

**Transcriptional control of Epithelial to Mesenchymal Transition by  
Regulatory Factors and Epigenetic Mechanisms**

**Inauguraldissertation**

zur Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

**Neha Tiwari**

aus Indien

Basel, 2011

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Gerhard Christofori

Prof. Dr. Nancy Hynes

Basel, 15<sup>th</sup> Nov 2011

Prof. Dr. Martin Spiess

Dekan

## Summary

The World Health Organization (WHO) states cancer to be a leading cause of death worldwide accounting for 7.6 million deaths (around 13% of all deaths) and is projected rising to over 11 million in 2030. This is an alarming call to researchers for putting more effort into the analysis of the underlying patho-mechanisms. In a very simplified manner, cancer represents the destruction of healthy tissues and organs by uncontrolled cell proliferation and subsequent formation of a tumor. One key feature of solid tumors that marks the mostly deadly feature of the disease is the acquisition of the potential to invade into the surrounding tissue and form secondary tumors at distant sites, a process called ‘metastasis’. To gain migratory and invasive properties, cancer cells undergo epithelial to mesenchymal transition (EMT) where epithelial cells lose epithelial properties, e.g. their polarized organization and cell-cell junctions, and thus undergo changes in cytoskeleton organization and cell shape and acquire mesenchymal characteristics. Importantly, besides the formation of metastatic lesions, EMT is also involved during development as well as wound healing.

To gain insights into the complex process of EMT and to identify new potential markers for ongoing metastasis, we established different *in vitro* EMT model systems. Global expression profiling during TGF- $\beta$ -induced EMT revealed genome-wide transcriptome reprogramming during EMT and identified Kruppel-like factor 4 (Klf4) and the SRY-Related HMG-Box Gene4 (Sox4) as one of the key transcription factors that were modulated and may possibly contribute to transcriptional changes during EMT.

We investigated the role of Klf4 and Sox4 during EMT by employing two different *in vitro* systems of EMT, using normal murine mammary gland (NMuMG) and Polyoma middle T-breast cancer (Py2T) cells, which undergo a progressive EMT upon transforming growth factor (TGF- $\beta$ ) treatment. We further validated the role of Sox4 in breast cancer carcinogenesis *in vivo* by orthotopic injection of Sox4-depleted cells into the mammary fat pad of nude mice. In addition, we also investigated whether such TGF- $\beta$ -induced EMT accompanies epigenetic reprogramming and revealed how Polycomb group (PcG) complex-mediated H3K27me3 modification modulates transcription of key genes underlying this process, thereby regulating EMT.

Klf4 is a zinc-finger protein, known to be abnormally expressed in various cancers. It can act as a tumor suppressor or as an oncogene in context dependent manner in different carcinomas. Klf4 is downregulated during TGF- $\beta$ -induced EMT. Our data reveal a tumor suppressor role for Klf4 in breast carcinogenesis. Klf4 is essential for the maintenance of an epithelial phenotype during EMT, and forced expression of Klf4 leads to blockage of epithelial differentiation. Furthermore, Klf4 is inhibitory to EMT-driven cell migration and also behaves as a survival factor during TGF- $\beta$ -induced EMT. Genome-wide location analysis by next generation ChIP-seq analysis revealed that Klf4 directly occupies the promoter of many key EMT genes such as N-cadherin, Vimentin,  $\beta$ -catenin and Mapk8. Moreover, one of these Klf4 targets, Mapk8, encoding Jnk1, is upregulated during EMT and a double-knockdown of Klf4 and Jnk1 is able to overcome Klf4 knockdown-induced EMT, migration and apoptosis. These observations underscore a role of Klf4 during EMT by targeting and regulating crucial EMT genes.

Sox4 is also known to be deregulated in many cancers. Sox4 is upregulated during TGF- $\beta$ -induced EMT. We show that Sox4 is required for maintaining mesenchymal identity and depletion of Sox4 prevents TGF- $\beta$ -induced EMT. Sox4 reduction further impairs the migratory capacity of cells. Moreover, Sox4 provides a survival advantage to cells during breast carcinogenesis. In addition, Sox4 contributes towards TGF- $\beta$ -induced tumorigenicity and metastatic spread. Gene expression profiling after Sox4 depletion in complementation with Chromatin immunoprecipitation analysis revealed many key EMT genes such as Spred1, Edn1, Palld, Cyr61, Ereg, Areg and Yap1 which are directly targeted by Sox4 for transcriptional regulation. Furthermore, Sox4 also controls many genes which are shown to regulate various other features of EMT as well as cancer development such as angiogenesis, adhesion, migration, morphogenesis, cell cycle and cytoskeleton re-modeling. Ezh2, a catalytic subunit of the Polycomb Repressive Complex 2 (PRC2), has been also found to be transcriptionally regulated by Sox4. To delineate the role of Ezh2 during EMT, a loss of function approach has been used to demonstrate that Ezh2 is required for proper acquisition of EMT and EMT-driven processes such as migration and apoptosis. Taken together, our data provides a role of Sox4 during EMT via transcriptional regulation of key genes, including the Polycomb component, Ezh2.

We also studied the role of two prominent epigenetic modifications- DNA methylation and histone 3 lysine 27 tri-methylation (H3K27me3) during TGF- $\beta$ -induced EMT in a mammary epithelial cell line. Our data revealed no evidence of a reprogramming of DNA methylation



during this process. To assess the role of H3K27me3 during EMT, we performed chromatin immunoprecipitation using H3K27me3-specific antibodies followed by next-generation sequencing (ChIP-seq) on 6 different stages of EMT progression. This analysis revealed that many key EMT genes are regulated by H3K27me3 mark including *Mcam*, *Pdgfrb* and *Itga5* which are repressed by this mark in epithelial cells and loose it during EMT as they get activated conversely, *Cdh1*, *Ocln* and *Cdx2* gain this mark during EMT and get repressed in mesenchymal cells. We further illustrated that the coordinated activities of *Ezh1* and *Ezh2* are required for H3K27me3-mediated repression of the gene expression and their co-depletion de-represses target genes and blocks EMT. This study provides novel insights into the important regulatory role of the Polycomb machinery during EMT.

In summary, our findings demonstrate how transcription factors, such as *Klf4* and *Sox4* and the epigenetic machinery, such as PcG proteins, regulate EMT by directly contributing to the transcriptional reprogramming underlying this process.

Summary.....	1
Table of contents.....	4
<b>1. Introduction</b>	
<b>1.1 Mechanisms of carcinogenesis and acquired capabilities of cancer.....</b>	<b>8</b>
<b>1.2 Epithelial to mesenchymal transition (EMT).....</b>	<b>11</b>
<i>1.2.1 EMT and E-cadherin.....</i>	<i>14</i>
<i>1.2.2 EMT and growth factors.....</i>	<i>15</i>
<i>1.2.3 EMT and transcription factors.....</i>	<i>16</i>
<i>1.2.4 EMT and metastasis.....</i>	<i>18</i>
<b>1.2.5 EMT and migration.....</b>	<b>19</b>
1.2.5.1 Single cell migration.....	20
1.2.5.2 Collective cell migration.....	23
<i>1.2.6 EMT and cancer stem cells.....</i>	<i>24</i>
<i>1.2.7 EMT and miRNAs.....</i>	<i>28</i>
<i>1.2.8 EMT and splicing factor.....</i>	<i>31</i>
<b>1.3 TGF-<math>\beta</math>: its role in tissue homeostasis and cancer cell invasion.....</b>	<b>33</b>
<i>1.3.1 Dual role of TGF-<math>\beta</math>.....</i>	<i>33</i>
<i>1.3.2 TGF-<math>\beta</math> signaling.....</i>	<i>33</i>
<b>1.3.3 Canonical TGF-<math>\beta</math> signaling.....</b>	<b>34</b>
1.3.3.1 Canonical TGF- $\beta$ signaling mediated cell cycle arrest.....	35
1.3.3.2 Canonical TGF- $\beta$ signaling mediated apoptosis.....	35
<b>1.3.4 Non-Canonical TGF-<math>\beta</math> signaling.....</b>	<b>36</b>
1.3.4.1 TGF- $\beta$ -induced MAPK activation.....	36
1.3.4.2 TGF- $\beta$ -induced JNK/p38 activation.....	36
1.3.4.3 Rho-like GTPases in TGF- $\beta$ mediated EMT.....	37
1.3.4.4 TGF- $\beta$ -induced PI3K activation.....	37
<b>1.4 Epigenetic regulation of gene transcription.....</b>	<b>37</b>
<b>1.4.1 Epigenetic modification of Chromatin.....</b>	<b>38</b>
1.4.1.1 Histone acetylation.....	39
1.4.1.2 Histone phosphorylation.....	41

1.4.1.3 Histone methylation.....	42
<b>1.4.2 Polycomb-mediated repression.....</b>	<b>44</b>
1.4.2.1 EMT and Polycomb.....	46
<b>1.4.3 DNA methylation.....</b>	<b>49</b>
1.4.3.1 EMT and DNA methylation.....	52
<b>1.5 MARA.....</b>	<b>52</b>
<b>1.5.1 Model.....</b>	<b>53</b>
<b>1.5.2 MARA output.....</b>	<b>54</b>
<b>2. Aims of the study.....</b>	<b>55</b>
<b>3. Results</b>	
<b>3.1 Klf4 directly regulates transcription of genes crucial for Epithelial to Mesenchymal transition.....</b>	<b>56</b>
<b>3.1.1 Abstract.....</b>	<b>56</b>
<b>3.1.2 Introduction.....</b>	<b>56</b>
<b>3.1.3 Results.....</b>	<b>59</b>
3.1.3.1 Identification of Klf4 as a repressor during EMT.....	59
3.1.3.2 Klf4 blocks epithelial differentiation.....	59
3.1.3.3 Klf4 prevents cell migration and provides a survival advantage to the cells during EMT.....	61
3.1.3.4 Klf4 overexpression prevents EMT.....	65
3.1.3.5 Klf4 regulates the expression of crucial EMT genes by directly binding to their promoter.....	67
3.1.3.6 Klf4 is regulated by canonical TGF- $\beta$ signaling but regulates non-canonical TGF- $\beta$ signaling during EMT.....	70
<b>3.1.4 Discussion.....</b>	<b>73</b>
<b>3.1.5 Supplemental data.....</b>	<b>75</b>
<b>3.1.6 Methods and materials.....</b>	<b>81</b>
<b>3.2 Sox4 regulates epithelial to mesenchymal transition by directly controlling transcription of underlying master genes.....</b>	<b>88</b>

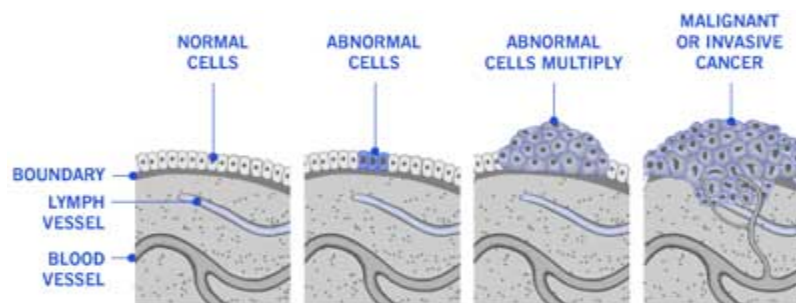
<b>3.2.1</b>	<b><i>Abstract</i></b> .....	<b>88</b>
<b>3.2.2</b>	<b><i>Introduction</i></b> .....	<b>88</b>
<b>3.2.3</b>	<b><i>Results</i></b> .....	<b>91</b>
	3.2.3.1 Identification of Sox4 as a transcriptional factor upregulated in EMT.....	91
	3.2.3.2 Sox4 depletion prevents EMT.....	93
	3.2.3.3 Sox4 provides a survival advantage during TGF- $\beta$ -induced EMT and supports migration.....	95
	3.2.3.4 Sox4 is required for TGF- $\beta$ -induced tumorigenesis and metastatic spread.....	97
	3.2.3.5 Ezh2 is transcriptionally regulated by Sox4.....	101
	3.2.3.6 Ezh2 knockdown phenocopies Sox4 depletion.....	103
<b>3.2.4</b>	<b><i>Discussion</i></b> .....	<b>105</b>
<b>3.2.5</b>	<b><i>Supplemental data</i></b> .....	<b>108</b>
<b>3.2.6</b>	<b><i>Methods and materials</i></b> .....	<b>112</b>
<b>3.3</b>	<b>Polycomb-dependent mechanisms regulate Epithelial to Mesenchymal transition</b> .....	<b>121</b>
<b>3.3.1</b>	<b><i>Abstract</i></b> .....	<b>121</b>
<b>3.3.2</b>	<b><i>Introduction</i></b> .....	<b>121</b>
<b>3.3.3</b>	<b><i>Results</i></b> .....	<b>124</b>
	3.3.3.1 Genomewide analysis of H3K27me3 mark reveals widespread epigenetic re-programming during TGF- $\beta$ -induced EMT.....	124
	3.3.3.2 Polycomb targets are key EMT genes.....	126
	3.3.3.3 Ezh1 and Ezh2 both contribute to the H3K27me3-mediated repression.....	126
	3.3.3.4 Co-depletion of Ezh1 and Ezh2 blocks epithelial differentiation....	130
	3.3.3.5 Ezh1 and Ezh2 ablation prevents cell migration but does not provides a survival advantage to the cells during EMT.....	132
	3.3.3.6 Dual role of Ezh1 and Ezh2 during Polycomb-mediated regulation of EMT.....	134

3.3.3.7 MeDIP analysis shows no significant difference in methylation patterns.....	137
<b>3.3.4 Discussion.....</b>	<b>138</b>
<b>3.3.5 Supplemental data.....</b>	<b>140</b>
<b>3.3.6 Methods and materials.....</b>	<b>144</b>
<b>4. References.....</b>	<b>153</b>
<b>5. Curriculum vitae.....</b>	<b>178</b>
<b>5.1 Personal information.....</b>	<b>178</b>
<b>5.2 Experience and education.....</b>	<b>178</b>
5.2.1 <i>Experience.....</i>	<i>178</i>
5.2.2 <i>Education.....</i>	<i>178</i>
<b>5.3 Workshops, courses and conferences.....</b>	<b>179</b>
5.3.1 <i>Workshops.....</i>	<i>179</i>
5.3.2 <i>Courses.....</i>	<i>180</i>
5.3.3 <i>Conferences.....</i>	<i>181</i>
<b>5.4 Achievements .....</b>	<b>181</b>
<b>5.5 Publications.....</b>	<b>182</b>
<b>6. Acknowledgements.....</b>	<b>184</b>

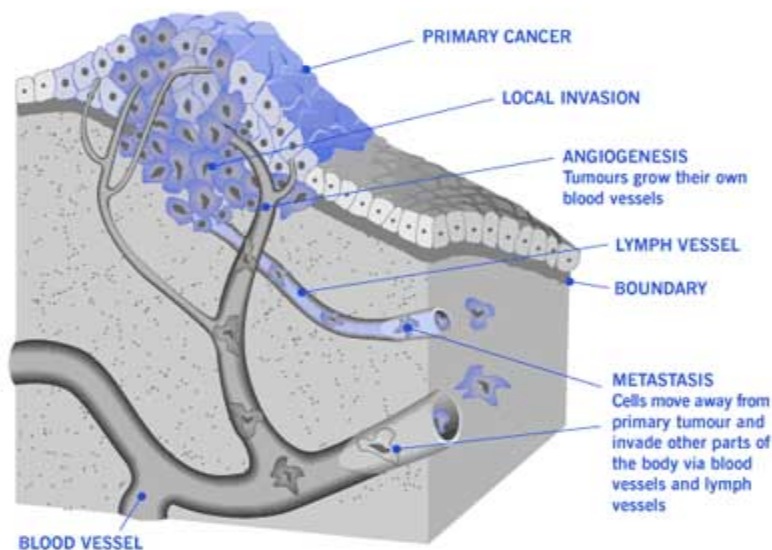
## 1.0 Introduction

### 1.1 Mechanism of carcinogenesis and acquired capabilities of cancer

Cancer was first described by Greek Hippocrates (ca. 460 BC – ca. 370 BC) who used the Greek word *carcinoma* (crab or crayfish) to describe it due to its crab-like tenacity. Cancer broadly refers to any one of a large number of diseases that are characterized by the development of abnormal cells that divide uncontrollably and have the ability to infiltrate and destroy normal body tissue. Cancer is the one of leading cause of death worldwide. However, thanks to basic research- due to improvements in screening and treatment, the survival rates are improving. The major challenge in cancer therapy is the metastatic spread of malignant cells and existence of cancer stem cells that survive any therapy and later leads to recurrence of the disease (**Figure 1**).



In 2000, Hanahan and Weinberg (Hanahan and Weinberg, 2000) described the following six essential physiological alterations which instruct malignant tumor growth:



1. **Sustaining proliferative signaling:** Growth factors are required for the proper propagation of normal cells and are usually provided exogenously. However, this dependency is greatly reduced in tumor cells which produce their own growth factors/ligands to which they respond via the expression of

**Figure 1: Systemic representation of cancer cell invasion.** Normal cells become metastatic and spread to distant organs through blood stream and lymph-vessels.

cognate receptors. Alternatively, cancer cells may send signals to stimulate normal cells within the supporting tumor-associated stroma, which reciprocate by supplying the cancer cells with various growth factors by over-expressing them (Bhowmick et al., 2004; Cheng et al., 2008). This liberation makes them free to grow in an uncontrolled manner.

2. **Evading growth suppressors:** Within a normal tissue, multiple anti-proliferative signals operate to maintain cellular quiescence and tissue homeostasis. In tumors, this signaling is mainly blocked by attacking the anti-proliferative signaling cascades which are commonly regulated by E2F transcription factors. These are essential for the expression of banks of genes required for cell cycle progression and p53, tumor suppressor (Burkhardt and Sage, 2008; Deshpande et al., 2005; Sherr and McCormick, 2002). Furthermore, Myc oncogene expression is also induced to override growth inhibition by these anti-growth molecules (Partanen et al., 2009). Recently, it has been shown that the Myc oncogene is regulated by another tumor suppressor, LKB1, whose expression is required for epithelial integrity (Shaw, 2009). The Myc oncogene only becomes active after suppression of LKB1 (Hezel and Bardeesy, 2008; Partanen et al., 2009). Another tumor suppressor protein Merlin, sequesters growth factor receptors and limits their ability to efficiently emit mitogenic signals and, thus, Merlin has to be repressed to provide a proliferative signal (Curto et al., 2007; Okada et al., 2005).
3. **Resisting cell death:** Several intrinsic factors (e.g. genomic instability) as well as extrinsic factors (e.g. hypoxia) induce apoptosis of untransformed cells. Tumor cells evolve a variety of strategies to limit or circumvent apoptosis. Most common is the loss of p53 tumor suppressor function, which eliminates this critical damage sensor from the apoptosis-inducing circuitry. Alternatively, tumors may achieve similar ends by increasing expression of anti-apoptotic regulators (Bcl-2, Bcl-xL) or of survival signals (Igf1/2), by downregulating proapoptotic factors (Bax, Bim, Puma), or by short-circuiting the extrinsic ligand-induced death pathway (Adams and Cory, 2007; Junttila and Evan, 2009; Willis and Adams, 2005).
4. **Enabling replicative immortality:** Untransformed cells are limited in their replicative potential by the length of their telomeres. Telomere maintenance is evident in virtually all types of malignant cells (Blasco, 2005; Shay and Wright, 2000). 85%–90% of them

succeed in doing so by upregulating expression of the telomerase enzyme or by the mechanism called ALT where telomere is maintained by recombination-based interchromosomal exchanges of sequence information (Bryan et al., 1995).

5. **Sustained angiogenesis:** The oxygen and nutrients supplied by the vasculature are crucial for cell function and survival. Tumors appear to activate the angiogenic switch by changing the balance of angiogenesis inducers such as VEGFs and FGFs (Baeriswyl and Christofori, 2009; Bergers and Benjamin, 2003) and countervailing inhibitors such as thrombospondin-1 or  $\beta$ -interferon (Singh et al., 1995).
6. **Tissue invasion and metastasis:** Normal cells are usually tethered with each other by adherens and tight junction proteins. Tumor cells act on these proteins and subject them to degradation. Consequently, cell-cell contact is lost and they can migrate and metastasize to the distant organs. In addition, many proteases also become active which pave a way for these cells to migrate by degrading extracellular matrix (Bex and van Roy, 2009; Cavallaro and Christofori, 2004; Fidler, 2003; Talmadge and Fidler, 2010).

In their recent review, they added two more hallmarks which are required for maintaining the malignancy (Hanahan and Weinberg, 2011).

1. **Reprogramming energy metabolism:** Even in the presence of ample oxygen, cancer cells prefer to metabolize glucose by glycolysis, which is a less efficient pathway for producing ATP. This effect was first observed by Otto Warburg (Hsu and Sabatini, 2008) and named after him as “Warburg effect”. To do so they upregulate GLUT1 transporters, which substantially increases glucose import into the cytoplasm; glycolytic enzymes and inhibitors of mitochondrial metabolism. One compelling idea to explain the Warburg effect is that the altered metabolism of cancer cells confers a selective advantage for survival and proliferation in the unique tumor microenvironment such as hypoxia (Semenza, 2010). Moreover, glycolysis allows the diversion of glycolytic intermediates into various biosynthetic pathways, which are required for assembly of new cells (Vander Heiden et al., 2009). In addition, in some tumors, two sub-populations of cancer cells are present which differ in their energy generating pathway. One population consists of glucose dependent cells that secrete lactate and other population basically used the lactate produced by their neighboring cells as an energy source (Feron, 2009; Kennedy and Dewhirst, 2010; Semenza,



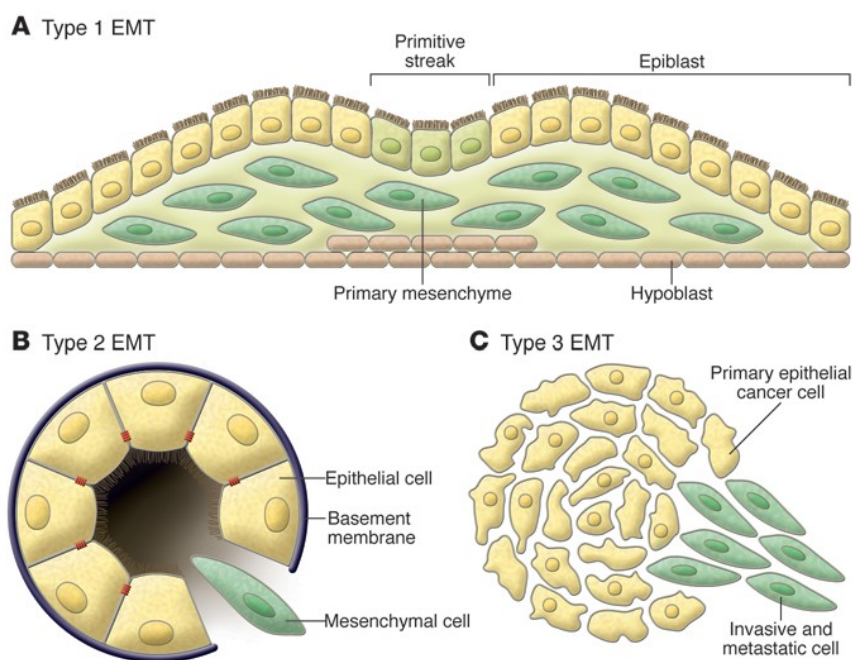
2008). In addition, increased glucose consumption, decreased oxidative phosphorylation, and accompanying lactate production are also distinguishing features of oncogene activation such as RAS and Myc (DeBerardinis et al., 2008).

- 2. Evading immune destructions:** In general, cells and tissues are constantly monitored by an ever-alert immune system, and that such immune surveillance is responsible for recognizing and eliminating the vast majority of cancer cells. To propagate happily, tumor cells have to omit themselves from the immune surveillance. It has been shown that tumors grow/arise rapidly in immunodeficient mice in comparison to immunocompetent mice (Kim et al., 2007; Teng et al., 2008). In addition, transplantation experiments have shown that cancer cells that originally arise in immunodeficient mice are not capable of inducing secondary tumors in immunocompetent hosts while the tumors cells from immunocompetent mice can do so (Kim et al., 2007; Teng et al., 2008). Cancer cells may paralyze infiltrating cytotoxic CD8<sup>+</sup> T-cells and NK cells by secreting TGF- $\beta$  or other immunosuppressive factors (Shields et al., 2010; Yang et al., 2010a). Furthermore, they can recruit cells that are actively immunosuppressive, including regulatory T cells (Tregs) and myeloid-derived suppressor cells (Mougiakakos et al., 2010; Ostrand-Rosenberg and Sinha, 2009). TGF- $\beta$  has been shown to inhibit the anti-tumoral activity of cytotoxic CD8<sup>+</sup> T cells, by inhibiting their ability to produce cytolytic factors such as pore-forming protein perforin, the caspase activating secreted factors granzyme A and B, and the pro-apoptotic cytokine Fas-ligand. In human glioma patients, TGF- $\beta$  decreases the expression of the activating immunoreceptor NKG2D on CD8<sup>+</sup> T cells and natural killer (NK) cells, which leads to reduced CD8<sup>+</sup> T cell, mediated cancer-directed cytotoxic response. Knockdown of TGF- $\beta$  synthesis in a glioma cell lines prevents NKG2D repression and enhanced glioma killing by cytotoxic T cells and NK cells (Thomas and Massague, 2005).

## 1.2 Epithelial to mesenchymal transition (EMT)

Epithelial-to-mesenchymal transition (EMT) is a basic cellular process in which epithelial cells lose epithelial properties, e.g. their polarized organization and cell-cell junctions, undergo changes in cytoskeleton and cell shape, acquire mesenchymal characteristics and

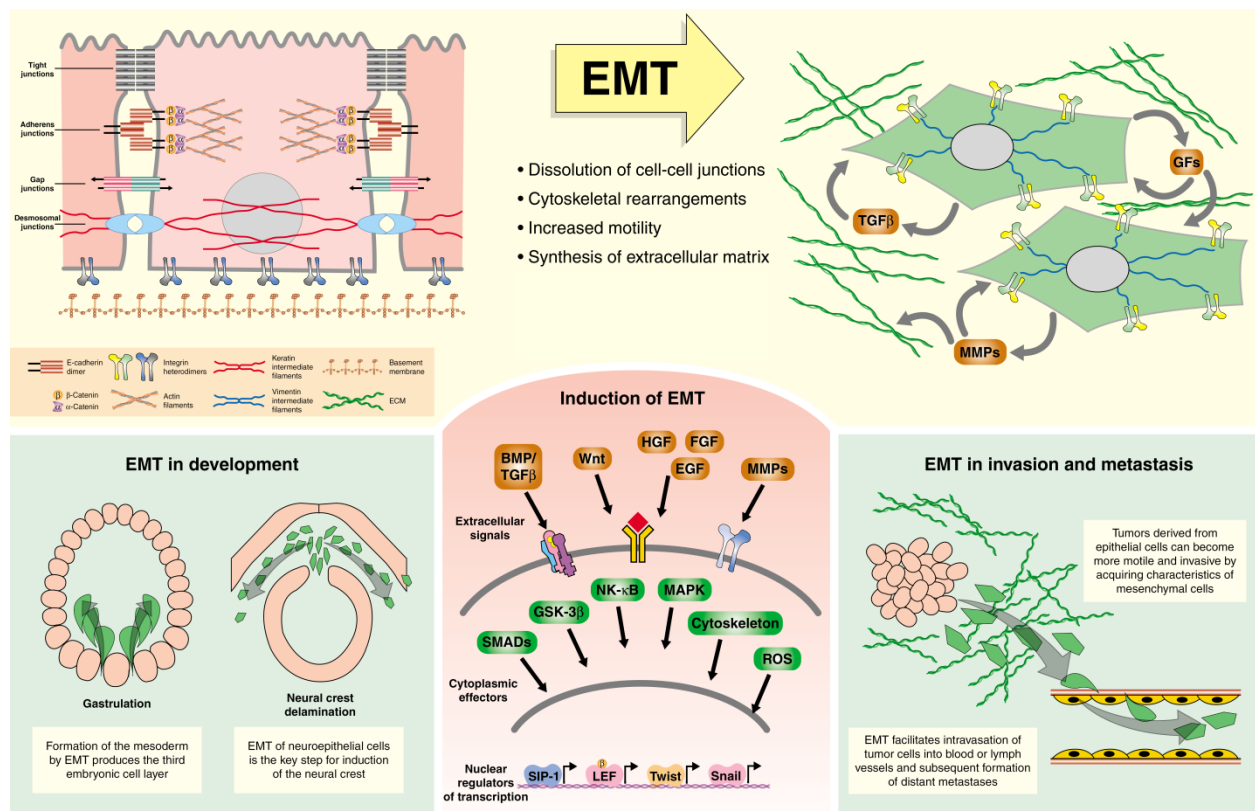
become migratory and invasive. EMT was first recognized as a distinct cell differentiation process in the late 70's, and has received increasing attention, as it not only occurs in normal development but is also an integral component of various pathological conditions (Hay, et.al, 1968). On the basis of its function, EMT can be classified into three different subtypes (**Figure 2**). Type 1 EMT is associated with implantation and embryonic gastrulation and gives rise to mesoderm and endoderm and to mobile neural crest cells. Type 2 EMT is involved in wound healing, tissue-regeneration, inflammation and fibrosis. Finally, Type 3 EMT comprises the transformation of epithelium associated with organs into the cancer cells which later leads to invasion and metastasis (Kalluri and Weinberg, 2009). Although the fate determination of Type 1 and Type 3 EMT is different, they follow the same pathophysiological pathway which involves the loss of epithelial marker E-cadherin and gain of mesenchymal marker N-cadherin. Such EMT associated “cadherin switch” is always annotated as a hallmark of this process (Hazan et al., 2004a; Maeda et al., 2005a).



**Figure 2: Three different subtypes of EMT.** (A) EMT during embryonic development (Type 1). (B) EMT during wound healing and fibrosis (Type 2). (C) EMT during pathological conditions such as cancer (Type 3). Adapted from Kalluri and Weinberg, 2009.

