further, extension of glycans occurs in the Golgi as illustrated. Adapted from http://biotechhelpline16.blogspot.com/2012/09/golgi-apparatus-golgi-apparatus-also.html

O-linked glycosylation occurs by the transfer of sugars onto hydroxyl groups of serine or threonine residues of specific proteins (Van den Steen et al., 1998). There are two main types of O-glycosylation that are distinguished by the specific sugar transferred, N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc). O-GlcNAc glycosylation is initiated by the transfer of GlcNAc to the hydroxyl groups of serine and threonine residues of cytoplasmic or nuclear proteins. The consensus sequence for O-GlcNAc glycosylation has not been defined. Two enzymes regulate protein O-GlcNAc glycosylation, O-GlcNAc transferase catalyzes the transfer of O-GlcNAc whereas O-GlcNAcase (N-acetylglucosamine hydrolase) removes O-GlcNAc from proteins. This reversible O-GlcNAc glycosylation of proteins has been proposed to

function as a response to nutrient and cellular stress, thus regulating processes including protein transcription, signal transduction, and cellular metabolism (Bektas and Rubenstein, 2011). Disruption of O-GlcNAc processes has been implicated in severe diseases including metabolic disorders, cancer, and neurodegeneration (Yang and Qian, 2017). In contrast to O-GlcNAc glycosylation, O-GalNAc glycosylation occurs in the Golgi and is initiated by the transfer of GalNAc onto serine and threonine residues of membrane and secreted proteins (Figure 2) (Brockhausen et al., 2009). A large family of glycoenzymes called polypeptide-N-acetylgalactosaminyl transferases (GALNTs) found in the Golgi catalyzes the initiation of O-GalNAc glycosylation (Figure 2). GALNTs are comprised of 20 members in humans and 19 in mouse that can be functionally redundant. Although the consensus sequence for the addition of O-GalNAc has not been defined, it has been suggested that the presence of a proline at the +3 or -3 position near serine or threonine is preferred for O-GalNAc glycosylation (Bennett et al., 2012). The linkage of a single GalNAc to the serine or threonine residues constitutes the simplest form called Tn antigen (Bennett et al., 2012). Complex O-GalNAc glycans have core 1, 2, 3 glycan structures added to the Tn antigen (Figure 2) (Bennett et al., 2012). The addition of core 1 O-GalNAc glycans to the Tn antigen forms T antigen. Both T and Tn antigens can be terminated with sialic acid, termed sialylated-Tn (sTn) or T (sT) antigens. Increased expression of short O-GalNAc glycans occurs in many tumors or poorly differentiated cancers (Chia et al., 2016).

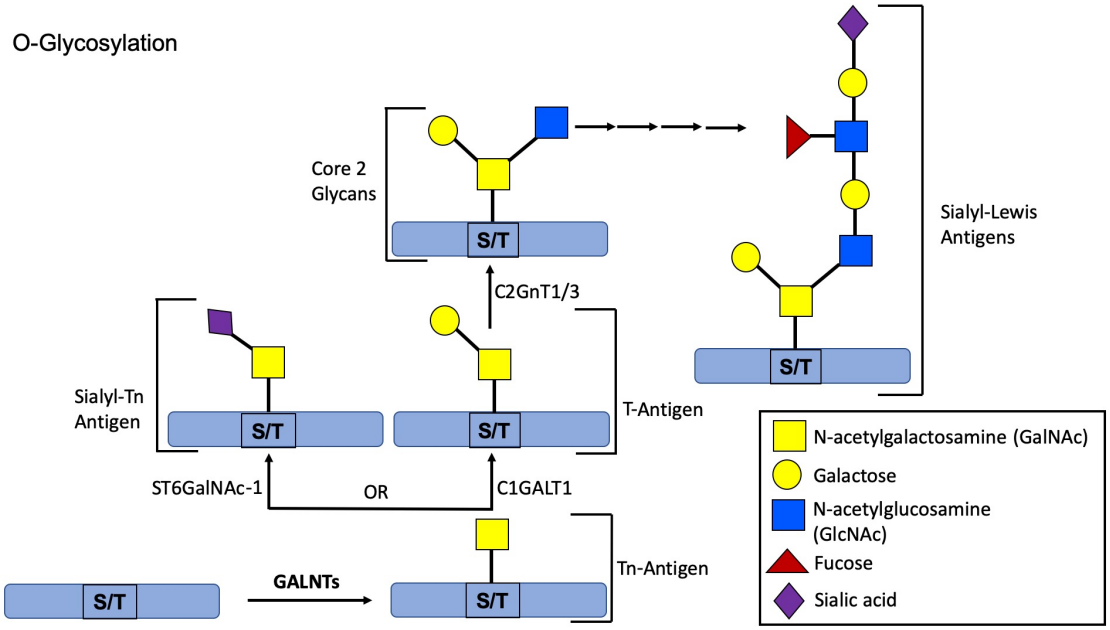

Figure 2: O-GalNAc glycosylation occurs in the Golgi. The transfer of N-acetylgalactosamine (GalNAc) on Serine or Threonine (S/T) residues of specific proteins is mediated by GALNTs and marks the initiation of O-glycosylation. The simplest form of O-glycosylation is called Tn antigen. Addition of Core-1 glycans to the Tn antigen by C1GALT1 enzyme forms the T antigen. These Tn and T antigens can also be terminated with sialic acid forming the sialyl-Tn (sTn) antigen. The Core-1 glycans are further elongated by the addition of fucose or galactose sugars or sialyl-Lewis antigens. Overexpression of T, Tn, sTn and sialyl-Lewis antigens are found in many cancers. Adapted from Chia et al. (2016).

O-GalNAc Glycosylation in Cancer

O-GalNAc glycosylation is the simplest form of O-glycosylation. Aberrant levels of O-GalNAc glycans on proteins have been implicated during tumorigenesis and cancer progression (Chia et al., 2016). There are two main mechanisms associated with altered glycosylation in tumors, incomplete synthesis or neo-synthesis (Pinho and Reis, 2015). Loss of initiating glycosyltransferases results in incomplete synthesis with increased short O-GalNAc glycans and disrupted protein structures. In contrast, synthesis of glycosyltransferases not normally expressed in a cell or tissue occurs in a neo-synthetic process and can lead to new carbohydrate structures on proteins that may trigger a

tumorigenic response (Pinho and Reis, 2015). In addition, mislocalization of GALNTs in the ER also increases the presence of short O-GalNAc glycans on proteins(Bard and Chia, 2016). Termination of proteins with sTn antigens are rarely found in healthy tissues, but are abundantly present in cancers including pancreatic, stomach, colorectum, and breast (Brockhausen, 1997; Julien et al., 2006; Munkley, 2016; Sewell et al., 2006). The presence of sTn antigens in carcinomas is correlated with increased tumor size, decreased cell-cell adhesion proteins, increased invasion, and poor prognosis (Kim and Varki, 1997; Pinho et al., 2007; Pinho and Reis, 2015). In addition, loss of glycosyltransferases or increased expression of sTn antigens has been associated with cancer related EMT (Barkeer et al., 2018). Further, overexpression of short O-glycans on MUC1 protein has been shown to cause cancer EMT in breast, ovary, prostate, and pancreatic carcinomas (Aoki et al., 1998; Avichezer et al., 1998; Burdick et al., 1997; Hough et al., 2000; Kufe, 2009, 2013; Maeshima et al., 1997; Medina et al., 1999; Strous and Dekker, 1992; van Hof et al., 1996; Willsher et al., 1993; Zhang and Weinberg, 2018). Altered MUC1 glycosylation increases expression of EMT-inducing transcription factors *Snai1* and *Zeb1* and decreases expression of a key epithelial marker, E-cadherin (Ponnusamy et al., 2013). Together, structural alterations to the glycans due to absence of the glycosyltransferases have been shown to be associated with cancer EMT (Carvalho-Cruz et al., 2018; da Fonseca et al., 2016).

O-GalNAc Glycosylation in Development

The roles of O-GalNAc glycosyltransferases during cancer EMT are well defined, but the functions of GALNTs in developmental EMT are poorly understood. Due to their functional redundancies that may result in compensatory mechanisms, it has been

difficult to identify the essential functions of GALNTs in developmental EMT. However, studies in model organisms have identified specific roles for a few O-GalNAc glycosyltransferases in development. In *Drosophila*, inactivating mutations in *Pgant35A* causes disruption in the formation of the respiratory tract leading to lethality (Tian and Ten Hagen, 2007). This phenotype exhibits EMT-like characteristics, including loss of apical-basal polarity, distorted barrier formation, and decreased tight junction proteins (Tian and Ten Hagen, 2007). In addition, loss of *Pgant3* in *Drosophila* results in reduced secretion of basement membrane components, leading to loss of cell-cell adhesion (Zhang et al., 2008). In mice, loss of *Galnt1* decreases secretion of basement membrane components including Laminins (Tian et al., 2012). This increased intracellular accumulation of basement membrane proteins results in decreased transcript levels of the FGF pathway thus leading to reduced proliferation and growth of submandibular glands (Tian et al., 2012). In Xenopus, knockdown of *Galnt11* results in defects in left-right patterning of the body (Boskovski et al., 2013). Together, these data suggest the importance of O-GalNAc glycosylation in the maintenance of cell-cell adhesion proteins, tight junction proteins, axial patterning, and secretion of basement membrane components, thus regulating developmental EMT.

MAP3K4, a Mitogen Activated Protein Kinase Kinase Kinase

Mitogen activated protein kinase (MAPK) signaling pathways control key cellular functions including proliferation, differentiation, survival, and apoptosis. MAPK signaling is stimulated by external cues, and kinase activity depends on several factors such as the kinetics of activation or inactivation, localization, and formation of protein interaction complexes (Raman et al., 2007). These MAPK signaling pathways consist of

evolutionarily conserved sequentially acting kinases whereby the MAPKKK/MAP3K phosphorylates and activates MAPKK/MAP2K that phosphorylates and activates the MAPK. Each component of the cascade consists of multiple family members. For example, the MAP3K family of serine/threonine kinases is composed of 20 members. These MAP3Ks selectively phosphorylate and activate different MAP2Ks. For example, MAP3K4 phosphorylates and activates MAP2K3/4/6 and 7. These dual specific serine, threonine and tyrosine MAP2Ks selectively phosphorylate and activate MAPKs. The MAPK family is composed of 14 members that are subdivided into seven groups: extracellular signal-regulated kinase 1 and 2 (ERK1/2), c-Jun amino-terminal kinase 1/2/3 (JNK 1/2/3), p38 (isoforms α, β, γ, and δ), ERK3/4, ERK5, ERK7, and nemo-like kinase (NLK) (Chen et al., 2001; Cooper et al., 1982; Coulombe and Meloche, 2007; Kyriakis and Avruch, 2001; Pearson et al., 2001). The MAPK subgroups are stimulated by specific external cues. For example, the ERK1/2 subgroup is phosphorylated and activated by MAP2K1 and MAP2K2 in response to growth stimuli (Kazlauskas and Cooper, 1988; Ray and Sturgill, 1988). In contrast, environmental stress, bacterial endotoxins, and proinflammatory cytokines stimulate both the JNK and p38 MAPKs (Bogoyevitch et al., 2010; Cuadrado and Nebreda, 2010). These stresses induce the phosphorylation of MAP2K4 and MAP2K7 that activates JNK, and the phosphorylation of MAP2K3 and MAP2K6 that activates p38 (Derijard et al., 1995; Han et al., 1996; Lawler et al., 1998; Stein et al., 1996). Further, MAP3K4 phosphorylates MAP2K4/7 and MAP2K3/6. Importantly, MAP3K4 expression has been shown to be important during developmental EMT, specifically in neural tube closure and skeletal patterning (Abell and Johnson, 2005; Chi et al., 2005).

MAP3K4 has several well-defined protein domains (Figure 3). The N-terminus consists of growth arrest and DNA damage inducible 45 binding (GADD45) domain and a kinase autoinhibitory domain. The C terminus includes a CDC42 and RAC interactive binding (CRIB)-like domain and a 37 kDa kinase domain. Conformational changes of these domains control MAP3K4 activity. In the inactive confirmation, MAP3K4 folds at the CRIB-like domain and the N-terminal kinase-inhibition domain blocks the C-terminal kinase domain (Mita et al., 2002). In the active confirmation, the N-terminal kinase inhibition domain releases the C-terminal kin]ase domain, leading to dimerization of MAP3K4 and phosphorylation of downstream MAP2Ks, MAP2K4/7 and MAP2K3/6 (Mita et al., 2002). GADD45 binds to the N-terminal regulatory domain, while Axin and TNF receptor-associated factor 4 (TRAF4) bind to the kinase domain of MAP3K4 (Abell et al., 2007; Abell and Johnson, 2005; Luo et al., 2003). Binding of these proteins facilitates dimerization and activation of MAP3K4. TRAF4 binding to the MAP3K4 kinase domain promotes MAP3K4 dimerization and synergistic activation of JNK (Abell and Johnson, 2005). Further, MAP3K4 can also bind to a scaffold protein, Axin, resulting in JNK activation (Luo et al., 2003). Binding of GADD45 α, β, and γ proteins to the N-terminal regulatory domain of MAP3K4 results in activation of p38 and JNK signaling pathways (Mita et al., 2002). In contrast, binding of glycogen synthase kinase-3 beta (GSK3β), a multifunctional serine or threonine kinase, to the MAP3K4 kinase domain inhibits the activity of MAP3K4 (Abell et al., 2007). GSK3β phosphorylates the N-terminus of MAP3K4, thus preventing the dimerization of MAP3K4 (Abell et al., 2007). Together, these studies have identified multiple binding partners of MAP3K4 that either activate or inhibit the downstream signaling cascade.

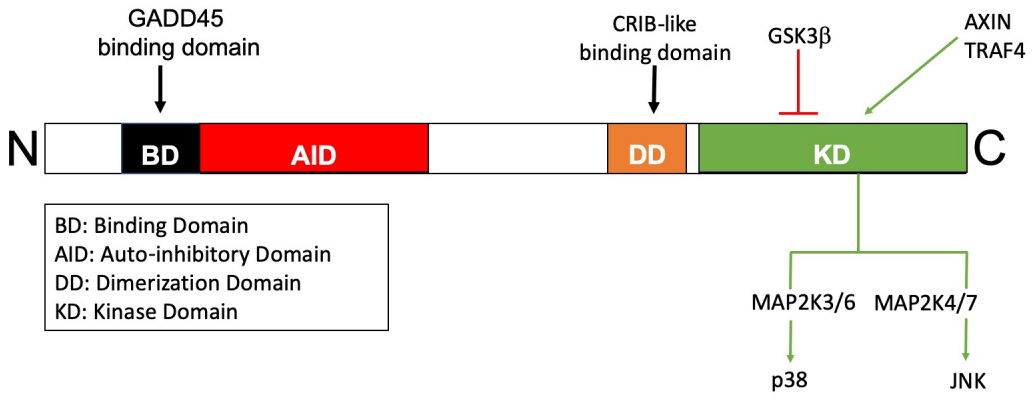

Figure 3: Diagram depicts the four protein domains of MAP3K4. Binding sites for MAP3K4 interacting proteins are shown.

MAP3K4 expression is strongest during development with highest expression in embryonic stem (ES) cells, trophoblast stem (TS) cells, neuroepithelium, and chorion (Abell et al., 2005, Abell et al., 2009). Three mouse models have been created to determine the physiological roles of MAP3K4. Two of these models have eliminated MAP3K4 protein expression (Bogani et al., 2009; Chi et al., 2005; Warr et al., 2012). Chi et al. created a targeted deletion of exon 3, resulting in the complete absence of MAP3K4 protein (Chi et al., 2005). The second model was generated using an ENU mutagenesis strategy to identify genes important for sex determination. This strategy induced a single nucleotide substitution of an A to T at position 1144 in MAP3K4. This mutation introduces a premature stop codon at amino acid 382 of 1597 residues of the MAP3K4 protein. Western blotting suggests that the truncated MAP3K4 protein is unstable (Bogani et al., 2009). Unlike the two models that eliminate MAP3K4 expression, the Kinase-Inactive MAP3K4 (KI4) model was produced through a targeted mutation of the active site lysine at position 1361 to arginine in the kinase domain of MAP3K4 (Abell and Johnson, 2005). This mutation inactivates MAP3K4 phosphotransferase activity

without affecting MAP3K4 protein expression. All three MAP3K4 mouse models show embryonic and perinatal lethality due to multiple, severe developmental defects. Consistent with the high expression of MAP3K4 in the developing neuroepithelium, all three mouse models exhibit neural tube defects, including exencephaly and spina bifida (*). These defects result from increased apoptosis of neuroepithelial cells, suggesting that MAP3K4 might serve an anti-apoptotic function during neural development. Further, MAP3K4 elimination models also display defects in sex determination (Bogani et al., 2009; Warr et al., 2012). This disorder was due to reduced expression of a sex determining gene *Sry* that caused the development of ovaries in a XY embryo. Importantly, loss of MAP3K4 results in low SOX9 expression leading to growth retardation of the gonads and disrupted testicular development (Bogani et al., 2009). The KI4 mice have not been examined for changes in sex determination. In contrast to the MAP34 elimination models, homozygous KI4 mice show defective skeletal patterning including rib malformations and scoliosis. In addition, KI4 mice exhibit significant hearing loss due to impairment in the mechanosensory hair cell development of the inner ear and distorted hair cell differentiation (Haque et al., 2016). As previously stated, MAP3K4 is highly expressed in the TS cells. Examination of homozygous KI4 conceptuses reveals defective decidualization and hyperinvasion, suggesting implantation defects (Abell et al., 2009). Importantly, all of these defects, including neural tube, skeletal and implantation defects, are due to perturbations in the EMT process. Together, these data suggest that MAP3K4 is important for the regulation of normal developmental EMT processes.

Trophoblast Stem Cells as a Model for Developmental EMT

Murine trophoblast stem (TS) cells can be isolated from embryonic day (E) 3.5 preimplantation blastocysts and from E6.5 extraembryonic ectoderm. Preimplantation (E3.5) blastocysts form a few days after conception and consist of approximately one hundred cells. These preimplantation embryos are composed of two cell types: the inner cell mass (ICM) and the outer trophectoderm (TE). Embryonic stem (ES) cells are isolated from the inner cell mass, which differentiates to form the body. The trophectoderm layer contains trophoblast stem (TS) cells, which differentiate to form the placenta. *In vivo*, the inner cell mass of the blastocyst secretes FGF4 that binds FGFR2 receptors in the trophectoderm to maintain TS cell stemness (Simmons and Cross, 2005). Similar to ES cells, the TS cells can be cultured indefinitely in the presence of specific growth factors including FGF4 and conditioned media produced from mitotically inactivated mouse embryonic fibroblasts (Tanaka et al., 1998). These cultured TS cells demonstrate self-renewal properties and express stemness markers including *Cdx2, Eomes, Sox2, Elf5a, Tfap2c,* and *Esrrβ* (Auman et al., 2002; Avilion et al., 2003; Donnison et al., 2005; Russ et al., 2000; Simmons and Cross, 2005; Strumpf et al., 2005). Upon removal of FGF4, the TS cells differentiate into specific lineages that form mature trophoblast subtypes of the placenta (Simmons and Cross, 2005). Differentiated subtypes include spongiotrophoblasts, syncytiotrophoblasts, and giant cells (Simmons and Cross, 2005). There are several markers for differentiated trophoblasts including *Gcm1, Tpbpa, SynA, SynB, Prl3d1, Prl2C2, Ctsq,* and *Csh2* (Simmons and Cross, 2005).

Murine TS cells have been used to define the molecular and epigenetic mechanisms regulating developmental EMT. The first EMT event occurs during

blastocyst implantation. The epithelial TS cells in the trophectoderm undergo EMT to become invasive giant cells that digest the uterine wall to establish placentation (Abell et al., 2011). In culture, differentiation of TS cells induces EMT as characterized by loss of the epithelial marker E-cadherin, gain in mesenchymal markers N-cadherin and vimentin, and most importantly, acquisition of invasiveness (Abell et al., 2011). The wild-type TS (TS^{WT}) cells isolated from E3.5 preimplantation blastocysts are epithelial with tight cell-cell adhesion and apical-basal polarity (Abell et al., 2011). The TS cells isolated from KI4 (TS^{KI4}) conceptuses exhibit a mesenchymal morphology with front-back polarity (Abell et al., 2011). Embryos with inactivated MAP3K4 display developmental disorders including neural tube, skeletal, and implantation defects that are due to perturbations in EMT (Abell et al., 2009; Abell and Johnson, 2005). Undifferentiated TS^{KI4} cells show key characteristics of EMT, including gain of mesenchymal morphology with increased expression of mesenchymal markers *Vim* and *Cdh2*, and EMT-inducing transcription factors *Twist1* and *Snai2* (Abell et al., 2011). In addition, TS^{KI4} cells display increased invasiveness relative to TS^{WT} cells. Importantly, TS^{KI4} cells maintain stemness properties and multipotency (Abell et al., 2011). This work suggests that loss of MAP3K4 activity in TS cells induces EMT in the context of stemness.

Reactivation of EMT occurs in epithelial cancer cells during cancer metastasis. These metastatic cancers contain subset of cells that exhibit self-renewing and stemness properties (Mani et al., 2008). Therefore, it has been hypothesized that EMT imparts stemness in these cancer cell populations (Mani et al., 2008). Induction of EMT by expressing constitutively active EMT inducing transcription factors such as *Twist* and *Snai* in primary mammary epithelial cells results in a mesenchymal state with increased

expression of stemness genes (Mani et al., 2008). Importantly, CL breast cancer cells display a mesenchymal morphology and exhibit EMT and stemness characteristics compared to all other breast cancer subtypes (Prat et al., 2010). CL breast cancers have decreased expression of epithelial markers including E-cadherin, EpCAM, and keratins, and increased expression of mesenchymal markers like Vimentin (Prat et al., 2010). Importantly, CL breast cancers have increased expression of tumor initiating genes (Prat et al., 2010). These CL breast cancer cells respond poorly to standard therapies (Prat et al., 2010). Together, these data suggest that acquisition of EMT in the epithelial cancers induces an invasive mesenchymal phenotype with increased expression of stemness and tumor initiating genes.

Both the TS^{KI4} cells and CL breast cancer cells exhibit the properties of stemness and EMT. DNA microarray data analysis from Abell et al. and Neve et al. found a maximum overlap of gene expression profiles between TS^{KI4} cells and CL breast cancer cells compared to other breast cancer subtypes (Abell et al., 2011; Neve et al., 2006). Analysis of gene expression changes in TS^{KI4} cells and CL breast cancer cells revealed 31 shared down regulated genes, suggesting that these genes may be important for maintaining an epithelial state (Abell et al., 2011). Lentiviral shRNA knockdown of 8 of these 31 genes in the epithelial TS^{WT} cells induced a mesenchymal morphology (Mobley et al., 2017). Importantly, reduced expression of one of these eight genes, *Galnt3* (UDP-GalNAc transferase 3) in the TS^{WT} cells resulted in the most dramatic mesenchymal phenotype. Based on these data, we predicted that GALNT3 may be important for epithelial state maintenance.

GALNT3, an O-GalNAc Glycosyltransferase

GALNT3 initiates O-linked glycosylation in the Golgi by transferring GalNAc to hydroxyl groups of serine or threonine residues on specific proteins (Bennett et al., 2012; Topaz et al., 2004). Importantly, GALNT3 promotes protein folding and localization of both membrane and secreted proteins (Beaman and Brooks, 2014). The structure of GALNT3 includes three protein domains: a transmembrane domain, a glycosyltransferase domain, and a carbohydrate-binding domain. Mutations or deletions in either the glycosyltransferase domain or carbohydrate-binding domain have been implicated in severe human pathologies (Esapa et al., 2012; Garringer et al., 2006; Garringer et al., 2007; Ichikawa et al., 2009). For example, deletion of a 16 bp nucleotide region in exon 1 of *GALNT3* results in a frameshift mutation leading to a premature stop codon in the protein. This GALNT3 protein lacks activity due to the deletion of the glycosyltransferase domain (Garringer et al., 2006). The absence of GALNT3 activity results in misfolding of FGF23, a hormone controlling the phosphate homeostasis, leading to a severe autosomal disorder called hyperphosphatemic familial tumoral calcinosis (HFTC) (Topaz et al., 2004). This disease is characterized by increased serum phosphate levels and calcium deposits in the skin and subcutaneous tissues (Topaz et al., 2004). The role of GALNT3 is well established in a disease state such as tumoral calcinosis (TC). GALNT3 has also been implicated in other diseases such as cancers.

Roles of GALNT3 in Cancer Progression

Expression levels of GALNT3 correlate with the differentiation state of cancer. GALNT3 is strongly expressed in well differentiated cancers including pancreas, colon, stomach, and breast adenocarcinoma (Sutherlin et al., 1997). Poorly differentiated

mesenchymal pancreatic cancers have reduced endogenous *Galnt3* levels relative to moderately/well differentiated pancreatic cancers (Maupin et al., 2010). Reduction of GALNT3 expression in well differentiated pancreatic cancers results in a poorly differentiated cancer phenotype displaying a mesenchymal morphology with increased invasiveness (Maupin et al., 2010). Pancreatic cancers with reduced *Galnt3* expression display increased EMT characteristics. Together, these data suggest that reduced *Galnt3* expression is associated with a poorly differentiated EMT-like state (Maupin et al., 2010). These findings are supported by work from Chugh et al. who showed that loss of GALNT3 in the moderately/well differentiated pancreatic cancers induces EMT with increased expression of Tn and T antigens. These changes in glycosylation may be due to a compensatory mechanism mediated by the expression of other *Galnts* (Chugh et al., 2016). Importantly, knockdown of *Galnt3* in a specific moderately/well differentiated pancreatic ductal adenocarcinoma, S2-013 cells, results in decreased proliferation and induces apoptosis. This phenotype maybe explained through the differential glycosylation of a GALNT3 specific substrate called GNAT-1, a membrane associate 3-subunit guanine nucleotide-binding protein. Loss of GALNT3 in S2-013 cells results in reduced the glycosylation of GNAT1, and altered protein conformation of the protein. siRNA knockdown of *Gnat1* in S2-013 cells also results in a phenotype similar to the loss of GALNT3 (Taniuchi et al., 2011). *Galnt3* expression has also been associated with poor prognosis and low survival rates in patients with lung adenocarcinomas and gastric cancers (Gu et al., 2004; Onitsuka et al., 2003). Patients with moderately/well differentiated lung adenocarcinomas and gastric cancers have better survival rates relative to low GALNT3 expressing poorly differentiated cancers (Gu et al., 2004; Onitsuka et

al., 2003). In addition, the secretory cells including bronchial epithelial cells, alveolar cells, and bronchial gland cells have endogenous expression of GALNT3. Loss of GALNT3 in these tissue types is associated with decreased Ki67 staining, a proliferation marker (Dosaka-Akita et al., 2002). In contrast to pancreatic, lung, and gastric cancers, GALNT3 expression is increased in poorly differentiated ovarian cancers. shRNA knockdown of *Galnt3* in the A2780 ovarian cancer cells results in decreased migration, invasion and proliferation. This phenotype in A2780 ovarian cancers was explained through the differential glycosylation of MUC1 (Wang et al., 2014). Together, these data suggest that the functional consequences of GALNT3 expression depend on cancer cell type, grade of cancer, morphology, and most importantly expression of protein substrates.

Roles of GALNT3 in Development

Although the role of GALNT3 in cancers has been well defined, the functional role of GALNT3 in development is restricted to a pathological disorder, TC (Balin et al., 2012). Importantly, patients with TC have increased phosphate levels and subcutaneous depositions of calcium associated with painful skin ulcerations and secondary infections in bone (Ichikawa et al., 2009). Mutations in GALNT3, FGF23, and Klotho factors each induce TC. To examine the role of GALNT3 in TC, Ichikawa et al. created a GALNT3 deficient mouse model. Although deletion of the *Galnt3* gene results in decreased intact FGF23 secretion and hyperphosphatemia, there was no visible evidence for calcifications (Ichikawa et al., 2009). Interestingly, there were sex specific phenotypic differences, where male mice show infertility, growth retardation, and increased bone mineral density (Ichikawa et al., 2009). In addition to these mouse models with deletion of GALNT3, ENU mutagenesis was used to create mice with a mutation in GALNT3 at Tryptophan

position 589 to Arginine in the carbohydrate binding domain. These mice called TCAL (Tumoral CALcinosis) displayed a TC phenotype characterized by hyperphosphatemia, subcutaneous calcium deposits, and hyperostosis (Esapa et al., 2012). The TCAL male mice also showed infertility with decreased sperm count similar to the GALNT3 deletion mice (Esapa et al., 2012). Together, the GALNT3 deletion and/or mutant mouse models were created to better understand TC. However, the associated developmental defects occurring in these GALNT3 mouse models remain poorly studied.

Epigenetic Regulation of *Galnt3* Expression

Although GALNT3 expression is altered in TS cells and CL breast cancers, the mechanisms controlling the expression of *Galnt3* gene during development and cancer EMT are undefined. As previously described, MAP3K4 is an important regulator of developmental EMT processes in TS cells (Abell et al., 2009; Abell and Johnson, 2005). This kinase, activates CBP, a histone acetyltransferase, to promote acetylation of specific histones H2A and H2B on the promoters of key epithelial genes (Abell et al., 2011). Importantly, loss of MAP3K4 activity in the TS^{KI4} cells decreases CBP activity and increases the expression and activity of a histone deacetylase, HDAC6 (Mobley et al., 2017). MAP3K4-dependent HDAC6 expression and activity lead to the deacetylation of the promoters of many epithelial proteins, thus inducing EMT. Both TS^{KI4} cells and CL breast cancer cells have decreased H2BK5 acetylation (Abell et al., 2011). Using anti-H2BK5Ac antibody chromatin immunoprecipitation coupled to sequencing, we identified a significant list of genes that are predicted to regulate EMT (Mobley et al., 2017). Among the list, we found 1817 genes that displayed decreased H2BK5Ac in the TS^{KI4} cells. Further, using RNA sequencing data for TS^{WT}, TS^{KI4} and TS^{KI4} cells with HDAC6

knockdown we found that 2624 transcripts were reduced in TS^{KI4} cells (Mobley et al., 2017). Altogether, comparison of the genes in H2BK5Ac ChIP sequencing and RNA sequencing data revealed 514 genes that were MAP3K4, HDAC6 and H2BK5Ac dependent. Importantly, eight of 31 shared genes between TS^{KI4} and CL breast cancer cells were also present in the 514 gene list. Among these eight genes, *Galnt3* was one of the top candidate genes. Based on these data, we predict that MAP3K4, HDAC6, and H2BK5Ac may epigenetically regulate *Galnt3* expression.

Specific Aim 1: Examine the role of GALNT3 in regulating EMT in trophoblast stem cells, blastocyst trophectoderm, and human mammary epithelial cells.

Our preliminary data show that *Galnt3* expression is reduced in both TS^{KI4} cells and CL breast cancer cells that display characteristics of stemness and EMT. Based on these preliminary findings, we predict that *Galnt3* is important in the maintenance of the epithelial phenotype. Here, we investigate the role of *Galnt3* on EMT in the blastocyst trophectoderm, TS cells and human mammary epithelial cells (HMECs). Lentiviral shRNA knockdown of GALNT3 in TS^{WT} cells induces a mesenchymal phenotype, decreasing epithelial marker expression and increasing the expression of mesenchymal markers and EMT-inducing transcription factors. Importantly, loss of GALNT3 promotes invasiveness. Further, lentiviral re-expression of GALNT3 in the TS^{KI4} cells restores the epithelial state, thus inducing MET. In the mesenchymal TS^{KI4} cells, re-expression of GALNT3 increases the expression of epithelial markers, decreases mesenchymal marker expression, and importantly decreases invasiveness. These data suggest that GALNT3 regulates EMT in cultured TS cells. In addition, we wanted to examine the impact of loss of *Galnt3* on EMT in the blastocyst trophectoderm. Lentiviral shRNA knockdown of *Galnt3* in the preimplantation blastocyst trophectoderm induces early attachment and outgrowth. Since EMT is reactivated during cancer metastasis, we examined the role of GALNT3 in breast cancer EMT. Mesenchymal CL SUM159 breast cancer cells have decreased *Galnt3* expression relative to HMECs. Importantly, shRNA knockdown of *Galnt3* in HMECs induces a mesenchymal state. Together, these data led to the hypothesis that GALNT3 is important for the epithelial state maintenance. Further, we

hypothesized that GALNT3 might promote the O-GalNAc glycosylation of key proteins important for epithelial state maintenance.

Specific Aim 2: Examine the impact of GALNT3 on O-GalNAc glycosylation levels in trophoblast stem cells, blastocyst trophectoderm, and human mammary epithelial cells.

GALNT3 is a key O-GalNAc glycosyltransferase that catalyzes the addition of O-GalNAc on serine or threonine residues of specific proteins. To measure O-GalNAc glycosylation levels, we used a lectin called Vicia-Villosa Lectin (VVL) that specifically binds to O-GalNAc modified proteins. Using this approach, we examined total O-GalNAc glycosylation of proteins with altered expression of GALNT3 in both TS cells and HMECs. O-GalNAc glycosylation is reduced in TS^{KI4} cells and CL SUM159 cells lacking GALNT3 relative to TS^{WT} cells and HMECs respectively. shRNA knockdown of *Galnt3* in TS^{WT} cells and HMECs reduced O-GalNAc glycosylation. In addition, re-expression of GALNT3 in the TS^{KI4} cells and CL SUM159 breast cancer cells increases total O-GalNAc glycosylation. To measure the cell surface and intracellular localization of O-GalNAc glycosylated proteins, paraformaldehyde fixed and detergent permeabilized immunostaining was performed, respectively. Immunofluorescence staining shows significant cell surface localization of O-GalNAc glycosylated proteins in TS^{WT} cells. Reduced expression of GALNT3 in TS^{KI4} cells results in decreased total O-GalNAc staining and O-GalNAc modified proteins are intracellularly retained in the Golgi. In addition, loss of *Galnt3* in the blastocyst trophectoderm results in decreased total O-GalNAc glycosylation. Importantly, loss of GALNT3 in the TS cells does not induce the compensatory expression of other GALNTs, suggesting that GALNT3 is the critical O-GalNAc glycosyltransferase in TS cells. Together, these data led to our hypothesis that

GALNT3 controls the O-GalNAc glycosylation levels in TS cells and HMECs, thus regulating the cellular phenotype. We further hypothesized that GALNT3 may promote O-GalNAc glycosylation of epithelial proteins for the maintenance of an epithelial state.

Specific Aim 3: Identify EMT-related targets of GALNT3.

The data in Specific Aim 2 suggest that GALNT3 controls the O-GalNAc glycosylation of multiple proteins in both TS^{WT} cells and HMECs. Specifically, we observe differential O-GalNAc glycosylation of a 130 Kda protein between TS^{WT} and TS^{KI4} cells. We identified the band as E-cadherin, a key epithelial protein. Importantly, decreased GALNT3 expression in TS^{KI4} cells results in the intracellular retention of E-cadherin in the Golgi. In addition, re-expression of GALNT3 in TS^{KI4} cells increases O-GalNAc glycosylation of E-cadherin, restoring the cell surface localization of E-cadherin and the epithelial state. Further, we identified a threonine at position 63 in the extracellular cadherin (EC) domain 1 of E-cadherin as the GALNT3 dependent O-GalNAc glycosylation site. Together, this aim identifies a previously unknown target of GALNT3. Importantly, O-GalNAc glycosylation of E-cadherin promotes the cell surface localization of E-cadherin that is critical for the maintenance of the epithelial state.

Chapter II

GALNT3 is Important for Epithelial State Maintenance in Trophoblast Stem Cells, Blastocyst Trophectoderm, and Human Mammary Epithelial Cells

Introduction

Epithelial cell-cell adhesion is promoted by both proteins located at the cell surface and those secreted into the extracellular matrix. These proteins are heavily modified by the addition of carbohydrates controlling their localization, stability, secretion, and proteolytic processing. Carbohydrate modifications occur in two ways: N-glycosylation, carbohydrates attached to amide groups of asparagine (N), and O-glycosylation, carbohydrates attached to hydroxyl groups of serine and threonine (S/T). Proteomic screens suggest the majority of membrane and secreted proteins are glycosylated, but the biological roles of these modifications remain mostly unknown.

Disruption of epithelial cell-cell adhesion occurs during epithelial to mesenchymal transition (EMT) when epithelial cells with tight cell-cell adhesion and apical-basal polarity convert to motile, mesenchymal cells with front-back polarity (Yang and Weinberg, 2008). EMT is critical during mammalian development including implantation, gastrulation, and neural crest formation (Thiery et al., 2009). Importantly, EMT is reversible through a mesenchymal to epithelial transition (MET) in which motile, mesenchymal cells are restored to a non-motile, epithelial state (Yang and Weinberg, 2008). In addition to its role in development, EMT is reactivated during tumor progression and cancer metastasis. Aberrant O-glycosylation has been implicated in cancer EMT (Chia et al., 2016). However, its role in developmental EMT is poorly understood.

In mammals, the first developmental EMT occurs in trophoblast stem (TS) cells during implantation, where epithelial TS cells in the trophectoderm (TE) transition into invasive giant cells (Thiery et al., 2009). Multipotent TS cells isolated from pre-implantation blastocysts can be cultured indefinitely in the presence of FGF4 (Tanaka et al., 1998). Upon FGF4 withdrawal, TS cells differentiate forming all the mature trophoblast subtypes of the placenta (Tanaka et al., 1998). Wild-type TS (TS^{WT}) cells are epithelial with tight cell-cell adhesion and apical-basal polarity. In contrast, TS cells isolated from mice with a targeted mutation that inactivates the kinase activity of MAP3K4 (TS^{KI4} cells) exhibit a mesenchymal morphology with front-back polarity (Abell et al., 2011). Embryos with inactivated MAP3K4 display developmental disorders including neural tube, skeletal, and implantation defects that are due to perturbations in EMT (Abell et al., 2009; Abell and Johnson, 2005). TS^{KI4} cells show key characteristics of EMT, including reduced expression of epithelial markers like E-cadherin with increased expression of mesenchymal markers *Vim* and *Cdh2*, and EMT-inducing transcription factors *Twist1* and *Snai2*, and promoted invasiveness (Abell et al., 2011). Importantly, TS^{KI4} cells and mesenchymal Claudin-low (CL) breast cancer cells share gene expression profiles and display properties of stemness and EMT (Abell et al., 2011).

Using DNA microarray data from Abell et al. (2011) and Neve et al. (2006), we have identified a gene, *Galnt3*, whose expression is reduced in both TS^{KI4} cells and CL breast cancer cells. UDP-GalNAc transferase 3 (GALNT3) is a member of a large family of homologous genes called GALNTs, each displaying selective tissue and target specificities (Beaman and Brooks, 2014)**.** GALNTs initiate O-GalNAc glycosylation in the Golgi through the transfer of N-acetylgalactosamine (GalNAc) onto S/T residues

forming a Tn antigen on specific protein substrates (Bennett et al., 2012). Abnormalities in the activity of GALNTs have been implicated in several pathologies. The role of specific GALNTs in developmental EMT has been difficult to determine using GALNT specific knockout mice, due in part to the potential functional redundancy of this large family of glycosyltransferases. GALNT3 initiates O-GalNAc glycosylation, but its function in stem cells is unknown. Reduced expression of GALNT3 in mesenchymal TS^{KI4} cells relative to TS^{WT} cells and in CL breast cancer cells relative to human mammary epithelial cells (HMECs) suggested that GALNT3 may be important in promoting an epithelial phenotype in these cell types.

Herein, we show that a MAP3K4/HDAC6/H2BK5Ac dependent gene, GALNT3 is a key regulator of EMT. shRNA knockdown of GALNT3 in TS cells, blastocyst trophectoderm (TE), and HMECs results in the acquisition of the mesenchymal state. In addition, loss of GALNT3 in TS^{WT} cells results in the reduced epithelial marker expression with gain of mesenchymal markers and EMT-inducing transcription factors. Importantly, GALNT3 re-expression in the mesenchymal TS^{KI4} cells restores the epithelial phenotype. Finally, loss of *Galnt3* in the trophectoderm of the blastocyst results in premature attachment and outgrowth. Together, our findings demonstrate that GALNT3 is critical for the epithelial state maintenance in both the TS cells and HMECs.

Materials and Methods

Animals

Kinase-inactive MAP3K4 (KI4) mice having a point mutation of Lysine 1361 to Arginine were created as described in Abell et al. 2005. These mice were maintained on a pure 129/SvEv background as described in Abell et al. 2009. Because this mutation is

embryonic lethal, mice heterozygous for the mutation were used to maintain the line. Mice were housed under standard animal house conditions. Wild-type mice used in these experiments were generated from heterozygous crosses of the KI4 mice. Male and female wild-type mice ages 2-6 months were used in timed breedings to generate blastocysts. Female mice were euthanized 3.5 days post mating and blastocysts were isolated as described under the heading Blastocyst analysis. The procedures and the project were approved by and performed in accordance with the University of Memphis (Protocol number 0825) Institutional Animal Care and Use Committee (IACUC) following NIH guidelines.

Cell lines and culture conditions

Mouse extraembryonic TS^{WT} cells (male) and TS^{KI4} cells (female) were isolated as previously described from heterozygote crosses of mice with a targeted mutation in MAP3K4 (K1361R) that inactivates MAP3K4 kinase activity (Abell et al., 2009). TS cells were cultured in the absence of feeders in 30% TS media (RPMI 1640, 20% heat inactivated fetal bovine serum (FBS), 1% penicillin and streptomycin (PS), 1% L-glutamine, 1% sodium pyruvate, and 100 μM β-mercaptoethanol) and 70% TS media conditioned by mitotically inactivated mouse embryonic fibroblasts (MEF-CM). TS cells were supplemented with FGF4 (37.5 ng/ml) and Heparin (1 μg/ml) to maintain their stemness. In differentiation experiments, TS cells were cultured in TS media lacking growth factors and MEF-CM. The human cell lines HMECs (adult female), SUM159 cells (adult female), and HEK293T (fetal female) were a kind gift from Dr. Gary Johnson (UNC, Chapel Hill). HMEC cells were cultured in HuMEC ready medium (Thermo Fisher Scientific) containing 5% FBS, 1% PS with HuMEC supplement and bovine

pituitary extract. SUM159 cells were cultured in Ham's F12 medium (Thermo Fisher Scientific) supplemented with 5% FBS, 1% PS with 5 μg/ml insulin and 1 μg/ml hydrocortisone. Dulbecco's modified essential medium supplemented with 10% FBS and 1% PS was used to culture HEK293T cells. All cell lines were cultured in a humidified atmosphere at 37ºC containing 7% CO_2 (mouse) or 5% CO_2 (human).

Plasmids

Lentiviral pLKO, TRCN0000055098 and TRCN0000055099 plasmids (Open Biosystems) were used to create control and murine GALNT3 knockdown cells. TRCN0000035456 and TRCN0000035458 plasmids (Dharmacon) were used to create human GALNT3 knockdown cells. To create a human FLAG-GALNT3, we used clone # 55179 BCII3565 in pDONR 223 (Vidal human ORFeome (Version 5.1)). These inserts were cloned into a lentiviral FLAG tagged destination vector (Jordan et al., 2013) using LR clonase II enzyme mix (Invitrogen).

Lentiviral production and infection

HEK293T cells (fetal female) were used to produce replication incompetent lentivirus as previously described (Abell et al., 2011). Briefly, HEK293T cells were cotransfected with either pLKO, shRNAs (Open Biosystems) or FLAG-GALNT3 constructs in combination with pMD2.G and psPAX2 (Addgene) using calcium phosphate. After 48-72 hours, viral supernatants were harvested by ultracentifugation at 21,000 rpm for 2 hours at 4C and viral pellets were resupended in 100 μl of TS media with growth factors or HuMEC media or Ham's F12 media. Infection of TS cells was performed as previously described (Abell et al., 2011). Transduced cells were selected using puromycin (2 μg/ml).

Real-time quantitative PCR

Isolation of RNA from TS cells was performed using RNeasy Plus minikit (Qiagen). High-Capacity reverse transcription kit (Thermo Fisher Scientific) was used to prepare cDNA from 3 μg RNA. The Bio-Rad CFX96 Touch qPCR system with iTaq (Bio-Rad) was used to measure gene expression changes. Expression levels were measured using $2^{-\Delta\Delta Ct}$ method and were normalized to mouse *Actb* or human *Gapdh* or human *Tbp*. Primers used are specified in Appendix E.

Western blot analysis

Whole-cell, cytoplasmic, and nuclear lysates of TS cells were harvested as previously described (Abell et al., 2009; Abell and Johnson, 2005). Briefly, whole-cell lysates were harvested in buffer A (20 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA and 1% Triton X) with protease inhibitors (1 mM PMSF and 17 μg/ml aprotinin) and phosphatase inhibitors (1 mM sodium vanadate and 1 mM sodium fluoride). Cytoplasmic lysates were harvested in 0.5% Triton in 1X phosphate buffered saline with protease and phosphatase inhibitors described above. Nuclear lysates were harvested in RIPA buffer (buffer A, 0.1 mg/ml sodium dodecyl sulphate and 0.1 mg/ml sodium deoxycholate) with protease and phosphatase inhibitors described above. Lysates were probed with the indicated antibodies specified in the Appendix A.

Immunofluorescence staining

Immunostaining was performed as previously described (Abell et al., 2009). Briefly, cells were cultured on glass coverslips for two or three days. Cells were fixed with 3% paraformaldehyde in 1X PBS for 10 min and then washed three times with 1X PBS. Cells were permeabilized with 0.1% Triton for 3 min. Cells were blocked with 10%

fetal bovine serum for 1 hour at RT. Coverslips were incubated with primary antibody overnight at 4C. Next, the cells were washed for 30 min and then incubated with DAPI (0.1 μg/ml), Dy-Alexa 488 (1:500) (Thermo Fisher Scientific), Alexa 594 (1:250) (Cell Signaling Technology) for one hour at RT. Coverslips were washed and mounted on slides with mounting media (90% glycerol and 10% 1mM Tris pH 7.5). Coverslips were imaged using an EVOS epifluorescence microscope or a Nikon A1 laser scanning confocal microscope.

Immunofluorescence imaging

Immunofluorescence imaging was performed using Nikon A1 laser scanning confocal microscopy and a 60X Plan Apochromat 1.4 numerical-aperture (NA) oil objective with lasers at 408, 488 and 594 nm. Individual Z-stacks and the 2D images were obtained by either Maximum Intensity Projection (MIP) or Extended Depth of Field (EDF) for all wavelengths.

Blastocyst analysis

Wild-type blastocysts were isolated from timed matings of pure 129/SvEv males with females in the pure 129/SvEv and 129/SvEv/C57B6 mixed backgrounds. Mid-stage blastocysts were hatched by transfer into drops of Acid-tyrodes (Millipore Sigma) for two min. The trophectoderm layer of the hatched, intact blastocysts was infected as previously described (Georgiades et al., 2007). Briefly, 100 μl drops of lentiviruses encoding shControl or shGalnt3 resuspended in KSOM-AA (Millipore Sigma) were used to infect hatched blastocysts for 6 hours. Post infection, blastocysts were transferred into KSOM-AA for 24 hours. Attachment of the blastocysts was assessed after transfer of infected blastocysts into 70% TS medium supplemented with FGF4.

For measurement of puromycin resistance gene expression, RNA was isolated five days post infection from attached and differentiated trophoblast cells in shControl and shGalnt3 using RNeasy micro kit (Qiagen). High-Capacity reverse transcription kit (Thermo Fisher Scientific) was used to prepare cDNA. PCR reactions for puromycin resistance gene encoded by the lentiviruses were performed. PCR products were run on a 2% agarose gels and imaged using Bio-Rad ChemiDoc.

Chromatin IP coupled to qPCR

ChIP was performed as previously described (Abell et al., 2011). Sonicated samples were immunoprecipitated (IP) with 5 μg of anti-HDAC6 antibody (Cell Signaling Technology) or 5 μg of anti-H2BK5Ac antibody (Abcam). The Bio-Rad CFX96 Touch qPCR system with SsoFast (Bio-Rad) was used to measure promoter enrichment. Primers used for ChIP coupled to qPCR (ChIP-PCR) are specified in Appendix E.

H2BK5Ac ChIP-seq read density plot analysis

Anti-H2BK5Ac ChIP-seq data from database: GSE92426 aligned to the mm9 reference genome was visualized as in Mobley et al. using Easeq. Aligned data files were deposited as Datasets and annotated using the mouse mm9 genome Geneset. The FillTrack utility in Easeq software was used to display the read density relative to input at the *Galnt3* transcription start site + 20 kb. The units of the y-axis are normalized, binned read counts calculated as specified in Lerdrup et al. (Lerdrup et al., 2016)

Transwell solute flux assay

Barrier assays were performed as previously described (Mobley et al., 2017). Briefly, TS cells were plated in 24 well 0.4 μm pore size PET membrane Transwells

(TS^{KI4Gab} cells) or polycarbonate membrane transwells (TS^{WTGsh} cells) (Corning Life Sciences). Cells were grown to confluence for three days and then dye diffusion was measured using the Synergy H1MD plate reader (BioTek).

Invasion assay

Invasion assays were performed as previously described (Abell et al., 2011). Briefly, TS cells were plated on growth factor reduced Matrigel coated 8 μM pore transwell (TW) chambers. After 48 hours, invasion assays were terminated. Non-invading cells from the top of the TW were removed using swabs and 1X PBS washes. Invasive cells at the bottom of the TW were fixed in 3% paraformaldehyde for 10 min and stained with DAPI (2 μg/ml) for 30 min. Five 10X fields for each TW were imaged using EVOS epifluorescence microscope. These images were manually counted, and graphs were plotted using Prism7 (version 7.0, GraphPad Prism software).

Statistical methods

The statistical details of the experiments including statistical tests used, value of n, and what n represents, can be found in the figure legends. A *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, and ****p-value < 0.0001 were considered statistically significant. Softwares used to calculate statistical analyses are detailed below.

For qPCR data, the Bio-Rad CFX Maestro software (version 3.0) was used to perform a Student's t test corrected for multiple comparisons. For all other statistics shown, Prism7 (version 7.0, GraphPad software) was used to perform a Student's t test corrected for multiple comparisons.

Breast cancer data statistical analysis

For the analysis of breast cancer cell line data from Neve et al., the molecular subtypes were assigned by Prat et al. based on gene expression signatures. Of the 52 cell lines analyzed, 49 were classified into molecular subtypes by hierarchical clustering by Prat et al: 25 as luminal, 15 as basal-like, and 9 as claudin-low. Using the "stats" package in "R", the following analyses were performed on the 49 classified cell lines (Team, 2013). The significance of differences in expression of *Galnt3* between luminal, basal-like, and claudin-low subtypes was assessed using the Student's t test with Bonferroni's correction for multiple comparisons (pairwise.t.test()). Linear regression analysis (lm()) was used to model the relationship between *Galnt3* and *Cdh1* gene expression across the 49 classified breast cancer cell lines.

Results

Loss of GALNT3 is associated with acquisition of the mesenchymal state

We previously demonstrated that TS^{KI4} cells lacking MAP3K4 kinase activity exist in a metastable state, having reduced epithelial features and increased mesenchymal characteristics while maintaining stemness (Abell et al., 2011). Unlike TS^{WT} cells that displayed an epithelial morphology with tight cell-cell adhesion and apical-basal polarity, TS^{KI4} cells displayed a mesenchymal-like morphology with the loss of cell-cell adhesion and gain of front-back polarity (Figure 4A). Similar to TS^{KI4} cells, CL breast cancer cells display properties of both EMT and stemness (Abell et al., 2011). Unlike HMECs that have an epithelial morphology, the CL breast cancer cell line SUM159 displayed a mesenchymal morphology, consistent with the EMT features of CL breast cancer cells (Figure 4B).

Analysis of DNA microarray data from Abell et al. suggested that *Galnt3* expression is reduced in TS^{KI4} cells relative to TS^{WT} cells. *Galnt3* transcript and protein were reduced in TS^{KI4} cells relative to TS^{WT} cells as measured by real time quantitative PCR (qPCR) and Western blotting, respectively (Figures 4C and 4D). Induction of EMT through differentiation of TS cells also decreased *Galnt3* transcript levels (Figure 4E). Examination of *Galnt3* expression in the breast cancer cell line microarray dataset from Neve et al. revealed that loss of *Galnt3* expression was strongly associated with the CL form of breast cancer ($p < 0.0001$) (Figure 4F). E-cadherin levels are lowest in the CL form of breast cancer as compared to all other breast cancer subtypes (Prat et al., 2010). Comparison of *Cdh1* and *Galnt3* expression in breast cancer cells from the Neve et al. dataset showed that *Galnt3* loss was associated with the loss of *Cdh1* ($p < 0.0001$), suggesting a strong correlation between *Galnt3* and *Cdh1* expression (Figure 4G). *Galnt3* transcript and protein were reduced in mesenchymal SUM159 cells relative to epithelial HMECs (Figures 4H and 4I). These data showed that GALNT3 expression was decreased in both mesenchymal-like TS^{KI4} cells and CL breast cancer cells, suggesting that loss of GALNT3 is associated with loss of the epithelial state in TS cells and CL breast cancer.