Moreover, the classical cadherin switch is assisted by substitution of many more epithelial markers such as tight junction proteins, claudins and occludins; desmosomal proteins;  $\alpha$  and  $\beta$ -catenins localization and cytokeratins with mesenchymal markers such as vimentin; fibronectin and matrix metalloproteinase secretion; integrin  $\alpha$  V, integrin  $\beta$  1 and smooth muscle actin

(Yang and Weinberg, 2008). Furthermore, in case of many carcinomas, tumor-associated stroma produces many growth factors, mainly HGF (Savagner et al., 1997), EGF (Lo et al., 2007), PDGF (Yang et al., 2006), FGF2 (Strutz et al., 2002) and TGF- $\beta$  (Zavadil and Bottinger, 2005), that result in induction of various transcription factors, notably Snail, Slug, zinc finger E-box binding homeobox 1 (ZEB1), Twist, Goosecoid, and FOXC2 (Jechlinger et al., 2002; Kokudo et al., 2008; Medici et al., 2008; Niessen et al., 2008; Shi and Massague, 2003; Thiery, 2002) which drive EMT. But the actual implementation is dependent on signal-transducing proteins, ERK, MAPK, PI3K, Akt, Smads, RhoB,  $\beta$ -catenin, lymphoid enhancer binding factor (LEF), Ras, and c-Fos as well as cell surface proteins such as  $\beta_4$  integrin,  $\alpha_5\beta_1$  integrin, and  $\alpha_v\beta_6$  integrin [(Tse and Kalluri, 2007); **Figure 3**]. Recently, microRNAs were also implicated to have a role in EMT. Non-coding miRNA such as miR200 and miR205 modulate the function of Zeb1 and Zeb2 transcription factors that are known to be repressors of E-cadherin gene expression (Gregory et al., 2008a; Gregory et al., 2008b; Korpala et al., 2008; Park et al., 2008; Paterson et al., 2008). In addition, the role of alternative splicing of transcripts in EMT is the most recent advancement in the field of cancer biology (see below 1.2.9).



**Figure 3: Overview of EMT process and its regulation.** During EMT, cells lose their cell-cell junctions and rearrange the cytoskeleton so that they can migrate, invade the neighbouring tissues and metastasize to the distant organs. Many signaling molecules have been implicated in the induction of EMT such as BMP/TGF- $\beta$ , Wt, HGF, FGF and EGF. Adapted from Derek C. Radisky, 2005.

### 1.2.1 EMT and E-cadherin

In mammals, adhesion between epithelial cells is generally mediated by three types of junctions: tight junctions (TJs), adherens junctions (AJs), and desmosomes, which together constitute the Intercellular Junctional Complex (Perez-Moreno et al., 2003). The transmembrane core of AJs consists of cadherins. Cadherins were first discovered as cell surface glycoproteins responsible for  $\text{Ca}^{2+}$ -dependent homophilic cell-cell adhesion during morula compaction in the preimplantation mouse embryo and during chick development (Gallin et al., 1983; Peyrieras et al., 1983; Yoshida and Takeichi, 1982). E-cadherin is the prototype and well-characterized member of this family, is primarily expressed in epithelia and required for cell-cell adhesion between two cells. The highly conserved, cytoplasmic tail of classical cadherins possesses a binding site for either  $\beta$ -catenin or  $\gamma$ -catenin via which they are connected to actin cytoskeleton (Aberle et al., 1996). On the other hand, E-cadherin also binds to p120-catenin and regulates surface trafficking, lysosomal degradation and localization of newly synthesized cadherins at the membrane (Ireton et al., 2002; Stehbens et al., 2006). Furthermore, p120-catenin has emerged as a major regulator and integrator of signaling by the Rho family of small GTPases (Anastasiadis, 2007), and this is at least partially dependent on its interaction with the cadherin and repressing the activity of RhoA (Wildenberg et al., 2006). On the other hand, three closely related PDZ-domain-containing proteins (ZO-1, ZO-2 and ZO-3) constitute the undercoat structure of TJs together with other peripheral membrane proteins such as cinglin, 7H6 antigen and symplekin. Occludin and claudins are the main players in organization of TJs and thought to constitute the backbone of TJs strands (Tsukita et al., 1999).

During EMT, cells lose their epithelial polarity and dissolve the adherent and tight junctions, favoring a more labile cell-cell adhesion and communicate with the extracellular matrix through focal adhesions. Thus, E-cadherin behaves as a tumor-suppressor, by keeping the cell-cell contacts intact and resisting their mobility (Egeblad and Werb, 2002; Friedl and Wolf, 2003;

Hood and Cheresch, 2002). It has been demonstrated that E-cadherin gene is silenced in many carcinomas and activation of E-cadherin is sufficient to reduce the aggressiveness of tumor cells (Vleminckx et al., 1991). Cadherins are generally regulated at both the mRNA and protein levels, by means of changes in subcellular distribution, translational or transcriptional events, and degradation. In various human carcinomas, functional loss of E-cadherin may result from the production of a defective protein, which could be a result of gene mutation, abnormal post-translational modifications (phosphorylation or glycosylation) or protein degradation (proteolysis) (Kang and Massague, 2004; Thiery and Morgan, 2004). It can also be regulated at the transcriptional level by silencing through its promoter hypermethylation. Although E-cadherin is downregulated during EMT, metastatic cells gain back E-cadherin, specifically during intravasation and seeding of metastatic cells (Kang and Massague, 2004; Thiery and Morgan, 2004). In addition to promoter hypermethylation, E-cadherin transcriptional repression may result from the activation of repressors, such as Snail, Slug, Zeb1, Zeb2 and Ets (Peinado et al., 2007).

### **1.2.2 EMT and growth factors**

The oncogenic pathways involving receptor tyrosine kinases (RTK) are shown to regulate adherent and tight junction proteins that are known to be involved in normal integrity of the epithelial cells such as E-cadherin, occludins, claudins and cytokeratins (Bos, 2005). It has been suggested that RTK activation participates in the EMT program by rendering the tight junction leaky and thus allowing access of TGF- $\beta$  to its receptor, one subunit of which would otherwise remain segregated in the tight junction (Bos et al., 2003). Growth factors that activate RTKs were the first identified as potent inducers of EMT by activating Mitogen activated protein kinase (MAPK) signaling through Extracellular-regulated kinase (Erk). For instance, hepatocyte growth factor (HGF) signaling leads to induction of various matrix metalloproteinase (MMPs) and extracellular matrix proteins (ECMs) which alters the cell-ECM and cell-cell interaction through regulation of integrins and cadherin expression (Berrier et al., 2000). Constitutive activation of Erk is shown to be required for complete EMT in epithelial tumor metastasis models (Chen et al., 1997; Ingber et al., 1995; Lele et al., 2006; Parker and Ingber, 2007; Zhang et al., 1996). HGF also influences the EMT process by regulating the key EMT transcription factor, Snail (Chrzanowska-Wodnicka and Burridge, 1996). Another EMT inducer, FGF also

determines the fate of epithelial cells by triggering the MAPK signaling and possibly, TGF- $\beta$  signaling (Bershadsky et al., 2006; Riveline et al., 2001). Recent studies in *Xenopus* showed that FGF signaling promotes mesodermal differentiation by enhancing embryonic TGF- $\beta$ /nodal signaling (Bershadsky et al., 2006; Riveline et al., 2001). Binding of FGF ligand to its receptor elicits MAPK/FGFR signaling that induces phosphorylation of the N-terminal of p53 and further interaction of p53 with TGF- $\beta$ -induced Smads in the nucleus.

In addition to MAPK signaling, phosphatidylinositol 3' kinase (PI3K) signaling also plays a key role in inducing EMT. Cells producing a constitutively active form of Akt, a downstream effector of PI3K signaling, produce the transcription factor Snail, which is known to repress E-cadherin transcription (Grille et al., 2003) and triggering EMT. Moreover, PI3K signaling can also be activated by integrins and small GTPases from Rho family, which controls cytoskeleton re-modeling, extended its necessity in EMT-driven processes (Xia et al., 2008; Zamir and Geiger, 2001). Furthermore, autocrine PDGFR signaling with oncogenic Ras, hyperactivates PI3K signaling that is required for survival during EMT (Jechlinger et al., 2002).

Epidermal growth factor (EGF) is a potent stimulator of EMT in several cell types, and the EGFR has been shown to directly interact with  $\beta$ -catenin, leading to the tyrosine phosphorylation of  $\beta$ -catenin and disruption of cadherin-dependent junctions (Klymkowsky, 2005; Nelson and Nusse, 2004). Endocytosis of E-cadherin results in the release of  $\beta$ -catenin to act on the Wnt pathway, resulting in Snail gene transcription and consequently E-cadherin repression (Lu et al., 2003). On the other hand, engaged E-cadherin complexes in the intact adherens junction directly inhibit the activity of the EGFR by inhibiting transphosphorylation of Tyr845 (Perrais et al., 2007).

Similarly, insulin growth factor (IGF) can also induce EMT by affecting the distribution of E-cadherin and internalization of E-cadherin in the vesicles located around the nucleus and degrading it. Similar to E-cadherin, IGF also redistributes  $\beta$ -catenin from the cell membrane to the nucleus, and induces the translocation of TCF3 from the cytoplasm to the nucleus (Morali et al., 2001). Thus, growth factors play a key role in initiating and maintaining EMT.

### **1.2.3 EMT and transcription factors**

It is well known fact that TGF- $\beta$  induces many transcription factors which are essential for EMT. Transcription factors like Snail1/2, Slug, Zeb1/2, Twist and FoxC2 are potent repressor of genes which are important for keeping epithelial polarity and organization, particularly cell-cell interaction proteins such as E-cadherin, claudin, occludins and ZO-family genes (Zhang et al., 2008). Snail expression can be induced in a Smad-3 or MAPK dependent manner (Matsumoto et al., 2001; Xue et al., 2008). Snail1 binds to E-box element of the E-cadherin promoter and recruits a complex containing HDAC1, HDAC2 and mSin3A, and thus represses gene transcription (Batlle et al., 2000; Cano et al., 2000; Hajra et al., 2002). Snail2 also works in the same way as Snail1 but it recruits a different combination of co-repressors, i.e. HDAC1/3 and CTBP (Bolos et al., 2003; Hajra et al., 2002; Hemavathy et al., 2000). However, Snail proteins do not affect the PAR complex, another polarity complex (Whiteman et al., 2008). Some of the Snail target genes regulate tissue specific EMT processes such as HNF-1 $\beta$  (Boutet et al., 2006; Boutet et al., 2007). Similarly, Zeb1 and Zeb2 [also known as Smad-interacting protein1 and 2 (SIP1 and 2)] also form repressive complexes with Smads and bind to the E-cadherin promoter to suppress it. The Zeb factors have been recently shown to be repressed by miRNA from miR-200 family. These miRNAs are downregulated during EMT and their forced expression is sufficient to block TGF- $\beta$ -induced EMT (Gregory et al., 2008a; Korpala et al., 2008; Park et al., 2008). In addition, Zeb2 is subjected to post-translation regulation by Polycomb complex 2 where sumoylation impairs its repressor activity (Long et al., 2005). Both Zeb proteins promote cell migration and induce invasion (Comijn et al., 2001; Spaderna et al., 2008; Vandewalle et al., 2005). Helix-loop-Helix (HLH) proteins are a large family of transcription factors controlling a wide variety of developmental and biological processes. HLH family can be divided into seven categories on the basis of their tissue distribution, dimerization capabilities and DNA binding specificity (Massari and Murre, 2000). E12 and E47 from class I, Twist1 and 2 from class II and Id1- 4 from class V had been elaborated in EMT. E12 and E47 are encoded by alternative splicing products of the E2A gene (Massari and Murre, 2000). They have been shown to repress E-cadherin by directly binding to its promoter. On the other hand, Ids are downregulated in response to TGF- $\beta$  and act as antagonist for E-cadherin repression by binding to E2A protein (Kondo et al., 2004). ID1 is usually suppressed during EMT by rapid activation of expression of the transcription repressor ATF3 by TGF- $\beta$  and the subsequent binding of an ATF3/Smad3/Smad4 complex to the Id1 promoter (Kang et al., 2003; Kowanetz et al., 2004).

Loss of Id1 expression is co-relates with a decrease in E-cadherin expression, and ectopic expression of Id2 or Id3 dose-dependently blocks TGF- $\beta$ -induced repression of E-cadherin expression, inhibits TGF- $\beta$ -induced ZO-1 delocalization and represses TGF- $\beta$ -induced smooth muscle actin expression (Kang et al., 2003; Kowanetz et al., 2004). Besides E-cadherin, E47 also represses desmoplakin expression and induces the expression of N-cadherin, Sparc and  $\alpha$ 5-integrin (Moreno-Bueno et al., 2006). Twist1 and 2, other members of this family, are a major regulator of mesoderm formation (Chen and Behringer, 1995). Ectopic expression of both leads to EMT (Ansieau et al., 2008; Yang et al., 2004). Hmga2, a downstream molecule of Smad-3 signaling, is also shown to be expressed in mesenchymal cells and induces the expression of Snail1/2, Twist, and represses Id2 expression (Thuault et al., 2006).

#### **1.2.4 EMT and metastasis**

Metastasis (from the Greek “change of place”) is a major cause of death among cancer patients. This process refers to the spread of cancer from its original site to other areas in the body. Cancer cells have the ability to invade the blood vessels and lymph-nodes and find their way into the bloodstream. Once in the blood, cancer cells can disseminate to virtually any part of the body and make a home for themselves. Metastasis is a multistage process. These stages have been defined as local invasion, intravasation into the circulation, survival and transport in the circulation, extravasation from the bloodstream, and growth in the metastatic site (Bogenrieder and Herlyn, 2003; Condeelis and Segall, 2003; Fidler and Balch, 1987; Fidler and Radinsky, 1996; Gopalkrishnan et al., 2001; Kauffman et al., 2003). Progression through these stages requires changes in cellular phenotype such as cellular motility, antiapoptotic capability, adhesion molecule expression, expression of matrix metalloproteinases and other proteases and expression of angiogenic factors and other paracrine or autocrine factors. Several authors have emphasized that if even one requisite step of the multistep metastatic process could be blocked, it would result in the abrogation of clinically relevant metastasis (Fidler and Balch, 1987; Fidler and Radinsky, 1996; Gopalkrishnan et al., 2001; Kauffman et al., 2003).

The tumor metastatic process has been compared to a marathon. Tumor cells have to invade the solid tissues around the primary tumor site. The tissue in which the tumor arose is complex, containing other cells such as fibroblasts, a protein filled matrix that provides a solid support and



immune cells and lymphatic drainage. Tumors have to invade past these barriers. To do so they develop the ability to move. Tumor cells do not float out of a tissue, they crawl. Basically, tumor cells react to factors in their environment; they put out a “finger” of the cell toward the attractant and ratchet the cell forward. To move, tumor cells must alter their adhesion to other cells and to the protein matrix in a very dynamic fashion. They may also have to create a pathway amongst the tissue, by degrading the protein matrix using enzymes (proteases).

Tumor cells can spread around the body using one of two major “highways”. All tissues are served by blood vessels (which provide oxygen and nutrients) and also lymphatic vessels which drain excess fluid to nearby lymph glands. For many cancer cells, their first opportunity to escape is to use the lymphatic drainage system. This is why for many cancers lymph nodes are biopsied or removed at surgery to see if the cancer has spread and oncologists use the information to determine the “stage” of the cancer. Cancer cells can enter the bloodstream either indirectly via the lymphatics or directly from a vessel in the primary tumor. The bloodstream is a very harsh environment with a high velocity of flow and full of immune cells. Moreover, cancer cells are used to being attached to the proteinaceous matrix, many tumor cells die when detached from their support and have to swim (detachment mediated death is called anoikis, another Greek word describing the death of leaves from as they detach from trees in Fall). The majority of tumor cells get stuck (arrest) in the first capillary bed that they float into. To get out from the bloodstream, they attach to the endothelial cells lining the blood vessels and the endothelial cells retract, they move apart, to permit the tumor cells to enter the tissue. This may be a normal reaction of endothelial cells to immune cells, cells of our immune systems migrate in and out of the bloodstream all the time to maintain surveillance. In fact, tumor cells can disguise themselves as lymphocytes by expressing similar molecules on their surface that fools the endothelial cells. These molecules may also determine their apparent ability to “home” to specific organs preferentially, as they may respond to gradients of chemicals differentially expressed there.

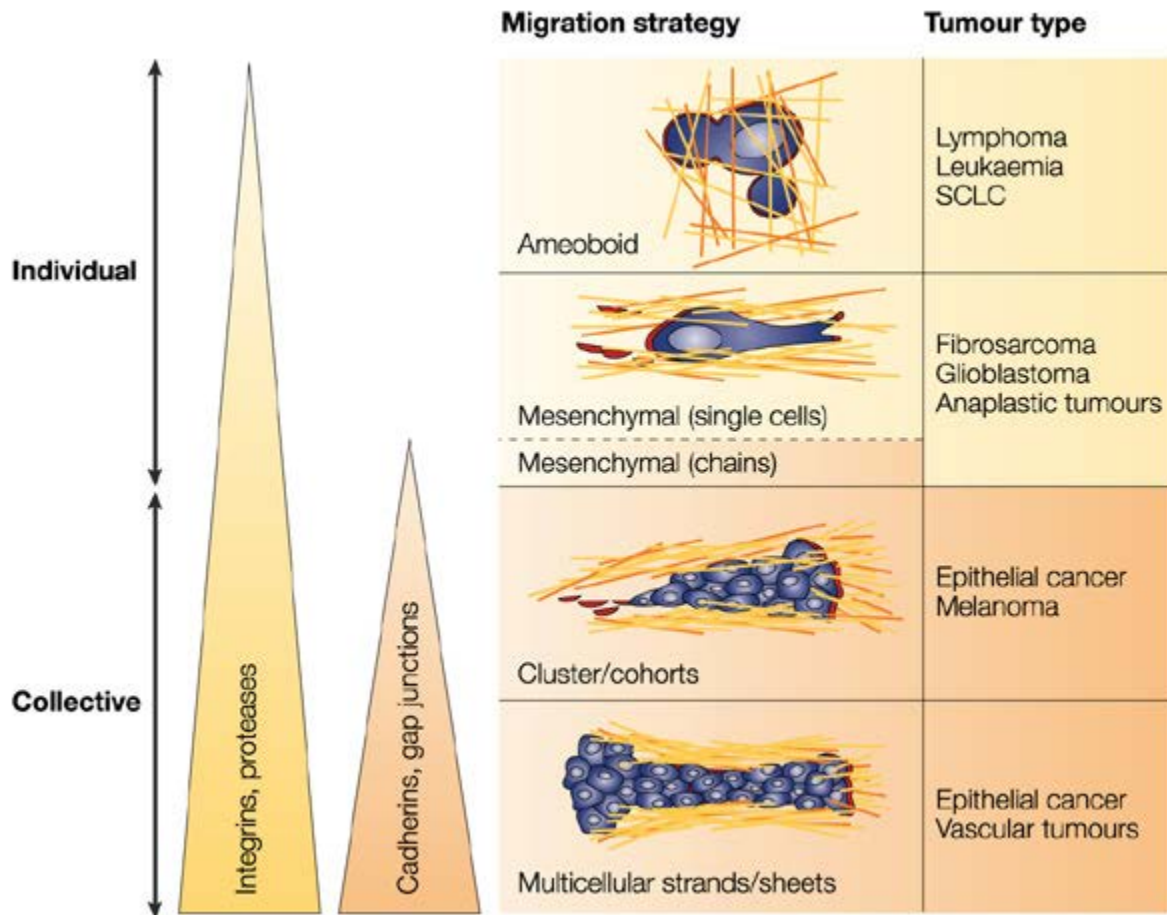
### **1.2.5 EMT and migration**

Migration is certainly a key step in metastasis and a universal process. Depending on the tumor type and the surrounding tissue, cell migration involves different cellular strategies to overcome the physical restraint provided by primary tissues and their epithelium. For most cells, including

epithelial, stromal, and neuronal cells, migration phases are confined to morphogenesis and cease with terminal differentiation toward intact tissue to become reactivated only for tissue regeneration or neoplastic processes. For other cell types, such as leukocytes, migration is integral to their function and maintained throughout their life span. Some cell types migrate only in the context of a defined substrate, such as epithelial cells moving along a basement membrane but not through interstitial tissues, whereas other cell types, including leukocytes, are versatile, as they interact with and migrate within virtually any substrate present in the body. These cell and tissue-type specific patterns of migration are acquired in various cancers can be subdivided into at least two main migration modes: single cell migration and collective cell migration (Figure 4). The essential molecules that control and specify these different types of migration include adhesion molecules of the  $\beta 1$  and  $\beta 3$  integrin families that mediate interaction with the extracellular matrix (ECM); matrix metalloproteinases (MMPs) and serine proteinases, such as uPA/uPAR, responsible for ECM degradation; cadherins and associated molecules that mediate stable intercellular adhesions; and signaling molecules that control the actin cytoskeleton, specifically the small GTPases RhoA, Rac and Cdc42 and their downstream effectors. Single cell migration is characterized by the presence of isolated and dispersed tumor cells in an adjacent tissue. This contrasts with what is observed during collective cell invasion, where the cancerous tissue pushes forward as a whole, thereby displacing the healthy surrounding cells (Friedl and Wolf, 2010; Ilina and Friedl, 2009; Matise et al., 2009; Pals et al., 2007; Yilmaz et al., 2007).

#### **1.2.5.1 Single cell migration**

Single cell migration can be further classified into two sub-groups: (a) Mesenchymal migration and (b) Amoeboid migration. Single cell mesenchymal migration has been identified in numerous cancers, including fibrosarcomas, glioblastomas and epithelial cancers. This kind of migration is also implanted in non-neoplastic neural crest cells (Jacques et al., 1998), myoblasts (El Fahime et al., 2000), infiltrating lobular or metaplastic breast carcinoma (Pitts et al., 1991), ovarian cancer (Sood et al., 2001) and melanoma of a vascular-type pattern (Seftor et al., 2002). However, single cell amoeboid migration is common in leukocytes and some tumor cells, such as in leukaemias, as well as in tumors occurring in organs lacking extensive cross-linked collagen network (Friedl and Wolf, 2003).



**Figure 4: Diversity of tumor cell migration.** Tumor cells can migrate individually or collectively. Ameoboid migration involves the deattachment of individual cells for invasion into the surrounding tissue but here the interaction between cells and ECM matrix is weak. The mesenchymal migration also follows similar strategy but the cells become spindle-shaped and possess integrins for invading the ECM. During collective cell migration, cells form a cohort of 5-6 cells. In cluster migration, cells loss their contact with primary tumor while in sheet migration, contact with primary tumor remains intact. Adapted from Peter Friedl & Katarina Wolf, 2003.

**(a) Mesenchymal Migration:** In mesenchymal migration, which represents the archetype of cell migration, cells complete a migration sequence consisting of (1) cell polarization driven by localized actin polymerization causing formation of a leading pseudopod; (2) attachment of this pseudopod to ECM ligands via  $\beta 1$  and  $\beta 3$  integrin clusters called focal adhesions, interaction sites that recruit cytoplasmic adaptor, signaling, and cytoskeletal

proteins as well as cell surface proteases such as MMPs and the uPA/uPAR complex; (3) local proteolysis of the ECM, widening the space for forward movement of the cell; (4) activation of contractile proteins, such as myosin II, and consequent shortening of membrane-anchored actin filaments; (5) contraction of the cell, leading to retraction of its rear end and consequent forward movement. This 5-step migration program is typical for single-cell migration of fibroblasts and keratinocytes as well as for single epithelial (cancer) cells that have undergone epithelial to mesenchymal transition (EMT), and represents a relatively slow process with migration velocities of 0.1 to 2  $\mu\text{m}/\text{min}$ . Furthermore, chain migration is also a part of mesenchymal migration where chains of single tumor cells aligned between stromal fibres, termed 'Indian files' and infiltrate the lobular or metastatic breast cancer, ovarian cancer and melanoma of vascular type-pattern.

**(b) Amoeboid Migration:** Interestingly, recent studies indicate that lymphocytes display a characteristic form of cell migration, which has been termed “amoeboid” migration, because it mimics that of the amoeba *Dictyostelium discoideum* (Friedl, 2004). In this migration type, integrin-mediated adhesion is partially dispensable and stable focal contacts are not formed, but cell movement is driven by short-lived relatively weak interactions with the stromal cell networks in the T- and B-cell areas of the lymphoid tissues (Bajenoff et al., 2006; Friedl, 2004; Mempel et al., 2006). The lack of focal contacts and high deformability of lymphocytes allow movement at high velocity (2-30  $\mu\text{m}/\text{min}$ ) (Bajenoff et al., 2006; Friedl, 2004; Mempel et al., 2006; Miller et al., 2003; Miller et al., 2002). Moreover, the fast deformability of lymphocytes allows them to overcome matrix barriers by physical mechanisms, that is, adaptation of shape to preformed matrix structures (contact guidance), extension of lateral footholds (elbowing), and squeezing through narrow spaces (contraction rings). Thus, lymphocyte migration is shape-change driven and lymphocytes use protease-independent physical mechanisms that allow easy cell traffic toward and between structurally different tissue compartments. Among higher eukaryotes, this migration type is found only in lymphocytes and other leukocytes, hematopoietic stem cells, and certain tumor cells (Friedl, 2004).

### 1.2.5.2 Collective cell migration

Collective movement, which relies on local proteolytic degradation of the extracellular matrix can occur by two-dimensional (a) sheet migration or three-dimensional group (Lauffenburger and Horwitz, 1996), or (b) strand/ cohort migration (Adams, 2001). However, collective migration of cells is a well-described phenomenon that occurs during embryological development (Davidson and Keller, 1999), or during the development of glands and ducts of mammary tissue (termed 'branching morphogenesis') (Klinowska et al., 1999; Simian et al., 2001) and the sprouting of endothelial cells during the formation of new blood vessels (Collen et al., 2003; Hiraoka et al., 1998).

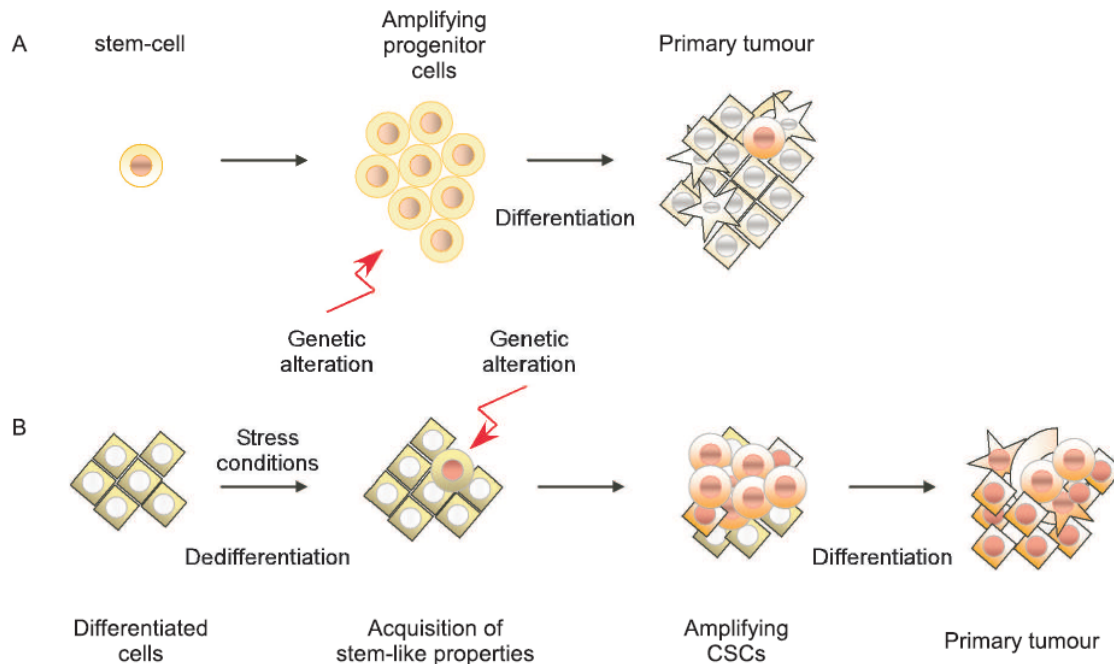
- (a) Multicellular strands/sheets migration:** This kind of migration consists of cell sheets that extend into tissue, yet maintain contact with the primary tumor to generate local invasion. Cells present at the migrating front produce MMPs that generate a path for the cells to follow. This type of movement is observed in invasive epithelial cancer such as oral squamous cell carcinoma, basal cell carcinoma, melanoma, mammary carcinoma and colon carcinoma (Nabeshima et al., 2000). The precise molecular mechanisms underlying this coordinated invasion are still largely unknown, mainly due to the difficulties to model this complex type of migration *in vitro*.
- (b) Cluster/ Cohort migration:** Collective cell migration, as seen during wound healing and during the invasion of epithelial cancer cell, uses the same integrin- and protease-dependent migration cycle as single-cell migration, but in this migration type the cell junctions within the invasive collectives are stabilized by cadherins and gap-junctional cell-to-cell communication (Pals et al., 2007). Cellular cohorts are heterogeneous in nature with leading, and presumably more motile, cells serving as a guide for following cells (path-generating cells) which basically generates migratory traction via pseudopod activity and reducing apico-basal polarity. Moreover, cells at the leading edge cluster  $\beta_1$  integrins in anterior protrusions towards the ECM substrate (Hegerfeldt et al., 2002; Klinowska et al., 1999), and show an increased expression and activity of MT1-MMP and MMP-2, leading to polarized ECM degradation (Nabeshima et al., 2000). The following cells maintain polarity and junctional proteins (Burrige and Chrzanowska-Wodnicka, 1996). However, communication between cells in a cohort is necessary to induce collective rearrangement of their cytoskeletal structures, thereby facilitating

collective migration in response to extrinsic guidance cues. Collective cell migration is morphologically expressed as either a branch-like extension protruding out of the primary tumor mass or as a separate cohort of cells moving through the tumor stroma (Friedl and Brocker, 2000). Collective cell migration *in vivo* may be advantageous for cell survival, such as robust protection of the innermost cells in the cluster from immunological attack. Additionally, collectively moving cells probably have enhanced cell–cell communication in response to micro-environmental cues, due to the heterogeneity of leading and following cells. This is beneficial for effective movement and invasion beyond the primary tumor, as the less motile cells are influenced by highly motile cells (BurrIDGE and Chrzanowska-Wodnicka, 1996; Friedl and Brocker, 2000).