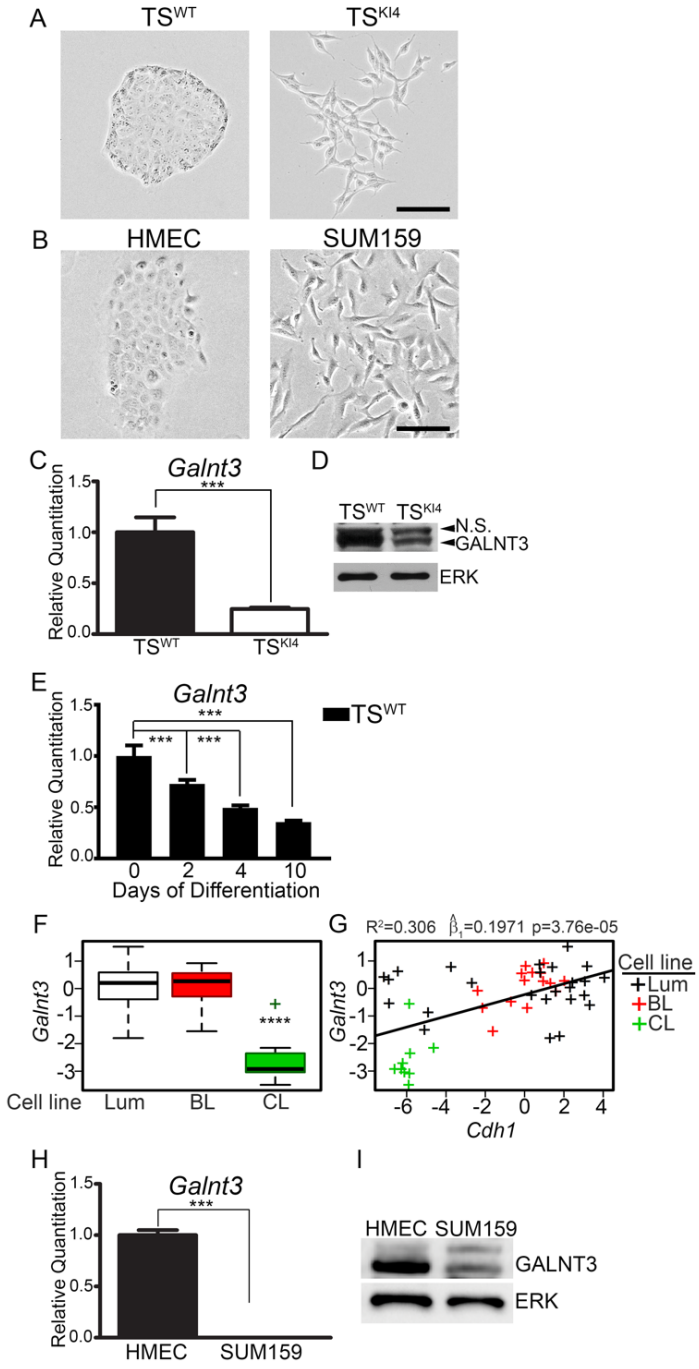

Figure 4. Loss of GALNT3 correlates with the gain of a mesenchymal state. (A-B) Phase microscopy images are representative of three independent experiments. Black bar represents 100 µm. (C) Reduced *Galnt3* transcripts in mesenchymal TS^{KI4} cells relative to TS^{WT} cells. qPCR data normalized to *Actb* are expressed as a fold-change relative to TS^{WT} cells and are the mean ± range of two independent experiments. (D) Western blots are representative of three independent experiments. (N.S., Non-Specific). (E) Decreased

Figure 4 (Continued)

Galnt3 transcripts with differentiation of TS^{WT} cells by FGF4 withdrawal for the indicated number of days. qPCR data normalized to *Actb* are expressed as a fold-change relative to undifferentiated TS^{WT} cells (0 day) and are the mean ± range of two independent experiments. (F) *Galnt3* expression is significantly reduced in Claudin-low (CL) breast cancer cell lines. Luminal (Lum), Basal-like (BL), and CL breast cancer subtypes. p value was calculated using Student's t test and Bonferroni's correction for multiple comparisons. **** p-value < 0.0001. (G) Loss of *Galnt3* is significantly correlated with reduced *Cdh1* levels. Plot shows the correlation between *Galnt3* and *Cdh1* transcripts in breast cancer subtypes. Each "+" represents a specific breast cancer line. p value was calculated using linear regression and $\hat{\beta}1$ represents estimate of the slope of linear regression line. (H) *Galnt3* transcripts are reduced in mesenchymal SUM159 cells relative to HMECs. qPCR data normalized to *Gapdh* are expressed as a fold-change relative to HMECs and are the mean ± SEM of three independent experiments. (I) Western blots are representative of three independent experiments. *** p-value < 0.001, Student's t test.

Transcriptional regulation of *Galnt3* expression by MAP3K4 and HDAC6

We have previously demonstrated that MAP3K4 inhibits HDAC6 expression and activity (Mobley et al., 2017). HDAC6 induces EMT in TS cells by binding to the promoters of epithelial maintenance genes, deacetylating histone H2BK5, and repressing gene expression. RNA-seq data from Mobley et al. suggested that *Galnt3* expression was HDAC6 dependent. HDAC6 expression at both transcript and protein levels was elevated in TS^{KI4} cells relative to TS^{WT} cells (Figure 5A and 5B). HDAC6 protein was also elevated in the nuclei of SUM159 cells relative to HMECs (Figure S1A). Knockdown of HDAC6 in TS^{KI4} cells with a short hairpin RNA (shRNA) ($TS^{KI4H6sh}$ cells) reduced HDAC6 transcript and protein and restored an epithelial morphology (Figure 5A-5C). *Galnt3* transcripts were increased in $TS^{KI4H6sh}$ cells relative to TS^{KI4} cells resulting in a 50% increase in GALNT3 protein (Figure 5D and 5E). These data suggested that HDAC6 inhibits the expression of *Galnt3*.

We predicted that HDAC6 regulates the expression of *Galnt3* by directly binding the *Galnt3* promoter and deacetylating H2BK5. As measured by anti-HDAC6 chromatin immunoprecipitation coupled to qPCR (ChIP-PCR), HDAC6 was enriched on the *Galnt3* promoter in TS^{KI4} cells relative to TS^{WT} cells (Figure 5F and Figure S1B). Anti-H2BK5Ac ChIP-PCR and published ChIP-seq data from Mobley et al. showed reduced H2BK5Ac on the *Galnt3* promoter in TS^{KI4} cells relative to TS^{WT} cells (Figure 5G and 5H and Figure S1C). Knockdown of HDAC6 in TS^{KI4} cells increased H2BK5Ac on the *Galnt3* promoter (Figure 5G and 5H). Similar to TS cells, anti-HDAC6 ChIP-PCR showed enrichment of HDAC6 on the *Galnt3* promoter in SUM159 cells relative to HMECs (Figures S1D and S1E). Further, anti-H2BK5Ac ChIP-PCR showed a

concomitant decrease in H2BK5Ac on the *Galnt3* promoter in SUM159 cells (Figure S1F). These data suggested that HDAC6 is bound to the *Galnt3* promoter in mesenchymal-like TS^{KI4} cells and CL SUM159 breast cancer cells, actively deacetylating H2BK5 and repressing *Galnt3* expression.

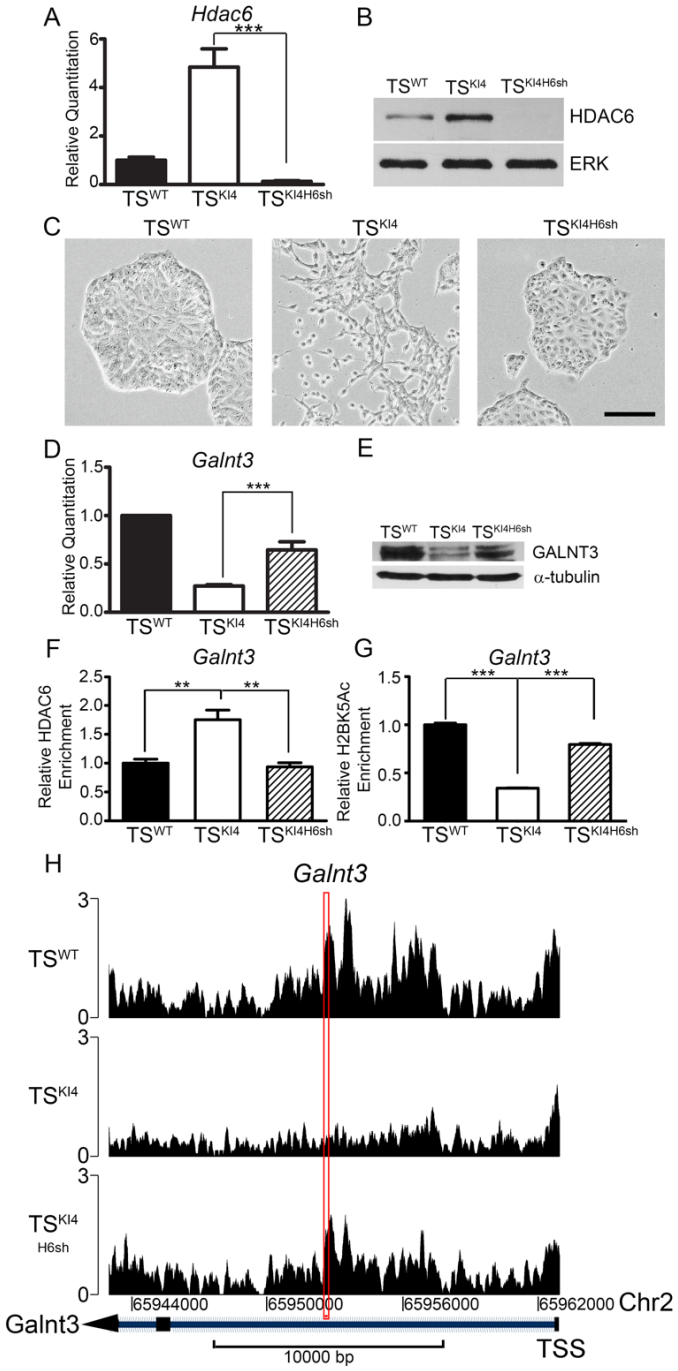

Figure 5. *Galnt3* expression is co-regulated by MAP3K4 and HDAC6. (A) *Hdac6* transcripts analyzed by qPCR in TS^{WT} cells and TS^{KI4} cells expressing control shRNA or TS^{KI4} cells expressing *Hdac6* shRNA ($TS^{KI4H6sh}$) are shown. qPCR data normalized to *Actb* are expressed as a fold-change relative to TS^{WT} cells and are the mean ± SEM of three independent experiments. (B) Western blots of whole cell lysates are representative

Figure 5 (Continued)

of three independent experiments. (C) Knockdown of HDAC6 in TS^{KI4} cells restores an epithelial morphology. Phase microscopy images are representative of three independent experiments. Black bar represents 100 μm. (D) *Hdac6* knockdown increases *Galnt3* transcripts in $TS^{KI4H6sh}$ cells relative to TS^{KI4} cells. qPCR data normalized to *Actb* are expressed as a fold-change relative to TS^{WT} cells and are the mean ± SEM of five independent experiments. (E) GALNT3 protein expression increases with knockdown of HDAC6 in $TS^{KI4H6sh}$ cells. Western blots are representative of two independent experiments. (F) Anti-HDAC6 ChIP-PCR shows increased HDAC6 enrichment on the *Galnt3* promoter in TS^{KI4} cells relative to TS^{WT} cells. Data shown are the mean ± SEM of four independent experiments. (G) Decreased H2BK5 acetylation on the *Galnt3* promoter in TS^{KI4} cells relative to TS^{WT} cells as measured by anti-H2BK5Ac ChIP-PCR. Data shown are the mean ± SEM of six independent experiments. (H) Knockdown of HDAC6 increases H2BK5 acetylation on the *Galnt3* promoter in $TS^{KI4H6sh}$ cells relative to TS^{KI4} cells. H2BK5Ac ChIP-seq read density plots at the *Galnt3* transcription start site + 20 kb are shown. Red rectangle indicates region amplified in ChIP PCR reactions in panels (F) and (G). Data are expressed as the normalized read count of the IP divided by the normalized read count of the TS^{WT} cell input. ** p-value < 0.01; *** p-value < 0.001, Student's t test. See also Figure S1.

Knockdown of GALNT3 in TS^{WT} cells induces EMT

To define the role of GALNT3 in TS cell EMT, we infected TS^{WT} cells with lentiviruses expressing two independent *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2} cells). GALNT3 expression was reduced > 98% at both message and protein levels in TS^{WTGsh1} and TS^{WTGsh2} cells relative to TS^{WT} cells (Figures 6A and 6B). Loss of GALNT3 expression in TS^{WTGsh1} and TS^{WTGsh2} cells induced a mesenchymal morphology similar to TS^{KI4} cells (Figure 6C). Transcripts of epithelial genes *Cdh1*, *Krt8, Krt18* and *Cldn6* were reduced in TS^{WTGsh1} and TS^{WTGsh2} cells relative to TS^{WT} cells (Figure 6D). E-cadherin and Claudin-6 protein levels were also diminished with GALNT3 knockdown (Figures 6E and 6F). In addition to the loss of epithelial markers, TS^{WTGsh1} and TS^{WTGsh2} cells gained the expression of mesenchymal markers *Cdh2* and *Vim* and the EMT-inducing transcription factors *Twist1*, *Zeb1*, and *Snai2* (Figures 6G and 6H). Consistent with changes in Claudin-6 transcript and protein, immunofluorescence confocal microscopy showed reduced Claudin-6 expression and loss of co-localization with ZO1 at cell-cell junctions in TS^{KI4} cells and TS^{WTGsh} cells (Figure S2A). Further, fluorescent dye exclusion assays showed increased permeability in TS^{KI4} cells and TS^{WTGsh} cells relative to TS^{WT} cells (Figure 6I). These data suggested that loss of GALNT3 reduces epithelial features and disrupts barrier formation.

A key feature of EMT is the acquisition of cell motility and/or invasiveness. Cell movement assessed through live cell imaging showed TS^{WT} cells remaining in a static position while rapidly dividing (data not shown). In contrast, TS^{KI4} cells showed frequent, rapid movements throughout the field (data not shown). Knockdown of GALNT3 in TS^{WT} cells increased motility similar to TS^{KI4} cells (data not shown). In

addition to examining motility, we measured invasiveness using growth factor reduced Matrigel-coated transwell assays. TS^{KI4}, TS^{WTGsh1} and TS^{WTGsh2} cells displayed increased invasiveness relative to TS^{WT} cells, suggesting that loss of GALNT3 increases invasiveness (Figure 6J). Although loss of GALNT3 induced EMT, cells lacking GALNT3 retained the expression of stemness markers and the ability to differentiate into all trophoblast subtypes, suggesting that stemness features were maintained in the absence of GALNT3 (Figures S2B-S2I). These data showed that TS cells lacking GALNT3, either through loss of MAP3K4 activity in TS^{KI4} cells or with shRNA knockdown of GALNT3 in TS^{WT} cells, expressed both stemness and EMT features. Together, these data suggested that loss of GALNT3 in TS cells induces EMT.

We also examined the role of GALNT3 in HMECs by transiently expressing two independent *Galnt3* shRNAs. We were unable to achieve stable reduction of GALNT3 expression in HMECs, suggesting that GALNT3 performs key functions. Transient expression of *Galnt3* shRNAs resulted in the partial reduction of *Galnt3* transcripts and morphological changes that were similar to TS^{KI4} cells (Figures S2J and S2K). Control infected HMECs formed epithelial colonies whereas GALNT3 knockdown cells were individual with front to back polarity and long cytoplasmic extensions (Figure S2K). Although total E-cadherin protein levels were unchanged with knockdown of GALNT3 in HMECs, cell movement was altered (Figure S2L and data not shown). Control infected HMECs moved as sheets of cells (data not shown). In contrast, HMECs with GALNT3 knockdown moved more as individual cells with front to back polarity that was similar to SUM159 cells (data not shown). Together, these data suggested that GALNT3 is

important for maintaining the epithelial state in both TS cells and HMECs as reduced expression of GALNT3 induced EMT characteristics.

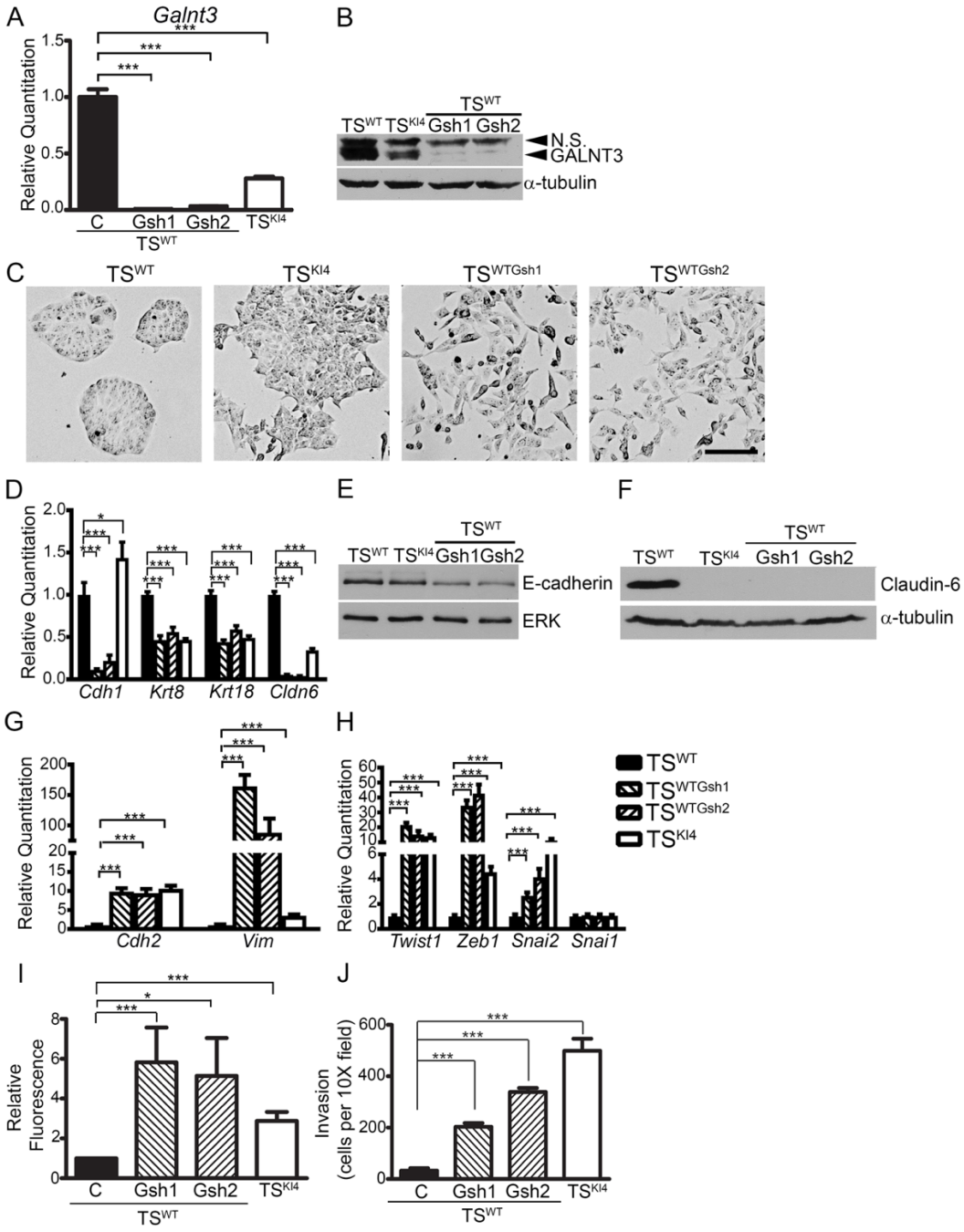

Figure 6. Knockdown of GALNT3 in TS^{WT} cells induces EMT. (A) Transcripts were measured using qPCR in TS^{WT} and TS^{KI4} cells expressing control shRNA (C) or TS^{WT} cells expressing two independent *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2}). (B) Western blots of whole cell lysates are representative of three independent experiments. (N.S., Non-Specific). (C) GALNT3 knockdown in TS^{WT} cells induces a mesenchymal

Figure 6 (Continued)

morphology. Representative phase microscopy images from three independent experiments are shown. Black bar represents 100 μm. (D) GALNT3 knockdown in TS^{WT} cells reduces epithelial marker expression in TS^{WTGsh1} and TS^{WTGsh2} cells. (E, F) Western blots are representative of three independent experiments. (G, H) Increased transcripts of mesenchymal markers (G) and EMT-inducing transcription factors (H) with reduced *Galnt3* expression are shown. (I) Impaired barrier formation in TS^{WTGsh1} and TS^{WTGsh2} cells with reduced GALNT3 expression. Diffusion of fluorescent dye across a confluent monolayer of cells is expressed as a fold-change in fluorescence relative to TS^{WT} cells. Data shown are the mean ± SEM of three independent experiments. (J) Loss of GALNT3 increases invasiveness in TS^{WTGsh1} and TS^{WTGsh2} cells through growth factor reduced Matrigel coated transwells. Data show cells per 10X field and are the mean ± range of two independent experiments performed in triplicate. (A, D, G, H) qPCR data normalized to *Actb* are expressed as a fold-change relative to TS^{WT} cells and are the mean ± SEM of three independent experiments. * p-value < 0.05; *** p-value < 0.001, Student's t test. See also Figures S2.

Loss of *Galnt3* in the blastocyst TE results in premature EMT

To further define the role of GALNT3 in TS cells, we examined the impact of loss of GALNT3 in pre-implantation blastocysts. Mid-stage E3.5 blastocysts were hatched and infected with lentiviruses expressing either control shRNA or *Galnt3* shRNA. This approach has previously been shown to only infect the TE layer of these intact blastocysts (Georgiades et al., 2007). Using semi-quantitative PCR, we validated that infected blastocysts expressed transcripts for the puromycin resistance gene encoded by these viruses (Figure 7A). shRNA knockdown of *Galnt3* in the TE of intact blastocysts resulted in premature attachment and outgrowth compared to control shRNA infections (Figures 7B and 7C). Attachment of blastocysts with *Galnt3* knockdown occurred eight hours earlier than either control shRNA infected or uninfected blastocysts, suggesting that loss of *Galnt3* leads to premature EMT (Figure 7B). Together, these data demonstrate the role of *Galnt3* in EMT of the TE, which is the origin of TS cells.

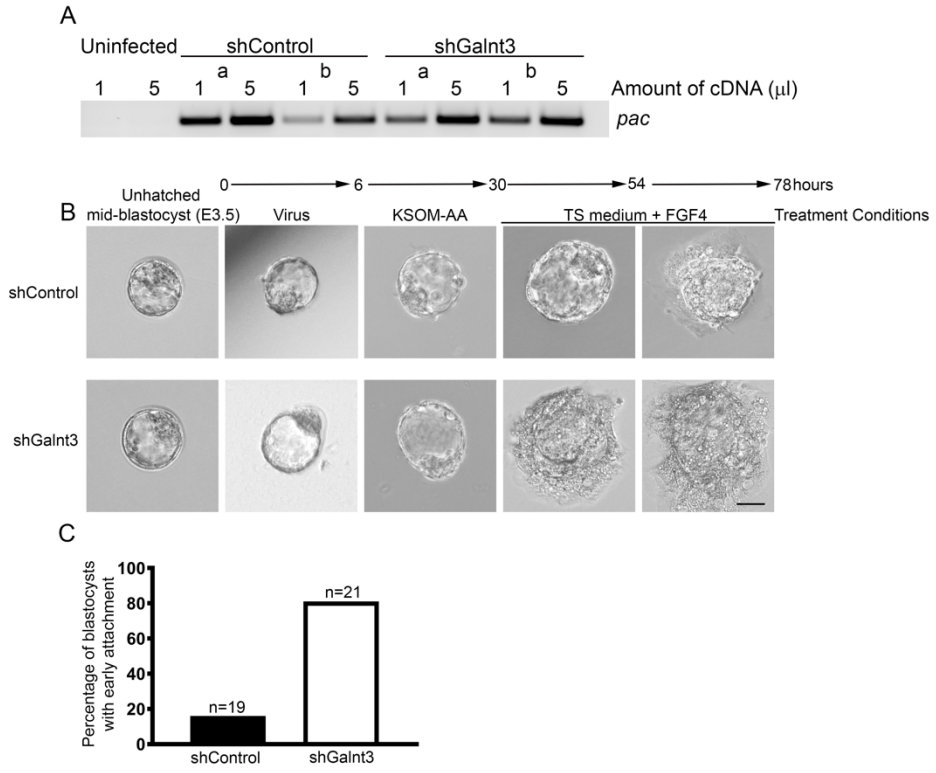

Figure 7. Reduction of *Galnt3* expression in the trophectoderm of E3.5 blastocysts results in premature EMT. (A) Semi-quantitative PCR gel for transcript expression of the puromycin resistance gene (*pac*) is shown for blastocysts that are uninfected or infected with lentiviruses expressing shControl or shGalnt3 constructs. a and b represent two independent blastocysts that were infected with either shControl or shGalnt3 lentiviruses. PCR gel shown is representative of seven independent blastocyst experiments. (B) *Galnt3* knockdown in blastocysts induced premature attachment and outgrowth compared to control infected blastocysts. Phase images show the timeline of infection and treatment of both shControl and shGalnt3 infected blastocysts. TS, Trophoblast Stem cells; FGF4, Fibroblast Growth Factor 4. Black bar represents 50 μm. (C) Reduced *Galnt3* expression results in early attachment of blastocysts. Quantitation of experiments as performed in (B). n, total number of blastocysts assessed for shControl or shGalnt3 infections.

Re-expression of GALNT3 in mesenchymal-like TS^{KI4} cells restores the epithelial state

Using a lentiviral construct for human *Galnt3*, we re-expressed GALNT3 in TS^{KI4} cells ($TS^{KI4Gab1}$ and $TS^{KI4Gab2}$ cells). *Galnt3* transcripts were increased in $TS^{KI4Gab1}$ and $TS^{KI4Gab2}$ cells relative to TS^{KI4} cells as measured with primers detecting both human and mouse *Galnt3* (Figure 8A). Similarly, GALNT3 protein levels in $TS^{KI4Gab1}$ and $TS^{KI4Gab2}$ cells were also increased relative to TS^{KI4} cells (Figures 8B and 8C). Importantly, re-expression of GALNT3 in TS^{KI4Gab} cells restored an epithelial morphology in TS^{KI4} cells that was similar to TS^{WT} cells (Figure 8D). In addition, transcripts of epithelial markers *Krt8*, *Krt18*, and *Cldn6* were also increased in $TS^{KI4Gab1}$ and $TS^{KI4Gab2}$ cells relative to TS^{KI4} cells (Figure 8E). Further, E-cadherin and Claudin-6 proteins were expressed in $TS^{KI4Gab1}$ and $TS^{KI4Gab2}$ cells at levels similar to TS^{WT} cells (Figures 8F and 8G). Fluorescent dye flux was reduced in TS^{KI4Gab} cells relative to TS^{KI4} cells, suggesting partial restoration of tight junctions and barrier function with re-expression of GALNT3 (Figure 8H). Together, these data suggested that GALNT3 re-expression in TS^{KI4} cells restored epithelial features.

Re-expression of GALNT3 in TS^{KI4} cells reduced the expression of mesenchymal markers and EMT-inducing transcription factors. Although *Cdh2* expression remained elevated, $TS^{KI4Gab1}$ and $TS^{KI4Gab2}$ cells showed reduced expression of the mesenchymal marker *Vim* and EMT-inducing transcription factors *Twist1*, *Zeb1*, and *Snai2* relative to TS^{KI4} cells (Figures 8I and 8J). Importantly, TS^{KI4Gab} cells had reduced cell motility as measured by live cell imaging and GALNT3 re-expression reduced invasiveness to TS^{WT} cell levels (data not shown and Figure 8K). Altogether, these data suggested that re-

expression of GALNT3 in TS^{KI4} cells induces MET, restoring the epithelial phenotype in TS cells.

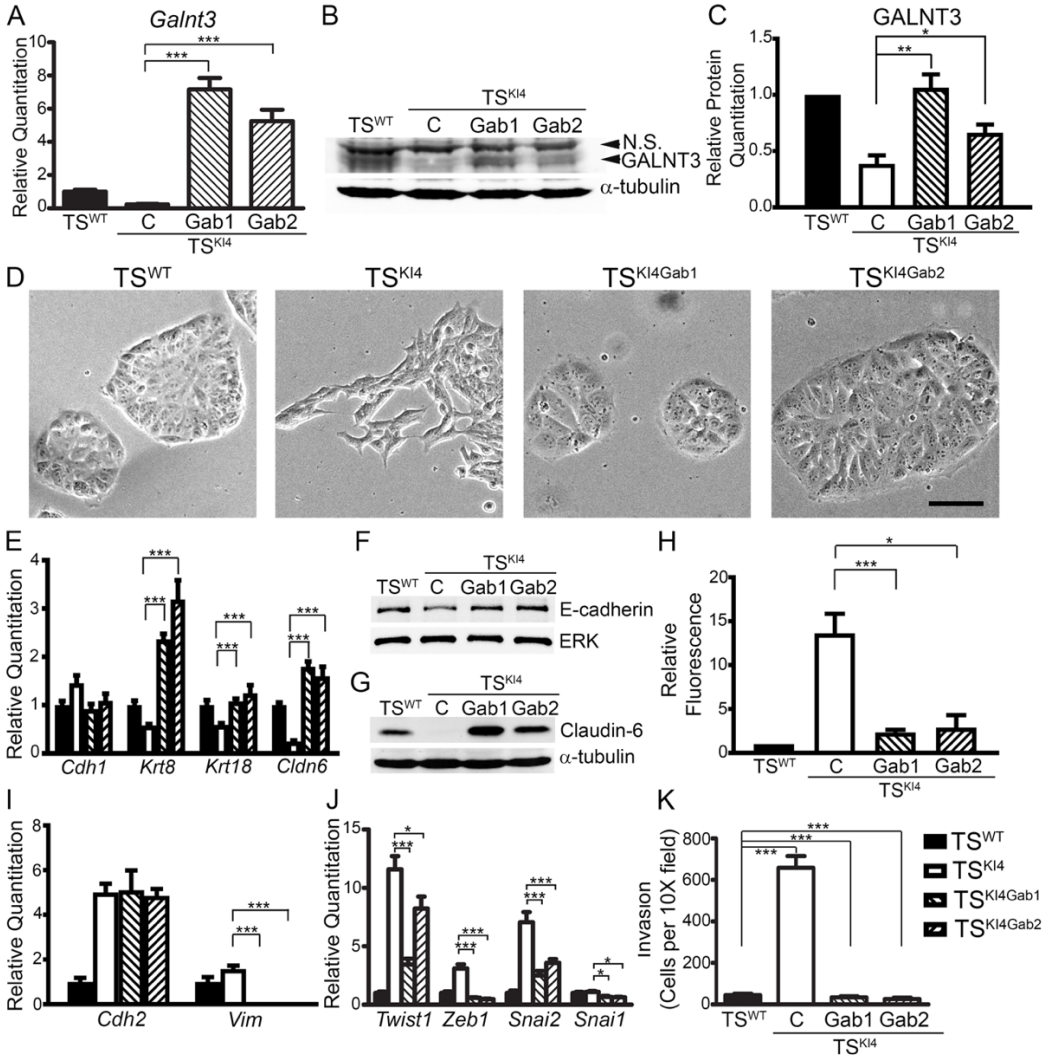

Figure 8. GALNT3 re-expression in TS^{KI4} cells restores an epithelial phenotype. (A) Increased *Galnt3* transcripts in TS^{KI4} cells expressing human *GALNT3*. Transcripts were measured using qPCR in TS^{WT} and TS^{KI4} cells infected with a control lentiviral construct (C) or TS^{KI4} cells infected with a lentiviral construct expressing human *Galnt3* ($TS^{KI4Gab1}$ and $TS^{KI4Gab2}$). (B, C) Re-expression of GALNT3 in $TS^{KI4Gab1}$ and $TS^{KI4Gab2}$ cells was measured by Western blotting. (B) Blots are representative of three independent experiments. (C) Densitometry was used to quantify three independent experiments. (D) GALNT3 re-expression restores an epithelial morphology. Representative phase microscopy images from three independent experiments are shown. Black bar represents 100 μm. (E) Increased transcripts of epithelial markers with re-expression of GALNT3. (F, G) Western blots are representative of three independent experiments. (H) GALNT3 re-expression promotes barrier formation. Diffusion is expressed as a fold-change in

Figure 8 (Continued)

fluorescence relative to TS^{WT} cells. Data shown are the mean ± SEM of three independent experiments. (I, J) Re-expression of GALNT3 reduces mesenchymal markers (I) and EMT-inducing transcription factors (J). (K) Re-expression of GALNT3 in TS^{KI4} cells reduces invasiveness through growth factor reduced Matrigel. Data show cells per 10X field and are the mean ± range of two independent experiments performed in triplicate. (A, E, I, J) qPCR data normalized to *Actb* are expressed as a fold-change relative to TS^{WT} cells and are the mean ± SEM of three independent experiments. * p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001, Student's t test.

Re-expression of GALNT3 in CL SUM159 cells fails to restore epithelial state

In contrast to TS^{KI4} cells, re-expression of GALNT3 in SUM159 cells to levels greater than HMECs failed to alter cell morphology (Figures 9A-9C). Although GALNT3 colocalized with the Golgi marker Giantin, epithelial proteins like E-cadherin and claudins were not expressed, and mesenchymal markers and EMT inducing transcription factors remained elevated (Figure 9D and data not shown). These data suggested that unlike TS cells, GALNT3 re-expression was insufficient to restore the epithelial state in SUM159 cells.

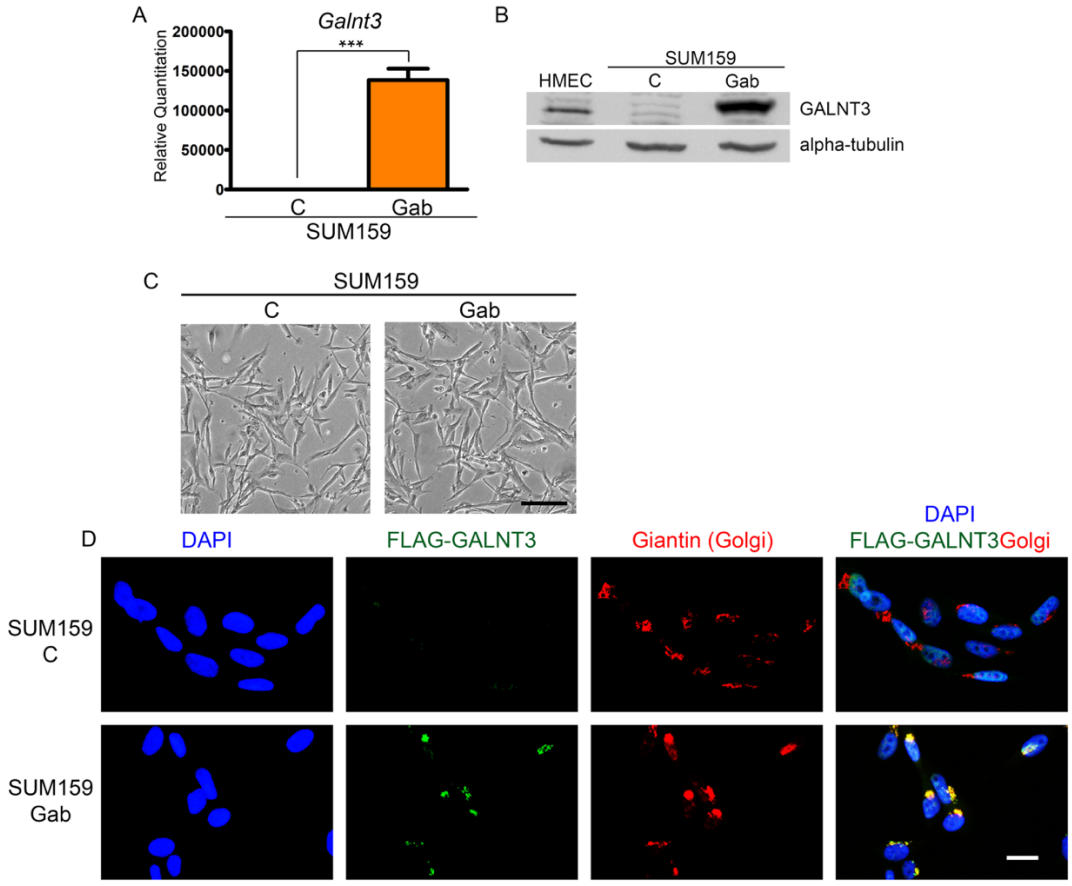

Figure 9. GALNT3 re-expressed in CL SUM159 breast cancer cells localizes to the Golgi. (A) *Galnt3* transcripts analyzed by qPCR in CL SUM159 breast cancer cells infected with a control virus (C) or a lentiviral construct expressing FLAG-tagged human *Galnt3* ($SUM159^{Gab}$). qPCR data normalized to *Tbp* are expressed as a fold-change relative to control infected CL SUM159 breast cancer cells and are the mean ± SEM of two independent experiments. ***p-value < 0.001, Student's t test. (B) Western blots for GALNT3 and α-tubulin are shown and are representative of three independent experiments. (C) Re-expression of GALNT3 in SUM159 cells does not alter morphology. Phase microscopy images are representative of two independent experiments. Black bar represents 100 µm. (D) FLAG-GALNT3 localizes to the Golgi in $SUM159^{Gab}$ cells. Images are representative of two independent experiments. DAPI (blue), FLAG-GALNT3 (green), Giantin (Golgi) (red). White bar represents 50 µm.

Discussion

Herein, we define a critical role of GALNT3 in epithelial TS cells derived from the preimplantation blastocyst. Loss of GALNT3 in TS cells induces EMT. Importantly, re-expression of GALNT3 in mesenchymal-like TS^{KI4} cells restores the epithelial state. In addition, we identify a conserved role for GALNT3 in HMECs, where loss of GALNT3 is associated with the gain of the mesenchymal state. Further, loss of *Galnt3* in the blastocyst trophectoderm results in premature attachment and outgrowth. Our work defines a key role for GALNT3 in promoting the epithelial state in TS cells and HMECs. This study identifies the role of a specific O-GalNAc glycosyltransferase in TS cells of the early blastocyst.

In many cancers, heterogeneity in protein structures due to aberrant glycosylation results in loss of cell-cell adhesion and gain of invasiveness. Changes in protein glycosylation occur in two ways: incomplete synthesis and neo-synthesis (Pinho and Reis, 2015). Loss of expression of initiating glycosyltransferases results in truncated glycans and disrupted protein structures. In contrast, synthesis of glycosyltransferases not normally expressed can lead to new carbohydrate structures on proteins that may trigger a tumorigenic response (Pinho and Reis, 2015). Thus, transcriptional changes in the expression of glycosyltransferases are a key factor associated with different types of cancers. We show GALNT3 expression is nearly absent in CL SUM159 cells, and knockdown of GALNT3 in HMECs results in phenotypic changes related to EMT. Changes in GALNT3 expression have been implicated during cancer EMT. In poorly differentiated pancreatic cancers, GALNT3 expression is reduced (Maupin et al., 2010).

These same mesenchymal-like pancreatic cancers have loss of E-cadherin. Further, knockdown of GALNT3 in differentiated pancreatic cancers induces EMT (Chugh et al., 2016). In contrast, GALNT3 expression is increased in ovarian cancers including A2780s cells (Wang et al., 2014). shRNA knockdown of GALNT3 in A2780s cells reduced proliferation and invasiveness and increased E-cadherin. The authors suggested that these effects were related to GALNT3-dependent MUC1 glycosylation (Wang et al., 2014). Importantly, CL SUM159 cells that lack GALNT3 have a >90% decrease in MUC1 expression relative to HMECs, suggesting MUC1 glycosylation does not promote EMT in CL SUM159 cells (Raghu, unpublished data). Differences in the impact of altered GALNT3 expression may be due to the functions of GALNT3 in normal versus cancerous tissues, and different stages and grades of cancer. Further, the absence or presence of GALNT3 targets like MUC1 or differential activation of pathways triggering GALNT3 mislocalization might lead to different responses in specific cancers (Chia et al., 2016). Compensatory changes in the expression of other GALNTs may also result in different outcomes. Wang et al. showed that *Galnt14* expression was strongly reduced with GALNT3 knockdown in A2780s. Our data emphasize a key role for GALNT3 in the maintenance of the epithelial state in TS cells, but its role in cancer cells may be dependent on cell type-specific expression of GALNT3 and its targets.

A key feature of EMT is the loss of tight junction proteins. shRNA knockdown of GALNT3 completely disrupts tight junction assembly. Importantly, re-expression of GALNT3 in the TS^{KI4} cells restores tight junction protein expression and barrier function. We have previously shown that HDAC6 directly deacetylates H2BK5 on the promoter of a key tight junction protein, *Cldn6*, repressing its expression (Mobley et al., 2017).

Similar to *Cldn6*, we show that HDAC6 represses *Galnt3* expression by directly deacetylating H2BK5 on the *Galnt3* promoter. These data suggest that HDAC6 controls the expression of several key genes that promote the epithelial state and repression of these genes results in EMT. Importantly, in TS^{KI4} cells that have elevated HDAC6 levels, re-expression of GALNT3 restores the expression and localization of Claudin-6. These data suggest that GALNT3 may regulate other key targets that control the expression and localization of Claudin-6. Together, these data demonstrate that the MAP3K4/HDAC6/H2BK5Ac co-regulated gene, *Galnt3*, plays a key role in regulating tight junction assembly.

In summary, we define a post-translational mechanism of EMT regulation through GALNT3 in TS cells and HMECs. Loss of GALNT3 in the TS cells, blastocyst trophectoderm, and HMECs induces EMT. We predict that GALNT3 promotes the O-GalNAc glycosylation of key targets that are important in regulating EMT in TS cells, and future work will focus on identifying these targets. In addition, we observe that loss of GALNT3 results in EMT in the preimplantation blastocysts, suggesting a possible role for GALNT3 during early development. Together, our work provides insights into GALNT3 mediated EMT events occurring during developmental EMT that may be reactivated during cancer progression.

Supplemental Figures

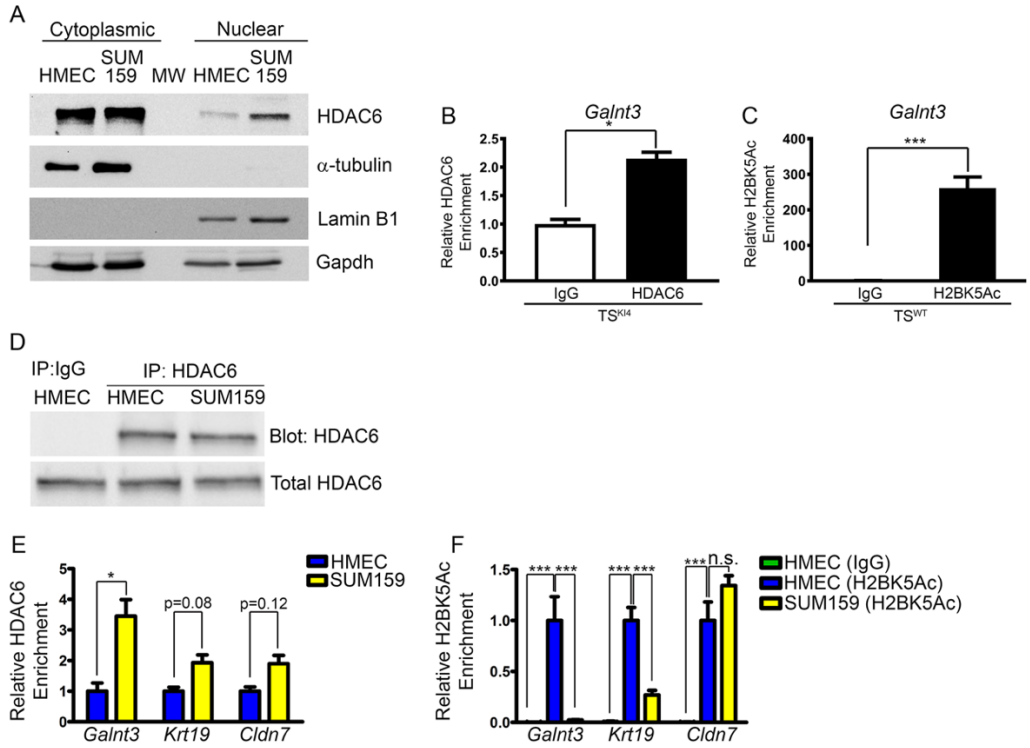

Figure S1. HDAC6 epigenetically inhibits the promoters of epithelial genes in SUM159 Claudin-low breast cancer cells. Related to Figure 5. (A) Increased nuclear HDAC6 in SUM159 Claudin-low (CL) breast cancer cells relative to HMECs. Western blots are representative of three independent experiments. (B) Anti-HDAC6 ChIP-PCR shows the efficiency of anti-HDAC6 antibody versus Immunoglobulin G (IgG) control in TS^{KI4} cells. Data shown are the mean ± SEM of two independent experiments. (C) Anti-H2BK5Ac ChIP-PCR shows the efficiency of anti-H2BK5Ac antibody versus IgG control in TS^{WT} cells. Data shown are the mean ± SEM of two independent experiments. (D) Immunoprecipitation (IP) efficiency of the anti-HDAC6 antibody relative to IgG control in HMECs as shown by Western blotting. Blots are representative of two independent experiments. (E) Anti-HDAC6 ChIP-PCR shows increased HDAC6 enrichment on the promoters of specific genes in SUM159 CL breast cancer cells relative to HMECs. Data shown are the mean ± SEM of three independent experiments. (F) Decreased H2BK5 acetylation on the promoters of specific genes in SUM159 CL breast cancer cells relative to HMECs as measured by anti-H2BK5Ac ChIP-PCR. HMEC chromatin was immunoprecipitated with either IgG or anti-H2BK5Ac antibody. Data shown are the mean ± SEM of three independent experiments. *p-value < 0.05, ***p-value < 0.001, not significant (n.s.), Student's t test.

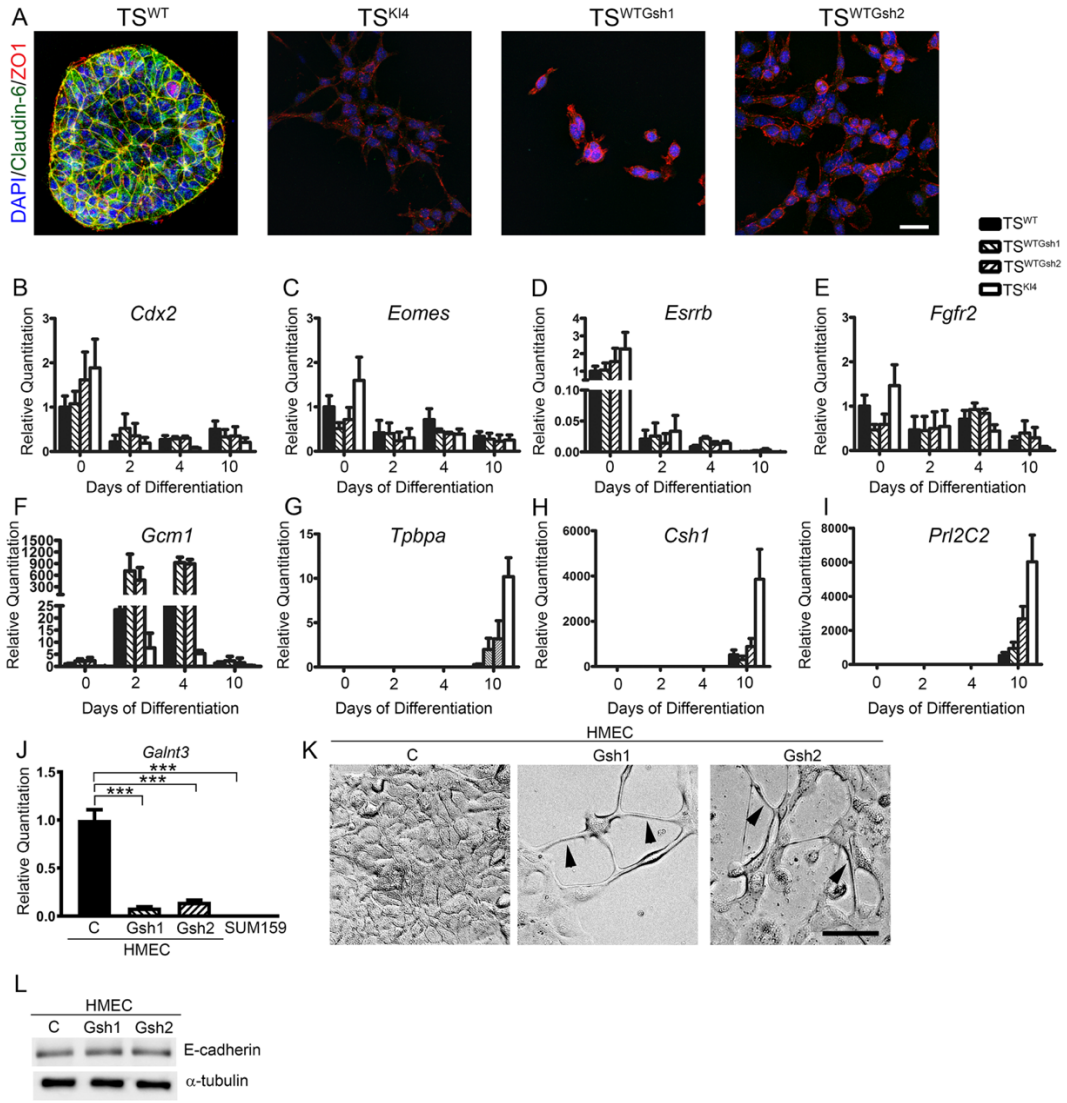

Figure S2. TS cells lacking GALNT3 maintain the expression of stemness markers and the ability to differentiate into all trophoblast subtypes. Related to Figure 6**.** (A) Reduced Claudin-6 and ZO1 protein expression and localization in cells lacking GALNT3. Confocal images in TS^{WT} and TS^{KI4} cells expressing control shRNA or TS^{WT} cells expressing two independent *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2}) are shown. Images are representative of three independent experiments. DAPI (blue), Claudin-6 (green), ZO1 (red). White bar represents 50 μm. (B-E) Undifferentiated TS^{WTGsh1} and TS^{WTGsh2} cells maintain stemness marker expression as measured by qPCR. Differentiation was induced by FGF4 withdrawal for the indicated number of days. (F-I) Differentiation of TS^{WTGsh1} and TS^{WTGsh2} cells induces the expression of trophoblast differentiation markers suggesting the maintenance of multipotency. Experiments were performed as in (B-E). (J) shRNA knockdown of *Galnt3* in HMECs. Transcript levels were measured using qPCR in HMECs expressing a control shRNA (C) or two independent *Galnt3* shRNAs

Figure S2 (Continued)

(Gsh1 and Gsh2) and SUM159 cells. qPCR data normalized to *Gapdh* are expressed as a fold-change relative to HMECs and are the mean ± SEM of three independent experiments. *** p-value < 0.001, Student's t test. (K) shRNA knockdown of GALNT3 in HMECs induces a mesenchymal morphology. Phase microscopy images of HMEC, $HMEC^{Gsh1}$ and $HMEC^{Gsh2}$ cells are shown. Images are representative of three independent experiments. Arrowheads show cytoplasmic, mesenchymal-like extensions that are induced with *Galnt3* knockdown in epithelial HMECs. Black bar represents 100 μm. (L) E-cadherin protein levels are unchanged with loss of GALNT3 in HMECs. Western blots are representative of three independent experiments. (B-I) qPCR data normalized to *Actb* are expressed as a fold-change relative to undifferentiated TS^{WT} cells (0 day) and are the mean ± range of two independent experiments.

Chapter III

GALNT3 Mediated O-GalNAc Glycosylation of E-cadherin Promotes the Epithelial State in Trophoblast Stem Cells and Human Mammary Epithelial Cells

Introduction

Epithelial to mesenchymal transition (EMT) is a reversible morphogenetic program, where the non-motile epithelial cells convert to motile mesenchymal cells. EMT is a critical developmental process important for tissue regeneration and wound healing (Thiery et al., 2009; Yang and Weinberg, 2008). Importantly, EMT is reactivated during cancer metastasis (Chapman, 2011; De Craene and Berx, 2013). We have previously shown that a MAP3K4/HDAC6/H2BK5Ac dependent gene, GALNT3, regulates EMT and controls the cellular phenotype in both TS cells and mammary cells (Chapter II). Lentiviral shRNA knockdown of *Galnt3* in the TS^{WT} cells and HMECs induces EMT (Chapter II). For the first time, we have shown the impact of GALNT3 on EMT in the blastocyst TE (Chapter II). Loss of *Galnt3* in the preimplantation blastocyst TE results in premature EMT. Together, these data suggest that GALNT3 is a key regulator of EMT in TS cells and HMECs.

The preimplantation blastocyst TE is the origin of TS cells. The TS cells isolated from wild-type mice display an epithelial morphology with tight cell-cell adhesion (Abell et al., 2011). In contrast, TS cells isolated from Kinase-Inactive MAP3K4 (KI4) mice display a mesenchymal morphology with reduced cell-cell adhesion (Abell et al., 2011). Further, KI4 embryos display severe developmental defects that are caused by perturbations in the process of EMT (Abell et al., 2009; Abell and Johnson, 2005). TS^{KI4} cells that have undergone EMT retain features of stemness including self-renewal and

multipotency (Abell et al., 2011). We have previously shown that the mesenchymal TS^{KI4} cells have a 70% decrease in the expression of GALNT3 relative to TS^{WT} cells (Chapter II). Importantly, *Galnt3* expression is significantly reduced in all the Claudin-low (CL) breast cancer cells relative to HMECs (Chapter II). These data indicate that GALNT3 is critical for epithelial state maintenance.

The *Galnt3* gene encodes for UDP-GalNAc Transferase 3. GALNT3 initiates the attachment of N-acetylgalactosamine (GalNAc) to hydroxyl groups of serine or threonine residues on specific proteins in the Golgi (Bennett et al., 2012). The function of GALNT3 is to promote folding and localization of membrane and secreted proteins. Mutations in GALNT3 result in the misfolding of a phosphate homeostasis hormone, FGF23, the only known target of GALNT3. Aberrant folding of FGF23 causes a severe autosomal disorder called Hyperphosphatemic Familial Tumoral Calcinosis (Topaz et al., 2004). We have shown that GALNT3 is a key regulator of EMT during development in the TS cells (Chapter II). However, the protein targets glycosylated by GALNT3 are unknown.

In Chapter II, we showed that GALNT3 is important in the maintenance of the epithelial state. Loss of GALNT3 expression in TS cells, blastocyst TE, and HMECs resulted in the acquisition of the mesenchymal state. Importantly, GALNT3 re-expression in the mesenchymal TS^{KI4} cells restored the epithelial phenotype. Herein, we identify proteins with GALNT3 dependent O-GalNAc glycosylation and defined a mechanism by which GALNT3 controls the cellular phenotype. Our findings show that changes in GALNT3 expression alters the total O-GalNAc levels in both TS cells and HMECs. Importantly, in TS^{WT} cells, GALNT3 promotes initiation of O-GalNAc on multiple proteins including E-cadherin resulting in the cell surface localization of E-cadherin,

linear adherens junction (AJ) assembly, and epithelial state maintenance. Loss of GALNT3 expression in TS^{KI4} cells reduces O-GalNAc glycosylation of E-cadherin and increases the intracellular retention of E-cadherin in the Golgi. GALNT3 re-expression in TS^{KI4} cells restores E-cadherin localization to the plasma membrane and leads to an epithelial state. Together, our findings demonstrate that GALNT3 protects the epithelial phenotype in TS cells and HMECs, by promoting O-GalNAc glycosylation of proteins including E-cadherin.

Materials and Methods

Animals

See Chapter II Materials and Methods.

Cell lines and culture conditions

See Chapter II Materials and Methods.

Plasmids

See Chapter II Materials and Methods. To create an E-cadherin lentiviral construct, we used pENTR-Cdh1 plasmid # 49776 Gateway entry vector. The insert was cloned into a lentiviral FLAG tagged destination vector (Jordan et al., 2013) using LR clonase II enzyme mix (Invitrogen). pENTR-Cdh1 was a gift from Jamie Davies (Addgene plasmid#49776; http://n2t.net/addgene:49776; RRID:Addgene_49776).

Site directed mutagenesis

Q5 Site-Directed Mutagenesis Kit (NEB) was used to create single substitutions in pENTR-Cdh1 at threonine at positions 40, 63 and 68 to alanine. Briefly, PCR amplification was performed with custom designed mutagenesis primers for substitutions created using NEBaseChanger software (NEB). The PCR amplified product was treated

with Kinase-Ligase-DpnI (KLD) enzyme mix for 5 min to degrade the parent templates. The KLD product was transformed into high-efficiency NEB 5-alpha competent E.coli. The transformed bacteria were grown for 16 hours and plasmids were purified using GeneJet miniprep kit (Thermo Fisher Scientific). The plasmids were validated for substitution mutations through Sanger sequencing at the Molecular Resource Center at the University of Tennessee Health Science Center. The custom designed primers are listed in the Appendix E.