### **1.2.6 EMT and cancer stem cells**

The cells within a tumor display functional heterogeneity, with different morphology, differentiation grade, proliferation rate, and invasiveness (Heppner and Miller, 1983). Recent studies suggest that the ability of a tumor to proliferate and propagate relies on a small population of stem-like cells, called cancer stem cells (CSCs). CSCs share fundamental characteristics with normal adult stem cells: they divide asymmetrically producing one stem cell and one progenitor cell (self-renewal). In normal stem cells, this allows the continuation of the stem cell compartment and starts the production of cells that undergo multi-lineage differentiation. Similarly, CSCs have the ability to perpetually self-renew and to produce tumors comprised of cells with different phenotypes. Since their discovery in leukaemia (Bonnet and Dick, 1997), the existence of a subpopulation of CSCs has been corroborated in several solid tumors, including breast, brain, colon, pancreas, prostate, lung, and head and neck tumors (Eramo et al., 2008; Glinsky, 2007; Li et al., 2007; Prince et al., 2007). Several of the molecular signaling pathways associated with normal stem cell development, such as Wnt, Shh and Notch, are also active in cancer development. Stem cells are notable for the presence of ATP-binding cassette transporters, which remove drugs from the cell (Dean et al., 2005). Cancer stem cells maintain this characteristic, which accounts for the ineffectiveness of chemotherapy to destroy these cells. Normal stem cells are under tight metabolic control and divide only under specific conditions. Cancer stem cells no longer have these controls. Furthermore, like cancer cells, they are also resistance to apoptosis.

While some studies suggest that CSC may arise from the transformation of their normal counterparts, recent observations rather suggest that they originate from fully differentiated cells through an adaptive trans-differentiation program such as EMT (Mani et al., 2008; Morel et al., 2008; Vesuna et al., 2009).



**Figure 5:** (A) The “cancer stem-cell theory” is based on the assumption that during tissue regeneration, the amplification of progenitor cells opens a window of time suitable for accumulating genetic alterations, leading to the emergence of cancer cell-stems (CSCs). CSCs would thus initiate and sustain tumour growth. (B) Alternatively, under stress conditions, fully differentiated cells reacquire stem-like properties, including self-renewal. This gain of function is influenced by cellular intrinsic properties as well as micro-environmental conditions. These cells could potentially be prone to transformation and give rise to CSCs. Both models are not exclusive. CSCs and cell dedifferentiation would thus constitute the initial and secondary tumour drivers, respectively.

Cells that have undergone an EMT were thus found to form mammospheres in low adherent conditions and to be highly tumorigenic when orthotopically xenografted at limit dilution in *nude* mice. They additionally display a  $CD44^{\text{high}} CD24^{\text{low}}$  antigenic phenotype that was previously allotted to mammary CSCs (Al-Hajj et al., 2003). Several lines of evidence exist where it has been shown that these cells are resistant to chemotherapy and do not undergo apoptosis. EMT

being by definition a reversible process, these cells continuously generate CD44<sup>low</sup> CD24<sup>high</sup> epithelial cells that interestingly lack a tumorigenic potential (Mani et al., 2008; Morel et al., 2008; Vesuna et al., 2009). In regards to the EMT-associated properties, the trans-differentiation process is thus considered as a biological process able to convert differentiated epithelial cells into CSCs. EMT being strongly impacted by micro environmental conditions, the balance between differentiated cells and CSCs was then proposed to be a highly dynamic process with important repercussions on therapeutic approaches, eradication of the entire primary tumor, including differentiated cells, being henceforth a requisite to prevent recurrence (Gupta et al., 2009a).

The key roles of CSCs in breast cancer biology suggest that new therapies must target these cells. The main objective of those therapies would be the eradication of the CSC compartment with no harm to other cell types. Eradication of breast CSCs may include different strategies as summarized in **Table 1**.

Different approaches have been used to overcome ABC transporter-mediated chemoresistance. The anthracycline-modified drug annamycin, which is not extruded by ABC transporters, is toxic to the resistant cell line MCF-7/VP (Perez-Soler et al., 1997). The plant alkaloid berberine decreases the expression of the ABCG2 transporter and reduces the “side population” of the MCF-7 cell line (Kim et al., 2008a; Kim et al., 2008b), suggesting that downregulation of ABC transporters may be useful for targeting breast CSCs. However, the ability to target drug transport in CSCs may be difficult since these cells express multiple ABC transporters (de Grouw et al., 2006). The use of inhibitors of ABC transporters simultaneously with anticancer drugs is an efficient approach to overcome resistance *in vitro* and in animal models (Ozben, 2006). However, clinical trials with this kind of inhibitors have shown that they produce serious side effects (Ozben, 2006). High-throughput screening identified the ionophore salinomycin as toxic to breast CSCs (Gupta et al., 2009b). Salinomycin induces caspase-independent apoptosis in human cancer cells of different origins that display multiple mechanisms of drug resistance, at concentrations that do not affect normal cell viability (Fuchs et al., 2009). Subsequent studies have shown that salinomycin induces a conformational change of the ABC transporter MDR1/ABCB1 that reduces its activity (Riccioni et al., 2010). Therefore, salinomycin is particularly effective at inducing apoptosis in leukemia cells that display ABC transporter-mediated drug-resistance (Fuchs et al., 2010). Targeting CSCs through their specific markers



was partially successful in acute myeloid leukemia (AML) (Sperr et al., 2005; Tsimberidou et al., 2006). Cytotoxic antibodies directed against CD33 (a common marker in leukemic stem cells) induced remission in some patients. However, the antibody produced cytopenia due to its effects on normal hematopoietic stem cells (Sperr et al., 2005; Tsimberidou et al., 2006). Similarly, a monoclonal antibody against CD44 induced terminal differentiation and apoptosis of AML cells in engrafted mice (Jin et al., 2006). Anti-CD44 antibodies conjugated with cytotoxic drugs or radiolabels have shown to reduce disease progression in breast cancer patients and animal models (Platt and Szoka, 2008). Other potential targets in breast CSC therapy include molecules that participate in self-renewal and cell fate. Inhibition of Hedgehog signaling in xenografts established from pancreatic cancer cell lines reduced the number of ALDH-overexpressing cells (Feldmann et al., 2008). The promoters of the MDR, hTERT, and Cox-2 genes are active in breast CSCs. Oncolytic adenoviruses driven by these promoters were effective in killing CD44+/CD24-/low cells *in vitro*, and reducing tumor growth *in vivo* (Bauerschmitz et al., 2008). Interruption of signals generated in the CSC microenvironment using antibodies or soluble ligands against adhesion receptors may be useful in CSC targeting.  $\alpha$ 6-integrin inactivation with antibodies or siRNA abrogated mammosphere-forming ability and tumorigenicity of breast cancer cells (Cariati et al., 2008). The IL-8 receptor CXCR1 inhibitor repertaxin reduced the breast CSC population, producing apoptosis in the tumor population, and reduced metastasis (Ginestier et al., 2010).

Metformin is an anti-diabetic drug that has found to reduce breast cancer incidence and improve survival of breast cancer patients with type 2 diabetics (Vazquez-Martin et al., 2010). Recent studies showed that the drug metformin selectively reduces the breast CSC population. In human breast cancer cell lines, metformin reduced the CD44+/CD24- population and their ability to form mammospheres (Hirsch et al., 2009). In a xenograft mice model, concurrent treatment with metformin and doxorubicin reduced tumor mass much more effectively than either drug alone (Hirsch et al., 2009). Metformin also targeted trastuzumab-resistant CSCs that overexpressed HER-2 (Vazquez-Martin et al., 2011). The mechanism involved in the metformin effects on CSCs is unclear, but seems to be associated with its activator effect on AMP-activated kinase (AMPK) (Vazquez-Martin et al., 2010). AMPK phosphorylates and inhibits Acetyl CoA carboxylase (ACACA), the limiting enzyme of the fatty acid synthesis. Thus, metformin may be

affecting cancer cell metabolism and functioning on lipid raft platforms (Vazquez-Martin et al., 2010).

**Table 1: Strategies for the eradication of CSCs.**

Target in breast CSCs	Strategy	Example
ABC transporters	Cytotoxic drugs that cannot be extruded by ABC transporters	Annamycin
	Reduce expression	Berberine siRNAs
	ABC transporters inhibitors	Salinomycin
Membrane markers	Antibodies conjugated with drugs or radioligands	Anti-CD44
Intracellular signalling molecules	Small molecule inhibitors	---
	Reduce expression	siRNAs
	Oncolytic virus activated by specific promoters	MDR promoter
Signals from the microenvironment	Small molecule receptor antagonists	Repertaxin
	Blocking antibodies	Anti- $\alpha 6$ integrin
	Blocking soluble ligands	Soluble HA
Others	Metabolic alteration?	Metformin

### 1.2.7 EMT and MicroRNAs

miRNAs are highly conserved, small 17–25-nucleotide non-coding RNA molecules which are able to control gene expression at the post-transcriptional level by specifically interacting with a target mRNA. It is estimated that miRNAs regulate 30% of all proteins in humans (Lewis et al., 2005). They play a pivotal role in regulation of key processes including cell differentiation, proliferation, apoptosis, angiogenesis, and the cell cycle (Esquela-Kerscher and Slack, 2006; Johnson et al., 2007; Wang and Olson, 2009).

Beyond their roles in physiological processes, many miRNAs have been shown to be aberrantly expressed in various pathologies including cancer (Calin and Croce, 2006; Cho, 2010a; Cho, 2010b; Shenouda and Alahari, 2009) and usually present in the chromosomal regions which are prone to deletion, amplification, or translocation, e.g., during the development of tumors (Calin et al., 2004). This has led to the identification of “miRNA signatures” that are characteristic for

certain tumors and allow their further classification (Calin and Croce, 2006; Lu et al., 2005). miRNAs can act as oncogenes or tumor-suppressors and influence the tumorigenesis process when down or upregulated, respectively. They can also affect the sensitivity of tumor cells to cytostatics or radiotherapy (Trang et al., 2008; Weidhaas et al., 2007). For many miRNAs, target genes have been identified which are relevant in tumorigenesis, tumor growth, tumor angiogenesis and metastasis such as p53 (Le et al., 2009), p63

**Table 2: miRNAs relevant in tumor cell invasion and metastasis (Adapted from A. Aigner, 2011)**

miRNA	Impact on metastatic processes	Upstream regulators	Target genes	Tumor types (misregulation/relevance in metastasis)
mir-9	suppressing Promoting	MYC	E-cadherin	Breast, brain, hepatocellular
miR-10a	Promoting	Retinoid acid	HOXB1, HOXB3	Pancreatic
miR-17-92	Promoting	Myc	CTGF, TSP-1	Leukemia, breast, colon
miR-21	Promoting	AP-1, BMP2, BMP4, BMP6, EGF/HER2, TGF-beta, ZEB1	ARCKS, HNRPK, NF1B, PDCD4, PTEN, TPM1, MARCKS, Maspin, RECK, SPRY2, TA $\rho$ 63, TIMP3, TPM1	Breast, bladder, colorectal, gastric, lung, ovarian, cholangiocarcinoma, pancreatic, prostate, hepatocellular, glioblastoma, uterine leiomyoma, cervical, B cell lymphoma, lymphocytic leukemia, esophageal
miR-27a	Promoting			Gastric
miR-27b	Promoting		ST14	Breast
miR-29a	Promoting		Tristetraprolin	Breast
miR-30d	Promoting		GNAI2	Hepatocellular
miR-107	Promoting		DRCE1	Gastric
miR-130a	Promoting		GAX, HOXA5	
miR-143	Promoting	NF- $\kappa$ B	FNDC3B	Hepatocellular
miR-151-5p	Promoting		RhoGDIA	Hepatocellular
miR-155	Promoting	Smad4	RhoA	Breast
miR-182	Promoting		FOXO3, Mitf	Melanoma
miR-210	Promoting			Breast
miR-224	Promoting		CD40	Pancreatic
miR-373	Promoting	BRMS1	CD44	Breast
miR-378	Promoting		Fus-1, SuFu	Breast, glioblastoma
miR-486	Promoting		CD40	Pancreatic
miR-520c	Promoting	BRMS1	CD44	Breast
miR-520h	Promoting	E1A	PP2A/C	Breast
miR-532	Promoting		RUNX3	Melanoma

miRNA	Impact on metastatic processes	Upstream regulators	Target genes	Tumor types (misregulation/relevance in metastasis)
miR-7	Suppressing	HoxD10	Pak1	Breast, urothelial
miR-9	Suppressing		TrkC	Colorectal
miR-16	Suppressing		CDK1, CDK2	Prostate
miR-17/20	Suppressing		IL8, Cyclin D1	Breast
miR-22	Suppressing			Ovarian
miR-23	Suppressing		c-met, uPA	Hepatocellular
miR-29b	Suppressing	MBP-1	Mcl-1, COL1A1, COL4A1, MMP-2	Prostate
miR-29c	Suppressing		Laminin gamma1, collagens	Nasopharyngeal
miR-30	Suppressing			Breast
miR-31	Suppressing		ITGA5, RDX, RhoA	Breast
miR-34a	Suppressing		c-met	Hepatocellular
miR-106b	Suppressing			Renal
miR-122	Suppressing	HNF1A, HNF3A, HNF3B	ADAM17	Hepatocellular
miR-125b	Suppressing		LIN28B	Hepatocellular
miR-126	Suppressing		Crk	Breast, lung, gastric
miR-130b	Suppressing	TAp63		HNSCC
miR-138	Suppressing			HNSCC
miR-139	Suppressing		ROCK2	Hepatocellular
miR-145	Suppressing		c-myc, mucin1	
miR-146ab	Suppressing	BRMS1	EGFR, IRAK-1, MMP16, NF-κB	Breast, glioma, pancreatic
miR-183	Suppressing		Ezrin	Breast, lung
miR-193b	Suppressing		uPA	Breast
miR-194	Suppressing		N-cadherin, DNMT3A, HBEGF, IGF1R, PTPN12, PTPN13, ITGA9, SOCS2, RAC1	Liver
miR-204	Suppressing		APRC1B, CTSC, FAP, MMPs, BMP1, CDH11, ITGB4	HNSCC
miR-205	Suppressing		ZEB1, ZEB2, LRP1	Bladder, breast, glioblastoma, lung
miR-206	Suppressing		NOTCH3	Breast, lung
miR-211	Suppressing	MITF	KCNMA1	Melanoma
miR-218	Suppressing		Robo1	Gastric
miR-221	Suppressing		c-kit	Prostate
miR-222	Suppressing		MMP1, SOD2	Tongue SCC
miR-335	Suppressing	BRMS1	MERTK, PTPRN2, SOX4, TNC	Breast
miR-338	Suppressing			Hepatocellular
miR-339-5p	Suppressing			Breast
let-7	Suppressing	LIN28, MYC	HMGA2, MYC, NOTCH, RAS, COL1A2	Breast, lung, pancreatic, hepatocellular
miR-10b	Promoting/suppressing	BRMS1, Twist	HoxD10, KLF4, RhoC, Tiam1	Breast, hepatocellular, ovarian, AML, esophageal, nasopharyngeal, SCC
miR-125a-3p/5p	Promoting/suppressing	EGFR	ARID3B	Lung, ovarian
miR-141/miR-200a,b,c/miR-429	Promoting/suppressing	ZEB1, ZEB2	ZEB1, ZEB2, E-cadherin, N-cadherin	Breast, lung, pancreatic
miR-661	Promoting/suppressing	c/EBPalpha	MTA1, Nectin-1, StarD10	Breast

(Manni et al., 2009; Papagiannakopoulos et al., 2008; Scheel et al., 2009), Hmga2, Myc and Ras (Bussing et al., 2008) as well many of them are regulated by upstream regulators such as EGFR (Wang and Olson, 2009), Myc (Chang et al., 2008; Ma et al., 2010; Sampson et al., 2007) and NF $\kappa$ b (Zhang et al., 2009) which are also shown to affect tumorigenicity (**Table 2**). Furthermore, some of these molecules are involved in processing of miRNA by regulating the major components of the miRNA processing machinery including Drosha-DGCR8, Dicer-TRBP2, and Ago proteins (Boominathan, 2010). microRNA-like miR-31 has been shown to involved in regulation of three different steps of metastasis, i.e., invasion, extravasation, survival, and colonization by targeting three different molecules, integrin- $\alpha$ 5 (ITGA5), radixin (RDX), and RhoA (Valastyan et al., 2010). Moreover, in some cases, auto-regulatory feedback loops have been observed, like for let-7 which inhibits metastasis by downregulating Myc, with Myc in turn transcriptionally and post-transcriptionally inhibiting let-7. A similar double-negative feedback loop is also observed for the miR-200 family, with ZEB1 and ZEB2 being a target of miR-200 which, at the same time, are transcriptional repressors of both miR-200 gene clusters (Burk et al., 2008; Wellner et al., 2009). Since ZEB1/ZEB2 are relevant in the EMT transition, this directly affects the epithelial vs. mesenchymal status of tumor cells and thus their migratory potential. However, other studies have demonstrated a miR-200-mediated increase in metastasis of mammary carcinoma cell lines (Dykxhoorn et al., 2009). An approach to reconcile those seemingly contradicting findings focuses on the different steps during metastasis. E-cadherin, a key molecule in EMT, is also regulated by several other miRNAs including miR-101 via Ezh2 (Varambally et al., 2008) or miR-9 (Ma et al., 2010). Together, these data underscore the importance of miRNA in EMT and carcinogenesis.

### **1.2.8 EMT and splicing factors**

Splicing aberrations have been associated with several diseases, including cancer, where altered splicing can lead to production of protein isoforms with oncogenic properties (Pajares et al., 2007). 41 breast cancer-specific markers have been identified that can discriminate between normal breast tissue and ductal breast tumors (Venables et al., 2008). Furthermore, luminal and basal B cell lines could be distinguished based solely on their splicing patterns. Therefore, it appears likely that alternative splicing analysis will dramatically increase the pool of potential biomarkers for cancer diagnostics.

Since EMT is considered an early event in the metastatic process, splicing changes associated with EMT in particular have the potential to become useful prognostic and diagnostic markers for breast cancer metastasis. Fibroblast growth factor-2 (FGFR2) was the first gene to be associated with splicing and EMT (Savagner et al., 1994). Alternative splicing of the third Ig-like domain determines the ligand-binding specificity of the receptor and generates the IIIb or the IIIc isoform of the FGFRs. The IIIb isoforms are usually expressed in epithelial cells, whereas the IIIc isoforms are normally expressed in mesenchymal cells. In contrast, the ligands for the IIIb isoforms are usually expressed in mesenchymal cells and the ligands for the mesenchymal-restricted IIIc isoforms in epithelial cells. In this way, FGFR signaling functions in a paracrine manner between the epithelial and mesenchymal cells. A switch from one isoform to another can thus lead to autocrine signaling. Exon switching in epithelial cells from the epithelial FGFR2 IIIb isoform to the mesenchymal FGFR2 IIIc isoform by alternative splicing has been described in rat models of prostate and bladder cancer (Oltean et al., 2006; Savagner et al., 1994). After the discovery of FGFR2, many other genes such as ENAH1 (also known as Mena) and CTNND1 (also known as p120) are also shown to be regulated by alternative splicing. ENAH1 belongs to Enabled/vasodilator-stimulated phosphoprotein (Ena/VASP) family and regulates actin organization. Human Mena (hMena) is overexpressed in human breast tumors, and a splice variant termed hMena<sup>+11a</sup> was recently isolated from a breast cancer cell line with an epithelial phenotype (Pino et al., 2008). Similar to ENAH1, p120 can be spliced into isoform 1 and 3. Isoform 3 is expressed in epithelial cells while isoform 1 is expressed in mesenchymal cells (Keirsebilck et al., 1998). Cell surface marker CD44 also undergoes splicing events and leads to the formation of CD44<sup>s</sup> and CD44<sup>v</sup> isoforms. The CD44<sup>s</sup> form is prominent in mesenchymal cells while CD44<sup>v</sup> in epithelial cells (Brown et al., 2011). Two recently described RNA binding proteins, ESRP1 and ESRP2 (epithelial splicing regulatory protein 1 and 2) have taken a central role in controlling splicing during EMT (Warzecha et al., 2010; Warzecha et al., 2009). In addition to the tight transcriptional regulatory control of EMT inducers, ESRP1 and ESRP2 exert an additional control to maintain epithelial homeostasis by promoting the splicing of epithelial-specific forms of EMT-associated genes including MENA, CD44, FGFR2 and CTNND1 (Warzecha et al., 2010; Warzecha et al., 2009). The RBFOX2 splicing factor has recently been demonstrated to regulate subtype-specific splicing in a panel of breast cancer cell lines (Lapuk et al., 2010). Shapiro and colleagues recently showed that EMT-associated splicing is likely to be

regulated by several splicing factors, including the ESRPs and members of the RBFOX, CELF, MBNL and hnRNP classes of splicing factors (Shapiro et al., 2011). These findings suggest that alternative splicing is an additional mechanism to control epithelial plasticity.

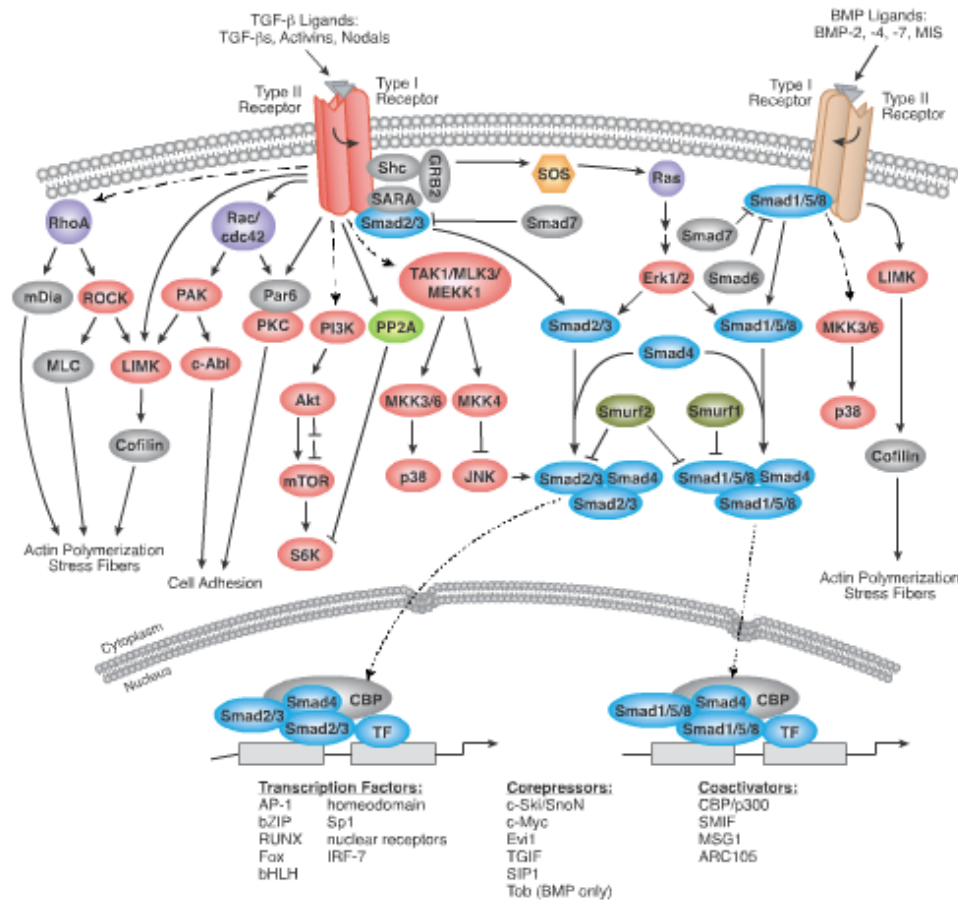
### **1.3 TGF- $\beta$ : its role in tissue homeostasis and cancer cell invasion**

#### **1.3.1 Dual role of TGF- $\beta$**

TGF- $\beta$  plays a central role in various biological processes including development, tissue homeostasis, immune system and cancer. In normal and premalignant cells, TGF- $\beta$  enforces homeostasis and suppresses tumor progression directly through cell-autonomous tumor-suppressive effects (cytostasis, differentiation, apoptosis) or indirectly through effects on the stroma (suppression of inflammation and stroma-derived mitogens) (Derynck et al., 2001; Massague, 2008; Siegel and Massague, 2003). However, when cancer cells lose TGF- $\beta$  tumor-suppressive responses, they can use TGF- $\beta$  to their advantage to initiate immune evasion, growth factor production, differentiation into an invasive phenotype, and metastatic dissemination or to establish and expand metastatic colonies (Pardali and Moustakas, 2007; Thiery, 2002; Yang and Weinberg, 2008).

#### **1.3.2 TGF- $\beta$ signaling**

The human TGF- $\beta$  family can be sub-divided into two groups. Activin, nodal, lefty, myostatin and TGF- $\beta$  belong to one group while bone morphogenetic proteins (BMPs), anti-muellerian hormone (AMH), and various growth factors and differentiated factors (GDFs) are belonging to the other. All these cytokines are well known to regulate various biological processes like cell proliferation, apoptosis, differentiation, angiogenesis and migration. Each ligand presents unique features of action, while they all share a common machinery to transmit intracellular signals, the TGF- $\beta$  receptor complex. TGF- $\beta$  can elicit its signaling either by using a canonical signaling pathway where it interacts with Smad proteins or via a non-canonical signaling where it interacts with non-smad proteins (**Figure 6**). These two modes of regulation result in immense complexity and variability of TGF- $\beta$  signaling and its ability to control various cellular processes.



**Figure 6:** Schematic diagram of TGF- $\beta$  signaling from Cell membrane to the nucleus

### 1.3.3. Canonical TGF- $\beta$ signaling

There are three variants of TGF- $\beta$ : TGF $\beta$ 1, TGF $\beta$ 2 and TGF $\beta$ 3 and all of them are synthesized as precursor molecules containing a propeptide region in addition to the TGF- $\beta$  homodimer (Ulloa and Tabibzadeh, 2001). After it is synthesized, the TGF- $\beta$  homodimer interacts with a Latency Associated Peptide (LAP) [a protein derived from the N-terminal region of the TGF- $\beta$  gene product] forming a complex called Small Latent Complex (SLC). This complex remains in the cell until it is bound by another protein called Latent TGF- $\beta$ -Binding Protein (LTBP), forming a larger complex called Large Latent Complex (LLC). It is LLC that gets secreted to the ECM (Laping et al., 2002). In most cases, before the LLC is secreted, the TGF- $\beta$  precursor is cleaved from the propeptide but remains attached to it by noncovalent bonds (Blobe et al., 2001). After its secretion, it remains in the extracellular matrix as an inactivated complex containing both the



LTBP and the LAP which needs to be further processed in order to release active TGF- $\beta$ . The attachment of TGF- $\beta$  to the LTBP is by disulfide bonds which allow it to remain inactive by preventing it from binding to its receptors. Because different cellular mechanisms require distinct levels of TGF- $\beta$  signaling, the inactive complex of this cytokine gives an opportunity for a proper mediation of TGF- $\beta$  signaling.

Active TGF- $\beta$  binds to two related receptors, namely TGF $\beta$ RI and TGF $\beta$ RII [(also known as activin receptor-like kinase 5 (ALK5)], which are serine/threonine kinases. Upon TGF- $\beta$  binding, TGF $\beta$ RI and TGF $\beta$ RII form hetero-tetrameric complexes of two identical TGF $\beta$ RI/TGF $\beta$ RII receptors heterodimers. This binding causes the specific phosphorylation of serine and threonine residues of TGF $\beta$ RI via TGF $\beta$ RII, which further leads to rotation of the receptors so that their cytoplasmic kinase domains are arranged in a catalytically favorable orientation. Furthermore, TGF $\beta$ RI phosphorylates effector proteins such as the receptor-associated Smad proteins (RSmads). Activated R-Smads subsequently interact with the common mediator Smad 4, translocate into the nucleus and control the transcriptional expression of various genes (Itoh et al., 2001). In addition, canonical TGF- $\beta$  signaling is tightly regulated by specialized factors such as Smad6 and Smad7 which control the signaling by competing with Smad4 for binding to receptor-activated Smad1 or via recruiting the Smurf to TGF- $\beta$  and BMP receptors for inactivation, respectively (Massague, 2008).

#### **1.3.3.1. Canonical TGF- $\beta$ signaling mediated cell-cycle arrest**

In epithelial cells, TGF- $\beta$  infers its cytostatic role via mobilization of cyclin-dependent kinase inhibitors (CDKI) such as p21CIP1 that targets cyclinE/A-cdk2 complexes for inhibition, or p15INK4B which inhibits cyclinD-cdk4/6 complexes. TGF- $\beta$  may also suppress c-Myc and it is demonstrated that coordinated activity of Smad3/4, retinoblastoma-like 1 (p107) and the transcription factors E2F4/5 and C/EBP $\beta$  may downregulate c-Myc (Chen et al., 2002; Gomis et al., 2006b).

#### **1.3.3.2. Canonical TGF- $\beta$ signaling mediated apoptosis**

Many apoptotic genes such as GADD45 $\beta$ ; Bim (Bcl-2 homolog domain-only factor); DAPK (death-associated protein kinase) and SHIP are regulated by TGF- $\beta$ . All these factors, except SHIP are controlled by canonical mitochondrial cytochrome C release and consequent activation

of induction of caspase-mediated apoptosis. The lipid phosphatase SHIP promotes apoptosis by inhibiting PI3K activity, thereby blocking its survival promoting signaling (Pardali and Moustakas, 2007).