Lentiviral production and infection

See Chapter II Materials and Methods.

Real-time quantitative PCR

See Chapter II Materials and Methods.

Immunoprecipitation

For E-cadherin immunoprecipitations, cells lysed in buffer A plus protease and phosphatase inhibitors as described in Chapter II were incubated for 1 hour with anti-E-cadherin antibody (Abcam). Then, lysates were incubated for 2 hours with Protein A sepharose beads (Invitrogen) and then centrifuged at 1000 rpm for 1 min at RT. Immunoprecipitates were washed four times with buffer A + PMSF and then boiled in 2X Laemli buffer for 7 min. Immunoprecipitates were separated by SDS-PAGE and proteins were transferred onto nitrocellulose overnight. Next day, membranes were washed in distilled water for 15 min and blocked with 1X carbo-free block (Vector Laboratories) for 30 min at room temperature. Biotinylated VVL (5 μg/ml) was used to probe membranes for 30 min at room temperature. Phosphate buffered saline (1X) with 0.5% Tween was used to wash membranes for 30 min. Membranes were probed with

streptavidin HRP (1:500) (Invitrogen) for 1 hour at room temperature in the dark. Clarity ECL (Bio-Rad) treated membranes were imaged and quantified using a Bio-Rad ChemiDoc and Image lab software respectively.

Vicia villosa lectin assays

For VVL pull down assays, cells lysed in buffer A plus protease and phosphatase inhibitors were incubated for 2 hours with VVL agarose beads (Vector Laboratories) and then centrifuged at 1000 rpm for 1 min at RT. Immunoprecipitates were washed four times with buffer A + PMSF and then boiled in 2X Laemmli buffer for 7 min. Immunoprecipitates were separated by 7.5% SDS-PAGE and proteins were transferred onto nitrocellulose. Samples were then probed with the indicated antibodies detailed in Appendix A. For E-cadherin immunoprecipitations, cells lysed in buffer A plus protease and phosphatase inhibitors were incubated for 1 hour with anti-E-cadherin antibody (Abcam). Then, lysates were incubated for 2 hours with Protein A sepharose beads (Thermo Fisher Scientific). Immunoprecipitates were washed four times with buffer A + PMSF and then boiled in 2X Laemmli buffer for 7 min. Immunoprecipitates were separated by 7.5% SDS-PAGE and proteins were transferred onto nitrocellulose overnight. Next day, membranes were washed in distilled water for 15 min and blocked with 1X carbo-free block (Vector Laboratories) for 30 min at room temperature. Biotinylated VVL (5 μg/ml) was used to probe membranes for 30 min at room temperature. Phosphate buffered saline (1X) with 0.5% Tween was used to wash membranes for 30 min. Membranes were probed with streptavidin HRP (1:500) (Thermo Fisher Scientific) for 1 hour at room temperature in the dark. Clarity ECL (Bio-Rad)

treated membranes were imaged and quantified using a Bio-Rad ChemiDoc and Image lab software, respectively.

VVL immunostaining

For VVL immunofluorescence staining, TS cells were cultured on glass coverslips for two days. For both intact and detergent permeabilized VVL staining, cells were fixed with 3% paraformaldehyde in 1X PBS for 10 min and then washed three times with 1X PBS. Coverslips for permeabilized VVL staining were treated with 0.1 % Triton for 3 mins and then washed three times with 1X PBS. Intact and permeabilized VVL coverslips were blocked with 1X carbo-free block (Vector laboratories) for 30 min at RT. Then, coverslips were incubated with 5 μg/ml of biotinylated VVL (Vector Laboratories) in 1X PBS for 30 min at RT. Next, cells were washed five times with 1X PBS-Tween (0.1%) and then incubated with DAPI (0.1 μg/ml) and 1:500 of Streptavidin-Alexa 488 (Thermo Fisher Scientific) or Streptavidin-Alexa 594 for one hour at RT. Coverslips were washed and mounted on slides with mounting media and imaged. For co-localization of Golgi and VVL, cells were first incubated with anti-Giantin antibody (Abcam) overnight at 4C and the next day VVL staining was performed.

Immunofluorescence imaging

Immunofluorescence imaging was performed using Nikon A1 laser scanning confocal microscopy and a 60X Plan Apochromat 1.4 numerical-aperture (NA) oil objective with lasers at 408, 488 and 594 nm. Individual Z-stacks and the 2D images were obtained by either Maximum Intensity Projection (MIP) or Extended Depth of Field (EDF) for all wavelengths. Confocal images for E-cadherin/Giantin were obtained using MIP. Confocal images for Claudin6/ZO1, E-cadherin, E-cadherin/δ-catenin, and

Actin/Vinculin were obtained using EDF. Confocal 3D images shown were obtained using NIS elements analysis software. Images displayed are a reconstruction of all the z-stacks with a step size of 0.413 μm. The pixel resolution of each section is 512 X 512. Imaging of *in vitro* biotinylated VVL staining experiments were performed using EVOS epifluorescence microscope at 20X and 40X magnification for 408, 488 and 594 nm wavelengths.

Blastocyst immunostaining

Wild-type 129/SvEv mice were euthanized 3.5 days post-mating. The uterus was isolated and washed in KSOM-Hepes buffer (Millipore Sigma). The utero-tubule junctions were cut and the uterine horns were flushed using a 25 gauge needle filled with 0.4 ml of M2 medium (Millipore Sigma). Using a mouth-controlled Pasteur pipette, blastocysts were collected and transferred into a drop of PBS. Blastocysts were hatched by serial passage through drops of Acid Tyrodes (Millipore Sigma). Hatched blastocysts were transferred into drops of PBS, and then fixed in 3% paraformaldehyde and VVL immunofluorescence staining was performed as described under Vicia villosa lectin assays. Blastocyst images were obtained using EVOS epifluorescence and Nikon A1 laser scanning confocal microscopes. Immunostaining of the trophectoderm of intact, infected blastocysts was performed 48 hours post-infection. The shControl or shGalnt3 infected blastocysts were fixed with 3% paraformaldehyde and blocked for 30 min using 1X carbofree block. Blastocysts were stained for 1 hour using VVL-FITC (Vector Laboratories) at room temperature. The VVL-FITC stained blastocysts were imaged at 20X using EVOS microscopy.

VVL staining quantitation

TS cells were plated for two days and cells were lifted using enzyme free PBS based cell dissociation solution (Millipore Sigma). Cells were centrifuged at 1000 rpm for 3 min and cell pellets were washed in 1X PBS. For both intact and detergent permeabilized cells, VVL staining was performed as described under Immunofluorescence Staining and all procedures were performed on ice. The stained cells were resuspended in 400 µl 2% FBS in 1X PBS and 100 µl of the resuspended cells were plated in a 96 well plate. Immunofluorescence staining for cell surface and permeabilized VVL (green) was measured using a LionHeart FX automated live cell imager (BioTek). For quantitation, beacons were set for each cell type using BioTek Gen5 (version 3.02) software. The LionHeart FX captured 3 X 3 20X montages for DAPI (Blue) and VVL (Green) for each cell type. Using BioTek Gen5 (version 3.02) software, montages were stitched together, and cellular analysis was performed. Parameters used for quantitation included total cell count, integrated green intensity (Total VVL intensity) and percentage of cells positive for green (VVL). Data were plotted using GraphPad Prism7.

Statistical methods

See Chapter II Materials and Methods.

For qPCR data, the Bio-Rad CFX Maestro software (version 3.0) was used to perform a Student's t test corrected for multiple comparisons. For all other statistics shown, Prism7 (version 7.0, GraphPad software) was used to perform a Student's t test corrected for multiple comparisons or ANOVA with Bonferroni's Multiple Comparison test was used to analyze VVL quantitation.

NIS-Elements

For blastocyst VVL-FITC quantitation, NIS analysis elements software was used to measure the intensity of VVL-FITC (green) in the trophectoderm. To measure the intensity of VVL, the Region of Interest Tool was used to select the trophectoderm area and the Perform Measurements Tool was used to determine the intensity. The VVL-FITC intensity for the blastocysts was analyzed and plotted using GraphPad Prism7 (version 7.0).

Results

Decreased O-GalNAc glycosylation with loss of GALNT3

GALNT3 is an O-GalNAc glycosyltransferase localized to the Golgi that modifies specific S/T with the sugar residue GalNAc. In most cell types, O-GalNAc is further modified by additional enzymes that extend the sugar, adding galactose, GlcNAc and/or sialic acid (Bennett et al., 2012). Vicia Villosa Lectin (VVL) binds with high affinity to unmodified O-GalNAc residues. Using this lectin, we measured the amount of unextended O-GalNAc labeled proteins in TS^{WT} cells relative to TS^{KI4} cells. Whole cell lysates were immunoprecipitated with agarose bound VVL, and blots were probed with biotinylated VVL. Unmodified O-GalNAc labeled proteins were detected in TS^{WT} cells at higher levels relative to TS^{KI4} cells that have reduced GALNT3 expression (Figure 10A). Further, shRNA knockdown of *Galnt3* in TS^{WT} cells resulted in the near complete loss of O-GalNAc labeled proteins (Figure 10A). Importantly, re-expression of GALNT3 in TS^{KI4} cells and in SUM159 cells increased the levels of O-GalNAc labeled proteins (Figure 10B and Figure S3A). Knockdown of HDAC6 in TS^{KI4} cells also increased O-GalNAc to levels similar to TS^{WT} cells (Figure 10C). VVL binding was also reduced in

HMECs expressing *Galnt3* shRNAs and in CL SUM159 cells relative to HMECs (Figures S3B and S3C). These data suggested that EMT in TS cells and HMECs was associated with loss of O-GalNAc labeled proteins.

We also used biotinylated VVL and immunofluorescence to examine the localization of O-GalNAc labeled proteins in intact or detergent permeabilized cells. Surprisingly, TS^{WT} cells showed strong cell surface expression of VVL labeled proteins in non-permeabilized cells that was weaker in TS^{KI4} cells or undetectable in TS^{WTGsh} cells lacking GALNT3 (Figure 10D and Figures S3D). TS^{KI4Gab} cells showed increased cell surface VVL staining relative to TS^{KI4} cells (Figure 10D). Quantitation revealed that >90% of TS^{WT} cells were positive for cell surface VVL staining, compared to 45% of TS^{KI4} cells, and 0.5% TS^{WTGsh} cells (Figure S3E). Similar to cultured TS cells, the TE layer of intact wild-type blastocysts displayed significant VVL staining, suggesting the presence of unextended O-GalNAc modified proteins (Figure 10E and S3F). Detergent-permeabilization revealed intracellular VVL staining that was tightly localized to puncta positioned near the nucleus in all cell conditions (Figure S3G). However, total VVL staining intensity in permeabilized cells was reduced in TS^{KI4} cells and nearly absent in TS^{WTGsh} cells, consistent with Western blotting data (Figures 10A, 10D and Figures S3D-S3G). Co-localization of VVL and the Golgi marker Giantin was greater in TS^{KI4} cells relative to TS^{WT} cells (Figure 10F). Together, these data suggested that loss of GALNT3 reduced O-GalNAc labeled proteins at the cell surface.

The presence of low levels of O-GalNAc labeled proteins in TS^{KI4} cells and TS^{WTGsh} cells suggested that other GALNTs may be alternatively expressed after loss of GALNT3. TS^{KI4} cells showed a modest increase in several of the GALNTs relative to

TS^{WT} cells that was reduced by re-expression of GALNT3 in TS^{KI4Gab} cells (Figure S3H). Importantly, knockdown of GALNT3 in TS^{WTGsh} cells did not induce transcripts of other GALNTs except *Galnt7* and *Galnt12* (Figure S3I). Together, these data suggested that GALNT3 is the dominant O-GalNAc glycosyltransferase in TS cells.

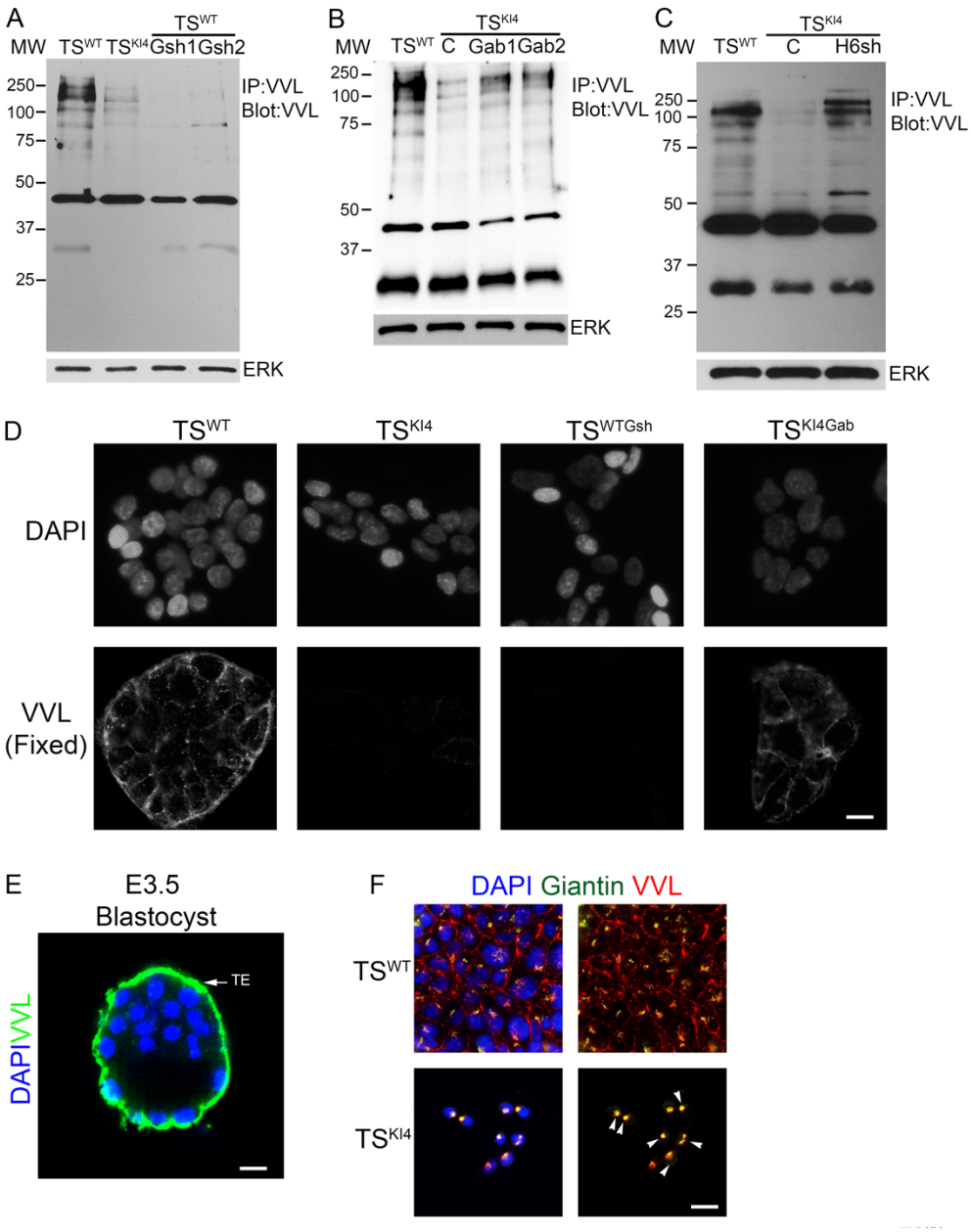

Figure 10. Reduced O-GalNAc glycosylation in cells lacking GALNT3 that is restored by re-expression of GALNT3. (A) Total protein O-GalNAc glycosylation is reduced with loss of GALNT3 in TS^{WTGsh1} and TS^{WTGsh2} cells. IP with VVL-agarose and blotting with biotinylated VVL is shown. Western blots of TS^{WT} and TS^{KI4} cells expressing control shRNA (C) or TS^{WT} cells expressing two independent *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2}) are representative of three independent experiments. (B) Re-expression of GALNT3 in TS^{KI4} cells ($TS^{KI4Gab1}$ and $TS^{KI4Gab2}$) increases protein O-GalNAc glycosylation. Experiments were performed as in (A). Western blots of TS^{WT} and TS^{KI4} cells expressing a control lentiviral vector (C) or or TS^{KI4} cells infected with a lentiviral vector expressing human *Galnt3* ($TS^{KI4Gab1}$ and $TS^{KI4Gab2}$) are representative of three

Figure 10 (Continued)

independent experiments. (C) Knockdown of HDAC6 in TS^{KI4} cells ($TS^{KI4H6sh}$) increases total protein O-GalNAc glycosylation similar to TS^{WT} cells. Experiments were performed as in (A). Western blots of TS^{WT} cells and TS^{KI4} cells expressing control shRNA (C) or TS^{KI4} cells expressing HDAC6 shRNA ($TS^{KI4H6sh}$) are representative of two independent experiments. (D) Absence of cell surface O-GalNAc glycosylation in TS^{KI4} and TS^{WTGsh} cells with reduced GALNT3. VVL immunofluorescence staining in non-permeabilized TS^{WT} and TS^{KI4} cells infected with a control virus, TS^{WT} cells infected with *Galnt3* shRNA (TS^{WTGsh}), and TS^{KI4} cells re-expressing GALNT3 (TS^{KI4Gab}) are shown. Images are representative of three independent experiments. (E) O-GalNAc glycosylation of the trophoectoderm (TE) of intact wild-type E3.5 hatched blastocyst is shown by confocal microscopy using biotinylated VVL. Image is representative of seven blastocysts. DAPI (blue), VVL (green). (F) Reduced GALNT3 expression in TS^{KI4} cells induces retention of O-GalNAc glycosylated proteins in the Golgi. DAPI (blue), VVL (red) and Giantin (Golgi) (green) staining of Triton permeabilized cells. Arrowheads show co-localization between VVL and Giantin in the Golgi. Images are representative of three independent experiments. IP, Immunoprecipitation; MW, Molecular weight in kDa; White bar represents 50 μm. See also Figure S3.

Loss of *Galnt3* in the blastocyst TE results in reduced O-GalNAc staining

To further define the role of GALNT3 mediated O-GalNAc glycosylation in TS cells, we examined the impact of loss of GALNT3 in pre-implantation blastocyst TE. The cell surface VVL staining of the TE was reduced by 57% in blastocysts infected with *Galnt3* shRNAs compared to control shRNA infections (Figures 11A and 11B). These data suggest that loss of *Galnt3* at the blastocyst TE reduces O-GalNAc glycosylation staining. Examination of VVL staining of uninfected blastocysts during attachment and outgrowth revealed the reduction of VVL staining with attachment and outgrowth (Figure 11C). Together, these data suggest that the loss of cell surface O-GalNAc labeled proteins on the TE may represent a normal process during blastocyst development.

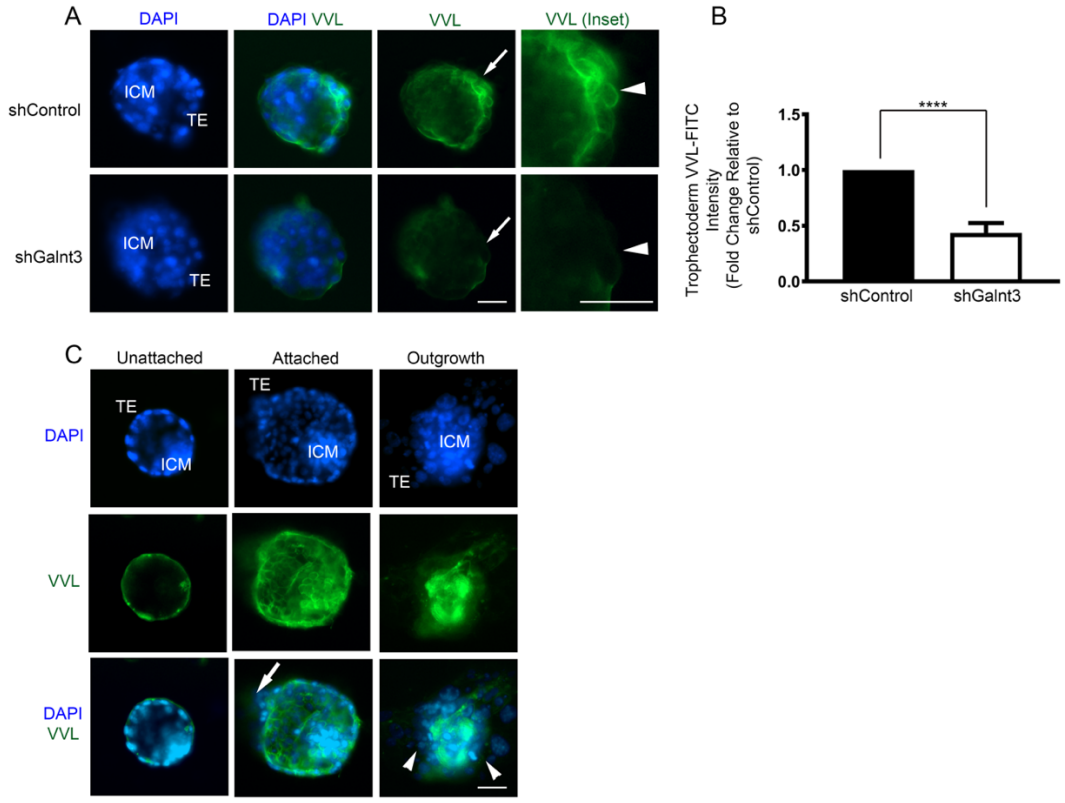

Figure 11. Knockdown of *Galnt3* in the trophectoderm of E3.5 blastocysts results in reduced O-GalNAc glycosylation staining. (A) Decreased O-GalNAc staining intensity on the trophectoderm of shGalnt3 blastocysts relative to control infected blastocysts. Epifluorescence images shown are representative of seven independent experiments. DAPI (blue), VVL-FITC (green). Arrows indicate area of enlarged insets. Arrowheads indicate VVL-FITC staining. White bar represents 50 μm. (B) Quantification of VVL-FITC intensity on the trophectoderm of shControl and shGalnt3 infected blastocysts is shown. Data are expressed as a fold change relative to control infected blastocysts and are the mean ± SEM of seven infected blastocysts. ****p-value < 0.0001, Student's t test. (C) Differentiation and outgrowth of the trophectoderm results in decreased O-GalNAc staining. Arrow indicates the area of trophoblast outgrowth initiation that is associated with decreased O-GalNAc staining. Arrowheads indicate the area for differentiated trophoblast cells correlated with loss of O-GalNAc staining. DAPI (blue), VVL-FITC (green). White bar represents 50 μm.

E-cadherin O-GalNAc glycosylation is dependent on GALNT3

Loss of GALNT3 reduced total levels of O-GalNAc glycosylated proteins in TS cells, suggesting that GALNT3 promotes the O-GalNAc glycosylation of multiple proteins in TS cells (Figures 10A-10C). Examination of VVL blots revealed reduced O-GalNAc glycosylation of a protein of ~140 kDa in TS^{KI4} and TS^{WTGsh} cells that is restored in TS^{KI4Gab} cells and $TS^{KI4H6sh}$ cells (Figures 10A-10C). Although the calculated mass of E-cadherin is 97.5 kDa, E-cadherin migrates around 135 kDa on polyacrylamide gels. In addition, E-cadherin was recently shown by mass spectrometry to be modified with a single O-GalNAc on the first extracellular cadherin domain (EC1) at threonine position 63, but the function of this modification and the O-GalNAc glycosyltransferases were undefined (Vester-Christensen et al., 2013). Further, our previously published work shows that E-cadherin is intracellularly retained in TS^{KI4} cells (Mobley et al., 2017). Based on the changes in O-GalNAc levels at ~140 kDa and altered localization of E-cadherin, we predicted that GALNT3 may be required for O-GalNAc glycosylation of E-cadherin. Immunoprecipitation of lysates with VVL-agarose and probing with anti-E-cadherin antibody revealed O-GalNAc labeled E-cadherin in TS^{WT} cells that was lost in TS^{KI4} cells and in TS^{WTGsh} cells (Figure 12A). Immunoprecipitation with anti-E-cadherin antibody and probing with VVL-biotin also showed O-GalNAc glycosylation of E-cadherin in TS^{WT} cells that was diminished in TS^{KI4} and TS^{WTGsh} cells (Figure 12B). A similar loss of O-GalNAc labeled E-cadherin was observed with knockdown of GALNT3 in HMECs and in SUM159 cells relative to HMECs (Figures 12C and 12D). Knockdown of HDAC6 in TS^{KI4} cells to levels that restore GALNT3 expression and GALNT3 re-expression in TS^{KI4Gab} cells increased O-GalNAc modification of E-cadherin (Figures

12A-B and Figure 12E). These data suggest that E-cadherin is O-GalNAc modified in a GALNT3 dependent manner. Based on the previous work from Vester-Christensen et al. using mass spectrometry, we predicted that threonine at position 63 on E-cadherin is O-GalNAc glycosylated in a GALNT3 dependent manner. We mutated threonine at positions 40, 63, and 68 in the E-cadherin EC1 domain to alanine. These mutated proteins were expressed in CL SUM159 cells stably transduced with GALNT3. Substitutions of threonine 63 to alanine in E-cadherin resulted in a 50% decrease in O-GalNAc glycosylation of E-cadherin relative to wild-type E-cadherin (Figures 12F and 12G). Together, these data suggest that GALNT3 promotes O-GalNAc glycosylation of E-cadherin in both TS cells and HMECs.

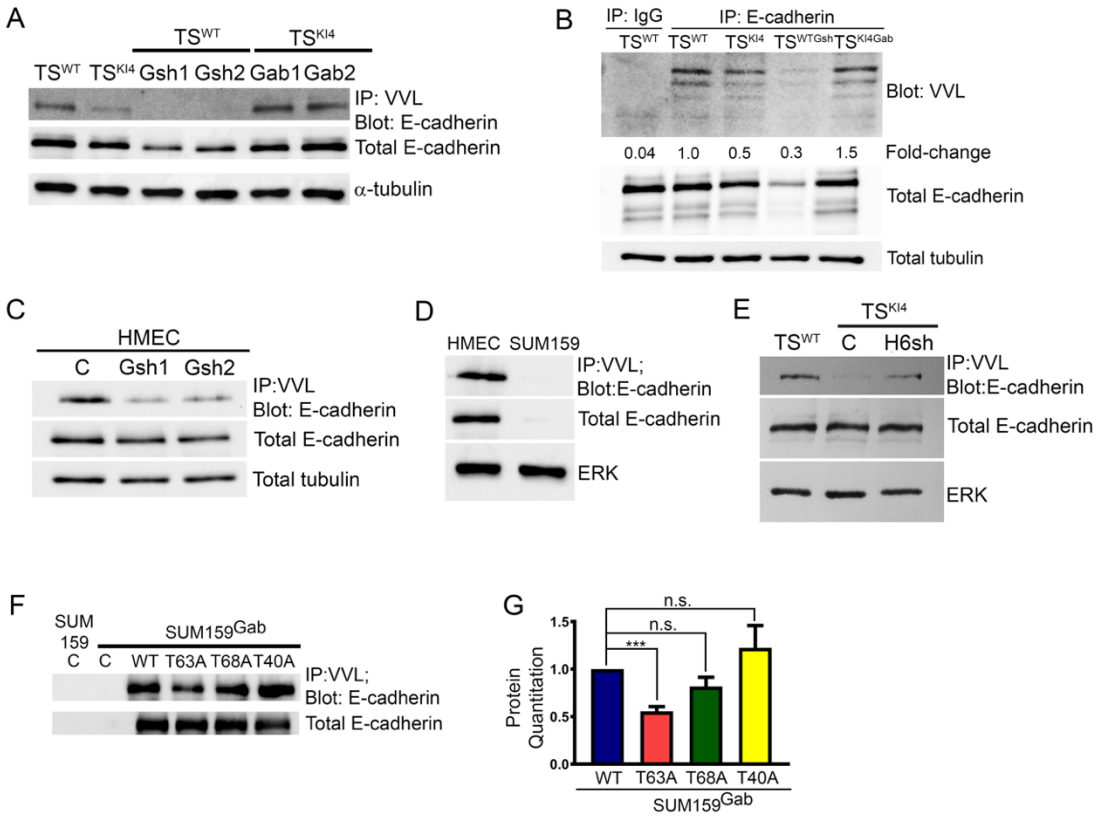

Figure 12. Reduced O-GalNAc glycosylation of E-cadherin in cells lacking GALNT3 that is restored by re-expression of GALNT3. (A) O-GalNAc glycosylation of E-cadherin is reduced in cells lacking GALNT3. IP with VVL-agarose and blotting with anti-E-cadherin antibody is shown for the indicated cells. Western blots are representative of three independent experiments. (B) Cells lacking GALNT3 show reduced O-GalNAc glycosylation of E-cadherin. IP with Immunoglobulin G (IgG) antibody or anti-E-cadherin antibody and blotting with VVL biotin is shown for TS^{WT} and TS^{KI4} cells infected with a control virus, TS^{WT} cells infected with *Galnt3* shRNA (TS^{WTGsh}), and TS^{KI4} cells re-expressing *Galnt3* (TS^{KI4Gab}). Western blots are representative of three independent experiments. (C) O-GalNAc glycosylation of E-cadherin is reduced in HMECs lacking GALNT3. IP with VVL-agarose and blotting with anti-E-cadherin antibody is shown for the indicated cells. Western blots are representative of three independent experiments. (D) CL SUM159 cells lacking GALNT3 have reduced O-GalNAc glycosylation of E-cadherin. IP with VVL agarose beads and blotting with anti-E-cadherin antibody is shown for the indicated cells. Western blots are representative of three independent experiments. (E) Knockdown of HDAC6 increases E-cadherin O-GalNAc glycosylation in $TS^{KI4H6sh}$ cells. Experiments were performed as in (A). Western blots are representative of two independent experiments. (F, G) Reduced GALNT3 dependent O-GalNAc glycosylation of E-cadherin with the substitution of threonine at position 63 to alanine. Control (C) SUM159 cells or SUM159 cells expressing FLAG tagged human *Galnt3* ($SUM159^{Gab}$) were infected with wild-type (WT) E-cadherin or

Figure 12 (Continued)

E-cadherin with substitution mutations in threonine at positions 40, 63, or 68 to alanine (T40A, T63A and T68A) are shown. IP with VVL-agarose and blotting with E-cadherin, and total E-cadherin blots are shown. Blots are representative of three independent experiments. (G) Bar graph shows the densitometry analyses of Western Blots from three independent experiments. O-GalNAc glycosylation of E-cadherin levels were normalized to total E-cadherin. Data are expressed as a fold change relative to WT E-cadherin infected SUM159^{Gab} and show the mean ± SEM of three independent experiments. ***p-value < 0.001, not significant (n.s.), Student's t test. IP, Immunoprecipitation

GALNT3 controls the localization of E-cadherin

E-cadherin expression at both transcript and protein levels was similar in TS^{WT} and TS^{KI4} cells (Figures 6E and 6F). However, confocal immunofluorescence microscopy showed a change in E-cadherin localization from the plasma membrane in TS^{WT} cells to the Golgi in TS^{KI4} cells as indicated by a 54% increase in co-localization of E-cadherin with a Golgi marker, Giantin (Figure 13A). Importantly, cell surface E-cadherin localization was restored in TS^{KI4Gab} cells, suggesting that GALNT3 controls E-cadherin localization (Figure 13A). Similarly, knockdown of HDAC6, a deacetylase that normally represses *Galnt3* expression, restores E-cadherin localization to the plasma membrane (Figure 13B). Knockdown of GALNT3 in HMECs resulted in E-cadherin retention in the Golgi (Figure 13C and data not shown). Together these data suggested that loss of GALNT3 in TS^{KI4} cells and HMECs with GALNT3 knockdown results in the intracellular retention of E-cadherin in the Golgi. Re-expression of GALNT3 in TS^{KI4} cells through either HDAC6 knockdown or a GALNT3 lentiviral construct restored the localization of E-cadherin to the plasma membrane.

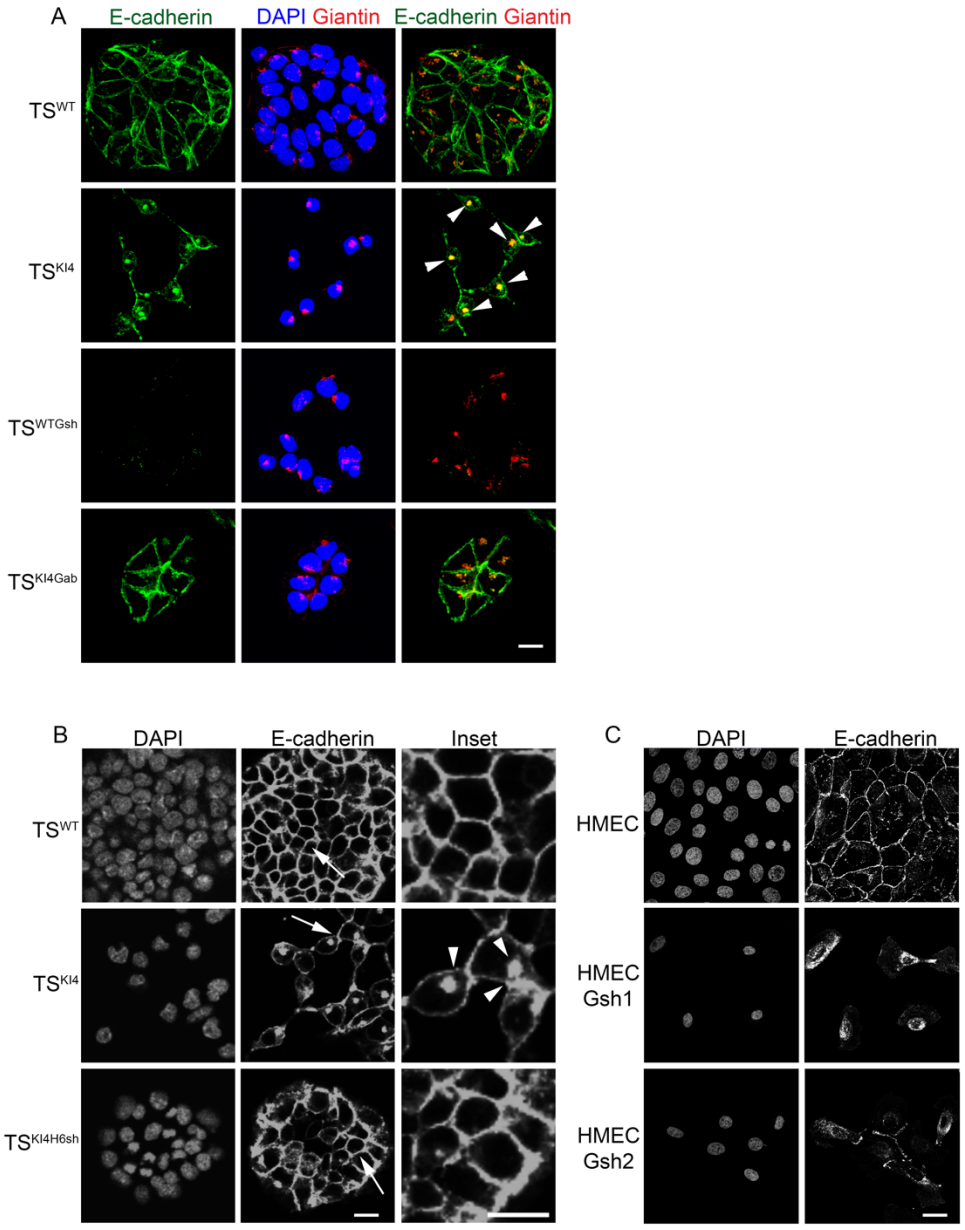

Figure 13. Loss of GALNT3 results in the intracellular retention of E-cadherin in the Golgi. (A) Reduced GALNT3 expression in TS^{KI4} cells results in the retention of E-cadherin in the Golgi. Confocal images in TS^{WT} and TS^{KI4} cells infected with a control virus, TS^{WT} cells infected with *Galnt3* shRNA (TS^{WTGsh}) and TS^{KI4} cells expressing human GALNT3 (TS^{KI4Gab}) are shown. DAPI (blue), E-cadherin (green) and Giantin (Golgi) (red). Arrowheads indicate co-localization of E-cadherin and Giantin in the Golgi. (B) HDAC6 knockdown in TS^{KI4} cells results in the partial restoration of E-cadherin to the cell surface. Confocal images in TS^{WT} cells and TS^{KI4} cells expressing control shRNA or TS^{KI4} cells expressing *Hdac6* shRNA ($TS^{KI4H6sh}$) are shown. Arrows show the area of enlarged insets. Arrowheads indicate punctate intracellular E-cadherin localization. (C) GALNT3 knockdown in HMECs results in the intracellular trapping of E-cadherin. Confocal images for HMECs infected with a control shRNA or two

Figure 13 (Continued)

independent *Galnt3* shRNAs ($HMEC^{Gsh1}$ and $HMEC^{Gsh2}$) are shown. (A-C) Images are representative of three independent experiments. White bar represents 50 μm.

Loss of GALNT3 disrupts the assembly of adherens junctions

Loss of E-cadherin often leads to disruption of AJ assembly (Gumbiner, 1996). TS^{KI4}, TS^{KI4Gab}, and $TS^{KI4H6sh}$ cells expressed α-catenin and δ-catenin protein at levels similar to TS^{WT} cells (Figures 14A and 14B). In addition, transient knockdown of GALNT3 in HMECs resulted in modest changes in α-catenin and δ-catenin protein levels relative to control infected HMECs (Figure 14C). Although total protein expression was unchanged in TS^{KI4} cells, localization of AJ proteins was altered (Figure 14D). Unlike TS^{WT} cells with AJ components E-cadherin, α-catenin, and δ-catenin co-localizing at cell-cell junctions, cell surface expression of all AJ components was reduced in TS^{KI4} cells, resulting in loss of co-localization (Figure 14D and data not shown). Re-expression of GALNT3 in TS^{KI4Gab} cells restored cell surface co-localization of AJ components (Figure 14D). In TS^{WTGsh} and SUM159 cells having a near complete loss of GALNT3, expression of δ-catenin and α-catenin were significantly reduced (Figures 14D-14F). Together, these data suggested that AJ assembly and stability in TS cells and HMECs are GALNT3 dependent.

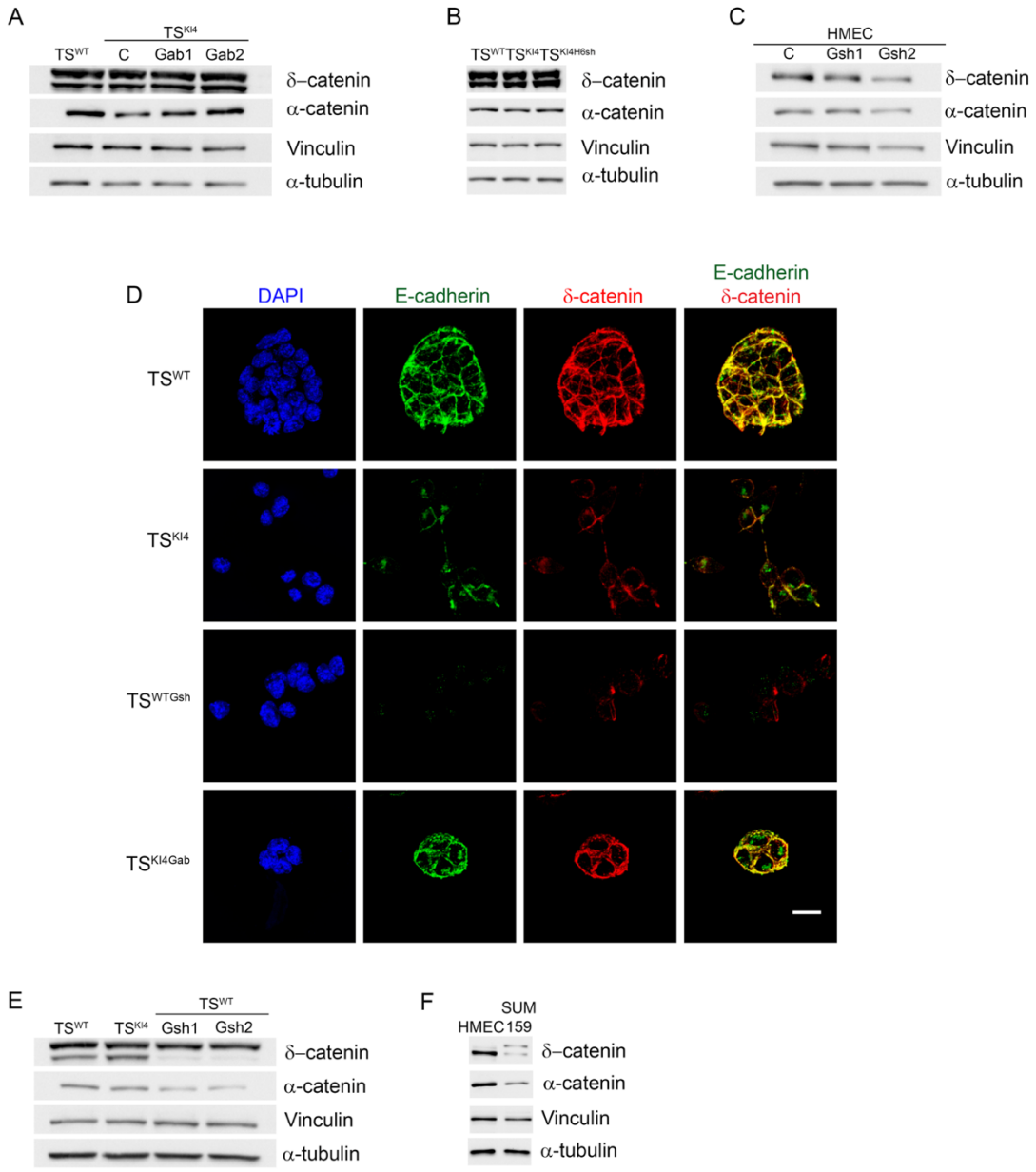

Figure 14. GALNT3 regulates adherens junction assembly. (A) Western blots of TS^{WT} and TS^{KI4} cells infected with a control lentiviral construct (C) or TS^{KI4} cells infected with a lentiviral construct expressing human *Galnt3* ($TS^{KI4Gab1}$ and $TS^{KI4Gab2}$) are shown. (B) Western blots of TS^{WT} and TS^{KI4} cells infected with a control shRNA or TS^{KI4} cells infected with *Hdac6* shRNA ($TS^{KI4H6sh}$) are shown. (C) Western blots of HMECs infected with a control shRNA (C) or *Galnt3* shRNAs ($HMEC^{Gsh1}$ and $HMEC^{Gsh2}$) are shown. (D) Loss of GALNT3 in TS^{KI4} and TS^{WTGsh} cells results in reduced co-localization of δ-catenin and E-cadherin. DAPI (blue), E-cadherin (green) and δ-catenin (red). Confocal microscopy images are representative of three independent experiments. White bar represents 50 μm. (E) Western blots of TS^{WT} and TS^{KI4} cells infected with a control shRNA or TS^{WT} cells with *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2}) are shown. (F)

Figure 14 (Continued)

SUM159 cells lacking GALNT3 have reduced δ-catenin, α-catenin and Vinculin protein levels relative to HMEC cells as shown by Western blotting. (A-C and E-F) Western blots are representative of three independent experiments. See also Figure S4.

Reduced GALNT3 in TS^{KI4} cells results in dynamic adherens junction assembly

AJs can be classified as two main forms, linear or punctate, based on their association with actin (Takeichi, 2014). TS^{WT} cells displayed linear cortical actin filaments that ran parallel to the cell borders forming static, linear AJs (Figures 15 and S4A). In TS^{KI4} cells, actin stress fibers were positioned perpendicular to cell borders, suggesting the presence of dynamic, punctate AJs (Figures 15 and S4A). Re-expression of GALNT3 in TS^{KI4Gab} cells restored cortical actin that ran parallel to the cell borders (Figures 15 and S4A). AJs are often linked to the actin cytoskeleton through the actin-binding protein Vinculin (Takeichi, 2014). Total Vinculin expression did not change significantly under any conditions tested (Figure 14). In TS^{WT} cells, AJs contained small Vinculin-positive puncta with larger Vinculin-positive adhesions positioned continuously around the entire TS^{WT} colony (Figures 15 and S4A). In TS^{KI4} and TS^{WTGsh} cells, Vinculin was present in discontinuous, finger-like projections (Figures 15 and S4A). Re-expression of GALNT3 in TS^{KI4Gab} cells restored continuous Vinculin-positive adhesions around the colony (Figures 15 and S4A). However, these Vinculin positive focal adhesions were more finger-like than those found in TS^{WT} colonies. Western blotting showed that shRNA knockdown or re-expression of GALNT3 did not alter HDAC6 levels (Figures S4B-S4D). The retention of elevated HDAC6 expression in TS^{KI4Gab} cells may explain the retention of finger-like Vinculin projections. Altogether, these data suggested that loss of GALNT3 leads to active remodeling and disassembly of AJs, and GALNT3 re-expression promotes stable AJ assembly. Importantly, re-expression of GALNT3 restores the localization of E-cadherin to the cell surface leading to stable adherens junction assembly and maintenance of the epithelial state.

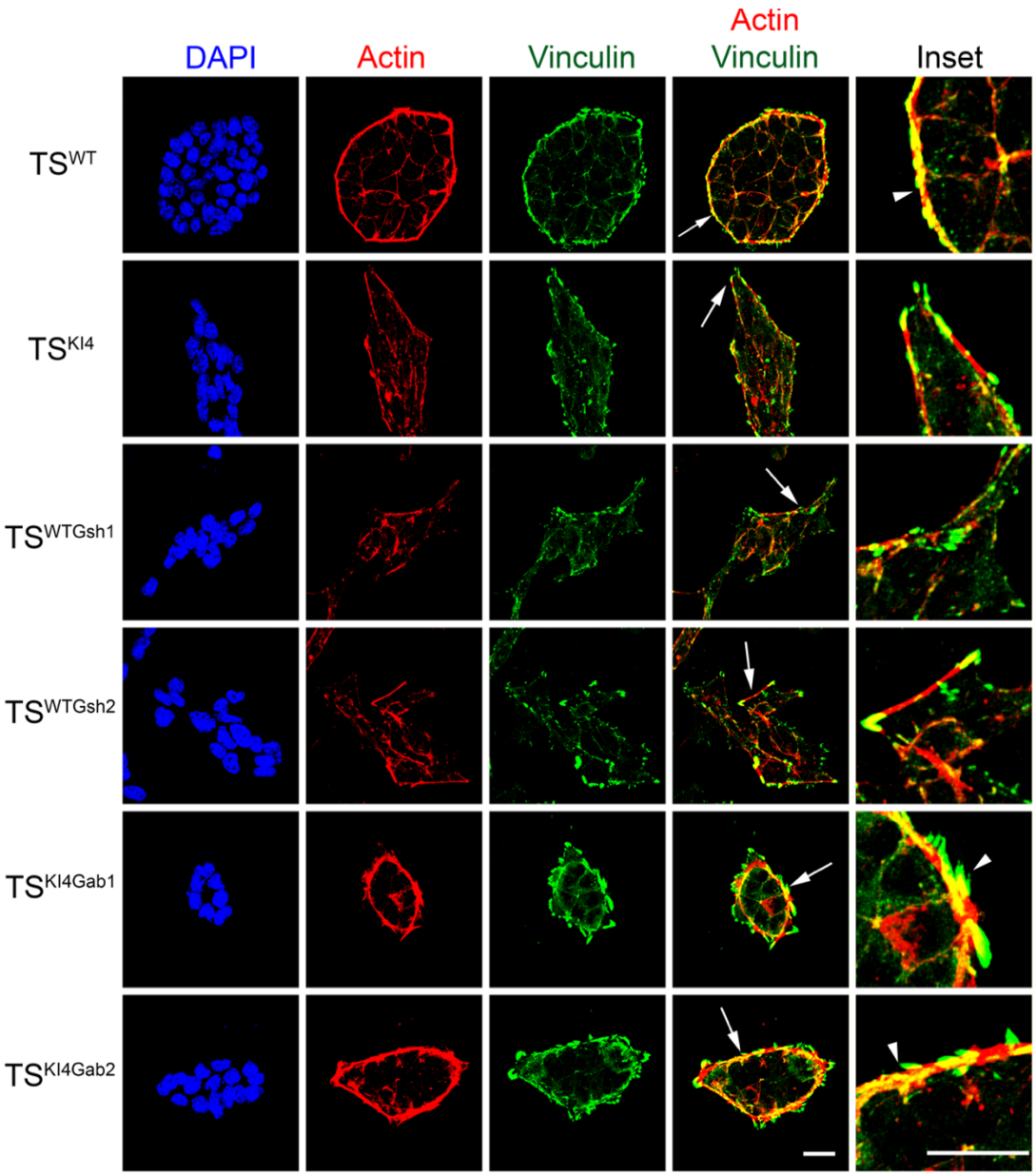

Figure 15. GALNT3 promotes adherens junction assembly and maintenance. Confocal images in TS^{WT} and TS^{KI4} cells infected with a control virus, TS^{WT} cells infected with *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2}), and TS^{KI4} cells expressing human GALNT3 ($TS^{KI4Gab1}$ and $TS^{KI4Gab2}$) are shown. DAPI (blue), Actin (red) and Vinculin (green). Arrows show the area of enlarged insets. Arrowheads indicate actin and vinculin co-localization. Images are representative of three independent experiments. White bar represents 50 μm. See also Figure S4.

Discussion

Herein, we define a critical role of O-GalNAc glycosylation in epithelial TS cells derived from the preimplantation blastocyst. We demonstrate that O-GalNAc glycosylation of E-cadherin in TS cells is dependent on GALNT3. Loss of GALNT3 results in the retention of E-cadherin in the Golgi, disruption of AJs, and induction of EMT. Importantly, re-expression of GALNT3 in mesenchymal-like TS^{KI4} cells restores E-cadherin localization to the cell surface and the epithelial state. Our work defines a key role for GALNT3 dependent O-GalNAc glycosylation in promoting the epithelial state in TS cells and HMECs. This study is the first to identify the role of a specific O-GalNAc glycosyltransferase in TS cells of the early blastocyst.

The role of O-GalNAc glycosylation in stem cells is poorly understood. Unextended O-GalNAc modified proteins have been detected in early preimplantation blastocysts, but the specific GALNT(s) that initiate this glycosylation were unknown (Poirier and Kimber, 1997). Differentiation of mouse embryonic stem cells to embryoid bodies or extraembryonic endoderm cells increases the expression of GALNTs, including *Galnt3*. These findings suggest that O-GalNAc glycosyltransferases may play key roles in early differentiation (Nairn et al., 2012). Here, we demonstrate GALNT3 dependent O-GalNAc glycosylation on the cell surface of the TE of E3.5 blastocysts. Cell surface O-GalNAc modified proteins are dramatically reduced in TS^{KI4} cells and TS^{WTGsh} cells lacking GALNT3 relative to TS^{WT} cells. Total O-GalNAc glycosylation levels in TS cells correlate with the absolute expression levels of GALNT3. For example, TS^{KI4} cells with a 70% decrease in GALNT3 show reduced total O-GalNAc glycosylation relative to TS^{WT}

cells. Further, >98% shRNA knockdown of GALNT3 in TS^{WT} cells results in a complete loss of total O-GalNAc glycosylation. In addition, loss of GALNT3 in blastocysts decreases total O-GalNAc levels. Importantly, re-expression of GALNT3 in TS^{KI4} cells increases total O-GalNAc glycosylation. Although GALNT3 strongly controls total O-GalNAc glycosylation in TS cells, alterations in GALNT3 levels in pancreatic and ovarian cancers result in more modest changes in total O-GalNAc glycosylation (Chugh et al., 2016; Maupin et al., 2010; Wang et al., 2014). Together, these data demonstrate that GALNT3 dependent O-GalNAc glycosylation maintains the epithelial TS cells of the TE during early development.

E-cadherin is a transmembrane protein localized to AJs that mediates calcium dependent adhesion. Loss of E-cadherin is a key step in EMT associated with tumor progression and resistance to conventional drug therapies (Pattabiraman et al., 2016). Previous studies have focused on the transcriptional regulation of E-cadherin. However, the mechanisms regulating the folding and localization of E-cadherin are poorly understood. Using mass spectrometry, Vester-Christensen et al. identified an O-GalNAc site on the EC1 domain of E-cadherin, but the functionality of the site was undefined. Our data show that a MAP3K4 dependent gene, *Galnt3,* regulates EMT by controlling the O-GalNAc glycosylation of multiple proteins including E-cadherin and the localization of E-cadherin. Differences in absolute levels of GALNT3 in TS^{KI4} cells and TS^{WTGsh} cells may result in a partial versus full EMT and differences in E-cadherin levels. While O-GlcNAc and N-glycosylation of E-cadherin results in cytoplasmic retention and disruption of AJs, we show that O-GalNAc modification of E-cadherin correlates with cell surface localization and proper AJ assembly, including α and

δ catenin localization (Jamal et al., 2009; Zhu et al., 2001). Together, these data demonstrate that GALNT3 promotes the epithelial TS cell phenotype in part by mediating O-GalNAc modification of E-cadherin.

In summary, we define a post-translational mechanism of EMT regulation through GALNT3 dependent O-GalNAc glycosylation of proteins including E-cadherin. Importantly, GALNT3 regulates the localization AJ proteins, indicating that O-GalNAc glycosylation promotes TS focal adhesion assembly. We predict that GALNT3 O-GalNAc glycosylates additional targets that are important in regulating EMT in TS cells, and future work will focus on identifying these targets. In addition, we detected Tn antigens at the surface of wild-type TS cells and the TE of E3.5 blastocysts, suggesting a possible role for unextended O-GalNAc glycosylation during early development. Further investigation will determine the identity of O-GalNAc modified proteins and their roles in early development. Together, our work provides insights into O-GalNAc glycosylation events occurring during developmental EMT that may be reactivated during cancer progression.