#### **1.3.4. Non-canonical TGF- $\beta$ signaling**

The TGF- $\beta$  receptor complex interacts with various non-Smad signaling proteins, including receptor tyrosine kinases (RTKs), cytoplasmic adapter proteins and G-protein-coupled receptors. Each of these interaction partners can be activated upon TGF- $\beta$  binding to the TGF- $\beta$  receptor complex and emit signals which can act independently or can interfere with canonical TGF- $\beta$  signaling. Among well-known signaling pathways which are triggered by non-canonical TGF- $\beta$  signaling are the MAPK, PI3K and Rho-like GTPase signaling pathways (Zhang, 2009).

##### **1.3.4.1. TGF- $\beta$ -induced MAPK activation**

TGF- $\beta$  can induce phosphorylation of tyrosine residues on both type I and type II receptors. The phosphorylated tyrosines are capable of recruiting the adapter proteins such as growth factor receptor binding protein 2 (Grb2) and Src homology domain 2 containing protein (Shc) which can further activate Erk through Ras, Raf and their downstream MAPK cascades. Erk activation is one of the non-smad pathways necessary for TGF- $\beta$  mediated EMT (Davies et al., 2005; Zavadil et al., 2001). It is required for disassembly of cell adherens junctions, modulation of cell-matrix interactions, increased motility and endocytosis (Zavadil et al., 2001).

##### **1.3.4.2. TGF- $\beta$ -induced JNK/p38 activation**

Another interaction partner of the activated TGF- $\beta$ -receptor complex is the adapter protein and E3 ligase TRAF6. Interaction of TRAF6 to the activated TGF- $\beta$  receptors can induce the poly-ubiquitination of TRAF6. Poly-ubiquitinated TRAF6 recruits TGF- $\beta$  activated kinase 1 (TAK1) to activate the c-Jun N-terminal kinase (JNK) and MAPK p38 through induction of MAPKs MKK4 and MKK3/6. The TRAF6-TAK1-JNK/p38 pathway is known to conjugate with canonical TGF- $\beta$  signaling to induce TGF- $\beta$  mediated apoptosis (Liao et al., 2001). Downregulation of TRAF6 and p38 prevent EMT and impair the EMT associated cytoskeleton re-modeling (Bakin et al., 2002; Yamashita et al., 2008; Yu et al., 2002).

### **1.3.4.3. Rho-like GTPases in TGF- $\beta$ mediated EMT**

Par6, a scaffold protein regulating epithelial cell polarity, interacts with TGF $\beta$ RI at tight junctions. TGF- $\beta$  stimulation induces the assembly and accumulation of the TGF- $\beta$  receptor complexes at tight junctions, where TGF $\beta$ RII phosphorylates Par6. Upon phosphorylation, Par6 recruits the E3 ubiquitin ligase Smurf1 to the activated receptor complex and mediates ubiquitination and turnover of RhoA, which finally enables TGF- $\beta$ -dependent dissolution of tight junctions, a prerequisite for EMT (Ozdamar et al., 2005). TGF- $\beta$  also induces the dissolution of tight junctions during EMT by recruiting Cdc42 to the receptor complex, and by triggering degradation of RhoA at cellular protrusions (Barrios-Rodiles et al., 2005; Wilkes et al., 2003).

### **1.3.4.4. TGF- $\beta$ -induced PI3K activation**

TGF $\beta$ RII was found to be constitutively associated with p85, the regulatory subunit of PI3K. Upon TGF- $\beta$  binding of the TGF- $\beta$  receptor complex, PI3K becomes activated leading the activation of mammalian target of rapamycin (mTOR). mTOR is a key regulator of protein synthesis via phosphorylation of S6 kinase (S6K) and eukaryotic initiation factor 4E-binding protein (4E-BP1). The activation of S6K and 4E-BP1 by mTOR enhances translational capacity and protein synthesis, which is important for TGF- $\beta$ -induced EMT (Lamouille and Derynck, 2007). Furthermore, activation of PI3K protects cells from TGF- $\beta$ -induced apoptosis and growth inhibition by regulating canonical TGF- $\beta$  signaling or via inhibition of the transcription factor FoxO, which is essentially involved in TGF- $\beta$  mediated cell-cycle arrest (Conery et al., 2004; Gomis et al., 2006a; Remy et al., 2004; Seoane et al., 2004).

## **1.4 Epigenetic regulation of gene transcription**

Transcriptional regulation in higher eukaryotes occurs in the context of a chromatinized DNA template. While genetic information provides the basic information to encode cellular contents, the epigenetic information defines how, when and where this information has to be used. While histones were historically assumed to be simply a DNA packaging material, recent years have witnessed several molecular pathways that modify histones and DNA in a dynamic manner. Indeed several recent studies established that promoter sequences in eukaryotic genomes show characteristic patterns of histone modifications that associate with active or inactive state of the linked gene (Suganuma and Workman, 2011). Such modifications of DNA that are heritable, but

do not involve changes in the DNA sequence are referred to as ‘epigenetic’ modifications. Several studies conclude that one of the main functions of epigenetic modifications is to regulate the accessibility of DNA for regulatory factors (Suganuma and Workman, 2011).

#### **1.4.1 Epigenetic modifications of chromatin**

The fundamental building blocks of chromatin are the nucleosomes, which consist of globular histone protein cores around which the DNA is wrapped 1.65 times, corresponding to 147bp of DNA. These histone cores are composed of an octamer with two copies of each of the four highly conserved canonical histones H2A, H2B, H3 and H4 (Luger et al., 1997). The individual histones have a positively charged globular histone-fold domain which binds to the negatively charged DNA via electrostatic interactions. The unstructured C- and N-terminal histone tails, parts of which protrude outside of the nucleosomes, are subject to various post-translational modifications. These modifications affect the affinity of histones to DNA, the interaction with neighboring nucleosomes, chromatin higher order packaging and/or the recruitment of chromatin binding proteins. If not further compacted, the nucleosomes are arranged in a linear fashion along the DNA molecule, leading to a so-called “beads-on-a-string” structure with 10-60 bp of “free” linker DNA between the individual nucleosomes. This form of chromatin is generally found at active genes. Upon incorporation of histone H1 the chromatin structure can be further compacted. H1 binds to the linker DNA between the nucleosomes and leads to a transcriptionally inert higher order structure termed 30nm fiber (Robinson and Rhodes, 2006).

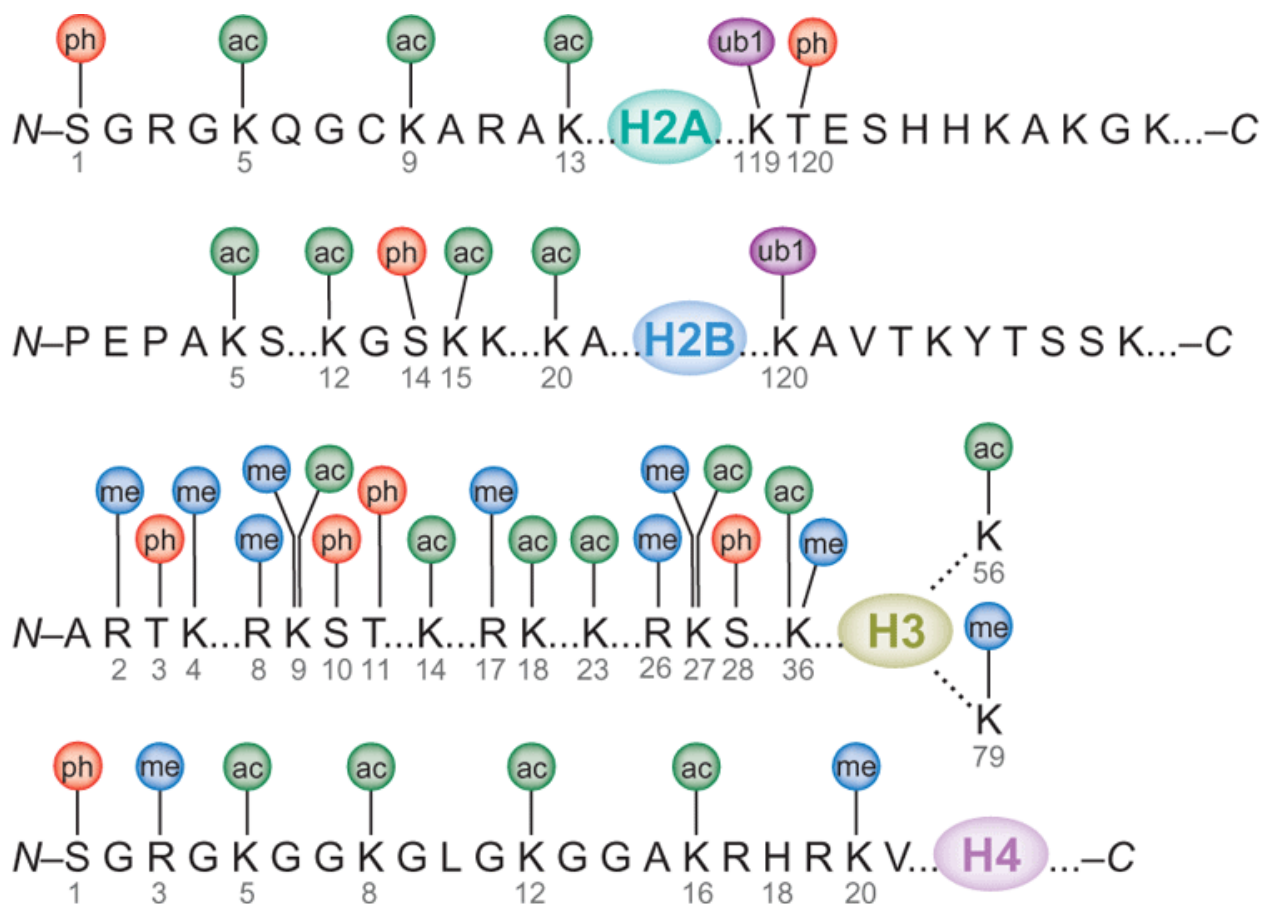
Histones contain over 60 residues under post-translational modification (PTMs) such as acetylation, methylation, ubiquitination, phosphorylation, sumoylation and others (Kouzarides, 2007). The N-terminal tails of histone H3 and H4 are the major substrates for such modifications as these protrude out of the nucleosomes cores. This further allows PTMs to be accessible to non-histone proteins that recognize these modifications and are recruited to chromatin. A second possible regulatory role of PTMs is direct interference with binding of DNA around nucleosomes via altering the electric charge of histones.

Historically, chromatin modifications and the overall chromatin structure were thought to be stable and simply providing a structural scaffold. The discovery of the yeast transcriptional coactivator Gcn5 to mediate histone acetylation (Brownell et al., 1996) and of the corepressor Rpd3 to mediate histone deacetylation (Taunton et al., 1996) provided the one of the first

evidences directly linking transcriptional regulation to histone modifications. Soon, Suv39h, a previously identified transcriptional regulator and suppressor of variegation in *Drosophila* (Tschiersch et al., 1994), and its yeast homolog Clr4 were identified as methylases specific for lysine 9 at histone H3 [H3K9; (Rea et al., 2000)]. These findings led to a new era of research focusing on chromatin-mediated transcription regulatory mechanisms. Importantly, the recent years witnessed the discovery of histone demethylases, which argued for a dynamic nature of PTMs (Shi et al., 2004). Some of the well-studied PTMs are discussed below.

**1.4.1.1 Histone Acetylation:** Lysines contribute largely to the overall positive charge of histone octamers and thus are crucial for binding the negatively charged phosphate backbone of DNA. The histone-DNA interactions need to be modified in order to facilitate DNA-template processes such as transcription, replication and DNA repair. All four core histones bear lysine residues which are subject to acetylation (**Figure 7**) and deacetylation by specific histone acetyltransferases (HAT) and deacetylases (HDAC) respectively (Kouzarides, 2007). Generally, lysine acetylation of histones H2B, H3 and H4 are highly correlated with transcriptionally active state (Pokholok et al., 2005; Schubeler et al., 2004; Wang et al., 2008). It is thought that acetylation of lysine neutralizes the positive charge which lowers the electrostatic interactions with the negatively charged phosphate backbone of DNA and thus weakens the DNA-histone interaction. Indeed, acetylation was shown to increase DNA accessibility, destabilize nucleosomes and lead to an increase of non-histone protein binding to DNA *in vitro* (Lee et al., 1993; Vettese-Dadey et al., 1996; Wolffe and Hayes, 1999). Thus, it is conceivable that acetylation of individual lysines conveys little specificity, but rather the cumulative effect of acetyl groups at multiple lysines would be important for regulating DNA accessibility.

The acetylation of lysine 16 of histone H4 (H4K16) stands as an exception, since it has been shown to directly interfere with higher order chromatin structure formation via preventing interactions between neighboring nucleosomes (Robinson et al., 2008; Shogren-Knaak et al., 2006). H4K16 acetylation has also been shown to play a specific role in *Drosophila* dosage compensation (Bell et al., 2008; Kind et al., 2008). A number of chromatin-associated proteins contain highly conserved Bromodomains which specifically bind to acetylated lysines, raising the possibility that acetylation may have much more regulatory potential than discovered so far (Taverna et al., 2007).



**Figure 7:** A number of distinct post-translational modifications are known to occur at the N-terminus of canonical histones including acetylation (ac), phosphorylation (ph), methylation (me) and monoubiquitination. Adapted from Bhaumik et.al 2007.

However, the data suggest that an “acetylation-code” (Kurdistani et al., 2004) may not exist. Genomewide data revealed that histone H3 and H4 lysine acetylation marks largely overlap and highly correlate with transcription (Wang et al., 2008). Furthermore, genetic substitution of individual lysines did not produce specific phenotypes in yeast (Dion et al., 2005). Together, these observations argue for a model where acetylation marks seems to have an additive function. Noteworthy, this does not hold true for certain specific functions for a few individual acetylated lysines such as H4K16ac in higher order chromatin compaction (Shogren-Knaak et al., 2006) and H3K56ac in nucleosome assembly during DNA repair (Das et al., 2009).

**1.4.1.2 Histone Phosphorylation:** In contrast to our knowledge about acetyltransferases and deacetylases, not so much is known about the enzymes involved in other histone modifications. Important progress has been made, however, towards understanding the role of histone phosphorylation in processes such as transcription, DNA repair, apoptosis and chromosome condensation (Cheung et al., 2000b).

Phosphorylation of Histone 3 Serine 10 (H3S10) is a previously described hallmark of condensed mitotic chromosomes, which however in contrast, is also known to associate with relaxed chromatin and active transcription in interphase cells (Baek; Nowak and Corces, 2004; Prigent and Dimitrov, 2003; Thomson et al., 1999; Zhang et al., 2006). A number of other signaling kinases including PIM1, IKKa, MSK1/2, PKB/Akt, and Rsk2 have previously been reported to modify H3S10 (Anest et al., 2003; Cerutti and Casas-Mollano, 2009; Dyson et al., 2005; Perez-Cadahia et al., 2009; Sassone-Corsi et al., 1999; Yamamoto et al., 2003; Zippo et al., 2007). Upon activation in response to upstream signals, these kinases were shown to phosphorylate H3S10 that accompanies proper gene expression responses (Anest et al., 2003; Cerutti and Casas-Mollano, 2009; Dyson et al., 2005; Perez-Cadahia et al., 2009; Sassone-Corsi et al., 1999; Yamamoto et al., 2003; Zippo et al., 2007) in a paradigm that further applies to a MAP kinase, JNK, that was recently shown to directly bind chromatin and phosphorylate H3S10 for gene activation (Tiwari et al, Nature Genetics, in press).

A number of mechanisms have been proposed by which such phosphorylation may result in transcriptional activation. The addition of negatively charged phosphate groups to histone tails was thought to neutralize their basic charge and thus reduce their affinity for DNA. Furthermore, several acetyltransferases have been shown to exert increased HAT activity on serine 10-phosphorylated substrates and that mutation of serine 10 decreased activation of Gcn5-regulated genes (Cheung et al., 2000c; Lo et al., 2000). This suggests that phosphorylation may contribute to transcriptional activation through the stimulation of HAT activity on the same histone tail. This is supported by the observed phosphoacetylation of histone H3 on *c-fos*- and *c-jun*-associated nucleosomes upon gene activation (Clayton et al., 2000). Importantly, high H3 Serine 10 phosphorylation levels during M phase block binding of the heterochromatic protein HP1 (Fischle et al., 2005), while reduced levels of the interphase histone H3S10 kinase, JIL-1, leads to spreading of the major heterochromatin markers H3K9me2 and HP1 to ectopic locations

(Zhang et al., 2006). These findings implied that H3S10 phosphorylation prevents ectopic recruitment or spreading of heterochromatic factors such as HP1 and thereby contributes to an active transcription state.

Phosphorylation of H2A at Serine 10 has been correlated with mitotic chromosome condensation (Cheung et al., 2000b). For example, mutation of serine 10 in *Tetrahymena* histones causes abnormal chromosomal condensation and defective chromosome separation during anaphase. The Ipl1/aurora kinase in yeast and nematodes and the NIMA kinase in *Aspergillus nidulans* were shown to regulate H3 serine 10 phosphorylation and the expression of these enzymes correlated with mitosis (De Souza et al., 2000; Hsu et al., 2000). Disregulation of Ipl1 or NIMA results in disruption of chromosome condensation or segregation during mitosis. This mark was shown to be dephosphorylated by the protein phosphatase Glc7/PP1 after mitosis.

The activation of DNA-damage signaling pathways was also shown to result in phosphorylation of histone H3. A conserved motif (ASQE, in the single-letter amino-acid code) found in the carboxyl terminus of yeast H2A and the mammalian H2A variant H2A.X was shown to be rapidly phosphorylated upon exposure to DNA-damaging agents (Downs et al., 2000; Rogakou et al., 1999). In yeast, Mec1 phosphorylates Serine 10 in response to damage and this is required for efficient non-homologous end-joining repair of DNA. This suggests that phosphorylation mediates an alteration of chromatin structure, which in turn facilitates repair.

**1.4.1.3 Histone Methylation:** Methylation of histones can either occur at lysine or arginine residues (**Figure 7**). In contrast to acetylation, mono- (me1), di- (me2) and tri-methylation (me3) states of the same residue are observed. These differential methylation states present another level of regulatory potential which indeed appears to be exploited. Several lysines display diverging functions and localization in the genome depending on their methylation state (Barski et al., 2007; Peters and Schubeler, 2005). Arginine methylation is performed by protein arginine methyltransferases (PRMTs) and is antagonized by PADI4 (Klose et al., 2006; Zhang and Reinberg, 2001). Lysine methylation is carried out by specific lysine methyltransferases (KTM), which all contain a conserved SET-domain with the exception of Dot1/KTM4 (Zhang and Reinberg, 2001). Lysine methylation can be removed by two distinct classes of histone demethylases (KDMs): the LSD1 enzyme and the JmjC protein family (Klose et al., 2006). In



contrast to acetylation, methylation cannot neutralize the nucleosomal charge and has been proposed to function via recruiting chromatin modifying and regulatory proteins. Till date, three protein domains have been found to specifically recognize methylated lysines: tudor-domains, chromo-domains and PHD-finger domains. Each domain has characteristic affinities for different lysines and methylation states which further depend on other domains of the respective protein and its interaction partners (Martin and Zhang, 2005; Taverna et al., 2007).

Genomewide location analysis experiments have revealed that active genes are enriched with methylation of lysine 4 of histone H3 (H3K4), H3K36 and H3K79 (Barski et al., 2007; Pokholok et al., 2005; Saunders et al., 2006; Schubeler et al., 2004). In yeast, Set1 and Set2 methylate H3K4 and H3K36 respectively which in turn directly interact with factors bound to the Pol II complex (Krogan et al., 2003a; Krogan et al., 2003b). Genetic evidence also predicts a recruitment of Dot1 (a H3K79 KMT) to chromatin via Pol II (van Leeuwen et al., 2002). H3K36me and H3K79me display a broader distribution within the gene body, starting just downstream of the H3K4me<sub>2/3</sub> peak (Bell et al., 2007; Wirbelauer et al., 2005). Consistent with a role for H3K36me in transcription, data from yeast denote that H3K36me prevents cryptic initiation via recruiting a histone deacetylase to the body of genes, which presumably leads to a less accessible chromatin structure (Carrozza et al., 2005).

H3K4 methylation is thought to function in transcriptional activation pathways since many chromatin remodeling and co-activator complexes contain H3K4me<sub>2/3</sub> recognition modules. For example a PHD-domain in the NURF chromatin remodeling complex specifically recognizes H3K4me<sub>3</sub> and might facilitate transcriptional activation via opening the chromatin structure around H3K4me<sub>2/3</sub> modified promoters (Wysocka et al., 2006).

Inactive loci display a different set of methylation marks mainly consisting of methylation of H3K9, H4K20, and H3K27. H3K9 and H4K20 di- and tri-methylation play essential roles in heterochromatin maintenance at pericentromeric repeat regions and are further present at repetitive, transposable and retroviral elements in mammalian genomes (Lehnertz et al., 2003; Mikkelsen et al., 2007; Peters et al., 2003). Very rarely, only certain regulatory regions have been identified to be methylated at H3K9 and/or H4K20 and are mostly CpG-poor promoters of large gene families such as the olfactory receptor clusters or zinc finger proteins (Mikkelsen et al., 2007; Vogel et al., 2006), suggesting that this is likely resulting from the repetitive nature of

their genomic organization. H3K9 methylation is carried out by 5 known KMTs with distinct specificities: Suv39h1 and Suv39h2 largely contribute to the methylation in constitutive heterochromatin in pericentric and telomeric regions. G9a, GLP and Setdb1 rather localize to euchromatin and have roles in silencing of repetitive and retroviral elements (Kouzarides, 2007). H3K9me<sub>2/3</sub> was also shown to recruit HP1 to chromatin, which is an integral component of heterochromatin and essential for repression (Lachner et al., 2001). Furthermore, direct interactions between DNMT1 and the H3K9 KMT G9a, and between DNMT3b and HP1 have been reported (Esteve et al., 2006; Lehnertz et al., 2003), suggesting that H3K9 methylation crosstalks to DNA methylation. The interaction between DNMT3b and HP1 was shown to be important for a correct establishment of pericentric heterochromatin (Lehnertz et al., 2003). According to a model by Feldman and coworkers, G9a complex methylates H3K9, which then recruits HP1. Subsequently, HP1 recruits DNMTs to mediate stable repression by DNA methylation (Cedar and Bergman, 2009; Feldman et al., 2006). Though attractive, this model needs further experimental validation. H3K27 di- and tri-methylation in turn is excluded from regions carrying H3K9 methylation and predominantly localizes to CpG-rich regions, which strongly implies different functions of these two repressive histone methylation marks.

#### **1.4.2 Polycomb-mediated repression**

The Polycomb Group (PcG) and trithorax Group (trxG) of proteins are involved in defining cellular memory and prevent changes in cell type specific transcription programs to maintain cell identity (Bantignies and Cavalli, 2006; Cao et al., 2005; Cao et al., 2002; Cao and Zhang, 2004a; Cao and Zhang, 2004b; Jacobs and van Lohuizen, 1999; Kuzmichev et al., 2004; Negishi et al., 2007; Vire et al., 2006). The antagonistic trithorax-group (TrxG) and Polycomb-group (PcG) proteins were first identified as Hox gene regulators guiding embryonic patterning in *Drosophila* (Ringrose and Paro, 2007; Schuettengruber et al., 2007; Schwartz and Pirrotta, 2007; Sparmann and van Lohuizen, 2006a). Remarkably, PcG proteins underwent an expansion during evolution of vertebrates. Many paralogs arose, which has been suggested to contribute to cell type specific gene regulation; an essential requirement for the observed increase in organismal complexity (Whitcomb et al., 2007).

Beyond their role in embryonic development, PcG proteins have been implicated in maintaining pluri-potency and cell identity via repression of key developmental regulators in embryos and

embryonic stem (ES) cells (Bernstein et al., 2006; Boyer et al., 2006; Lee et al., 2006). PcG proteins have further been shown to play a role in cellular differentiation, cell fate plasticity (Caretta et al., 2004; Ezhkova et al., 2009; Klebes et al., 2005; Lee et al., 2005) and proliferation (Martinez et al., 2006). In addition, their mis-regulation is shown to correlate with neoplastic development (Sparmann and van Lohuizen, 2006a; van Leeuwen et al., 2002). Self-renewal of ES cells lacking Ezh2, Eed or Suz12 is unaffected in culture, arguing that Polycomb is not essential for their propagation. Importantly however, all of these cells show transcriptional upregulation of many differentiation genes that are otherwise normally repressed by Polycomb. Induction of differentiation of mutant ES cells results in either death or incomplete differentiation to term, which is in support of the early embryonic lethal phenotypes *in vivo* (Chamberlain et al., 2008; Erhardt et al., 2003; O'Carroll et al., 2001; Pasini et al., 2007).

Biochemical purification experiments revealed that at least two separate multimeric PcG complexes or Polycomb repressive complexes (PRC) with distinct enzymatic activities exist (Schwartz and Pirrotta, 2007; Sparmann and van Lohuizen, 2006a). PRC2 contains Ezh2 which mediates di/tri-methylation of lysine 27 of Histone H3 [H3K27me<sub>2/3</sub>; (Czermin et al., 2002; Muller et al., 2002), the hallmark of Polycomb repressed genes. PRC1 on the other hand contains four conserved core components and mediates H2A lysine 119 mono-ubiquitination [H2AK119u<sub>1</sub>; (Wang et al., 2004a)]. PRC1 is thought to cooperate with PRC2 for binding at target genes enriched in H3K27me<sub>3</sub>, and to mediate repression by inhibiting chromatin remodeling, blocking transcription and/or by mediating chromatin compaction (Levine et al., 2004; Margueron et al., 2008). It is unclear how PRC complexes regulate gene expression, but their overlapping genomic distributions suggest concerted actions (Boyer et al., 2006; Ku et al., 2008). It was also shown that PcG-silenced genes may be clustered into sub-nuclear compartments termed PcG-bodies which may provide a means to coordinate repression of multiple genes (Grimaud et al., 2006; Lanzuolo et al., 2007; Terranova et al., 2008; Tiwari et al., 2008).

Interestingly, in mammals Polycomb-mediated repressive H3K27me<sub>3</sub> was shown to very often coincide with activating H3K4me<sub>2/3</sub> in close proximity and this was termed as 'bivalent chromatin' (Azuara et al., 2006; Bernstein et al., 2006; Bracken et al., 2006; Mikkelsen et al., 2007). Such bivalency was thought to "poise" genes for activation at a later time point (Bernstein et al., 2006; Spivakov and Fisher, 2007). However, no solid experimental evidence exists for

such proposal so far. In *Drosophila*, PcG proteins are recruited to regulatory elements called PREs (Polycomb Response Elements) to silence nearby genes and argues for a DNA sequence-dependent recruitment of Polycomb (Ringrose and Paro, 2007). In mammals, despite large scale mapping efforts, no such DNA sequence elements for PcG recruitment could be identified so far. Furthermore, *Drosophila* PREs are very weakly conserved at the sequence level. Together this may reflect that many different sequence-specific transcription factors can recruit Polycomb and thus, PRE sequence identification in mammals will not be feasible. In support of this, the transcription factor Snail1 was shown to recruit Polycomb proteins to the E-cadherin promoter coinciding with its transcriptional silencing in cancer cells (Herranz et al., 2008).

In addition to PcG targeting in mammals, the propagation of H3K27me3 and PcG complexes during replication is equally not understood. PRC2 was shown to directly bind to H3K27me3, thereby ensuring propagation of the mark during mitosis (Hansen et al., 2008). PRC1 was also shown to remain bound to the DNA during replication in an *in vitro* system (Francis et al., 2009). Despite these findings, the propagation of H3K27me3 and PcG complexes *in vivo* remains to be investigated in detail. A number of reports suggest interplay between the Polycomb pathway and DNA methylation. Ezh2 was shown to recruit DNMT1 in cancer cell lines by a direct interaction (Vire et al., 2006), which however stays controversial due to non-reproducibility in non-transformed cells. In addition, several recent studies suggested preferential aberrant DNA methylation in human cancer cell lines and primary cancers at promoters that are Polycomb targets in unrelated human ES cells in culture (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). These studies suggest that a potential targeting pathway of DNMTs via Polycomb may exist, however no direct conclusive evidences exist so far.

#### **1.4.2.1 Polycomb and EMT**

The components of both Polycomb repressive complex 1 and Polycomb repressive complex 2 are implicated in EMT. PRC1 complex comprises Bmi-1, Ring1, Hph1/2/3, and Hpc1/2/3 (Cao and Zhang, 2004a; Cao and Zhang, 2004b; Sparmann and van Lohuizen, 2006a) and PRC2 complex comprises Ezh2, Eed, Suz12, RbAp46/48 and Aebp2 (Cao and Zhang, 2004a; Cao and Zhang, 2004b; Schuettengruber et al., 2007; Sparmann and van Lohuizen, 2006a). The PRC1 protein B lymphoma Mo-MLV insertion region 1 homolog (Bmi-1) was the first reported PcG protein to be associated with cancer development and later on with EMT. Bmi-1 can inhibit c-Myc induced

apoptosis via Ink4a/Arf and regulate cell proliferation and senescence (Jacobs et al., 1999a; Jacobs et al., 1999b). Bmi-1 directly binds to Ink4a-Arf locus and represses it. Ink4a and Arf play a role to restrict cellular proliferation in response to aberrant mitogenic signaling. Ink4a is a cyclin-dependent kinase inhibitor and it can activate the Rb pathway. Arf can inhibit Mdm2 function to induce p53. In many types of tumors, the Ink4a/Arf locus is found to be mutated, deleted or epigenetically silenced. Furthermore, Bmi-1 depletion leads to death of embryonic carcinoma stem cells as well as neuroblastoma cells (Cui et al., 2007). Recently, it has been demonstrated that over-expression of Bmi-1 leads to EMT, and its blockage prevents EMT driven-invasion. To elicit its function, Bmi-1 targets the tumor suppressor Pten and downregulates its expression. This all contributes toward the activation of the PI3K/Akt pathway, stabilization of Snail and downregulation of E-cadherin (Song et al., 2009). In addition, Bmi-1 is essential for Twist-1-induced EMT and both of them are cooperatively required for repression of E-cadherin and p16INK4a, which is a cyclin kinase inhibitor, in head and neck carcinoma. Thus, Twist1 and Bmi1 were needed together to promote EMT and tumor-initiating capabilities (Yang et al., 2010b). Another PRC1 protein, Ring1, was also shown to interact with Bmi-1, and overexpression of Ring1 represses Engrailed and increases expression of c-Jun and c-Fos. Furthermore, Ring1 induces anchorage-independent growth of Rat1a and NIH3T3 cells, and overexpression of Ring1 can form tumors in nude mice (Ben-Saadon et al., 2006).