Supplemental Figures

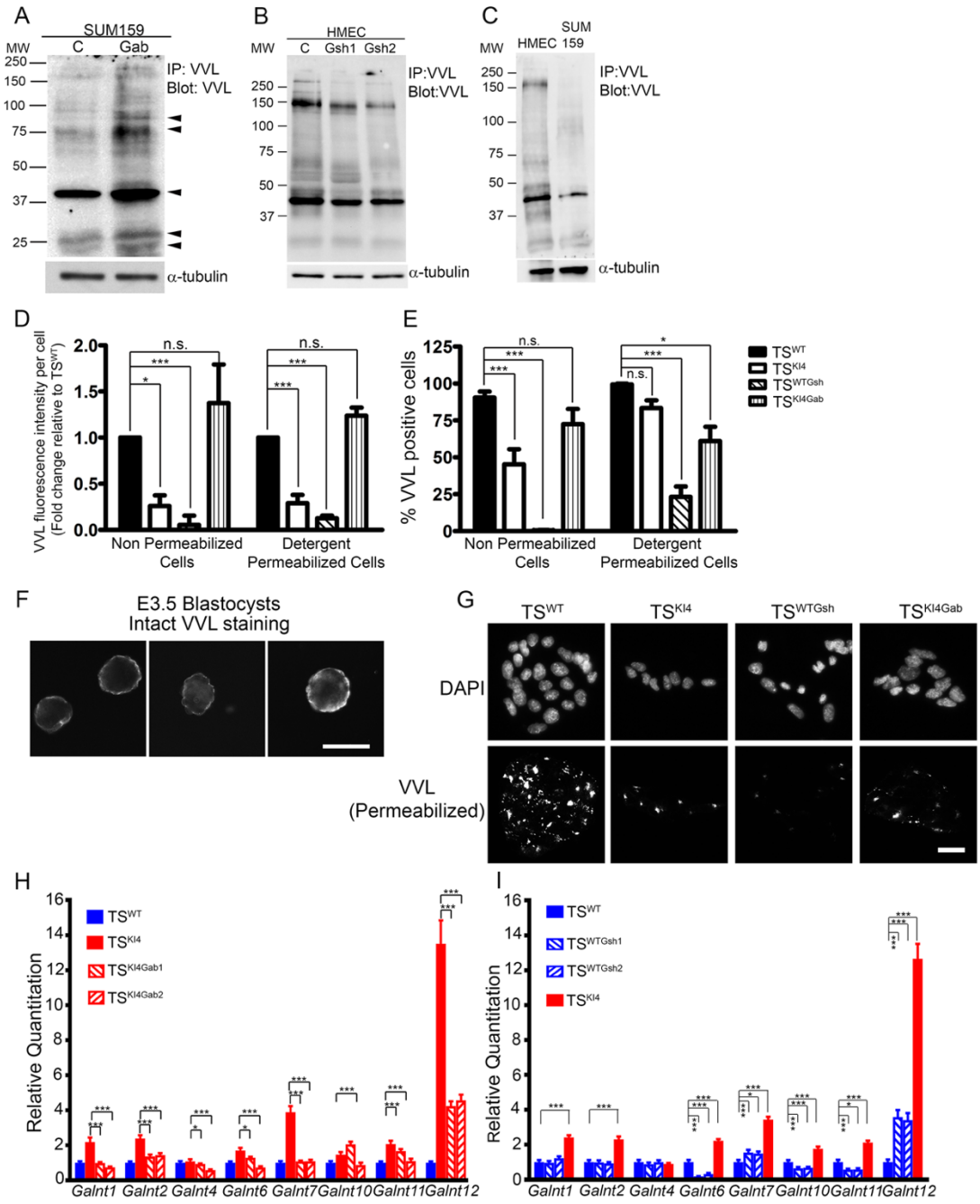

Figure S3. Decreased GalNAc modified proteins in TS cells and mammary epithelial cells lacking GALNT3. Related to Figure 10. (A) Increased protein O-GalNAc glycosylation with re-expression of GALNT3 in SUM159^{Gab} cells. Immunoprecipitation (IP) with VVL-agarose and blotting with biotinylated VVL is shown. Blots are representative of two independent experiments. Arrowheads indicate differentially

Figure S3 (Continued)

increased O-GalNAc modified proteins in SUM159^{Gab} cells relative to control infected SUM159 cells. (B) shRNA knockdown of GALNT3 in HMECs reduces total protein O-GalNAc glycosylation. IP with VVL agarose beads and blotting with biotinylated VVL in HMECs infected with control (C) or HMECs infected with *Galnt3* shRNAs (HMEC^{Gsh1} and HMEC^{Gsh2}) is shown. Blots are representative of three independent experiments. (C) SUM159 cells lacking GALNT3 have reduced total protein O-GalNAc glycosylation relative to HMECs. Experiments were performed as in (A). Blots are representative of three independent experiments. (D) Decreased expression of GALNT3 correlates with reduced VVL fluorescence intensity. Quantitation of VVL intensity is shown for TS^{WT} and TS^{KI4} cells infected with a control virus, TS^{WT} cells infected with *Galnt3* shRNA (TS^{WTGsh}), and TS^{KI4} cells re-expressing *Galnt3* (TS^{KI4Gab}). Data are expressed as a fold change relative to TS^{WT} cells and are the mean ± SEM of three independent experiments. (E) Decreased expression of GALNT3 correlates with reduced numbers of VVL positive cells. Data are expressed as the percent of cells that are VVL positive and are representative of three independent experiments. (D, E) *p-value < 0.05, ***p-value < 0.001, not significant (n.s.), ANOVA test. (F) Immunofluorescence images of biotinylated VVL staining of non-permeabilized E3.5 wild-type hatched blastocysts are shown. Images were obtained using EVOS epifluorescence microscopy. Representative images of seven blastocysts are shown. White bar represents 100 μm. (G) Intracellular retention of O-GalNAc glycosylated proteins in cells lacking GALNT3. Immunofluorescence staining with biotinylated VVL was performed on 0.1% Triton treated cells. Epifluorescence images are representative of three independent experiments. White bar represents 50 μm. (H) TS^{KI4Gab1} and TS^{KI4Gab2} cells re-expressing GALNT3 express transcripts for other *Galnts* at levels similar to TS^{WT} cells. qPCR data for TS^{WT} and TS^{KI4} cells infected with control shRNA or TS^{KI4} cells re-expressing *Galnt3* (TS^{KI4Gab1} and TS^{KI4Gab2}) are shown. (I) Transcript expression of other *Galnts* measured by qPCR in TS^{WT} and TS^{KI4} cells infected with control shRNA or TS^{WT} cells infected with *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2}) are shown. (H, I) qPCR data normalized to *Actb* are expressed as a fold change relative to TS^{WT} cells and are the mean ± range of two independent experiments. *p-value < 0.05; ***p-value < 0.001, Student's t test. MW, Molecular weight in kDa

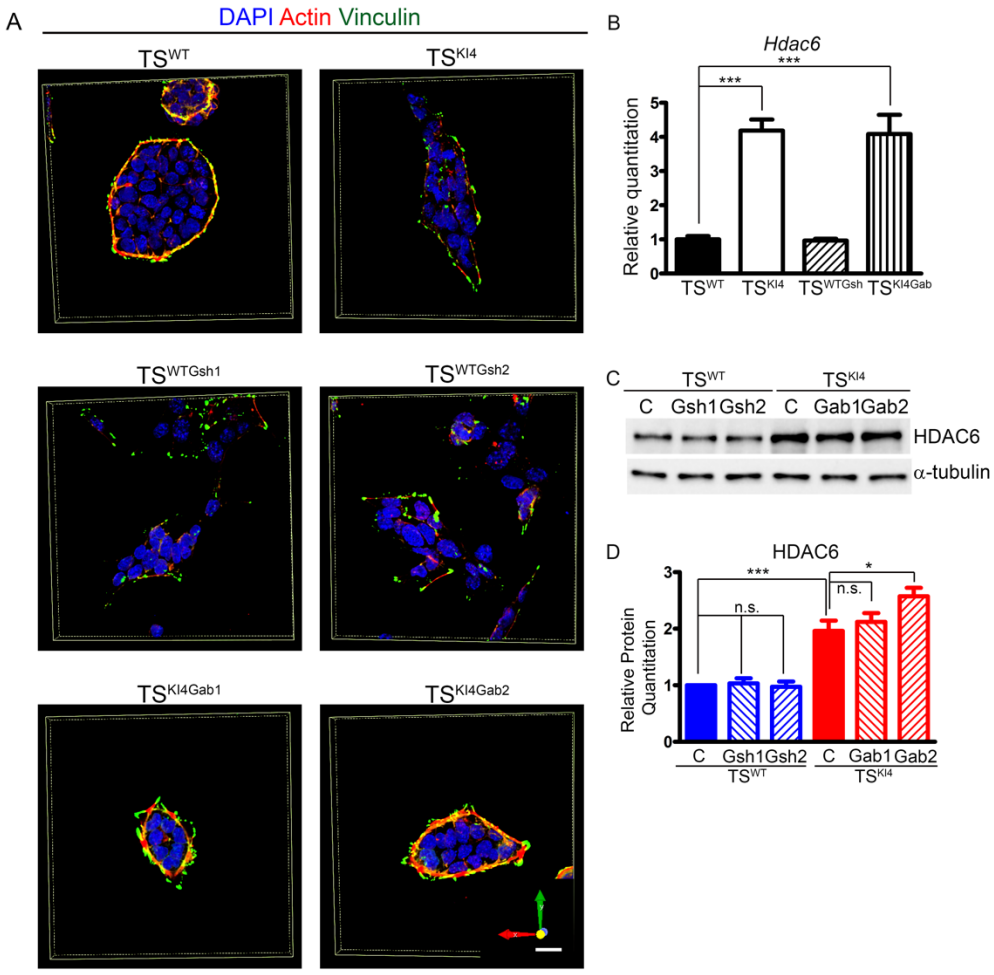

Figure S4. Loss of co-localization of Actin and Vinculin in cells lacking GALNT3. Related to Figure 15. (A) Confocal 3D images are shown for TS^{WT} and TS^{KI4} cells infected with a control virus, TS^{WT} cells infected with *Galnt3* shRNAs (TS^{WTGsh1} and TS^{WTGsh2}), and TS^{KI4} cells expressing human *Galnt3* ($TS^{KI4Gab1}$ and $TS^{KI4Gab2}$). DAPI (blue), Actin (red) and Vinculin (green). Images are representative of three independent experiments. X, Y and Z axes represent the spatial resolution and field of view of the image. Box represents the Voxel size that constitutes the X,Y pixel dimensions (512 X 512) with the section thickness that is represented by the z plane. White bar represents 50 µm. (B) *Hdac6* transcripts analyzed by qPCR in TS^{WT} and TS^{KI4} cells infected with a control virus (C), TS^{WT} cells infected with *Galnt3* shRNA (TS^{WTGsh}) and TS cells expressing human *Galnt3* (TS^{KI4Gab}). qPCR data normalized to *Actb* are expressed as a fold-change relative to TS^{WT} cells and are mean ± SEM of two independent experiments. ***p-value < 0.001, Student's t test. (C-D) Western blots for HDAC6 and α-tubulin are shown. (C) Blots are representative of three independent experiments. (D) Densitometry was used to quantify three independent experiments. Data normalized to α-tubulin loading control are expressed as a fold change relative to TS^{WT} cells and are the mean ± SEM of three independent experiments. *p-value < 0.05, ***p-value < 0.001, not significant (n.s.), ANOVA test.

Chapter IV

Conclusion

We define a mechanism controlling the expression of GALNT3 and the role of this enzyme in promoting the epithelial state. GALNT3 is abundantly expressed in both epithelial TS cells and HMECs. Mesenchymal-like TS^{KI4} cells and CL breast cancer cells have reduced GALNT3 expression, suggesting a correlation between GALNT3 levels and the cellular states. shRNA knockdown of GALNT3 in epithelial TS cells induces a mesenchymal state with a gain in EMT characteristics. Re-expression of GALNT3 in TS^{KI4} cells restores the epithelial morphology, thus resulting in MET. GALNT3 expression is epigenetically inhibited by HDAC6 dependent deacetylation of the *Galnt3* promoter. Inhibition of HDAC6 restores GALNT3 expression and the epithelial state. In addition to controlling the epithelial state in cultured TS cells and HMECs, GALNT3 plays a critical role in regulating the cellular phenotype in the preimplantation blastocyst TE. Loss of GALNT3 in the blastocyst TE results in early attachment and outgrowth, inducing premature EMT. Together, these data suggest that GALNT3 is an important regulator of cellular phenotypic states in the TS cells, blastocyst TE, and mammary cells.

GALNT3 belongs to a family of 19 functionally redundant GALNTs that initiate O-GalNAc glycosylation. All GALNTs are expressed in the epithelial TS cells and HMECs. shRNA knockdown of GALNT3 in the TS cells and HMECs results in a complete loss of O-GalNAc glycosylation without altering the expression of other GALNTs. This lack of compensation by other GALNTs suggests that GALNT3 is the dominant O-GalNAc glycosyltransferase in TS cells and HMECs. In addition to reduced O-GalNAc modification of proteins, loss of GALNT3 dependent O-GalNAc

glycosylation results in protein mislocalization. We observe that GALNT3 initiates O-GalNAc glycosylation of multiple proteins including E-cadherin in the epithelial TS cells and HMECs. Reduced O-GalNAc glycosylation of E-cadherin results in the intracellular retention of E-cadherin, thus inducing EMT. These data suggest that GALNT3 dependent O-GalNAc modification of E-cadherin promotes in cell surface localization that is critical for the maintenance of an epithelial state.

We predict that GALNT3 controls the O-GalNAc modification of multiple proteins. The future goal of this project is to identify additional targets of GALNT3 and to define their roles in EMT. We also observe O-GalNAc modified proteins in the TE of the wild-type blastocyst, suggesting a possible role of these proteins in early development. Future work will focus on identifying these proteins and examine their roles in early developmental EMT. Altogether, our work provides new insights on the role of O-GalNAc glycosylation in developmental EMT that is reactivated during cancer metastasis.

References

Abell, A.N., Granger, D.A., and Johnson, G.L. (2007). MEKK4 stimulation of p38 and JNK activity is negatively regulated by GSK3beta. J Biol Chem *282*, 30476-30484.

Abell, A.N., Granger, D.A., Johnson, N.L., Vincent-Jordan, N., Dibble, C.F., and Johnson, G.L. (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, A.N., and Johnson, G.L. (2005). MEKK4 is an effector of the embryonic TRAF4 for JNK activation. J Biol Chem *280*, 35793-35796.

Abell, A.N., Jordan, N.V., Huang, W., Prat, A., Midland, A.A., Johnson, N.L., Granger, D.A., Mieczkowski, P.A., Perou, C.M., Gomez, S.M.*, et al.* (2011). MAP3K4/CBP-regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. Cell Stem Cell *8*, 525-537.

Acloque, H., Adams, M.S., Fishwick, K., Bronner-Fraser, M., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. J Clin Invest *119*, 1438-1449.

Aoki, R., Tanaka, S., Haruma, K., Yoshihara, M., Sumii, K., Kajiyama, G., Shimamoto, F., and Kohno, N. (1998). MUC-1 expression as a predictor of the curative endoscopic treatment of submucosally invasive colorectal carcinoma. Dis Colon Rectum *41*, 1262-1272.

Auman, H.J., Nottoli, T., Lakiza, O., Winger, Q., Donaldson, S., and Williams, T. (2002). Transcription factor AP-2gamma is essential in the extra-embryonic lineages for early postimplantation development. Development *129*, 2733-2747.

Avichezer, D., Taylor-Papadimitriou, J., and Arnon, R. (1998). A short synthetic peptide (DTRPAP) induces anti-mucin (MUC-1) antibody, which is reactive with human ovarian and breast cancer cells. Cancer Biochem Biophys *16*, 113-128.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

Balin, S.J., Wetter, D.A., Andersen, L.K., and Davis, M.D. (2012). Calcinosis cutis occurring in association with autoimmune connective tissue disease: the Mayo Clinic experience with 78 patients, 1996-2009. Arch Dermatol *148*, 455-462.

Bard, F., and Chia, J. (2016). Cracking the Glycome Encoder: Signaling, Trafficking, and Glycosylation. Trends Cell Biol *26*, 379-388.

Barkeer, S., Chugh, S., Batra, S.K., and Ponnusamy, M.P. (2018). Glycosylation of Cancer Stem Cells: Function in Stemness, Tumorigenesis, and Metastasis. Neoplasia *20*, 813-825.

Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., and Garcia De Herreros, A. (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat Cell Biol *2*, 84-89.

Beaman, E.M., and Brooks, S.A. (2014). The extended ppGalNAc-T family and their functional involvement in the metastatic cascade. Histol Histopathol *29*, 293-304.

Bektas, M., and Rubenstein, D.S. (2011). The role of intracellular protein O-glycosylation in cell adhesion and disease. J Biomed Res *25*, 227-236.

Bennett, E.P., Mandel, U., Clausen, H., Gerken, T.A., Fritz, T.A., and Tabak, L.A. (2012). Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology *22*, 736-756.

Bogani, D., Siggers, P., Brixey, R., Warr, N., Beddow, S., Edwards, J., Williams, D., Wilhelm, D., Koopman, P., Flavell, R.A.*, et al.* (2009). Loss of mitogen-activated protein kinase kinase kinase 4 (MAP3K4) reveals a requirement for MAPK signalling in mouse sex determination. PLoS Biol *7*, e1000196.

Bogoyevitch, M.A., Ngoei, K.R., Zhao, T.T., Yeap, Y.Y., and Ng, D.C. (2010). c-Jun N-terminal kinase (JNK) signaling: recent advances and challenges. Biochim Biophys Acta *1804*, 463-475.

Boskovski, M.T., Yuan, S., Pedersen, N.B., Goth, C.K., Makova, S., Clausen, H., Brueckner, M., and Khokha, M.K. (2013). The heterotaxy gene GALNT11 glycosylates Notch to orchestrate cilia type and laterality. Nature *504*, 456-459.

Boutet, A., Esteban, M.A., Maxwell, P.H., and Nieto, M.A. (2007). Reactivation of Snail genes in renal fibrosis and carcinomas: a process of reversed embryogenesis? Cell Cycle *6*, 638-642.

Brockhausen, I. (1997). Biosynthesis and functions of O-glycans and regulation of mucin antigen expression in cancer. Biochem Soc Trans *25*, 871-874.

Brockhausen, I., Schachter, H., and Stanley, P. (2009). O-GalNAc Glycans. In Essentials of Glycobiology, nd, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, and M.E. Etzler, eds. (Cold Spring Harbor (NY)).

Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., and Relaix, F. (2003). The formation of skeletal muscle: from somite to limb. J Anat *202*, 59-68.

Burdick, M.D., Harris, A., Reid, C.J., Iwamura, T., and Hollingsworth, M.A. (1997). Oligosaccharides expressed on MUC1 produced by pancreatic and colon tumor cell lines. J Biol Chem *272*, 24198-24202.

Byler, S., Goldgar, S., Heerboth, S., Leary, M., Housman, G., Moulton, K., and Sarkar, S. (2014). Genetic and epigenetic aspects of breast cancer progression and therapy. Anticancer Res *34*, 1071-1077.

Byler, S., and Sarkar, S. (2014). Do epigenetic drug treatments hold the key to killing cancer progenitor cells? Epigenomics *6*, 161-165.

Cachat, E., Liu, W., Hohenstein, P., and Davies, J.A. (2014). A library of mammalian effector modules for synthetic morphology. J Biol Eng *8*, 26.

Cano, A., Perez-Moreno, M.A., Rodrigo, I., Locascio, A., Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M.A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nat Cell Biol *2*, 76-83.

Carvalho-Cruz, P., Alisson-Silva, F., Todeschini, A.R., and Dias, W.B. (2018). Cellular glycosylation senses metabolic changes and modulates cell plasticity during epithelial to mesenchymal transition. Dev Dyn *247*, 481-491.

Chapman, H.A. (2011). Epithelial-mesenchymal interactions in pulmonary fibrosis. Annu Rev Physiol *73*, 413-435.

Chen, Z., Gibson, T.B., Robinson, F., Silvestro, L., Pearson, G., Xu, B., Wright, A., Vanderbilt, C., and Cobb, M.H. (2001). MAP kinases. Chem Rev *101*, 2449-2476.

Cheung, M., and Briscoe, J. (2003). Neural crest development is regulated by the transcription factor Sox9. Development *130*, 5681-5693.

Cheung, M., Chaboissier, M.C., Mynett, A., Hirst, E., Schedl, A., and Briscoe, J. (2005). The transcriptional control of trunk neural crest induction, survival, and delamination. Dev Cell *8*, 179-192.

Chi, H., Sarkisian, M.R., Rakic, P., and Flavell, R.A. (2005). Loss of mitogen-activated protein kinase kinase kinase 4 (MEKK4) results in enhanced apoptosis and defective neural tube development. Proc Natl Acad Sci U S A *102*, 3846-3851.

Chia, J., Goh, G., and Bard, F. (2016). Short O-GalNAc glycans: regulation and role in tumor development and clinical perspectives. Biochim Biophys Acta *1860*, 1623-1639.

Chugh, S., Meza, J., Sheinin, Y.M., Ponnusamy, M.P., and Batra, S.K. (2016). Loss of N-acetylgalactosaminyltransferase 3 in poorly differentiated pancreatic cancer: augmented aggressiveness and aberrant ErbB family glycosylation. Br J Cancer *114*, 1376-1386.

Cooper, J.A., Bowen-Pope, D.F., Raines, E., Ross, R., and Hunter, T. (1982). Similar effects of platelet-derived growth factor and epidermal growth factor on the phosphorylation of tyrosine in cellular proteins. Cell *31*, 263-273.

Coulombe, P., and Meloche, S. (2007). Atypical mitogen-activated protein kinases: structure, regulation and functions. Biochim Biophys Acta *1773*, 1376-1387.

Cuadrado, A., and Nebreda, A.R. (2010). Mechanisms and functions of p38 MAPK signalling. Biochem J *429*, 403-417.

Da Fonseca, L.M., da Silva, V.A., Freire-de-Lima, L., Previato, J.O., Mendonca-Previato, L., and Capella, M.A. (2016). Glycosylation in Cancer: Interplay between Multidrug Resistance and Epithelial-to-Mesenchymal Transition? Front Oncol *6*, 158.

De Craene, B., and Berx, G. (2013). Regulatory networks defining EMT during cancer initiation and progression. Nat Rev Cancer *13*, 97-110.

Derijard, B., Raingeaud, J., Barrett, T., Wu, I.H., Han, J., Ulevitch, R.J., and Davis, R.J. (1995). Independent human MAP-kinase signal transduction pathways defined by MEK and MKK isoforms. Science *267*, 682-685.

Dias, K., Dvorkin-Gheva, A., Hallett, R.M., Wu, Y., Hassell, J., Pond, G.R., Levine, M., Whelan, T., and Bane, A.L. (2017). Claudin-Low Breast Cancer; Clinical & Pathological Characteristics. PLoS One *12*, e0168669.

Dong, C., Wu, Y., Wang, Y., Wang, C., Kang, T., Rychahou, P.G., Chi, Y.I., Evers, B.M., and Zhou, B.P. (2013). Interaction with Suv39H1 is critical for Snail-mediated E-cadherin repression in breast cancer. Oncogene *32*, 1351-1362.

Donnison, M., Beaton, A., Davey, H.W., Broadhurst, R., L'Huillier, P., and Pfeffer, P.L. (2005). Loss of the extraembryonic ectoderm in Elf5 mutants leads to defects in embryonic patterning. Development *132*, 2299-2308.

Dosaka-Akita, H., Kinoshita, I., Yamazaki, K., Izumi, H., Itoh, T., Katoh, H., Nishimura, M., Matsuo, K., Yamada, Y., and Kohno, K. (2002). N-acetylgalactosaminyl transferase-3 is a potential new marker for non-small cell lung cancers. Br J Cancer *87*, 751-755.

Esapa, C.T., Head, R.A., Jeyabalan, J., Evans, H., Hough, T.A., Cheeseman, M.T., McNally, E.G., Carr, A.J., Thomas, G.P., Brown, M.A*., et al.* (2012). A mouse with an N-Ethyl-N-nitrosourea (ENU) Induced Trp589Arg Galnt3 mutation represents a model for hyperphosphataemic familial tumoural calcinosis. PLoS One *7*, e43205.

Esteban, M.A., Bao, X., Zhuang, Q., Zhou, T., Qin, B., and Pei, D. (2012). The mesenchymal-to-epithelial transition in somatic cell reprogramming. Curr Opin Genet Dev *22*, 423-428.

Freeze, H.H., Schachter, H., and Kinoshita, T. (2015). Genetic Disorders of Glycosylation. In Essentials of Glycobiology, rd, A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer*, et al.*, eds. (Cold Spring Harbor (NY)), pp. 569-582.

Garringer, H.J., Fisher, C., Larsson, T.E., Davis, S.I., Koller, D.L., Cullen, M.J., Draman, M.S., Conlon, N., Jain, A., Fedarko, N.S.*, et al.* (2006). The role of mutant UDP-N-acetyl-alpha-D-galactosamine-polypeptide N-acetylgalactosaminyltransferase 3 in regulating serum intact fibroblast growth factor 23 and matrix extracellular phosphoglycoprotein in heritable tumoral calcinosis. J Clin Endocrinol Metab *91*, 4037-4042.

Garringer, H.J., Mortazavi, S.M., Esteghamat, F., Malekpour, M., Boztepe, H., Tanakol, R., Davis, S.I., and White, K.E. (2007). Two novel GALNT3 mutations in familial tumoral calcinosis. Am J Med Genet A *143A*, 2390-2396.

Georgiades, P., Cox, B., Gertsenstein, M., Chawengsaksophak, K., and Rossant, J. (2007). Trophoblast-specific gene manipulation using lentivirus-based vectors. Biotechniques *42*, 317-318, 320, 322-315.

Gu, C., Oyama, T., Osaki, T., Li, J., Takenoyama, M., Izumi, H., Sugio, K., Kohno, K., and Yasumoto, K. (2004). Low expression of polypeptide GalNAc N-acetylgalactosaminyl transferase-3 in lung adenocarcinoma: impact on poor prognosis and early recurrence. Br J Cancer *90*, 436-442.

Gumbiner, B.M. (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. Cell *84*, 345-357.

Han, J., Lee, J.D., Jiang, Y., Li, Z., Feng, L., and Ulevitch, R.J. (1996). Characterization of the structure and function of a novel MAP kinase kinase (MKK6). J Biol Chem *271*, 2886-2891.

Haque, K., Pandey, A.K., Zheng, H.W., Riazuddin, S., Sha, S.H., and Puligilla, C. (2016). MEKK4 Signaling Regulates Sensory Cell Development and Function in the Mouse Inner Ear. J Neurosci *36*, 1347-1361.

Heerboth, S., Lapinska, K., Snyder, N., Leary, M., Rollinson, S., and Sarkar, S. (2014). Use of epigenetic drugs in disease: an overview. Genet Epigenet *6*, 9-19.

Hennessy, B.T., Gonzalez-Angulo, A.M., Stemke-Hale, K., Gilcrease, M.Z., Krishnamurthy, S., Lee, J.S., Fridlyand, J., Sahin, A., Agarwal, R., Joy, C.*, et al.* (2009). Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. Cancer Res *69*, 4116-4124.

Herranz, N., Pasini, D., Diaz, V.M., Franci, C., Gutierrez, A., Dave, N., Escriva, M., Hernandez-Munoz, I., Di Croce, L., Helin, K.*, et al.* (2008). Polycomb complex 2 is

required for E-cadherin repression by the Snail1 transcription factor. Mol Cell Biol *28*, 4772-4781.

Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S.*, et al.* (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol *8*, R76.

Hough, C.D., Sherman-Baust, C.A., Pizer, E.S., Montz, F.J., Im, D.D., Rosenshein, N.B., Cho, K.R., Riggins, G.J., and Morin, P.J. (2000). Large-scale serial analysis of gene expression reveals genes differentially expressed in ovarian cancer. Cancer Res *60*, 6281-6287.

Housman, G., Byler, S., Heerboth, S., Lapinska, K., Longacre, M., Snyder, N., and Sarkar, S. (2014). Drug resistance in cancer: an overview. Cancers (Basel) *6*, 1769-1792.

Ichikawa, S., Sorenson, A.H., Austin, A.M., Mackenzie, D.S., Fritz, T.A., Moh, A., Hui, S.L., and Econs, M.J. (2009). Ablation of the Galnt3 gene leads to low-circulating intact fibroblast growth factor 23 (Fgf23) concentrations and hyperphosphatemia despite increased Fgf23 expression. Endocrinology *150*, 2543-2550.

Iwano, M., Plieth, D., Danoff, T.M., Xue, C., Okada, H., and Neilson, E.G. (2002). Evidence that fibroblasts derive from epithelium during tissue fibrosis. J Clin Invest *110*, 341-350.

Jamal, B.T., Nita-Lazar, M., Gao, Z., Amin, B., Walker, J., and Kukuruzinska, M.A. (2009). N-glycosylation status of E-cadherin controls cytoskeletal dynamics through the organization of distinct beta-catenin- and gamma-catenin-containing AJs. Cell Health Cytoskelet *2009*, 67-80.

Jordan, N.V., Prat, A., Abell, A.N., Zawistowski, J.S., Sciaky, N., Karginova, O.A., Zhou, B., Golitz, B.T., Perou, C.M., and Johnson, G.L. (2013). SWI/SNF chromatin-remodeling factor Smarcd3/Baf60c controls epithelial-mesenchymal transition by inducing Wnt5a signaling. Mol Cell Biol *33*, 3011-3025.

Julien, S., Adriaenssens, E., Ottenberg, K., Furlan, A., Courtand, G., Vercoutter-Edouart, A.S., Hanisch, F.G., Delannoy, P., and Le Bourhis, X. (2006). ST6GalNAc I expression in MDA-MB-231 breast cancer cells greatly modifies their O-glycosylation pattern and enhances their tumourigenicity. Glycobiology *16*, 54-64.

Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. J Clin Invest *119*, 1420-1428.

Kazlauskas, A., and Cooper, J.A. (1988). Protein kinase C mediates platelet-derived growth factor-induced tyrosine phosphorylation of p42. J Cell Biol *106*, 1395-1402.

Kim, Y.J., and Varki, A. (1997). Perspectives on the significance of altered glycosylation of glycoproteins in cancer. Glycoconj J *14*, 569-576.

Kufe, D.W. (2009). Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer *9*, 874-885.

Kufe, D.W. (2013). MUC1-C oncoprotein as a target in breast cancer: activation of signaling pathways and therapeutic approaches. Oncogene *32*, 1073-1081.

Kyriakis, J.M., and Avruch, J. (2001). Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. Physiol Rev *81*, 807-869.

Lawler, S., Fleming, Y., Goedert, M., and Cohen, P. (1998). Synergistic activation of SAPK1/JNK1 by two MAP kinase kinases in vitro. Curr Biol *8*, 1387-1390.

Lerdrup, M., Johansen, J.V., Agrawal-Singh, S., and Hansen, K. (2016). An interactive environment for agile analysis and visualization of ChIP-sequencing data. Nat Struct Mol Biol *23*, 349-357.

Luo, W., Ng, W.W., Jin, L.H., Ye, Z., Han, J., and Lin, S.C. (2003). Axin utilizes distinct regions for competitive MEKK1 and MEKK4 binding and JNK activation. J Biol Chem *278*, 37451-37458.

Maeshima, A., Miyagi, A., Hirai, T., and Nakajima, T. (1997). Mucin-producing adenocarcinoma of the lung, with special reference to goblet cell type adenocarcinoma: immunohistochemical observation and Ki-ras gene mutation. Pathol Int *47*, 454-460.

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

Mansour, S.G., Puthumana, J., Coca, S.G., Gentry, M., and Parikh, C.R. (2017). Biomarkers for the detection of renal fibrosis and prediction of renal outcomes: a systematic review. BMC Nephrol *18*, 72.

Maupin, K.A., Sinha, A., Eugster, E., Miller, J., Ross, J., Paulino, V., Keshamouni, V.G., Tran, N., Berens, M., Webb, C.*, et al.* (2010). Glycogene expression alterations associated with pancreatic cancer epithelial-mesenchymal transition in complementary model systems. PLoS One *5*, e13002.

Medina, M., Velez, D., Asenjo, J.A., Egea, G., Real, F.X., Gil, J., and Subiza, J.L. (1999). Human colon adenocarcinomas express a MUC1-associated novel carbohydrate epitope on core mucin glycans defined by a monoclonal antibody (A10) raised against murine Ehrlich tumor cells. Cancer Res *59*, 1061-1070.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49-55.

Mita, H., Tsutsui, J., Takekawa, M., Witten, E.A., and Saito, H. (2002). Regulation of MTK1/MEKK4 kinase activity by its N-terminal autoinhibitory domain and GADD45 binding. Mol Cell Biol *22*, 4544-4555.

Mobley, R.J., Raghu, D., Duke, L.D., Abell-Hart, K., Zawistowski, J.S., Lutz, K., Gomez, S.M., Roy, S., Homayouni, R., Johnson, G.L.*, et al.* (2017). MAP3K4 Controls the Chromatin Modifier HDAC6 during Trophoblast Stem Cell Epithelial-to-Mesenchymal Transition. Cell Rep *18*, 2387-2400.

Munkley, J. (2016). The Role of Sialyl-Tn in Cancer. Int J Mol Sci *17*, 275.

Nairn, A.V., Aoki, K., dela Rosa, M., Porterfield, M., Lim, J.M., Kulik, M., Pierce, J.M., Wells, L., Dalton, S., Tiemeyer, M.*, et al.* (2012). Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis. J Biol Chem *287*, 37835-37856.

Nakajima, Y., Yamagishi, T., Hokari, S., and Nakamura, H. (2000). Mechanisms involved in valvuloseptal endocardial cushion formation in early cardiogenesis: roles of transforming growth factor (TGF)-beta and bone morphogenetic protein (BMP). Anat Rec *258*, 119-127.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.P., Tong, F.*, et al.* (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell *10*, 515-527.

Okada, H., Strutz, F., Danoff, T.M., Kalluri, R., and Neilson, E.G. (1996). Possible mechanisms of renal fibrosis. Contrib Nephrol *118*, 147-154.

Onitsuka, K., Shibao, K., Nakayama, Y., Minagawa, N., Hirata, K., Izumi, H., Matsuo, K., Nagata, N., Kitazato, K., Kohno, K.*, et al.* (2003). Prognostic significance of UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-3 (GalNAc-T3) expression in patients with gastric carcinoma. Cancer Sci *94*, 32-36.

Parker, J.S., Mullins, M., Cheang, M.C., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z.*, et al.* (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol *27*, 1160-1167.

Pattabiraman, D.R., Bierie, B., Kober, K.I., Thiru, P., Krall, J.A., Zill, C., Reinhardt, F., Tam, W.L., and Weinberg, R.A. (2016). Activation of PKA leads to mesenchymal-to-epithelial transition and loss of tumor-initiating ability. Science *351*, aad3680.

Pearson, G., Robinson, F., Beers Gibson, T., Xu, B.E., Karandikar, M., Berman, K., and Cobb, M.H. (2001). Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. Endocr Rev *22*, 153-183.

Peinado, H., Ballestar, E., Esteller, M., and Cano, A. (2004). Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. Mol Cell Biol *24*, 306-319.

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

Pinho, S., Marcos, N.T., Ferreira, B., Carvalho, A.S., Oliveira, M.J., Santos-Silva, F., Harduin-Lepers, A., and Reis, C.A. (2007). Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells. Cancer Lett *249*, 157-170.

Pinho, S.S., and Reis, C.A. (2015). Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer *15*, 540-555.

Poirier, F., and Kimber, S. (1997). Cell surface carbohydrates and lectins in early development. Mol Hum Reprod *3*, 907-918.

Ponnusamy, M.P., Seshacharyulu, P., Lakshmanan, I., Vaz, A.P., Chugh, S., and Batra, S.K. (2013). Emerging role of mucins in epithelial to mesenchymal transition. Curr Cancer Drug Targets *13*, 945-956.

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

Prat, A., and Perou, C.M. (2011). Deconstructing the molecular portraits of breast cancer. Mol Oncol *5*, 5-23.

Raman, M., Chen, W., and Cobb, M.H. (2007). Differential regulation and properties of MAPKs. Oncogene *26*, 3100-3112.

Ray, L.B., and Sturgill, T.W. (1988). Insulin-stimulated microtubule-associated protein kinase is phosphorylated on tyrosine and threonine in vivo. Proc Natl Acad Sci U S A *85*, 3753-3757.

Roche, J. (2018). Erratum: Roche, J. The Epithelial-to-Mesenchymal Transition in Cancer. Cancers, 2018, 10, 52. Cancers (Basel) *10*.

Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C*., et al.* (2000). Eomesodermin is

required for mouse trophoblast development and mesoderm formation. Nature *404*, 95-99.

Saitoh, M. (2018). Involvement of partial EMT in cancer progression. J Biochem *164*, 257-264.

Sarkar, S., Horn, G., Moulton, K., Oza, A., Byler, S., Kokolus, S., and Longacre, M. (2013). Cancer development, progression, and therapy: an epigenetic overview. Int J Mol Sci *14*, 21087-21113.

Sato, M., Muragaki, Y., Saika, S., Roberts, A.B., and Ooshima, A. (2003). Targeted disruption of TGF-beta1/Smad3 signaling protects against renal tubulointerstitial fibrosis induced by unilateral ureteral obstruction. J Clin Invest *112*, 1486-1494.

Sewell, R., Backstrom, M., Dalziel, M., Gschmeissner, S., Karlsson, H., Noll, T., Gatgens, J., Clausen, H., Hansson, G.C., Burchell, J.*, et al.* (2006). The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer. J Biol Chem *281*, 3586-3594.

Shental-Bechor, D., and Levy, Y. (2008). Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. Proc Natl Acad Sci U S A *105*, 8256-8261.

Simmons, D.G., and Cross, J.C. (2005). Determinants of trophoblast lineage and cell subtype specification in the mouse placenta. Dev Biol *284*, 12-24.

Sparks, S.E., and Krasnewich, D.M. (1993). Congenital Disorders of N-Linked Glycosylation and Multiple Pathway Overview. In GeneReviews((R)), M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, K. Stephens, and A. Amemiya, eds. (Seattle (WA)).

Stanley, P., Taniguchi, N., and Aebi, M. (2015). N-Glycans. In Essentials of Glycobiology, rd, A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer*, et al.*, eds. (Cold Spring Harbor (NY)), pp. 99-111.

Stein, B., Brady, H., Yang, M.X., Young, D.B., and Barbosa, M.S. (1996). Cloning and characterization of MEK6, a novel member of the mitogen-activated protein kinase kinase cascade. J Biol Chem *271*, 11427-11433.

Strous, G.J., and Dekker, J. (1992). Mucin-type glycoproteins. Crit Rev Biochem Mol Biol *27*, 57-92.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Sutherlin, M.E., Nishimori, I., Caffrey, T., Bennett, E.P., Hassan, H., Mandel, U., Mack, D., Iwamura, T., Clausen, H., and Hollingsworth, M.A. (1997). Expression of three UDP-N-acetyl-alpha-D-galactosamine:polypeptide GalNAc N-acetylgalactosaminyltransferases in adenocarcinoma cell lines. Cancer Res *57*, 4744-4748.

Takeichi, M. (2014). Dynamic contacts: rearranging adherens junctions to drive epithelial remodelling. Nat Rev Mol Cell Biol *15*, 397-410.

Tan, X., Li, Y., and Liu, Y. (2006). Paricalcitol attenuates renal interstitial fibrosis in obstructive nephropathy. J Am Soc Nephrol *17*, 3382-3393.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075.

Taniuchi, K., Cerny, R.L., Tanouchi, A., Kohno, K., Kotani, N., Honke, K., Saibara, T., and Hollingsworth, M.A. (2011). Overexpression of GalNAc-transferase GalNAc-T3 promotes pancreatic cancer cell growth. Oncogene *30*, 4843-4854.

Team, R.C. (2013). R: A language and environment for statistical computing. R. Foundation for Statistical Computing (Vienna, Austria: http://www. R-project. org).

Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871-890.

Tian, E., Hoffman, M.P., and Ten Hagen, K.G. (2012). O-glycosylation modulates integrin and FGF signalling by influencing the secretion of basement membrane components. Nat Commun *3*, 869.

Tian, E., and Ten Hagen, K.G. (2007). A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is required for epithelial tube formation. J Biol Chem *282*, 606-614.

Tong, Z.T., Cai, M.Y., Wang, X.G., Kong, L.L., Mai, S.J., Liu, Y.H., Zhang, H.B., Liao, Y.J., Zheng, F., Zhu, W.*, et al.* (2012). EZH2 supports nasopharyngeal carcinoma cell aggressiveness by forming a co-repressor complex with HDAC1/HDAC2 and Snail to inhibit E-cadherin. Oncogene *31*, 583-594.

Topaz, O., Shurman, D.L., Bergman, R., Indelman, M., Ratajczak, P., Mizrachi, M., Khamaysi, Z., Behar, D., Petronius, D., Friedman, V.*, et al.* (2004). Mutations in GALNT3, encoding a protein involved in O-linked glycosylation, cause familial tumoral calcinosis. Nat Genet *36*, 579-581.

Van den Steen, P., Rudd, P.M., Dwek, R.A., and Opdenakker, G. (1998). Concepts and principles of O-linked glycosylation. Crit Rev Biochem Mol Biol *33*, 151-208.

Van Hof, A.C., Molthoff, C.F., Davies, Q., Perkins, A.C., Verheijen, R.H., Kenemans, P., den Hollander, W., Wilhelm, A.J., Baker, T.S., Sopwith, M.*, et al.* (1996). Biodistribution of (111)indium-labeled engineered human antibody CTMO1 in ovarian cancer patients: influence of protein dose. Cancer Res *56*, 5179-5185.

Vester-Christensen, M.B., Halim, A., Joshi, H.J., Steentoft, C., Bennett, E.P., Levery, S.B., Vakhrushev, S.Y., and Clausen, H. (2013). Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. Proc Natl Acad Sci U S A *110*, 21018-21023.

Wang, Z.Q., Bachvarova, M., Morin, C., Plante, M., Gregoire, J., Renaud, M.C., Sebastianelli, A., and Bachvarov, D. (2014). Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. Oncotarget *5*, 544-560.

Warr, N., Carre, G.A., Siggers, P., Faleato, J.V., Brixey, R., Pope, M., Bogani, D., Childers, M., Wells, S., Scudamore, C.L.*, et al.* (2012). Gadd45gamma and Map3k4 interactions regulate mouse testis determination via p38 MAPK-mediated control of Sry expression. Dev Cell *23*, 1020-1031.

Wheelock, M.J., Shintani, Y., Maeda, M., Fukumoto, Y., and Johnson, K.R. (2008). Cadherin switching. J Cell Sci *121*, 727-735.

Willsher, P.C., Xing, P.X., Clarke, C.P., Ho, D.W., and McKenzie, I.F. (1993). Mucin 1 antigens in the serum and bronchial lavage fluid of patients with lung cancer. Cancer *72*, 2936-2942.

Yang, J., and Weinberg, R.A. (2008). Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Dev Cell *14*, 818-829.

Yang, X., and Qian, K. (2017). Protein O-GlcNAcylation: emerging mechanisms and functions. Nat Rev Mol Cell Biol *18*, 452-465.

Yarema, K.J., and Bertozzi, C.R. (2001). Characterizing glycosylation pathways. Genome Biol *2*, REVIEWS0004.

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

Zeisberg, E.M., Tarnavski, O., Zeisberg, M., Dorfman, A.L., McMullen, J.R., Gustafsson, E., Chandraker, A., Yuan, X., Pu, W.T., Roberts, A.B.*, et al.* (2007). Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. Nat Med *13*, 952-961.

Zeisberg, M., Bottiglio, C., Kumar, N., Maeshima, Y., Strutz, F., Muller, G.A., and Kalluri, R. (2003). Bone morphogenic protein-7 inhibits progression of chronic renal

fibrosis associated with two genetic mouse models. Am J Physiol Renal Physiol *285*, F1060-1067.

Zhang, L., Zhang, Y., and Hagen, K.G. (2008). A mucin-type O-glycosyltransferase modulates cell adhesion during Drosophila development. J Biol Chem *283*, 34076-34086.

Zhang, Y., and Weinberg, R.A. (2018). Epithelial-to-mesenchymal transition in cancer: complexity and opportunities. Front Med *12*, 361-373.

Zhu, W., Leber, B., and Andrews, D.W. (2001). Cytoplasmic O-glycosylation prevents cell surface transport of E-cadherin during apoptosis. EMBO J *20*, 5999-6007.

Appendix A

Antibodies

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Alexa Fluor Phalloidin 594	Thermo Fisher Scientific	Cat#A12381; RRID: AB_2315633
Alexa Fluor 488-conjugated Streptavidin	Jackson ImmunoResearch Labs	Cat#016-540-084; RRID: AB_2337249
Donkey anti-mouse peroxidase conjugate	Jackson ImmunoResearch Labs	Cat#715-035-151; RRID: AB_2340771
Donkey anti-rabbit peroxidase conjugate	Jackson ImmunoResearch Labs	Cat#711-035-152; RRID: AB_10015282
Goat anti-mouse polyclonal Daylight 488	Thermo Fisher Scientific	Cat#35503; RRID: AB_1965946
Goat anti-rabbit polyclonal Daylight 488	Thermo Fisher Scientific	Cat#35553; RRID: AB_1965947
Mouse monoclonal anti-lClaudin-6	Santa Cruz Biotechnology	Cat#sc-393671
Mouse monoclonal anti-E-cadherin	BD Biosciences	Cat#610181; RRID: AB_397580
Mouse monoclonal anti-Flag	Thermo Fisher Scientific	Cat#MA1-91878; RRID: AB_1957945
Mouse monoclonal anti-GAPDH	Thermo Fisher Scientific	Cat#AM4300; RRID: AB_2536381
Mouse monoclonal anti-Tubulin	Santa Cruz Biotechnology	Cat#sc-53646; RRID: AB_630403
Mouse monoclonal anti-Tubulin	Sigma-Aldrich	Cat#T6793; RRID: AB_477585
Mouse monoclonal anti-Vinculin	Abcam	Cat#ab18058; RRID: AB_444215
Rabbit polyclonal Alexa 594	Cell Signaling Technology	Cat#8889; RRID: AB_2716249
Rabbit monoclonal anti-Alpha 1 catenin	Abcam	Cat#ab51032; RRID: AB_868700
Rabbit monoclonal anti-Delta 1 catenin	Abcam	Cat#ab92514; RRID: AB_10565040
Rabbit monoclonal anti-E-cadherin	Abcam	Cat#ab212059
Rabbit monoclonal anti-H2BK5Ac	Active Motif	Cat#39123; RRID: AB_2615079
Rabbit monoclonal anti-HDAC6	Cell Signaling Technology	Cat#7612; RRID: AB_10889735

Antibodies (Continued)

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Rabbit monoclonal anti-IgG	Abcam	Cat#ab172730; RRID: AB_2687931
Rabbit monoclonal anti-Lamin B1	Abcam	Cat#ab133741; RRID: AB_2616597
Rabbit polyclonal anti-ERK2	Santa Cruz Biotechnology	Cat#sc-154; RRID: AB_2141292
Rabbit polyclonal anti-GALNT3	Abgent	Cat#AP9208C; RRID: AB_10612485
Biotinylated Vicia Villosa Lectin	Vector Laboratories	Cat#B-1235; RRID: AB_2336855
Peroxidase-streptavidin	Jackson ImmunoResearch Labs	Cat#016-030-084; RRID: AB_2337238

Appendix B

Critical commercial assay

REAGENT or RESOURCE	SOURCE	IDENTIFIER
iTaq Universal SYBR Green Supermix	Bio-Rad	Cat#1725125
Rneasy Plus Mini Kit	Qiagen	Cat#74134
Rneasy Micro Kit	Qiagen	Cat#74004
High capacity cDNA Reverse Transcription Kit	Thermo Fisher Scientific	Cat#4368813
Gateway LR Clonase II Enzyme mix	Invitrogen	Cat#11791-020
Corning BioCoat Matrigel Invasion Chamber: With GFR Matrigel Matrix	Thermo Fisher Scientific	Cat#08-774-193
SsoAdvanced Universal SYBR Green Supermix	Bio-Rad	Cat#1725274
GeneJet Plasmid Miniprep Kit	Thermo Fisher Scientific	Cat#K0502
Transwells 6.4mm with 0.4μm pore polycarbonate	Corning	Cat#3413
Transwells 6.4mm with 0.4μm pore polyethylene terephthalate	Corning	Cat#353095
Q5 Site-Directed Mutagenesis Kit	New England BioLabs	Cat#E0554S

Appendix C

Chemicals, Peptides and Recombinant Proteins

REAGENT or RESOURCE	SOURCE	IDENTIFIER
2-mercaptoethanol (BME)	Thermo Fisher Scientific	Cat#31350010
Agarose bound Vicia Villosa Lectin beads	Vector Laboratories	Cat#AL-1233; RRID: AB_2336854
Enzyme Free Cell Dissociation Solution PBS Based (1X)	Millipore Sigma	Cat#S-014-C
Fetal Bovine Serum-Gibco	Thermo Fisher Scientific	Cat#10437-028
Hyclone Glutamine	Thermo Fisher Scientific	Cat#SH3003401
Heparin	Sigma-Aldrich	Cat#H3149-10KU; CAS: 9041-08-1
Hyclone Penicillin-Streptomycin	Thermo Fisher Scientific	Cat#SV30010
Recombinant Human FGF4	Peprotech	Cat#100-31
Recombinant Protein A Sepharose 4B	Thermo Fisher Scientific	Cat#101141
Sodium pyruvate	Thermo Fisher Scientific	Cat#SH30239.01
Trypsin (0.05%)	Thermo Fisher Scientific	Cat#SH30236.01
Insulin, Human Recombinant Zinc Solution	Thermo Fisher Scientific	Cat#12585014

Appendix D

Plasmids

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse *Galnt3* shRNA (Gsh1)	Open Biosystems	TRCN0000055098
Mouse *Galnt3* shRNA (Gsh2)	Open Biosystems	TRCN0000055099
Human *Galnt3* shRNA (Gsh1)	Dharmacon	TRCN0000035456
Human *Galnt3* shRNA (Gsh2)	Dharmacon	TRCN0000035458
Entry Vector Human *Galnt*3	Vidal Human ORFeome (Version 5.1)	Clone# 55179 BCII3565 in pDONR223
pENTR-*Cdh1*	(Cachat et al., 2014)	Addgene Plasmid#49776; RRID: Addgene_49776
Lentiviral FLAG tagged destination vector	Kind Gift from Dr. Gary Johnson (Jordan et al., 2013)	N/A

Appendix E

Primers for Real-time quantitation PCR for mRNA

Gene Name	Forward Primer	Reverse Primer
Actb mouse	AGC CAT GTA CGT AGC CAT CC	CTC TCA GCT GTG GTG GTG AA
Cdh1 mouse	CAG GTC TCC TCA TGG CTT TGC	CTT CCG AAA AGA AGG CTG TCC
Cdh2 mouse	AGC GCA GTC TTA CCG AAG G	TCG CTG CTT TCA TAC TGA ACT TT
Cdx2 mouse	AAG CCA AGT GAA AAC CAG GAC A	GGC AGC CAG CTC ACT TTT C
Csh1 mouse	TTA TCT TGG CCG CAG ATG TG	TCT GTC TGT TAT CCA AGT TTT ATC GAA
Eomes mouse	GCG CAT GTT TCC TTT CTT GAG	GGT CGG CCA GAA CCA CTT C
Esrrβ mouse	GGA CTC GCC GCC TAT GTT C	CGT TAA GCA TGT ACT CGC ATT TG
Ets2 mouse	CCT GTC GCC AAC AGT TTT CG	TGG AGT GTC TGA TCT TCA CTG A
Fgfr2 mouse	GCC TCT CGA ACA GTA TTC TCC T	ACA GGG TTC ATA AGG CAT GGG
Galnt1 mouse	TCC TGC TGC TTT ACT TCA GTG	GAC GAC TGG TTT CCC CAT TTC
Galnt2 mouse	GGA GAT GAG AAA GCA CAG GGT	TGC CTC CTG GTT AAA GTC TGG
Galnt3 mouse	TGC AAA TAG GAG CGC CCA TTA	GGC GAT CAA AAA CCG GCT TC
Galnt3 human	AGC GTT GGT CAG CCT CTA TG	TCC TTC TGG ATG TTG TGC CG
Galnt4 mouse	CGG TCA AGG AGG CAA TCA GT	CCG TCT TTG GGG CAG TTT TG
Galnt6 mouse	AGG RGC TGT GAA CAA CAT TAG G	CCC AGA AGG GTT TCA GCT CAG
Galnt7 mouse	TAT GCT CCA CTT GTA GCC CC	CCC CTC TGG CAA AAC CAT CT
Galnt10 mouse	AAG GAG GCT ATC AGG AGG GAC	AGA GAG CGA TTC AGG GAG ATT

Primers for Real-time quantitation PCR for mRNA (Continued)

Gene Name	Forward Primer	Reverse Primer
Galnt11 mouse	GTG ACT CAG CCC CTT AGG AAT	CTG TGG ATC TAA TAC CCG ACC T
Galnt12 mouse	TGT GGT GGC ACT CTG GAA AC	CGG GAA TAG GGA GCT TGC TT
Gcm1 mouse	AGA GAT ACT GAG CTG GGA CAT T	CTG TCG TCC GAG CTG TAG ATG
Hdac6 mouse	TCC ACC GGC CAA GAT TCT TC	CAG CAC ACT TCT TTC CAC CAC
Krt8 mouse	CGG CTA CTC AGG AGG ACT GA	TGA AAG TGT TGG ATC CCC CG
Krt18 mouse	GTG GAT GCC CCC AAA TCT CA	CCT CAA TCT GCT GAG ACC AGT
Pac mouse	TGA CCG AGT ACA AGC CCA C	ACA CCT TGC CGA TGT CGA G
Prl2C2 mouse	TCA ACC ATG CTC GTC GAT ACT G	GGC AAC ATT CTT CCA CAA TAA G
Snai1 mouse	CAC ACG CTG CCT TGT GTC T	GGT CAG CAA AAG CAC GGT T
Snai2 mouse	TGG TCA AGA AAC ATT TCA ACG CC	GGT GAG GAT CTC TGG TTT TGG TA
Tbp human	GAG AGT TCT GGG ATT GTA CCG	ATC CTC ATG ATT ACC GCA GC
Tpbpα mouse	TTC CTA GTC ATC CTA TGC CTG G	GGT CAT TTT CGC TAC TGT GAA GT
Twist 1 mouse	GGA CAA GCT GAG CAA GAT TCA	CGG AGA AGG CGT AGC TGA G
Vim mouse	TCC ACA CGC ACC TAC AGT CT	CCG AGG ACC GGG TCA CAT A
Zeb1 mouse	ACC GCC GTC ATT TAT CCT GAG	CAT CTG GTG TTC CGT TTT CAT CA

Primers for ChIP coupled to Real-time quantitative PCR

Gene Name	Forward Primer	Reverse Primer
Cldn7 Human	AAG AAA CCT GAG CAC CAG GG	CGG TGT TCG AGG GAA ATC CA
Galnt3 mouse	CTC GGG GCA AAC TCA TCT GT	AGG TGA CGC TTT CCA AAC CA
Human Galnt3	GCT TGG TGC AGT AAT TTT TTT CTTT	TTT CCA TCC TTG ATT CCT CTT TG

Primers for ChIP coupled to Real-time quantitative PCR

Gene Name	Forward Primer	Reverse Primer
Krt19 Human	AAG CTA ACC ATG CAG AAC CTC AA	CAG GGC GCG CAC CTT

Primers for Site Directed Mutagenesis

Gene Name	Forward Primer	Reverse Primer
Cdh1 mouse T40A	CGA GGT CTA CGC CTT CCC GGT	GAA CTG AAG CCG GGA CTG
Cdh1 mouse T63A	TGA AGG ATG CGC TGG CCG GCC	AAT CTC ACT CTG CCC AGG ACA TGG
Cdh1 mouse T68A	CCG GCC AAG GGC AGC CTT CTT	CCA GTG CAT CCT TCA AAT CTC ACT C

Primer for Sequencing

Gene Name	Forward Primer
Cdh1 sequencing primer	CAG GTC TCC TCA TGG CTT TGC