Components of the PRC2 complex are extensively studied in cancer, especially Ezh2, which is a catalytic subunit of PRC2. Among tumors of epithelial origin, Ezh2 was first observed to be significantly associated with metastatic prostate cancer. It has been shown that Ezh2 is overexpressed with Bmi-1 in prostate metastatic precursor cells (Berezovska et al., 2006). Berezovska *et al.* demonstrated a marked enrichment of the population of circulating human prostate carcinoma metastasis precursor cells with dual-positive high-Bmi1/Ezh2-expressing cells. Importantly, depletion of Bmi1 or Ezh2 in prostate carcinoma metastasis precursor cells diminishes their tumorigenic, metastatic and proliferation potential when injected into mice (Berezovska et al., 2006). Adrb2, a target of Ezh2 and a G-protein-coupled receptor (GPCR) of the  $\beta$ -adrenergic signaling pathway, has been implicated in EMT (Yu et al., 2007). Overexpression of Ezh2 can repress Adrb2 at both the transcript and the protein levels and the recently discovered Ezh2 inhibitor, DZNep (Tan et al., 2007), can prevent Adrb2 repression by Ezh2. Downregulation of Adrb2 blocks EMT and prevents tumor growth in a xenograft model of

prostate cancer. In addition, Ezh2 also targets E-cadherin for repression. This Ezh2-mediated repression could be inhibited by the HDAC inhibitor SAHA. Furthermore, co-expression of E-cadherin can attenuate Ezh2-mediated invasion in prostate and breast epithelial cells, indicating that E-cadherin is an important target of Ezh2 in cancer progression (Rhodes et al., 2003). These results are similar to ES cell data where Ezh2 is required for ES cell pluripotency and differentiation by repressing E-cadherin. Moreover, Zeb2, another EMT inducer, is subjected to post-translation regulation by PRC2 where sumoylation impairs its repressor activity (Long et al., 2005). It is also shown that ZEB1 represses E-cadherin and induces EMT by recruiting the Swi/Snf chromatin-remodeling protein Brg1, the ATPase subunit of the Swi/Snf chromatin-remodeling complex (Sanchez-Tillo et al., 2010). Recently, it has been reported that Snail1-mediated repression of E-cadherin is PRC2-dependent in ES cells and that depletion of Suz12, one of the components of PRC2, leads to partial de-repression of the E-cadherin promoter by Snail1. This study also showed that Snail1 interacts with Suz12 and Ezh2 and increases the binding of Suz12 at the E-cadherin promoter (Herranz et al., 2008).

E-cadherin is not the only gene regulated by PRC2. Another EMT gene,  $\beta$ 4 integrin has been shown to be controlled by Polycomb-mediated histone modifications. The expression of  $\beta$ 4 integrin is lost during EMT and this loss correlates with a decrease in active histone modifications H3K9Ac and H3K4me3 and an increase in the repressive histone modification H3K27me3. Furthermore, reversal of EMT leads to re-expression of  $\beta$ 4 integrin and restoration of active marks. These results argue for a dynamic nature of epigenetic regulation during fate changes of epithelial cells (Yang et al., 2009).

The most recent work by McDonald and colleagues studied the role of another repressive mark H3K9me2 and activating mark H3K4me3 during EMT. They showed a global reduction in heterochromatin, measured by the repressive mark H3K9me2, and increased levels of euchromatin, measured by the active mark H3K4me3, during EMT. The authors further demonstrate that these changes are largely dependent on lysine-specific demethylase-1 (Lsd1) and depletion of Lsd1 prevents TGF- $\beta$ -induced migration and chemoresistance. Interestingly, all these chromatin changes are associated with so-called LOCKs, which is an acronym for 'Large Organized Heterochromatic K9 modifications' regions (McDonald et al., 2011a).

### 1.4.3. DNA methylation

The best understood epigenetic modification of DNA in mammals is methylation of cytosine at position C5 in CpG dinucleotides. The mammalian DNA methylation machinery has two constituents, the DNA methyltransferases (DNMTs), which establish and maintain DNA methylation patterns, and the methyl-CpG binding proteins (MBDs), which 'read' these methylation marks (Goll and Bestor, 2005; Wade, 2001).

Passive DNA demethylation has long been known to occur by a reduction in activity or absence of DNA methyltransferases (DNMTs). Recently, three enzyme families have been implicated in active DNA demethylation via DNA repair (Bhutani et al., 2011). First, 5-methylcytosine (5mC) can be hydroxylated by the ten-eleven translocation (TET) family of enzymes to form 5-hydroxymethylcytosine (5hmC) or further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Second, 5mC (or 5hmC) can be deaminated by the AID/APOBEC family members to form 5-methyluracil (5mU) or 5-hydroxymethyluracil (5hmU). Third, replacement of these intermediates (i.e., 5mU, 5hmU, or 5caC) is initiated by the UDG family of base excision repair (BER) glycosylases like TDG or SMUG1, culminating in cytosine replacement and DNA demethylation.

DNA methylation is crucially involved in regulating many cellular processes including embryonic development, transcription, chromatin structure, X chromosome inactivation, genomic imprinting and chromosome stability. In normal cells, DNA methylation occurs predominantly in repetitive genomic regions, including satellite DNA and parasitic elements (such as long interspersed transposable elements (LINES), short interspersed transposable elements (SINES) and endogenous retroviruses). Approximately 60-90% of all CpG sequences in the genome are methylated, while unmethylated CpG dinucleotides are mainly clustered in the CpG rich sequence termed 'CpG islands' which are mostly associated with promoter regions of active genes. CpG dinucleotides are under-represented in the mammalian genome due to the conversion of methylated cytosine to thymidine via spontaneous deamination, followed by stabilisation of the genome during replication in the germline. CpG islands, particularly those associated with (housekeeping mainly) promoters, are generally unmethylated, although an increasing number of exceptions are being identified. Specific cis-acting sequences seem to protect them from a global wave of de novo methylation occurring during reprogramming event

at the time of implantation. DNA methylation usually represses transcription directly, by inhibiting the binding of specific transcription factors. For example, the abnormal methylation that occurs on the fragile X mental retardation 1 (FMR1) promoter prevents binding of nuclear respiratory factor (NRF1; also known as  $\alpha$ -PAL) and inhibits transcription of this gene, causing fragile X syndrome (Bardoni and Mandel, 2002). It could also cause transcriptional suppression indirectly, by recruiting methyl-CpG-binding proteins and their associated repressive chromatin remodeling activities such as histone deacetylases (HDACs) and histone methyl transferase (HMTs) that alters the local chromatin environment (Eden et al., 1998). Methylation can also affect nucleosomal positioning (Razin and Cedar, 1977). It could also lead to a general decrease in DNase I sensitivity and alteration in higher order structure (Eshet et.al, 1986) that makes these regions less accessible to the transcription machinery. Recently, it was shown that proximal sequence elements are both necessary and sufficient for regulating DNA methylation (Lienert et al., 2011). However, general principles of how DNMTs are targeted to specific regions are still not known; even though there are some observations suggesting that this may involve interactions between DNMTs and one or more chromatin-associated proteins. DNA methylation patterns are established by de novo DNMTs active on unmethylated DNA (de novo methylation) and then this pattern is maintained through replication by a semiconservative mechanism for copying the methylation residues from parental strands onto the newly synthesized nascent DNA by maintenance DNMTs specific for the hemi-methylated DNA resulting from replication (maintenance methylation). The phenotypes of some selected mouse mutants of epigenetic regulatory factors are listed below in **Table 3**.

There are several indications that DNA methylation and histone methylation are intimately linked and may control each other, but it stays unclear which one comes first due to reports from both ways (Cedar and Bergman, 2009; Vaissiere et al., 2008). Aberrations in DNA methylation patterns such as promoter CpG island hypermethylation of tumor suppressor genes, which results in their silencing, are accepted as being a common feature of human cancer (Berdasco and Esteller, 2010). Such CpG island promoter hypermethylation affects genes from a wide range of cellular pathways, such as cell cycle, DNA repair, toxic catabolism, cell adherence (including Cdh1), apoptosis, and angiogenesis, among others (Esteller, 2008a), and may occur at various stages in the development of cancer. During the last few decades, an increasing number of drugs



**Table 3.** Key epigenetic regulatory proteins, their function and mutant phenotype [Adapted from Jaenisch R and Bird A 2003 (Jaenisch and Bird, 2003)].

Protein	Function	Mutant Phenotype
<i>Methyltransferases</i>		
Dnmt1	Maintenance of methylation	Embryonic lethal, loss of imprinting and X-linked gene expression, ES cells viable
Dnmt1o	Oocyte-specific isoform	Loss of material imprint
Dnmt2	Non-CpG methylation in <i>Drosophila</i>	No phenotype
Dnmt3a, Dnmt3b	De novo methyltransferases, establishment of methylation	Embryonic lethal, ICF syndrome
Dnmt3L	No catalytic activity, colocalizes with Dnmt3a and Dnmt3b	Abnormal maternal imprinting
<i>Methyl binding proteins</i>		
MeCP2	Methyl binding proteins, recruit HDACs	RTT
MBD1	Methyl binding proteins, recruit HDACs	
MBD2	Methyl binding proteins, recruit HDACs	Behaviour abnormalities
MBD3	Methyl binding proteins, recruit HDACs	Lethal
MBD4	Repair enzyme	Increased mutation frequency
<i>Histone-modifying proteins</i>		
HDAC1	Histone deacetylase	Embryonic lethal
Suvar39	Lys9 methylation in histone H3	Embryonic lethal, chromosomal instability, increased tumor risk

targeting DNA methylation have been developed for epigenetic therapy and a few are in clinical trials (Yang et al., 2010c). One example of this success is the earliest and the most successful epigenetic drug to date, 5-Azacytidine, which is currently, recommended as the first-line treatment of high-risk myelodysplastic syndromes (MDS).

#### **1.4.3.1 DNA methylation and EMT**

In recent years, it has become evident that genomic regulation is not the only mechanism that governs anomalous gene expression during tumor progression. Epigenetic modification, especially DNA methylation, seems to play a widespread role during carcinogenesis. Anomalous hypermethylation of promoter CpG islands is an important means of repressing tumor suppressor genes. Initially, E-cadherin promoter methylation was observed in primary prostate cancers and breast cancer lesions where E-cadherin silencing can be de-repressed by treatment with the demethylating agent 5'-aza-2'-deoxycytidine (5'-aza) (Graff et al., 1995; Yoshiura et al., 1995). Following those primary studies, E-cadherin repression and DNA hypermethylation was observed in many different kinds of carcinomas.

To define the mechanism underlying E-cadherin promoter methylation, two models were proposed. According to the first model, Snail expression is co-related with the silencing of E-cadherin and the hypermethylation of its promoter, suggesting a role for Snail in the coordination of both processes (Cheng et al., 2001). On the other hand, the second model proposed that the silencing of E-cadherin is not always correlated with promoter hypermethylation, indicating the involvement of additional genetic and epigenetic modifications are required for E-cadherin downregulation. In addition, DNA methylation and histone deacetylation are correlated with transcriptional repression. Several co-repressor complexes have recently been described that can be recruited to the E-cadherin promoter by different repressors such as Snail-mediated recruitment of a mSin3a/Hdac1/2 complex (Peinado et al., 2004) or of  $\delta$ Efl/Sip1 (Shi et al., 2003) with ctbp complex, containing several Hdacs and Dnmts. In addition, hypermethylation of 8 tumor suppressor genes (*APC*, *BINI*, *BRCA1*, *CST6*, *GSTP1*, *P16*, *P21* and *TIMP3*) was observed in plasma of the breast cancer patients which allowed the development of blood based tests as a predictive and prognostic biomarkers for breast cancer (Radpour et al., 2011).

### **1.5 MARA**

MARA stands for “Motif Activity Response Analysis”. MARA models gene expression profiles across a set of samples in terms of the predicted transcription factor binding site (TFBS) and infers which transcription factors (TFs) are the key regulators that drive gene expression changes across the samples, how each regulator changes its activity across the samples, and what are the genome-wide targets of each regulator. It benefits from two key resources. First, MARA uses ‘promoteromes’, i.e. genome-wide promoter annotations, in human and mouse constructed from genome-wide transcription start site data that was obtained using next-generation sequencing (Balwierz et al., 2009). Second, it uses Bayesian probabilistic methods for predicting functional TFBSs by integrating rigorous models of TFBS evolution with comparative genomic sequence analysis. The MotEvo algorithm (van Nimwegen, 2007) is used to predict functional TFBSs for all regulatory motifs across all promoters in human and mouse. The result of this analysis was, for both human and mouse, a large matrix  $N$ , where  $N_{pm}$  is the predicted total number of functional binding sites in promoter  $p$  for motif  $m$ .

### 1.5.1 Model

$$E_{ps} = \tilde{c}_s + c_p + \sum_m N_{pm} A_{ms},$$

The model follows the approach of Bussemaker et al. (Gao et al., 2004) of assuming that the expression at each promoter  $p$  can be modeled as a linear function of the binding sites  $N_{pm}$ . Specifically, if we denote by  $E_{ps}$  the log-expression of promoter  $p$  in sample  $s$ , we assume the following model where  $c_p$  is a term reflecting the basal activity of promoter  $p$ ,  $c_s$  reflects the total expression in sample  $s$ , and  $A_{ms}$  is the (unknown) activity of motif  $m$  in sample  $s$ . That is, using the predicted site-counts  $N_{pm}$  and the measured expression levels  $E_{ps}$  we use this approximation to infer the activities  $A_{ms}$  of all motifs across all samples.

### 1.5.2 MARA Outputs

The results include an ordered list of regulatory motifs and corresponding TFs, sorted by a  $z$ -value that quantifies the importance of the motif in explaining observed expression changes. For each motif, MARA also provides an activity-profile across all input samples which quantify how much the corresponding TFs are driving the expression of their targets in each sample. In addition, MARA predicts a list of target promoters for each motif, as well as the binding sites in

the promoters through which the corresponding the binding TFs are predicted to act. Finally, to functionally characterize the role of each regulatory motif, MARA reports the functional gene ontology categories that are enriched among the targets, and provides links to the String database network representations of the predicted targets of each regulator.

## 2.0 Aims of the study

Epithelial-mesenchymal transition (EMT) is a critical process underlying the onset of cancer cell invasion and has been identified at the invasive front of murine and human tumors. Hence, the identification of reliable prognostic markers and the intervention with EMT driven invasion is of major interest in cancer diagnostics and therapy. To achieve these goals, we need to gain new and detailed insights into the molecular processes regulating EMT. For this purpose, we induced EMT in different *in vitro* model systems and followed changes in gene expression profile to investigate:

- Master regulatory genes underlying EMT
- Cell biological processes controlled by these master regulators
- Requirement of these master genes in tumor progression and metastasis formation *in vivo*
- Possibility of these master regulators to act as potential markers for invasion and metastasis in human tumors
- If these master regulators are under control of epigenetic mechanisms

Here, we studied two transcription factors, namely Klf4 and Sox4 for their role in EMT by investigating the cell biological functions they control and their contribution to tumor progression and metastasis formation *in vivo*. In the last part of the thesis, using genomewide approaches, we studied the contribution of the Polycomb machinery and DNA methylation to the transcriptional reprogramming that accompanies EMT.

## **3.0 Results**

### **3.1 Klf4 directly regulates transcription of genes crucial for Epithelial to Mesenchymal transition**

Neha Tiwari, Phil Arnold, Erik Van Nimwegen and Gerhard Christofori

*Manuscript in preparation*

#### **3.1.1 Abstract**

Krupple-like factor 4 (Klf4) is a zinc-finger transcription factor that is abnormally expressed in diseased state including cancer. However, its regulatory role during carcinogenesis remains unclear. Using a genome-wide expression profiling approach, we uncovered Klf4 among the transcription factors that were significantly downregulated during TGF- $\beta$ -induced Epithelial to Mesenchymal Transition (EMT) in NMuMG cells. Although Klf4 expression is diminished during EMT, its activity, as predicted by MARA analysis, goes up arguing for the regulatory role of Klf4. Gain of function experiments demonstrate that Klf4 is not only inhibitory to EMT and migration but also acts as a survival factor during this process. To reveal how its DNA binding activity may be related to the phenotype, we performed chromatin immunoprecipitation followed by deep-sequencing (ChIP-seq) to identify promoters genomewide that are occupied by Klf4. Interestingly, this analysis revealed Klf4 binding to the promoter of key EMT genes, including N-cadherin, Vimentin,  $\beta$ -catenin and Mapk8 (encoding Jnk1). In line with these observations, the depletion of endogenous KLF4 deregulated a number of Klf4 target genes, suggesting a direct transcriptional control by Klf4. One such target gene, Jnk1, is transcriptionally upregulated upon Klf4 depletion as well as during TGF- $\beta$ -induced EMT that also accompanies substantial decrease in Klf4 levels. Knockdown of Jnk1 not only counteracted TGF- $\beta$ -induced EMT but also Klf4 depletion-induced apoptosis and cell migration. These observations reveal a critical role of Klf4 as a tumor suppressor by direct targeting key EMT gene promoters for transcriptional regulation.

#### **3.1.2 Introduction**

Solid tumors are epithelial in origin. A loss of epithelial-cell markers and gain of mesenchymal-cell markers has been observed in patient tumor samples, particularly at the leading edge or

invasive front of solid tumors such as non-small cell (NSCLC), pancreatic, colorectal, and hepatocellular cancers and this process defined as 'Epithelial to Mesenchymal Transition' (EMT) constitutes a key step during carcinogenesis. Such changes in phenotypic epithelial-like and mesenchymal-like cellular markers have been associated with the degree of tumor progression (Brabletz et al., 2005; Christofori, 2006; Grunert et al., 2003; Huber et al., 2005; Thiery and Sleeman, 2006). The mechanism underlying EMT and contribution of critical transcription factors to this process remains elusive.

Krupple-like factor 4 (Klf4) is a zinc-finger transcription factor which is usually expressed in growth arrested cells, colon, small intestine, testis and lung (Shields et al., 1996). Klf4  $-/-$  mice die shortly after birth due to loss of skin barrier function, as measured by penetration of external dyes and rapid loss of body fluids (Segre et al., 1999). The defect was not corrected by grafting of Klf4  $-/-$  skin onto nude mice (Segre et al., 1999). This protein has been implicated to be an activator as well as repressor at the promoter of various genes in a context-dependent manner (Rowland et al., 2005). Various previous studies as well as Oncomine and Nextbio databases report Klf4 to be downregulated in cancers of different types (Dang et al., 2000; Katz et al., 2005; Ohnishi et al., 2003; Shie et al., 2000; Ton-That et al., 1997; Wei et al., 2005; Zhao et al., 2004). On the contrary, it is also been shown to be upregulated in a large fraction of mammary carcinoma and squamous-cell carcinomas of the oropharynx (Foster et al., 2000). Klf4 is proposed to execute its oncogenic activity by directly binding to and repressing the p53 promoter. In contrast, it can also act as a tumor suppressor by binding to p21 promoter for its activation (Rowland et al., 2005). Klf4 is further implicated to play a role in EMT (Yori et al., 2010), however the underlying mechanisms remain largely unclear. Interestingly, a number of recent studies have identified Klf4 as one of the crucial factors, in addition to Oct4, Sox2 and c-Myc, for the derivation of induced pluripotent cells (Takahashi and Yamanaka, 2006). It was further shown that the transcription factors Oct4, Sox2, c-Myc, and Klf4 can induce epigenetic reprogramming of a somatic genome to an embryonic pluripotent state (Wernig et al., 2007).

The mitogen-activated protein kinases (Mapks) are a family of serine-threonine protein kinases that participate in various signaling pathways and are crucial for cellular responses to extracellular stimuli. One such MAP kinase, Jnk, is activated primarily by cytokines and stress

such as UV. Exposure to UV light leads to Jnk1 phosphorylation at serine-63 and serine-73 residues (Derijard et al., 1994). Such activation of Jnk is required for UV-induced apoptosis in primary murine embryonic fibroblasts and the absence of Jnk caused a defect in the mitochondrial death signaling pathway, including the failure to release cytochrome c (Tournier et al., 2000). NF $\kappa$ B-mediated signaling cascade can downregulate Jnk1 via Gadd45b and regulates TNFR-induced apoptosis (De Smaele et al., 2001). Not only this, Jnk1 has also been previously implicated in EMT (Alcorn et al., 2008; Santibanez, 2006a; van der Velden et al., 2010). In line with these observations, Jnk1 can phosphorylates paxillin, a focal adhesion adaptor, required for maintaining the labile adhesions required for rapid cell migration (Huang et al., 2003). Jnk1-null mice exhibited progressive loss of microtubules within axons and dendrites due to hypophosphorylation of Map2, which reduced the ability of Map2 to bind microtubules and promote tubulin polymerization (Chang et al., 2003).

Using a genomewide transcription profiling approach, we identified a comprehensive list of genes at six consecutive morphological states (1 day, 4 days, 7 days, 10 days and 20 days) during TGF- $\beta$ -induced EMT in normal mammary epithelial cells [NMuMG; (Lehembre et al., 2008)]. The resulting datasets were computationally analysed to identify differentially regulated genes at each stage. We next searched for motifs for various transcription factors at the promoter of regulated genes (<http://test.swissregulon.unibas.ch/cgi-bin/mara>) and predicted a number of transcription factors that may possibly regulate a subset of genes in the early, intermediate and later stages (data not shown). Klf4 was discovered as one of such factor that might contribute to transcriptional regulation of certain genes during TGF- $\beta$ -induced EMT in NMuMG cells. This was especially interesting given the previous suggestions for an involvement of Klf4 in cellular transformation and motivated us to investigate the role of this transcription factor in EMT.

Here we show that Klf4 not only blocks epithelial differentiation but also provides a survival advantage to the cells. Moreover, Klf4 functions inhibitory to cell migration. A combinatorial approach of genomewide transcription profiling after Klf4 depletion in the presence or absence of TGF- $\beta$ , and identification of genomewide target promoters of Klf4 occupancy using ChIP-seq analysis, revealed many key EMT genes that are under direct transcriptional control by Klf4, such as N-cadherin, Vimentin and  $\beta$ -Catenin. Furthermore, we also show how through one of



such targets, Jnk1, Klf4 regulates EMT. Along with this, we also found Klf4 to be regulated by canonical TGF- $\beta$  signaling via Smad4.

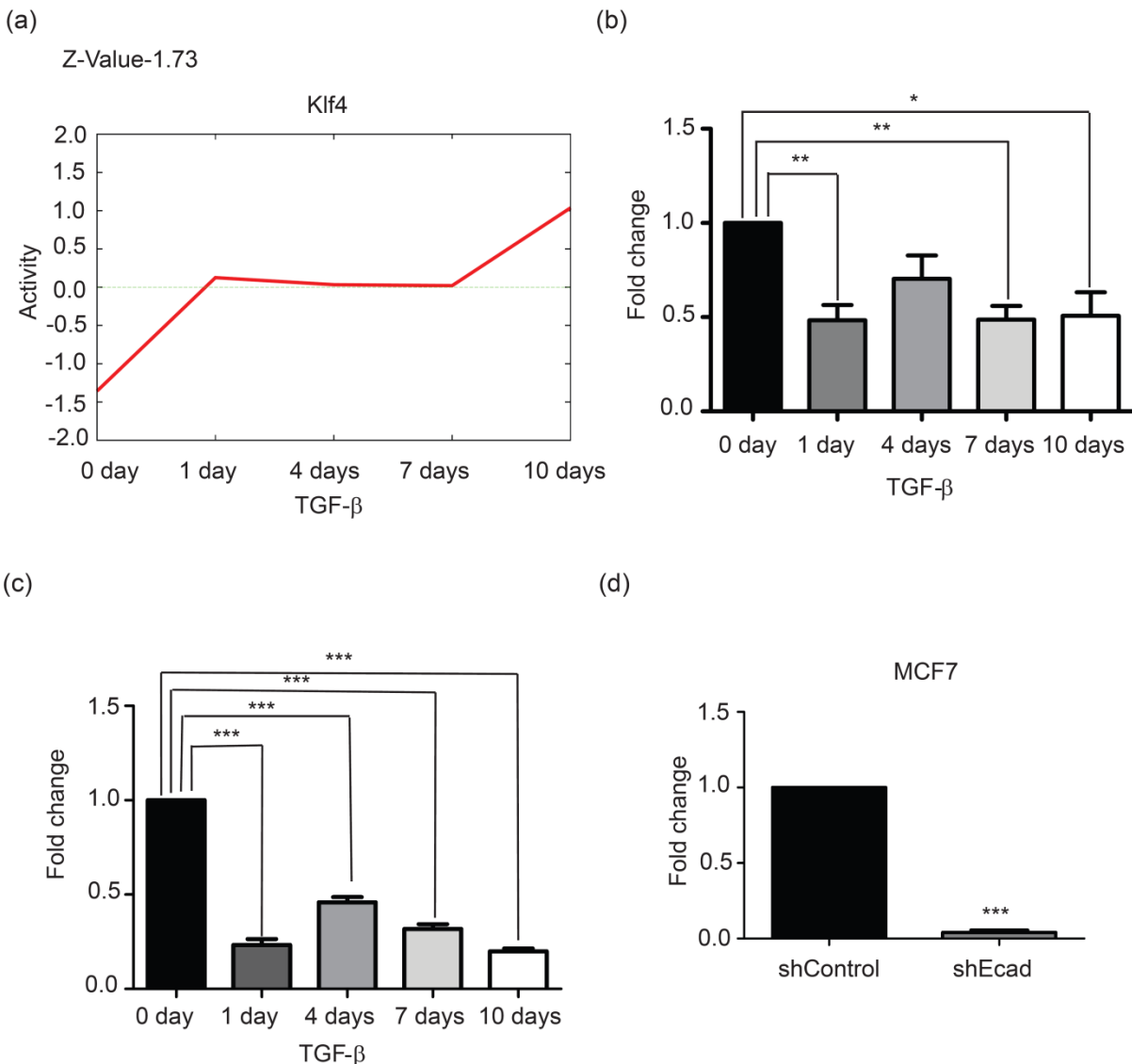
### **3.1.3 Results**

#### **3.1.3.1 Identification of Klf4 as a repressor during EMT**

To decipher gene regulatory mechanisms underlying epithelial to mesenchymal transition during the early, intermediate and later stages of breast carcinogenesis, we employed an established model of EMT in the untransformed normal murine mammary gland cell line NMuMG (Lehembre et al., 2008; Miettinen et al., 1994), which undergoes progressive EMT upon TGF- $\beta$  treatment and acquires a complete mesenchymal morphology by the end point, usually involving 10 days for this full morphological transition (Figure 1) (Lehembre et al., 2008; Piek et al., 1999b). We performed a genome wide expression analysis of non-treated cells as well as those treated with TGF- $\beta$  for 1, 4, 7, 10 and 20 days to identify differentially expressed genes at the respective stages. We then employed the MARA tool (<http://test.swissregulon.unibas.ch/cgi-bin/mara>) on the resulting gene list to search for transcription factor motifs for various regulatory factors at the promoters of modulated genes. This analysis revealed a set of transcription factors that possibly regulate the transcription of a subset of genes in the early, intermediate and later stages by a direct control of their promoter activity (data not shown). Klf4 was identified as one of such factors whose motif was found enriched at a number of genes that were modulated in transcription during TGF- $\beta$ -induced EMT. MARA analysis predicted an increase in Klf4 activity during EMT (Figure 1a). Since quantitative RT-PCR revealed that Klf4 itself is downregulated during EMT (Figure 1b), we speculated for a repressive function of Klf4 during this transition. Western blot analysis further confirmed RNA level analysis that Klf4 levels are severely downregulated during EMT (Supplemental Figure S1). We further validated these observations in a murine breast cancer line, Py2T cell line, which is also responsive to TGF- $\beta$  and undergoes EMT similar to NMuMG cells (Waldmeier et. al, unpublished) and in a human breast cancer line, MCF7, which undergoes EMT upon stable depletion of E-Cadherin [(Lehembre et al., 2008); Figure 1c and 1d]. Together these findings suggested a suppressive role of Klf4 during EMT.

### 3.1.3.2 KLF4 blocks epithelial differentiation

To directly assess the role of Klf4 in EMT, we first performed loss of function studies. We knocked down Klf4 efficiently both in the absence and presence of TGF- $\beta$  using a pool of two Klf4 siRNAs (Figure 2a). Depletion of Klf4 in the absence of TGF- $\beta$  did not lead to any noticeable changes in cell morphology other than a slight increase in cell size accompanied with



**Figure 1: Klf4 acts as repressor during TGF- $\beta$ -induced EMT.** (a) MARA analysis predicts Klf4 activity during EMT in NMuMG cells. (b) RT-qPCR analysis for quantification of Klf4 expression levels in NMuMG cells during TGF- $\beta$  mediated EMT. (c) RT-qPCR for quantification of Klf4 levels upon EMT induction in Py2T cells. (d)

Expression levels of Klf4 as quantified by RT-qPCR in control and E-cadherin-specific shRNA treated MCF7 cells. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

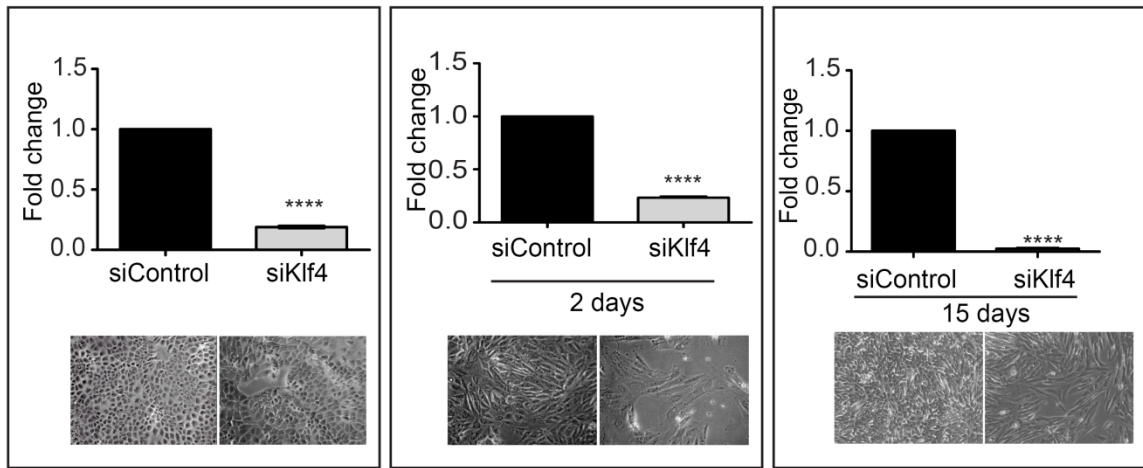
minor flattening (Figure 2a). Although no major morphological changes were apparent after Klf4 knockdown in the absence of TGF- $\beta$ , immunofluorescence analysis revealed an increase in the expression of mesenchymal markers like N-cadherin and Vimentin (Figure 2b). In addition, focal adhesion formation was greatly increased with the induction in stress fibers formation. On the other hand, epithelial markers such as adherent junction protein E-cadherin were slightly downregulated and distorted from the membrane. Interestingly, Klf4 knockdown in the presence of TGF- $\beta$ , sped up the EMT process and the cells became more elongated, spindle-shaped and fibroblast-like even after 2 days of TGF- $\beta$  treatment and gave morphological appearance similar to wild type NMuMG cells treated with TGF- $\beta$  for 10 days (Figure 2a). This accompanied appropriate changes in key EMT markers such as increase in vimentin and N-cadherin expression and complete loss of E-cadherin from the membrane within 2 days of TGF- $\beta$  treatment in Klf4 knockdown cells, in line with the classical cadherin switch observed during EMT (Hazan et al., 2004; Maeda et al., 2005). Furthermore, phalloidin staining revealed an increased stress fiber formation upon Klf4 knockdown (Figure 2b). Notably, Klf4 depletion in NMuMG cells treated with TGF- $\beta$  for 15 days had similar effect and led to acceleration of EMT process, demonstrating that Klf4 is not only inhibitory to initiation but also for the maintenance of EMT (Figure 2a). All of these results were further confirmed in a different breast cancer model of Py2T cells (Supplemental Figure S2a, Supplemental Figure S2b). Together, these results argue for a role of KLF4 in maintaining epithelial morphology and an inhibitory action in EMT.

### **3.1.3.3 Klf4 prevents cell migration and provides survival advantage to the cells during EMT**

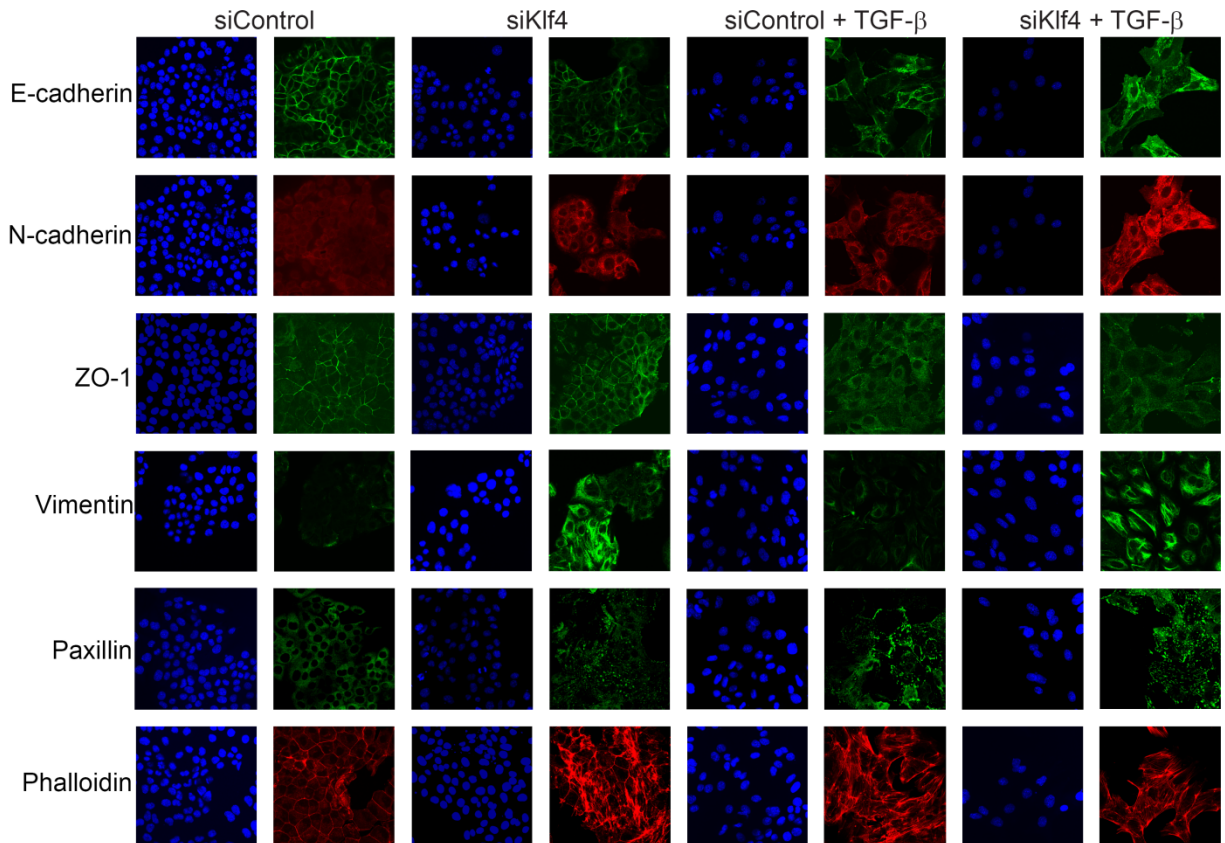
Tumor cells dissolve their cell-cell contacts and undergo EMT so as to leave the primary tumor and invade into the surrounding tissue. This process further facilitates the migration of tumor

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

(a)



(b)



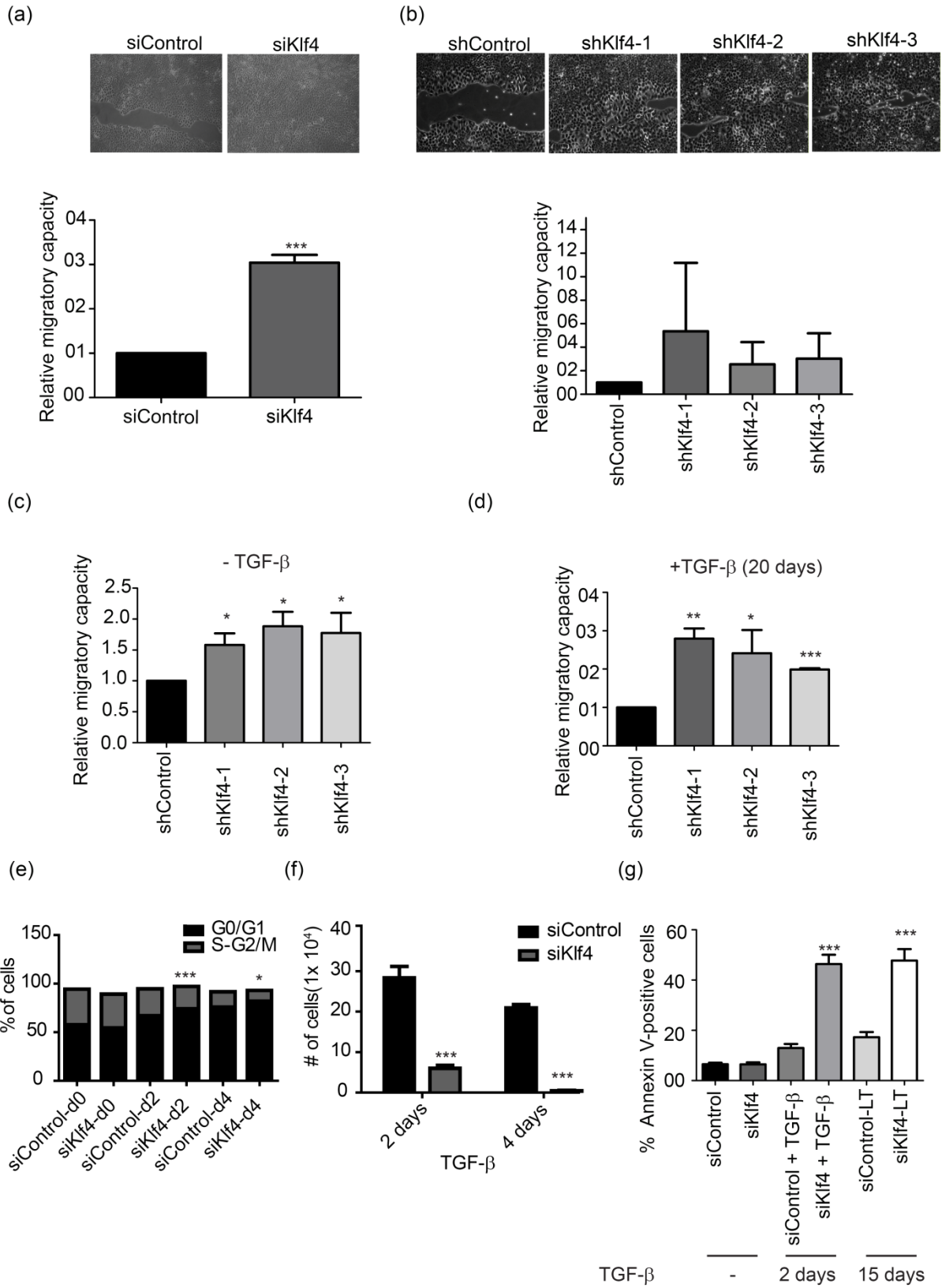
**Figure 2: Depletion of KLF4 induces epithelial differentiation.** (a) qRT-PCR was performed after siRNA mediated knockdown of Klf4 in the absence and presence of TGF- $\beta$  to assess the knockdown efficiency. Cells were

treated with TGF- $\beta$  for 2 days and 15 days. Morphology of NMuMG cells transfected with either control or Klf4-specific siRNA in the absence as well as presence of TGF- $\beta$ , as evaluated by phase contrast microscope. Original magnification was 10X. (b) Immunofluorescence stainings for adherent junction protein, E-cadherin and N-cadherin, tight junction protein ZO-1, intermediate filamentous protein vimentin, focal adhesion protein paxillin and actin cytoskeleton protein phalloidin were performed after Klf4 knockdown in the absence and presence of TGF- $\beta$ . Original magnification was 40X. Cells were treated with TGF- $\beta$  for 2 days. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.001$  indicated with (\*\*\*)

cells to neighboring blood vasculature and the distant organs to form metastasis (Brabletz et al., 2005; Christofori, 2006; Grunert et al., 2003; Huber et al., 2005; Thiery and Sleeman, 2006; Yilmaz and Christofori, 2010). Since migration constitutes such a key step in forming metastasis via EMT, we assessed the migratory capacity of Klf4 knockdown cells. Wound healing assay revealed that Klf4-depleted NMuMG cells migrated significantly faster than control counterpart even in the absence of TGF- $\beta$  (Figure 3a). We further validated these results by performing a scratch wound healing assay in cells stably knocked down for Klf4 and found similar results (Figure 3b). Moreover, transwell migration assays with these cells revealed a significantly higher chemo-tactic migration of Klf4 knockdown cells compared to control cells both in the absence and presence of TGF- $\beta$  (Figure 3c and 3d). These studies were repeated in Py2T cells which further confirmed and validated the above findings (data not shown).

Klf4 has been shown to inhibit proliferation and promote differentiation of skin and colonic epithelium (Segre et al., 1999; Shie et al., 2000). Thus, we next investigated whether Klf4 depletion can affect the proliferation and survival during EMT in NMuMG cells. It was surprising that the loss of Klf4 led to a decrease in growth of NMuMG cells after treatment with TGF- $\beta$ . On the other hand, we didn't observe any changes after Klf4 knockdown in the absence of TGF- $\beta$ . To determine whether this effect was due to alterations in proliferation, or changes in apoptosis profile, we performed a proliferation assay using propidium iodide staining as well as an apoptosis assay using Annexin V staining. In comparison to control siRNA treated cells, a significant G0/G1 blockade was observed upon Klf4 depletion in the presence of TGF- $\beta$  for 2 days and 4 days (Figure 3e). We observed a similar effect in proliferation assay (Figure 3f). In addition, there was increased apoptosis upon Klf4 depletion after treatment with TGF- $\beta$  for 2

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition



**Figure 3: KLF4 inhibits migration and provides a survival advantage during TGF- $\beta$ -induced EMT.** (a) Following siRNA-mediated ablation of Klf4 in the absence of TGF- $\beta$ , wound healing assays were performed after 19 hours of wound creation. (b) Wound healing assay was performed in stably transfected Klf4-ablated cells. Pictures were taken at 0 hour and 19 hours after wound creation. (c) shRNA-mediated Klf4 depleted cells were subjected to Boyden chamber migration assays for 20 hours using 20% FBS as a chemoattractant in the absence of TGF- $\beta$ . (d) Boyden chamber migration assays were carried out in stably transfected Klf4 depleted cells in the presence of TGF- $\beta$ . 20% FBS was used as a chemoattractant. (e) Cell cycle analysis was done after transient knockdown of Klf4 during EMT for 0, 1 and 4 days by FACS. Propidium Iodide (PI) was used for the staining. (f) Proliferation assays were performed after siRNA-mediated knockdown of Klf4 during a TGF- $\beta$  time-course. Cells were counted by using a Neubauer counting chamber. (g) Annexin V staining was performed by FACS to quantify cell death in Klf4 knockdown cells during TGF- $\beta$ -induced EMT for 0, 2 and 15 days. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

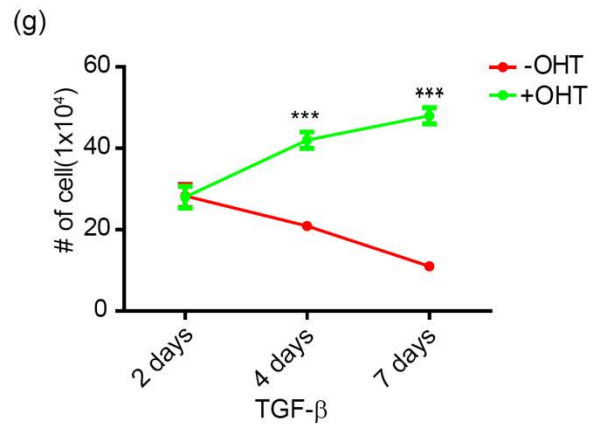
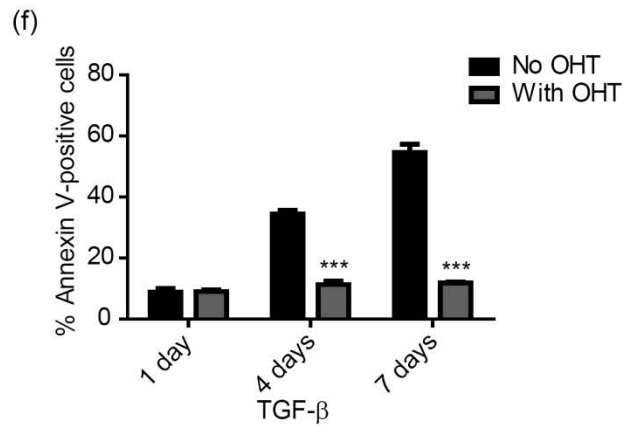
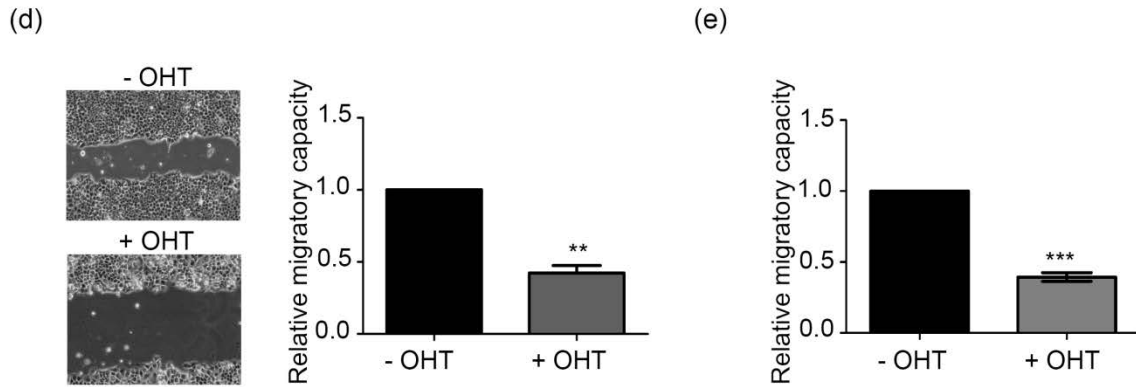
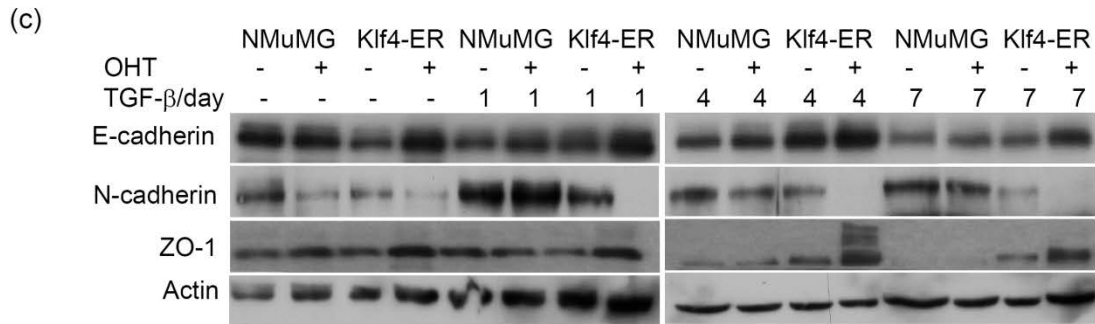
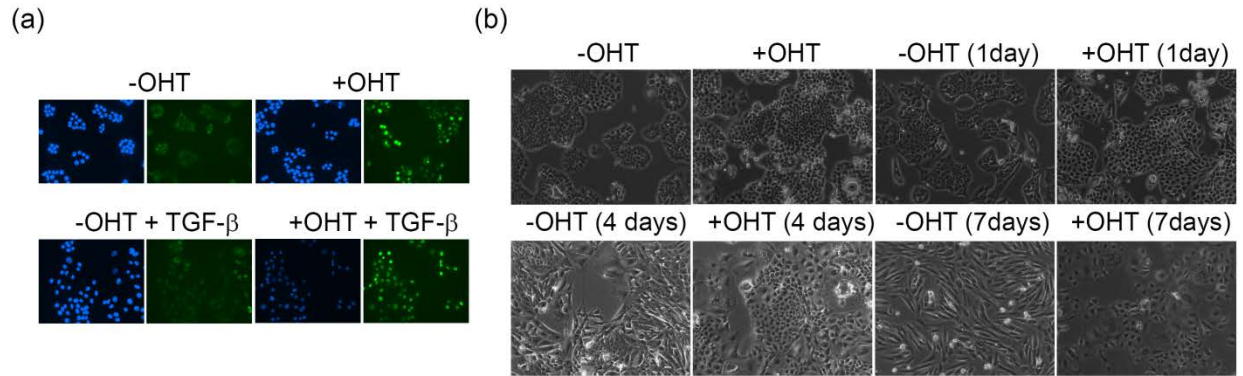
days or 15 days (Figure 3g). Thus, the TGF- $\beta$  sensitive growth of Klf4 knockdown cells depends on both decreased proliferation and increased apoptosis during EMT. We further extended our studies to Py2T cells and found that unlike NMuMG cells, these cells do not undergo apoptosis during TGF- $\beta$  treatment suggesting that the apoptotic phenotype is biased toward untransformed cells. We did observe a significant increase in apoptosis after Klf4 knockdown in Py2T cells, however, the fractions were not as substantial as NMuMG cells (Supplemental figure S3a). Furthermore, we could not notice any significant difference in cell cycle profiling after Klf4 depletion during TGF- $\beta$  time-course (Supplemental figure S3b).

Together, these results reveal an inhibitory role for Klf4 in migration as well as supporting role for Klf4 in proliferation and G<sub>1</sub>/S phase progression of non-transformed mammary epithelial cells and in cell survival during TGF- $\beta$ -induced EMT.

#### **3.1.3.4 Klf4 over-expression prevents EMT**

Given the above findings, we next asked whether the forced expression of Klf4 could have an inhibitory effect on EMT. We generated NMuMG cells expressing a Klf4-ER fusion protein (Foster et al., 2005) where the Klf4-ER fusion protein translocates to the nucleus upon treatment

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition





**Figure 4: Over-expression of KLF4 prevents EMT and migration and supports cell survival during EMT.** (a) Immunofluorescence staining by using a myc antibody to visualize the localization of Klf4 after 4-OHT treatment in the absence as well as presence of TGF- $\beta$  (2 days) in NMuMG cells stably transfected with a myc-Klf4-ER construct. Cells were treated with 1 $\mu$ M OHT for 2 days and compared with non-treated cells. (b) Morphological changes are observed in non-treated and 1 $\mu$ M 4-OHT treated myc-Klf4-ER transfected cells during TGF- $\beta$  time-course for 0, 1, 4 and 7 days. Original magnification was 10X. (c) Immunoblot to assess the expression of adherent junction proteins E-cadherin and N-cadherin and tight junction protein ZO-1 during TGF- $\beta$ -induced EMT in NMuMG and Klf4-ER transfected cells in the presence and absence of 4-OHT. Actin is used as loading control. (d) Wound healing assay to assess the migratory capacity after Klf4 over-expression. Cells were either treated with 1 $\mu$ M 4-OHT or not. Experiment was done for 19 hours. (e) Trans-well migration assay was carried out in the absence and presence of 1 $\mu$ M 4-OHT by using 20% FBS as a chemoattractant for 20 hours. (f) Annexin-V staining was performed by FACS to determine the percentage of dead cells during TGF- $\beta$ -induced EMT in Klf4-ER transfected cells in the presence and absence of 1 $\mu$ M 4-OHT. (g) Proliferation assays were performed during TGF- $\beta$  time-course to assess the growth rate in Klf4-ER transfected cells in the absence and presence of TGF- $\beta$ . Cells were counted by using a Neubauer counting chamber. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

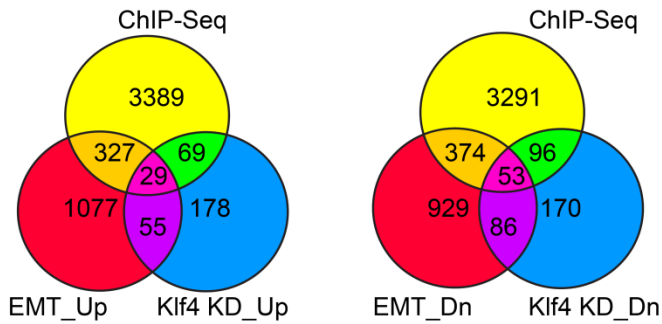
with the Tamoxifen derivative, 4- hydroxy tamoxifen (4-OHT) (Figure 4a). We induced Klf4 over-expression and followed morphological changes in the presence and absence of TGF- $\beta$ . Interestingly, despite the presence of TGF- $\beta$ , Klf4 overexpressing cells retained their epithelial phenotype (Figure 4b) which further accompanied gain in the epithelial markers like E-cadherin and ZO-1 and loss of mesenchymal markers such as N-cadherin as shown by immunoblot analysis (Figure 4c). Scratch wound healing assay and transwell migration assay in Klf4 overexpressing cells suggested that Klf4 functions inhibited cell migration (Figure 4d and 4e). In addition, Klf4 over-expression also blocked TGF- $\beta$ -induced apoptosis (Figure 4f) and cells grew better than control counterpart as shown by the proliferation assay (Figure 4g). These findings were further validated in Py2T cells and the results were compatible with the above findings (Supplemental figure S4a and S4b). These results suggested that, being a transcription factor, Klf4 probably elicited a transcriptional program that helped maintain epithelial phenotype and acted inhibitory to EMT.

### **3.1.3.5 KLF4 regulates the expression of crucial EMT genes by directly binding to their promoters**

Our observations and the previous established role of Klf4 as a transcriptional regulator led us to attempt identifying genes that are regulated by Klf4 through which it acts in maintaining epithelial phenotype and in repressing EMT. Towards this, we carried out siRNA mediated knockdown of Klf4 in NMuMG cells in the absence and presence of TGF- $\beta$  (2 days) and performed a global transcription profiling. This led to the identification of a number of genes that were significantly up and downregulated in Klf4-deficient cells (Figure 5a and Supplemental figure S5a-d). However, we speculated that this list may also contain genes which are indirect target. In order to reveal the transcriptional changes that are directly linked to Klf4 occupancy at their promoters we attempted to identify the genome-wide targets of Klf4. This was achieved by Chromatin Immunoprecipitation (ChIP) using a Klf4-specific antibody followed by next generation sequencing (ChIP-seq) in NMuMG cells. Computational analysis led to the identification of gene promoters directly bound by Klf4 in these cells. We next compared this list with the transcription profile for all the stages from the EMT time-course generated earlier (Figure 5a and 5b & Supplemental figure S5a-d). Interestingly, among the direct targets of Klf4 that are regulated during EMT, we identified N-cadherin. Quantitative real time RT-PCR showed that N-cadherin is significantly upregulated upon Klf4 depletion (data not shown). ChIP-qPCR using N-cadherin (Cdh2) promoter-specific primers confirmed its promoter to be bound by Klf4 (Figure 5c). In addition, Klf4 occupies promoters of many other crucial EMT genes like vimentin (Vim),  $\beta$ -catenin (Cttnb1), Inhibitor of differentiation 1 (Id1), Inhibitor of differentiation 2 (Id2), Vascular endothelial growth factor A (VEGFA) and Endothelin1 (Edn1) (Figure 5c and 5d). These results strongly suggest that Klf4 regulates transcription of key EMT genes by directly binding to their promoters. ChIP-qPCR data is further validated in Py2T cells, where also Klf4 binds at the promoter of key EMT genes such as N-cadherin, vimentin and  $\beta$ -catenin (Supplemental figure S5e).

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

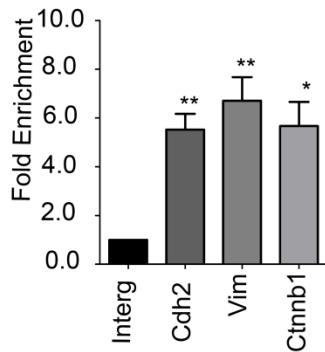
(a)



(b)

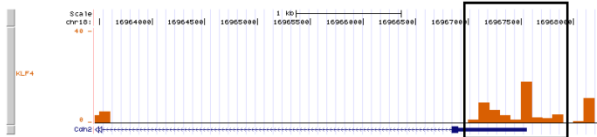
Upregulated	Downregulated	
1110067D22R	Anln	Kif23
Arrdc4	Arf2	Lbr
Arvcf	Calm3	Lgals3
Ccnj	Cav1	Mad2l1
Cdh2	Ccna2	Mcm10
Cyp1b1	Ccnd3	Mcm7
Edn1	Cdc25b	Mpp7
Ext1	Cdc6	Mthfd1
Fhl2	Cdt1	Nab1
Hist1h1c	Cept1	Nfib
Homer1	Chac1	Osgin1
Kdelr3	Chaf1b	PIK4
Kras	Chek1	Pola2
Lamb3	Chtf18	Ppl
Lats2	Cpt1a	Ppt2
Loxl4	Dut	Rmnd5a
Lpp	E2f8	S100a4
Lrrc58	Ect2	Sgol1
Marcks	Ehf	Shmt1
Mllt11	Emp1	Slc16a1
Mmp14	Enpp4	Slc25a1
Obfc2a	Gbe1	Spag5
Ppp3ca	Gdf15	Tcf19
Prkci	Hmgb2	Trim37
Rdh10	Incenp	Tspan4
Snx15	Kif11	Uap111
Vcl	Kif22	
Vegfa		
Wwc2		

(c)

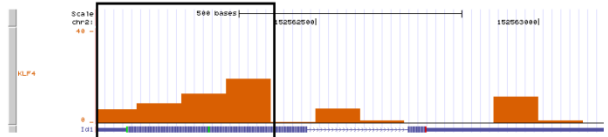


(d)

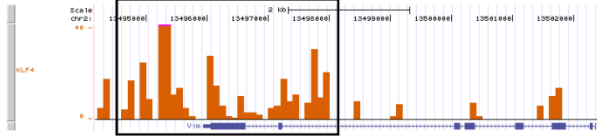
Cdh2 Promoter



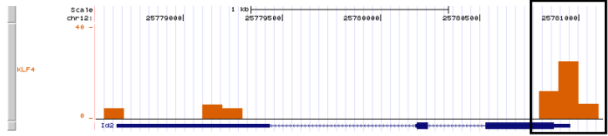
Id1 Promoter



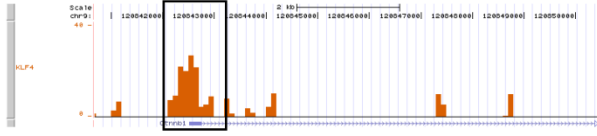
Vim Promoter



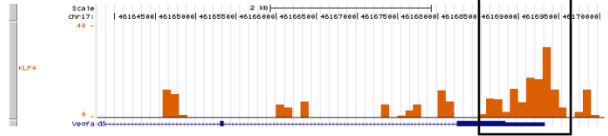
Id2 Promoter



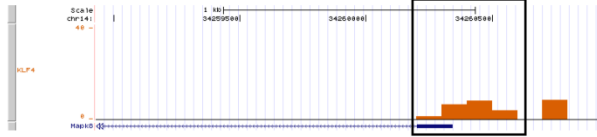
Ctnnb1 Promoter



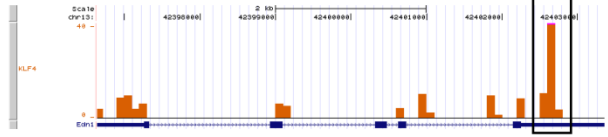
Vegfa Promoter



Mapk8 Promoter



Edn1 Promoter

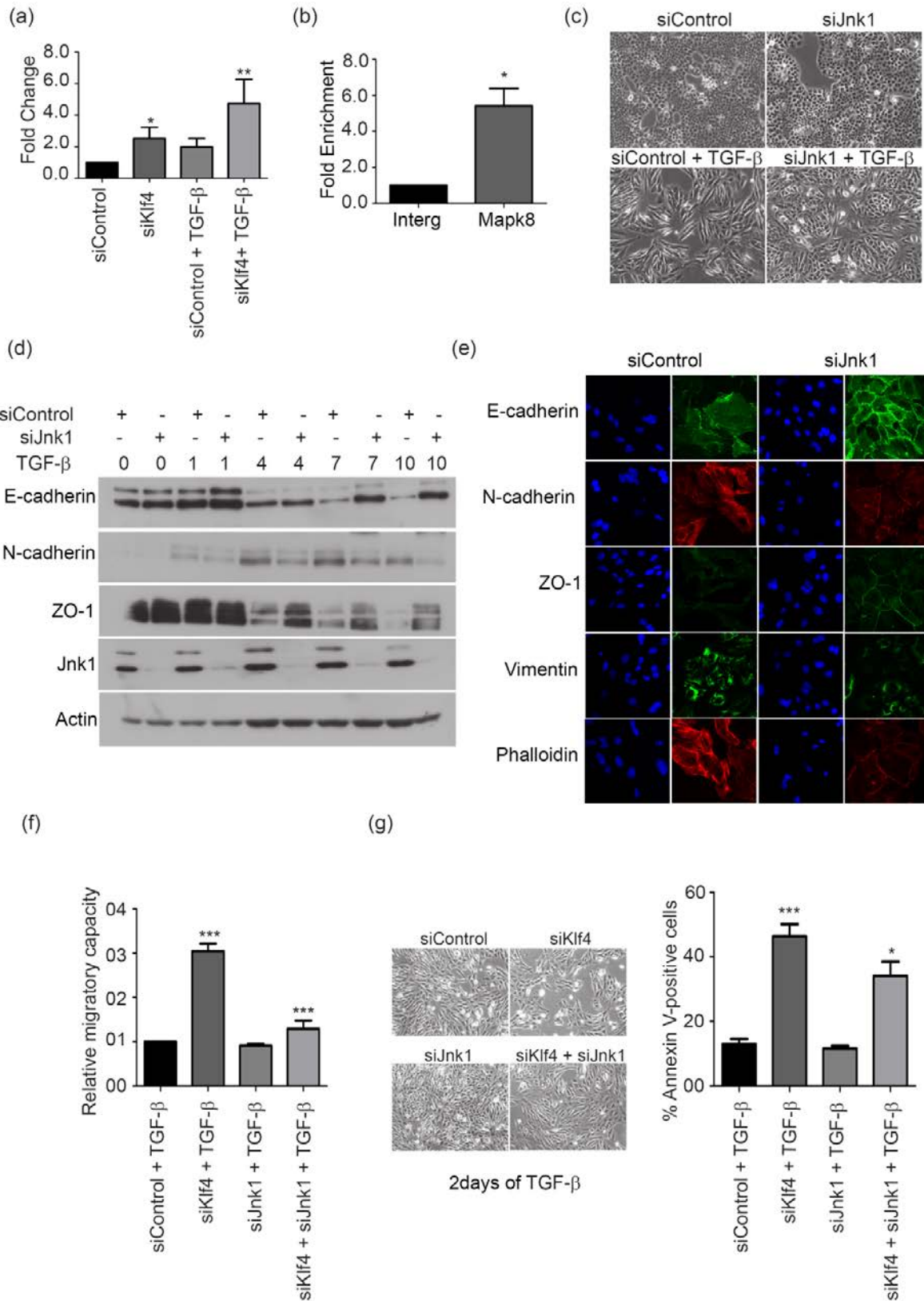


**Figure 5: Klf4 directly targets promoters of crucial EMT genes for transcriptional regulation.** (a) Venn-diagram showing the commonly regulated genes in Klf4 ChIP-Seq data, EMT time-course affymetrix array expression profile and Klf4 knockdown expression array. Left panel is for genes upregulated during EMT and right panel is for genes downregulated during EMT. All the analysis were done in NMuMG cells. (b) List of commonly regulated genes as mentioned above. (c) ChIP-PCR to see the occupancy of Klf4 at N-cadherin (Cdh2), vimentin (Vim) and  $\beta$ -catenin (Ctnnb1) promoters. ChIP-PCR was followed by Klf4 chromatin immunoprecipitation by using a Klf4 specific antibody in NMuMG cells. Data is normalized to an intergenic region, which supposed to be transcription factors free. (d) Wiggle-tracks to show the binding of Klf4 at N-cadherin (Cdh2), Vimentin (Vim),  $\beta$ -Catenin (Ctnnb1), Jnk1 (Mapk8), Inhibitor of differentiation 1 (Id1), Inhibitor of differentiation 2 (Id2), Vascular growth factor A (VEGFA) and endothelin1 (Edn1) promoters. These files were collected after Klf4 ChIP-Seq on Klf4 ChIP material. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

### **3.1.3.6 Klf4 is regulated by canonical TGF- $\beta$ signaling but regulates non-canonical TGF- $\beta$ signaling during EMT**

It has been shown before that Klf4 directly binds to p53 promoter, and Klf4 depletion in breast cancer cells restores p53 levels and causes p53-dependent apoptosis (Rowland et al., 2005). Recently, p53 was also shown to be involved in EMT and modulate stem cell properties via miR-200 (Chang et al., 2011). Jnk, a group of mitogen-activated protein kinases that are activated by cytokines or environmental stress, participate in various signaling pathways, including apoptotic pathways. Activated Jnk1 can modify p53 post-translationally by phosphorylation and can activate it (Hu et al., 1997). We found Jnk1 to be transcriptionally significantly upregulated in Klf4-depleted NMuMG and Py2T cells (Figure 6a and Supplemental Figure S6a) and downregulated in Klf4 over-expressing NMuMG and Py2T cells (Supplemental Figure S6b and S6c). Given the central role of Jnk1 in p53-dependent apoptosis, we next investigated whether this could be involved in EMT. To assess whether Jnk1 is a direct target of Klf4, Chromatin Immunoprecipitation (ChIP) assay was performed using a Klf4-specific antibodies followed by ChIP-qPCR revealed Jnk1 promoter to be directly bound by Klf4 (Figure 6b and Supplemental figure S6d) in both cell lines. Transient knockdown of Jnk1 using a pool of two siRNA led to a significant reduction in Jnk1 levels (Supplemental Figure S6e). Interestingly, similar to Klf4 over-expressing cells, Jnk1 depletion blocked epithelial differentiation and leads to retention of

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition



**Figure 6: KLF4 directly binds Jnk1 promoter for its transcriptional regulation and depletion of Jnk1 phenocopies the over-expression of KLF4 during TGF- $\beta$ -induced EMT.** (a) RT-qPCR for measurement of Jnk1 expression levels upon Klf4 knockdown. The cells were treated with TGF- $\beta$  for 2 days. (b) Following Chromatin Immunoprecipitation (ChIP) assay using Klf4-specific antibodies in NMuMG cells. Realtime PCR was performed using Jnk1 promoter-specific primers to test Klf4 occupancy at this region. (c) Phase contrast microscopy was performed for studying morphology of NMuMG cells transfected with siControl and siJnk1 is monitored before and after TGF- $\beta$  treatment. Cells were treated with TGF- $\beta$  for 2 days. (d) Expression levels of the epithelial markers, E-cadherin and ZO-1 and the mesenchymal marker N-cadherin were measured by immunoblot analysis in cells treated with either control of Jnk1-specific siRNA during a TGF- $\beta$  time-course. Actin was used as a loading control. Jnk1 siRNA knockdown efficiency was checked by using Jnk1 antibody. (e) The localization and expression levels of indicated EMT markers after 7 days of TGF- $\beta$  treatment in siControl and siJnk1 transfected cells was assessed by immunofluorescence. (f) Wound healing assays were performed to check the migratory capacity after single knockdown of Klf4 and Jnk1 and double-knockdown of Klf4 and Jnk1. Assays were done for 19 hours. (g) Apoptosis was measured by Annexin V staining in combination with FACS analysis in control (siControl) and Klf4-depleted (siKlf4), Jnk1-depleted (siJnk1) and siKlf4 with siJnk1 in NMuMG cells treated with TGF- $\beta$  for days for 2 days. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

epithelial morphology even in the presence of TGF- $\beta$  in NMuMG cells (Figure 6c). These morphological changes also accompanied corresponding changes in the expression of epithelial markers such as E-cadherin and ZO-1 and loss of mesenchymal markers like N-cadherin during TGF- $\beta$ -induced EMT (Figure 6d). These data were further validated by immunofluorescence analysis after TGF- $\beta$  treatment for 7 days for various EMT markers and showing that Jnk1 depletion led to a retention of the adherence junction and tight junction proteins E-cadherin and ZO-1, respectively at the membrane. This was also accompanied by a reduction in the mesenchymal marker N-cadherin. We did not observe any differences in the intermediate filamentous mesenchymal marker vimentin, although, TGF- $\beta$  mediated cytoskeleton re-modeling is highly prevented as shown by reduced phalloidin staining for stress fibers visualization (Figure 6e). Furthermore, double-knockdown of Klf4 and Jnk1 revealed that Klf4 knockdown-induced EMT could be rescued by Jnk1 depletion and all markers reverted back to normal levels (data not shown). Furthermore, the double-knockdown of Klf4 and Jnk1 not only prevented Klf4 knockdown-induced migration (Figure 6f) but also Klf4 knockdown-induced apoptosis (Figure 6g). These experiments strongly argue for a role of Klf4 in regulating EMT by a direct

transcriptional control of Jnk1.

To gain insight whether Klf4 is regulated by canonical or non-canonical TGF- $\beta$  signaling, we analyzed the expression of Klf4 in NMuMG cells stably knocked-down for Smad4 and found Klf4 to be significantly upregulated in Smad4 depleted cells upon TGF- $\beta$  treatment. This implicated that Klf4 is a target of canonical TGF- $\beta$ -signaling, while Klf4 is a downstream target of Smad4 (Supplemental figure S6f).

In summary, these results strongly argue that Klf4 regulates non-canonical TGF- $\beta$  signaling by targeting Jnk1 but itself is regulated by canonical TGF- $\beta$  signaling.

### **3.1.4 Discussion**

EMT is an orchestrated series of events in which cell-cell and cell-extracellular matrix (ECM) interactions are altered to release epithelial cells from the surrounding tissue, the cytoskeleton is reorganized to confer the ability to move through a three-dimensional ECM, and a new transcriptional program is induced to maintain the mesenchymal phenotype. Essential for embryonic development, EMT is nevertheless potentially destructive if deregulated, and it is becoming increasingly clear that inappropriate utilization of EMT mechanisms is an integral component of the progression of many tumors of epithelial tissues.

In this study, using genome-wide expression profiling during ongoing EMT process, we were able to predict transcription factors which may regulate EMT by modulating the expression of their target genes. Klf4 was identified as one of such factors whose targets were over-represented among genes that were transcriptionally modulated during the EMT process in our system. Our data revealed an inhibitory role of Klf4 during EMT as the absence of Klf4 in breast cancer cell lines lead to an acceleration of EMT and higher migration potential. In addition, in contrast to a study in colon cancer where Klf4 inhibits proliferation by blocking G1/S progression of cell cycle (Chen et al., 2001), we found that a part of the phenotypic effect of Klf4 could also be attributed to its property to infer a survival advantage to cells undergoing EMT. It is also possible that the decrease in proliferation observed after Klf4 depletion result from E-cadherin silencing. E-cadherin also has both growth inhibitory (Perrais et al., 2007; St Croix et al., 1998) and growth promoting roles similar to Klf4 (Fournier et al., 2008; Liu et al., 2006b) and this

function is suggested to mainly depend on E-cadherin expression levels. Blocking E-cadherin engagement abrogated the elevated proliferation levels at intermediate seeding densities, but no inhibition of proliferation was observed when cells were confluent, which again points to the fact that similar to E-cadherin, low expression of Klf4 in mesenchymal state is required for maintaining the basal level of proliferation. Moreover, genome-wide ChIP-seq analysis also demonstrated that Klf4 directly occupies the promoters of crucial EMT genes, including N-cadherin, vimentin,  $\beta$ -catenin, Jnk1, Id1, Id2, Edn1 and VEGFA, a number of which are deregulated upon Klf4 depletion. Since, Klf4 directly binds to the  $\beta$ -catenin promoter and ablation of Klf4 leads to a decrease in its expression levels (data not shown), it is possible that increased migration observed after Klf4 depletion is due to a decrease in the expression of  $\beta$ -catenin which further contributes to low levels of E-cadherin and liberates cells for the migration. In addition, the mesenchymal-specific isoform of p120 is also increased after Klf4 knockdown during EMT (data not shown), resulting in increased migratory capacity of these cells. Increased expression of the mesenchymal-specific isoform of p120 is has been shown to be responsible for increased migration (Mo and Reynolds, 1996; Yanagisawa et al., 2008). We further show that one of the Klf4 genomic targets, Jnk1, can not only counteract Klf4 knockdown-induced EMT but also migration and apoptosis.

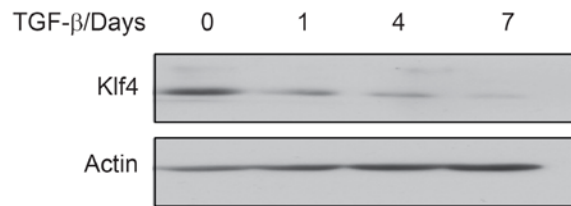
Klf4 is known for its dual role as a transcriptional activator and repressor. Our study emphasizes that Klf4 acts as a repressor in breast carcinogenesis. High Klf4 expression is correlated with metastasis free survival in the Uppsala breast cancer database (Bergh et al., 1995) and TRANSBIG database [(Desmedt et al., 2007); data not shown]. Certain studies have suggested an involvement of Klf4 (Yori et al., 2010) and Jnk1 (Alcorn et al., 2008; Santibanez, 2006b; Velden et al., 2011) in EMT, however, the underlying mechanisms and the relationship between these two remained unknown. In contrast to a previous report (Yori et al., 2010), our unbiased genomewide binding assay as well as single gene analysis did not reveal any Klf4 binding at the E-cadherin promoter. Importantly however, our genome-wide ChIP-Seq data analysis identified the promoters of N-cadherin as well as many other key EMT genes such as vimentin and  $\beta$ -catenin to be a direct target of Klf4 for transcriptional regulation. Moreover, Klf4 also regulates



many genes required for angiogenesis, including VEGFA and Edn1, further implicating it in the angiogenic switch during tumor formation and metastasis.

Together, these data reveal fundamental principles of how transcription factors like Klf4 regulate cell-fate changes such as during EMT by directly modulating transcription of underlying genes. These findings have potential to provide fundamental knowledge that has relevance for a broad range of research disciplines encompassing basic to translational research and will possibly pave the way for the development of therapeutic approaches involving such factors as targets in the long-term.

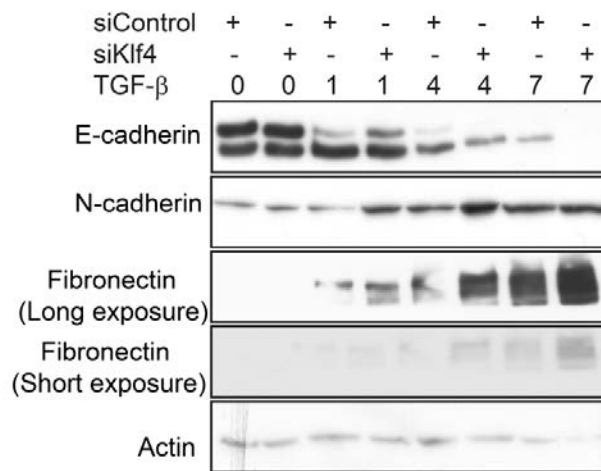
### 3.1.5 Supplemental figures



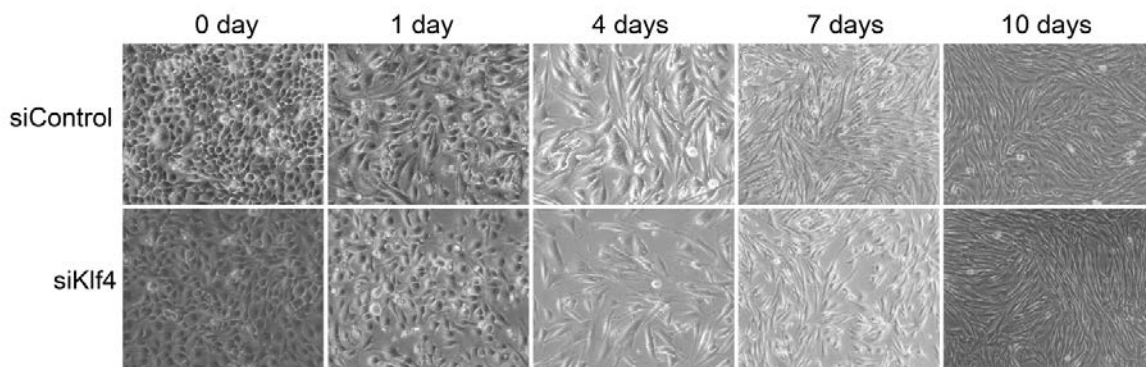
**Figure S1: Klf4 is downregulated during EMT.** Immunoblot for Klf4 during TGF- $\beta$ -induced EMT in NMuMG cells.

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

(a)



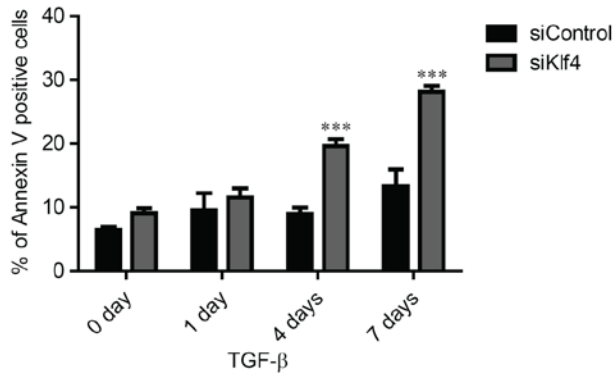
(b)



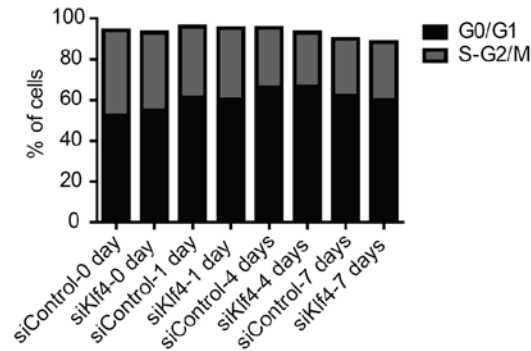
**Figure S2: Klf4 depletion accelerates EMT in Py2T cells.** (a) Western blot analysis for the epithelial protein E-cadherin and the mesenchymal proteins N-cadherin and fibronectin during TGF- $\beta$ -mediated EMT in Py2T cells after Klf4 reduction. Actin is used as a loading control. (b) Morphological changes were monitored after Klf4 ablation during a TGF- $\beta$  time-course in Py2T cells. Original magnification was 10X.

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

(a)



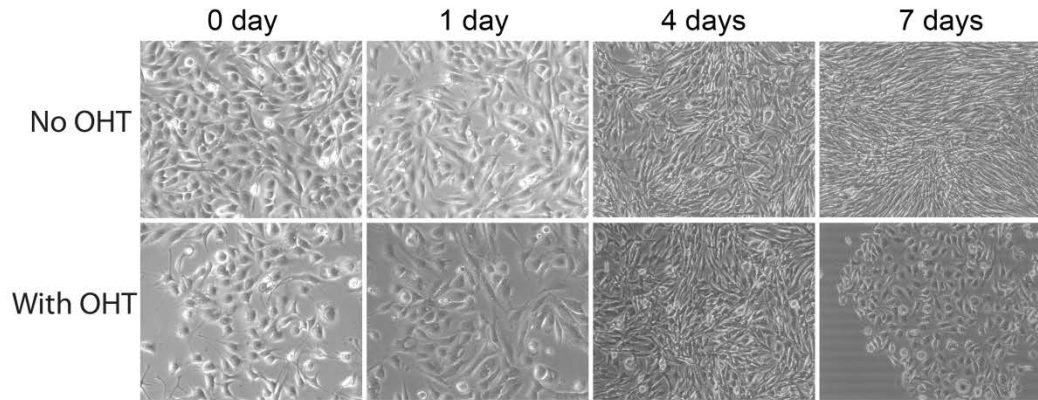
(b)



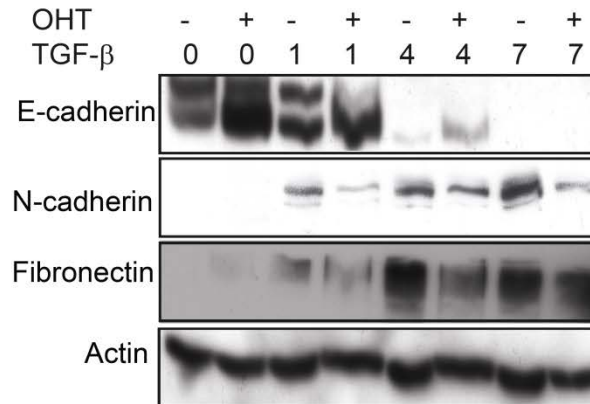
**Figure S3: Klf4 provides survival advantage to the cells during EMT in Py2T cells:** (a) Annexin V staining was performed to assess the percentage of dead cell during TGF- $\beta$  conducted EMT after Klf4 ablation. Cells were treated with TGF- $\beta$  for 0, 1, 4 and 7 days. (b) Propidium iodide staining was carried out to assess the changes in cell cycle phases during TGF- $\beta$  time-course in Py2T cells after Klf4 reduction. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.001$  indicated with (\*\*\*)

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

(a)

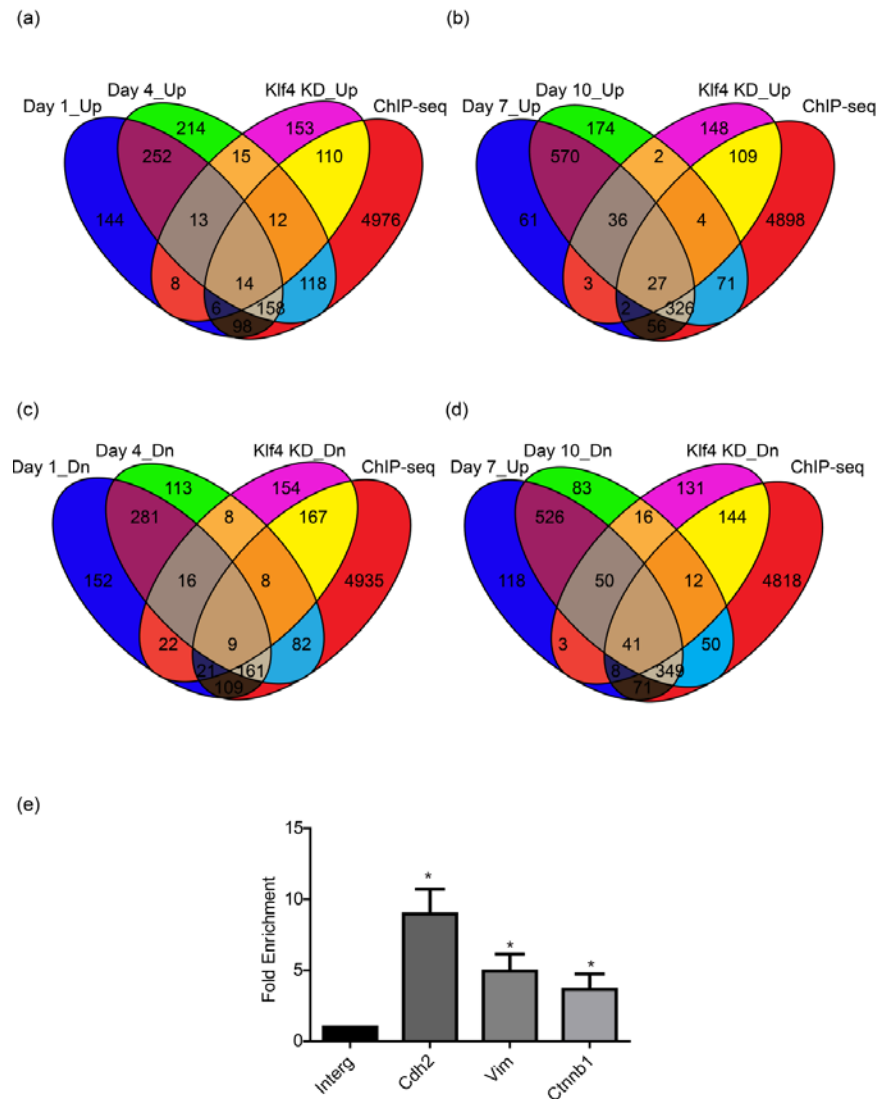


(b)



**Figure S4: Klf4 over-expression prevents epithelial differentiation in Py2T cells:** (a) Morphological changes were observed after Klf4 induction by using 1 μM 4-OHT during TGF-β-induced EMT in Py2T cells. Original magnification was 10X. (b) Expression levels of various EMT markers were assessed by doing an immunoblotting analysis after Klf4 induction with 1 μM 4-OHT. Cells were treated with TGF-β for 0, 1, 4 and 7 days. The epithelial marker E-cadherin and the mesenchymal markers N-cadherin and fibronectin were analyzed for their expression levels.

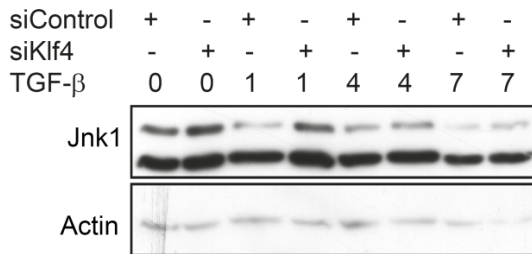
Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition



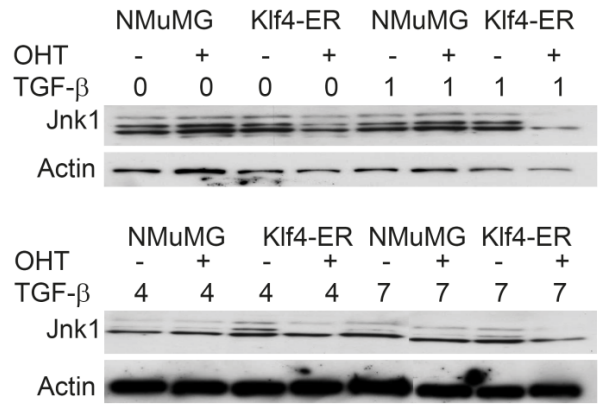
**Figure S5: Klf4 directly binds at the promoters of key EMT genes:** (a-d) Detailed comparison between ChIP-seq data after performing Klf4 ChIP, affymetrix array data after conducting chip in Klf4 knockdown cells in the absence and presence of TGF- $\beta$  (2days) and affymetrix array data after carrying out chip during TGF- $\beta$  mediated EMT (Day 0, 1, 4, 7 and 10). Lists generated after Klf4 knockdown array analysis, both in the presence and absence of TGF- $\beta$ , were merged together for upregulated and downregulated genes. Panel (a) and (b) comprise the genes which were upregulated during EMT as well as upregulated in Klf4 knockdown array. Panel (c) and (d) comprise the genes which were downregulated during EMT as well as downregulated in Klf4 knockdown array. All experiments were carried out in NMuMG cells (e) Klf4 ChIP was performed in Py2T cells, followed by quantitative-PCR to test the occupancy of Klf4 at Cdh2 (N-Cadherin), Vim (Vimentin) and Ctnnb1 ( $\beta$ -Catenin) promoters. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*).

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

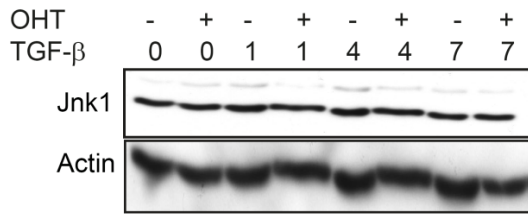
(a)



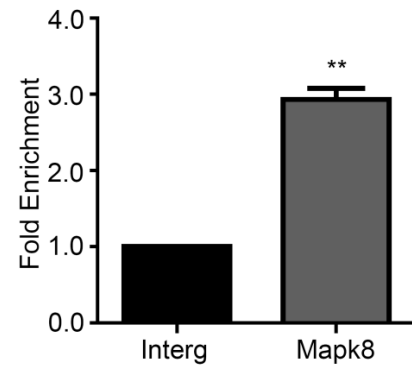
(b)



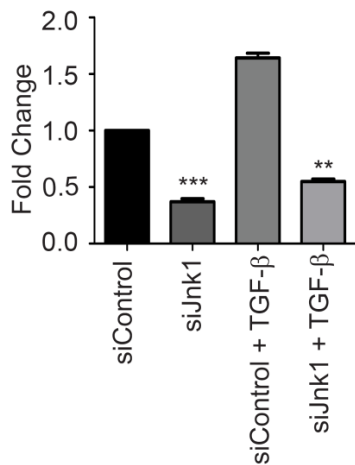
(c)



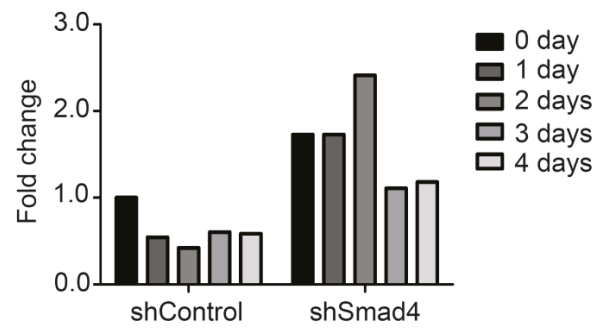
(d)



(e)



(f)



**Figure S6: Klf4 regulates non-canonical TGF- $\beta$  signaling via Jnk1 but itself is regulated by canonical TGF- $\beta$  signaling:** (a) Immunoblot analysis of the expression of Jnk1 during TGF- $\beta$ -induced EMT in Py2T cells after Klf4 depletion. (b-c) Immunoblot was performed to check the Jnk1 expression after Klf4 induction with 1 $\mu$ M 4-OHT in NMuMG (b) and Py2T (c) cells after treating them with TGF- $\beta$  for 0, 1, 4 and 7 days. (d) Quantitative-RT-PCR was carried out after Klf4 chromatin immunoprecipitation in Py2T cells to check the occupancy of Klf4 at Mapk8 promoter which encode for Jnk1. (e) Quantitative-RT-PCR was performed to assess the knockdown efficiency of Jnk1 siRNAs pool in the absence and presence of TGF- $\beta$  in NMuMG cells. Cells were treated with TGF- $\beta$  for 2 days. (f) Q-PCR was conducted after stable knockdown of Smad4 to monitor the expression level of Klf4 during TGF- $\beta$  mediated EMT in NMuMG cells. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq$  0.05 indicated with (\*), p-value  $\leq$  0.01 indicated with (\*\*), p-value  $\leq$  0.001 indicated with (\*\*\*)

### 3.1.6 Methods and Materials

#### Reagents and antibodies

**Reagents:** TGF- $\beta$  (240-B, R&D systems). DMEM (D5671, Sigma-Aldrich), PBS (D8537, Sigma-Aldrich), trypsin (T4174, Sigma-Aldrich), Opti-MEM (11058, Gibco), FBS (F7524, Sigma-Aldrich), Glutamine (G7513, Sigma-Aldrich), Pencillin/streptomycin (P4333, Sigma-Aldrich), Lipofectamine RNAiMax (11668-019, Invitrogen), Alexa Fluor-488, 568, 633 (Invitrogen), Polybrene (AL-118, Sigma-Aldrich), Puromycin (P7255, Sigma-Aldrich), Fugene HD (12998300, Roche), Trizol (T9424, Sigma-Aldrich), M-MLV reverse transcriptase (M314C 28692233, Promega), SYBR-green PCR MasterMix (Eurogentec) and Bradford reagent (500-0006, Biorad), Protease inhibitor cocktail (P2714, Sigma-Aldrich). **Antibodies: *Western Blot:*** E-Cadherin (610182, Transduction Laboratories), N-Cadherin (M142, Takara), ZO-1 (617300, Zymed), Fibronectin (F-3648, Sigma-Aldrich), Actin (SC-1616, Sanata Cruz Biotechnology). ***Immunofluorescence:*** E-Cadherin (13-1900, Zymed), N-Cadherin (610921, Transduction Laboratories), ZO-1 (617300, Zymed), Phalloidin (A12380, Invitrogen), Paxillin (13520, Transduction Laboratories), Vimentin (V2258, Sigma). **Apoptosis and Cell cycle:** Annexin-V (559934, BD Biosciences) and PI (P4170, Sigma-Aldrich), **Small interfering RNAs:** siControl (Stealth RNAi™ siRNA Negative Controls, 12935-100, Invitrogen), siKlf4 (SASI\_Mm01\_00114972 and SASI\_Mm01\_00114974, Sigma-Aldrich) and siJnk1 (SASI\_Mm01\_00061987 & SASI\_Mm01\_00061988, Sigma-Aldrich). **Small hairpin RNA:**

shControl (Mission Non-target shRNA control vector, SHC002) and shKlf4 (SHCLNG-NM\_010637 Mouse, TRCN0000095370; TRCN0000095371 and TRCN0000095372).

### Cell lines and cell culture

A subclone of NMuMG cells (NMuMG/E9; hereafter NMuMG) expressing E-cadherin has been previously described (Maeda et al., 2005). MCF7 shControl and MCF7-shEcad have been described before (Lehembre et al., 2008). Py2T cells were derived from Polyoma middle T breast cancer tumor model (unpublished data, Waldameir et.al). NMuMG, MCF7shControl, MCF7-shEcad, Py2T, 293T and PLAT-E cells were cultured in DMEM supplemented with 10% FBS, 2mM glutamine, 100U penicillin and 0.2mg/ml streptomycin. All the cells were cultured at 37°C with 5% CO<sub>2</sub> in humid incubator. For TGF- $\beta$  time-course experiments, cells were treated with 2ng/ml TGF- $\beta$  for indicated time point and it was replaced every 2 days. For siRNA transfections, Lipofectamine RNAiMax was used according to the manufacturer's instructions.

### Quantitative RT-PCR

Total RNA was prepared by using a Tri Reagent according to the manufacturer's instructions. RNA was further reverse transcribed with ImProm-II Reverse Transcriptase, and transcripts were quantified by PCR using SYBR-green PCR Mastermix in a real time PCR system (Step One Plus, Applied Biosystems). Human or mouse ribosomal L19 primers were used for normalization. PCR assays were performed in duplicates, and fold induction was calculated against control-treated cell lines using the comparative Ct method ( $\Delta\Delta$  Ct). Following primers were used:

Primer name	Sequences
mRpl19 Forward primer	ctcgttgccggaaaaaca
mRpl19 Reverse primer	tcatccaggtcaccttctca
mKlf4 Forward primer	cctcgctctcctcgtcct
mKlf4 Reverse primer	tcgtcttcgaactcgtcgt
hRpl19 Forward primer	gatgccggaaaaacaccttg
hRpl19 Reverse primer	tggctgtacccttccgctt



Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

hKlf4 Forward primer	ccaaatcttttggggactttt
hKlf4 Reverse primer	ctggcccctcaactcctc
mE-cadherin Forward primer	cgaccctgcctctgaatcc
mE-cadherin Reverse primer	tacacgctgggaaacatgagc
mN-Cadherin Forward primer	caatgacgtccaccctgttct
mN-Cadherin Reverse primer	ctgccatgactttctacggaga
mFibronectin1 Forward primer	cccagacttatggtggcaatt
mFibronectin1 Reverse primer	atattccgactcgagtctga
mJnk1 Forward primer	aactgttccccgatgtgct
mJnk1 Reverse primer	tctcttgcctgactggcttt

### Immunoblot

Cells were lysed for 1 hour on ice in RIPA-Plus buffer (50mM Tris-HCl, pH8.0), 150mM NaCl, 10% glycerol, 1% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, 2mM CaCl<sub>2</sub>, 1mM dithiothreitol, 1mM sodium fluoride, 0.2mM sodium orthovanadate, 1x protease inhibitor cocktail and further quantified by using Bradford reagent. 50 µg of cleared protein lysates were separated by SDS-PAGE and electroblotted on PDVF membranes, and proteins were visualized with the appropriate primary and secondary antibodies and ECL on superRX films. Depending on the species origin of antibodies, immunoblots were either probed sequentially or on multiple membranes. Adobe Photoshop has been used to excise the relevant portion of the immunoblots from the original scans of X-ray films exposed to chemoluminescence visualization of specific proteins.

### Immunofluorescence

siControl, siKlf4 and siJnk1 cells were plated on coverslips and treated with TGF-β for mentioned time. The cells were fixed with 4% paraformaldehyde in HBSS and further permeablized with 0.2% Triton for 5 minutes at room temperature. These cells were blocked by using 3.5% goat serum for 15 minutes and incubated with primary antibodies against E-cadherin, N-cadherin, Fibronectin, ZO-1, vimentin, Paxillin and Phalloidin for 1 hour and then incubated

with fluoro-chrome-labeled secondary antibody for 1 hour at room temperature. The coverslips were counterstained with DAPI and imaged with a confocal laser-scanning microscope. Data were processed with Adobe Photoshop 7.0 software.

### **Production of lentivirus for knockdown studies**

Murine Klf4 shRNAs and control shRNA were purchased from Sigma-Aldrich as described above. For lenti-virus production, 293T cells were transfected with the shRNA expressing lenti-viral vector in combination with the packaging vectors including envelope protein; HDM-pVSV/G, codon-optimized HIV gag-pol; HDM-Hpgr2, transactivator of transcription; HDM-Tat1b and pRC-CMV-RaII by Fugene HD. After 48 hours of transfection, viral supernatant was harvested, filtered (0.46  $\mu$ m), supplemented with polybrene (8ng/ml) and used to infect target cells. Infections were performed once a day for two consecutive days. Infected cells were positively selected using Puromycin (5ug/ml).

### **Production of retrovirus for over-expression studies**

Klf4-ER (kindly provided by Prof. J.M. Ruppert) cloned into the retroviral expression vector pbabe. Retroviral particles were produced by transfecting PLAT-E cells with the retroviral expression vectors using Fugene HD. After two days of virus production, retroviral-containing supernatants were harvested, filtered (0.45  $\mu$ m) and added to target cells in presence of polybrene (8 ng/ml). Infections were performed once a day for two days, and later on selection was performed by using puromycin. Single clones were picked for the experiments.

### **Wound healing assay**

In vitro wound healing assays were done on confluent siControl, siKlf4, siJnk1 and siKlf4 + siJnk1 cells. The media on the confluent cells was replaced with DMEM with 2% fetal bovine serum media, and an area of cells was scraped off using a 200ul pipette tips. Light microscopic images were taken at time 0 and at 20 hours. The data was further analyzed by using an ImageJ program.

### **Migration assay**

Cell migration was assessed in shControl and shKlf4 cells in a transwell migration assay (pore size: 8  $\mu\text{m}$ ; Falcon BD). 104 cells were seeded in 2% FBS/DMEM (Sigma) in the upper chamber and the lower chamber was filled with 20% FBS/DMEM. After 20 hours of incubation at 37°C, cells in the upper chamber were carefully removed with a cotton swap, and the cells that had traversed the membrane were fixed in 4% paraformaldehyde/PBS, stained with DAPI. Pictures of the membrane were taken at a 10x magnification using a fluorescent microscope (Nikon Diaphot 300). Quantification was done using the software ImageJ.

### **Apoptosis assay (Annexin assay)**

Cells were washed twice with cold PBS and resuspended in 1X Annexin V binding buffer at a concentration of  $1 \times 10^6$  cells/ml. 5  $\mu\text{l}$  of Cy5 Annexin V was added to the 100  $\mu\text{l}$  of cells ( $1 \times 10^5$ ) and incubated for 15 min on ice in the dark. Stained cells were filtered through 40 $\mu\text{m}$  mesh and analyzed on a FACSCanto II using DIVA software.

### **Cell growth curve**

$1 \times 10^4$  cells were seeded in each well of 24-well plate, and cell numbers were assessed for mentioned days by using a Neubauer counting chamber.

### **PI Staining**

Cells were trypsinized and fixed in 70% ice-cold Ethanol for overnight. Washed twice with PBS and resuspended in sodium citrate buffer with 5 $\mu\text{g/ml}$  PI for overnight. Stained cells were analyzed by FACSCanto II using DIVA software.

### **Microarray processing and data analysis**

RNA was isolated from NMuMG cells transfected with control siRNA or Klf4 siRNA and treated with TGF- $\beta$  for 0 and 2 days using Tri Reagent (Sigma-Aldrich). RNA quality and quantity was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies). The manufacturer's protocols for the GeneChip platform by Affimetrix were followed. Methods included synthesis of the first- and second-strand cDNA followed by synthesis of cRNA by *in vitro* transcription, subsequent synthesis of single-stranded cDNA, biotin labeling and

fragmentation of cDNA and hybridization with the microarray slide (GeneChip® Mouse Gene 1.0 ST array), posthybridization washings and detection of the hybridized cDNAs using a streptavidin-coupled fluorescent dye. Hybridized Affimetrix GeneChips were scanned using an Affimetrix GeneChip 3000 scanner. Image generation and feature extraction were performed using Affimetrix GCOS Software and quality control was performed using Affimetrix Expression Console Software. Raw microarray data were normalized with Robust Multi-Array (RMA) and analyzed using Partek® Genomics Suite Software (Partek Inc.). One-way analysis of variance (ANOVA) and asymptotic analysis were used to identify significantly differentially expressed genes. The gene ontology (GO) tool from Partek® Genomics Suite Software as well as the David gene ontology software were used for further analysis.

### **Chromatin Immunoprecipitation**

ChIP experiments were performed as previously described (Weber et al, 2007). In brief, crosslinked chromatin was sonicated to achieve an average fragment size of 500 bp. Starting with 100 µg of chromatin and 5 µg of anti-HA antibody, 1 µl of ChIP material and 1 µl of input material were used for quantitative real-time PCR using specific primers covering the motif of Klf4 in the promoter of target genes. Primers covering an intergenic region are used as a control. The efficiencies of PCR amplification were normalized for between the primer pairs. Following primers were used for ChIP –PCR.

Name	Sequences
Cdh2 FP	ggttgctgtagccatgtt
Cdh2 RP	cattccctttctgctttg
p53 FP	ggtcaagtggagaagggtga
p53 RP	gcacctcgagagaaggacac
Junb FP	ttccttcttgtaactctg
Junb RP	agatcacctactttccacca
Vim FP	gtttgcattgagttccattg
Vim RP	gtcacgaacagcagagagaaga
Ctnnb1 FP	agtctcaaagtgtgggacat

Klf4 directly regulates transcription of genes crucial for  
Epithelial to Mesenchymal transition

Ctnnb1 RP	ttatgctgaccttccacactg
Interg FP	attctcctgcaaaggaaacaaa
Interg RP	ggggctcagagtaggttatgtg

### **Statistical analysis**

Statistical analysis and graphs were generated using the GraphPad Prism software (GraphPad Software Inc, San Diego, CA). All statistical analysis was done by unpaired, two-sided t-test. Normality testing was performed using the Kolmogorov-Smirnov test with Dallal-Wilkinson-Lillie for p-values.

### **3.2 Sox4 regulates Epithelial to Mesenchymal Transition by directly controlling transcription of underlying master genes**

Neha Tiwari, Phil Arnold, Mikhail Pachkov, Nathalie Meyer-Schaller, Erik Van Nimwegen and Gerhard Christofori

*Manuscript submitted*

#### **3.2.1 Abstract**

Epithelial-to-mesenchymal transition (EMT) is a key process during organismal development and in the progression of solid epithelial tumors to invasive and metastatic cancers. It involves loss of various epithelial markers, which are required for normal integrity of cells and at the same time gain of various mesenchymal markers which make the cells more migratory and invasive. Using a genome-wide expression profiling approach, we uncovered Sox4 among the transcription factors that were significantly upregulated during TGF- $\beta$ -induced EMT in NMuMG cells. Sox4 depletion during TGF- $\beta$ -induced EMT leads to a retention of epithelial morphology and decrease in cell migration properties. In addition, Sox4 provides a survival advantage to cells during EMT. Importantly, Sox4 is also required for breast cancer metastasis and tumorigenesis *in vivo*. Interestingly, Chromatin Immunoprecipitation experiments reveal that Sox4 directly binds to the promoter and regulates expression of Ezh2, a Polycomb Group (PcG) Complex gene that encodes for the enzyme that trimethylates H3K27 for gene repression. This results in the activation of EMT, ultimately leading to enhanced invasiveness of epithelial cells. Taken together, our results provide a novel mechanism of EMT regulation by transcription factor Sox4, where in addition to controlling the expression of crucial EMT genes including Spred1, Edn1, Pald, Cyr61, Ereg, Areg and Yap1, it also contributes to EMT by regulating transcription of the epigenetic machinery component Ezh2.

#### **3.2.2 Introduction**

Epithelial–mesenchymal transition (EMT) is a cellular mechanism long known to constitute the core of normal embryonic development. Several critical development events, including gastrulation, neural crest formation and heart morphogenesis, rely on the plastic transition between epithelium and mesenchyme (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and

Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006). In the last few years it has been uncovered that similar but physio-pathological transitions occur during the progression of epithelial tumors, endowing cancer cells with increased motility and invasiveness (Hanahan and Weinberg, 2011; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006). Multiple oncogenic pathways mediated by peptide growth factors, TGF- $\beta$ , Src, Ras, Ets, integrin, Wnt/ $\beta$ -catenin and Notch signaling, are implicated in induction of EMT (Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Nieto, 2010; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006) .

Sox4 ( SRY-Related HMG-Box Gene 4) is a member of the Sox (SRY-related HMG-box) family of transcription factors and it has been shown to have a role in embryonic development and in cell-fate determination during organogenesis including heart (Restivo et al., 2006; Schilham et al., 1996), pancreas (Lioubinski et al., 2003; Wilson et al., 2005) and brain (Cheung et al., 2000a; Hong and Saint-Jeannet, 2005). Sox4 has been shown to be a lymphocyte specific-transcriptional activator (van de Wetering et al., 1993) and facilitates B and T cell differentiation (Cheung et al., 2000a; Hong and Saint-Jeannet, 2005; Lioubinski et al., 2003; Schilham et al., 1997; Schilham et al., 1996; van de Wetering et al., 1993; Wilson et al., 2005). Moreover, Sox4 gene expression is up-regulated in many tumor types, with experimental evidence suggesting that this contributes to cellular transformation (Liu et al., 2006a; Shin et al., 2004), control of apoptosis (Aaboe et al., 2006; Ahn et al., 2002; Liu et al., 2006a; Pramoonjago et al., 2006) and/or a metastatic phenotype (Liao et al., 2008; Tavazoie et al., 2008). Restoration of miR335 expression in malignant cells suppressed lung and bone metastasis in human cancer cells through targeting the progenitor cell transcription factor Sox4 and extracellular matrix component tenascin C (Tavazoie et al., 2008). Sox4 has also been shown to directly regulate key cellular regulators like EGFR, TNC, HSP70, FZD5, DLL1, PTCH1, various transcription factors like MLL, FOXA1, ZNF281 and NKX3-1 and components of the RNAi machinery including Dicer, Argonaute 1, and RNA Helicase A in prostate cancer (Scharer et al., 2009). Moreover, Sox4 has been shown to regulate Wnt signaling by directly binding to  $\beta$ -catenin and interacting with various Tcf family members (Sinner et al., 2007). Recently, organogenesis has been suggested to rely on SoxC transcription factors, which comprise Sox4, Sox11 and Sox12, for the survival of

neural and mesenchymal progenitors and in part by activating Tead2, which is a mediator of Hippo signaling (Bhattaram et al., 2010).

Ezh2 (Enhancer of Zeste, Drosophila, Homolog 2) is a histone methyltransferase that belongs to the polycomb group (PcG) complex 2 and is known to be involved in epigenetic regulation of genes involved in cell fate decisions. It specifically trimethylates nucleosomal histone H3 at lysine 27 (H3K27me3), an epigenetic mark associated with gene silencing (Cao et al., 2002; Ernst et al., 2010). Ezh2 has also been shown to be phosphorylated by Akt which further suppress its methyltransferase activity and de-represses the silenced genes (Cha et al., 2005). Ezh2 also has been shown to be upregulated in various cancers (Varambally et al., 2002). Furthermore, the expression and function of Ezh2 in cancer cell lines are inhibited by microRNA 101 (Varambally et al., 2008). It has been proposed that the genomic loss of MIRN101 in cancer leads to the overexpression of Ezh2 and concomitant deregulation of epigenetic pathways, resulting in prostate cancer progression (Varambally et al., 2008). Furthermore, Ezh2 can induce EMT by downregulation of DAB2IP, which is a Ras GTPase activating protein that acts as a tumor suppressor, and increases the metastatic potential of prostate cancer (Chen et al., 2005; Min et al., 2010). Ezh2 knockout mice suffer from early embryonic death (O'Carroll et al., 2001). Interestingly, a conditional knockout of Ezh2 in basal keratinocytes resulted in thickened stratum corneum and granular layer and precocious acquisition of epidermal barrier function in the embryo (Ezhkova et al., 2009) while a conditional knockout of Ezh2 in B-cells led to improper IGH re-arrangement in B-cell (Su et al., 2003), suggesting a possible role for Ezh2 in the determination of cell-type identity.

Using global gene expression profiling, we identified a comprehensive list of genes at six consecutive different morphological states (1 day, 4 days, 7 days, 10 days and 20 days) during TGF- $\beta$ -induced EMT in normal mammary epithelial cells, NMuMG (Lehembre et al., 2008). This data was further subjected to computational analysis and based on motif search for various transcription factors at the promoter of regulated genes (<http://test.swissregulon.unibas.ch/cgi-bin/mara>), we predicted a number of transcription factors that might regulate transcription of a subset of genes in the early, intermediate and later stages of EMT (data not shown). Interestingly, Sox4 was discovered as one of such factor that might possibly regulate expression



of certain crucial genes during TGF- $\beta$ -induced EMT in NMuMG cells. Given the findings for Sox4 in cell-fate decision during development as well as its implicated role in cellular transformation during carcinogenesis (Liu et al., 2006a; Shin et al., 2004), the above findings motivated us to investigate the role of this transcription factor during Epithelial to Mesenchymal transition.

Our data reveal that not only Sox4 induces EMT, but also acts as a survival factor during this process. Sox4 knockdown inhibits cell migration. We also show that loss of Sox4 function in Py2T cells, derived from the MMTV-mouse PyMT model of breast cancer, significantly impairs their ability to form primary tumors and metastatic lesions in the lungs, lymph-nodes and liver of transplanted nude mice. Furthermore, Sox4 loss in B16-F10 metastatic model also leads to a decrease in metastasis. Expression profiling analysis upon Sox4 depletion in the absence and presence of TGF- $\beta$  revealed many Sox4 target genes, including Ezh2. Interestingly, Sox4 directly binds to the Ezh2 promoter controlling its transcriptional regulation. Along these lines, we show that Ezh2 induces EMT and act as a survival factor, thus revealing Ezh2 as one of the Sox4-target gene, critical for EMT.

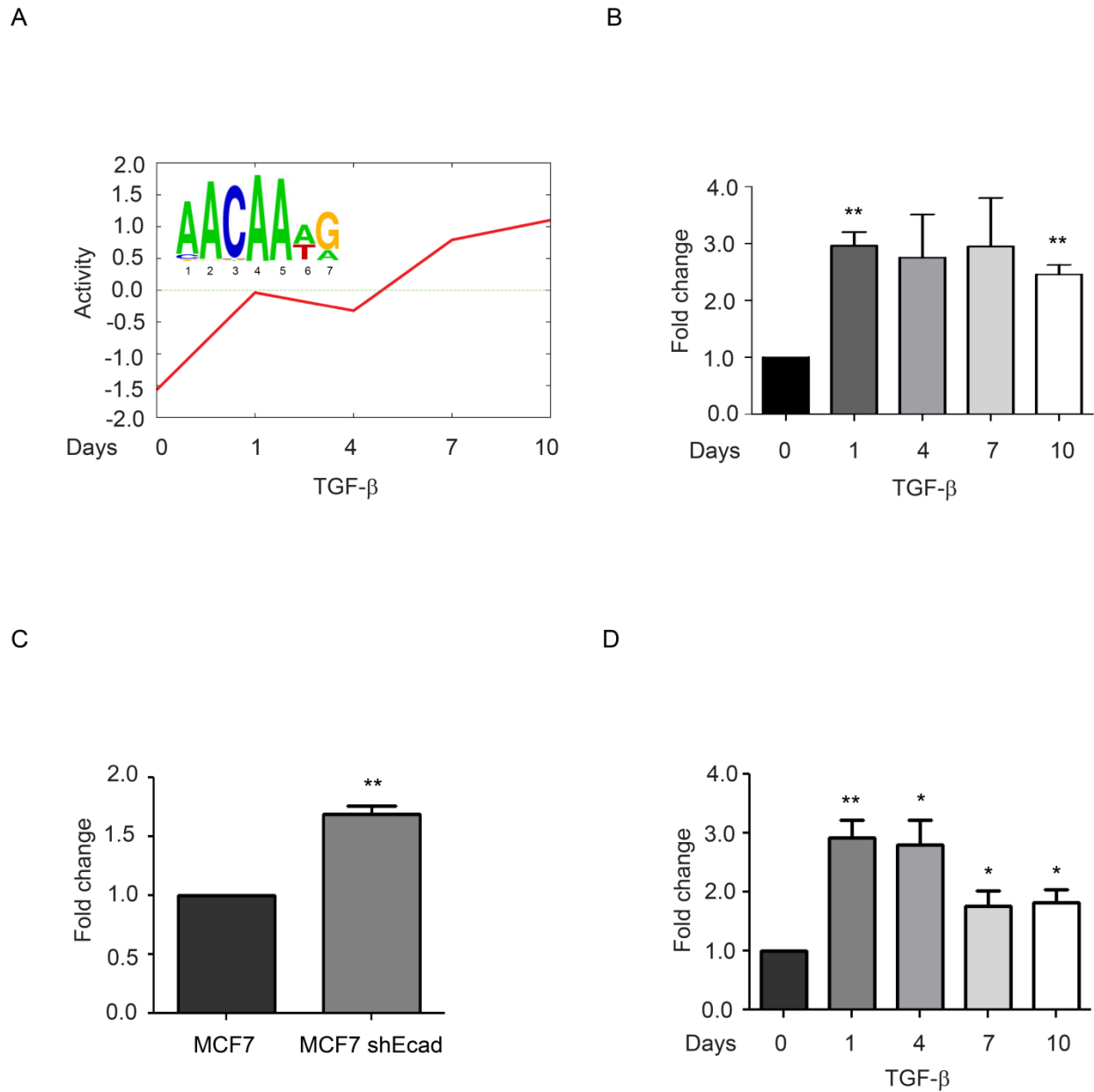
### **3.2.3 Results**

#### **3.2.3.1 Identification of Sox4 as a transcription factor upregulated in EMT**

To identify the crucial genes underlying EMT during early, intermediate and later stages, we used the established untransformed normal murine mammary gland cell line NMuMG (Lehembre et al., 2008; Miettinen et al., 1994). We treated NMuMG cells with TGF- $\beta$  for 1, 4, 7, 10 and 20 days and consequently these cells underwent progressive EMT and acquired a complete mesenchymal morphology by the end point, as shown previously (Lehembre et al., 2008). These morphological changes were accompanied by a “cadherin switch”, a hallmark of EMT, as well as with the gain of other mesenchymal markers and loss of epithelial markers (Hazan et al., 2004; Maeda et al., 2005). We next performed a genome-wide expression profiling of the above stages and identified genes that were differentially expressed at early, intermediate and later stages of the EMT process. We then employed MARA (Motif Activity Response Analysis) to identify crucial transcription factors whose targets are transcriptionally modulated

Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes

during EMT based on their motif occurrence at target promoters. This analysis revealed a number of interesting transcription factors that may possibly regulate EMT via transcriptional regulation of underlying genes (data not shown), and Sox4 was identified as one of the top



**Figure 1: Sox4 expression and activity is upregulated during TGF- $\beta$ -induced EMT.** (A) MARA analysis predicts Sox4 activity during EMT in NMuMG cells. (B) RT-qPCR analysis for quantification of Sox4 expression levels in NMuMG cells during TGF- $\beta$  mediated EMT. (C) Expression levels of Sox4 as quantified by RT-qPCR in

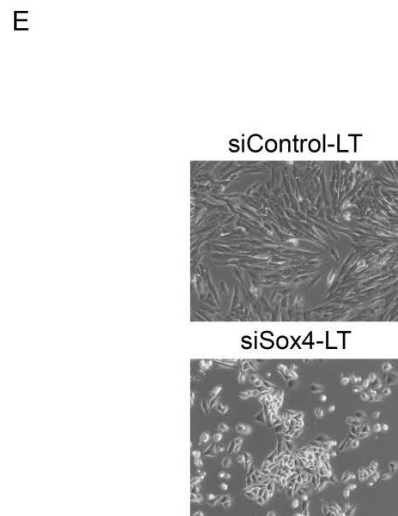
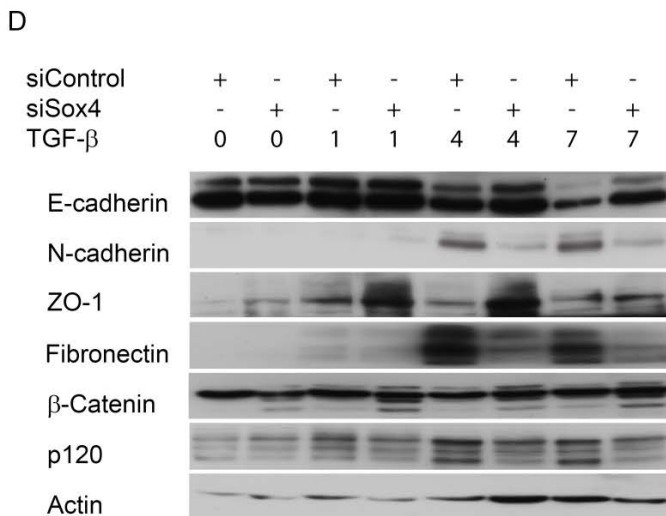
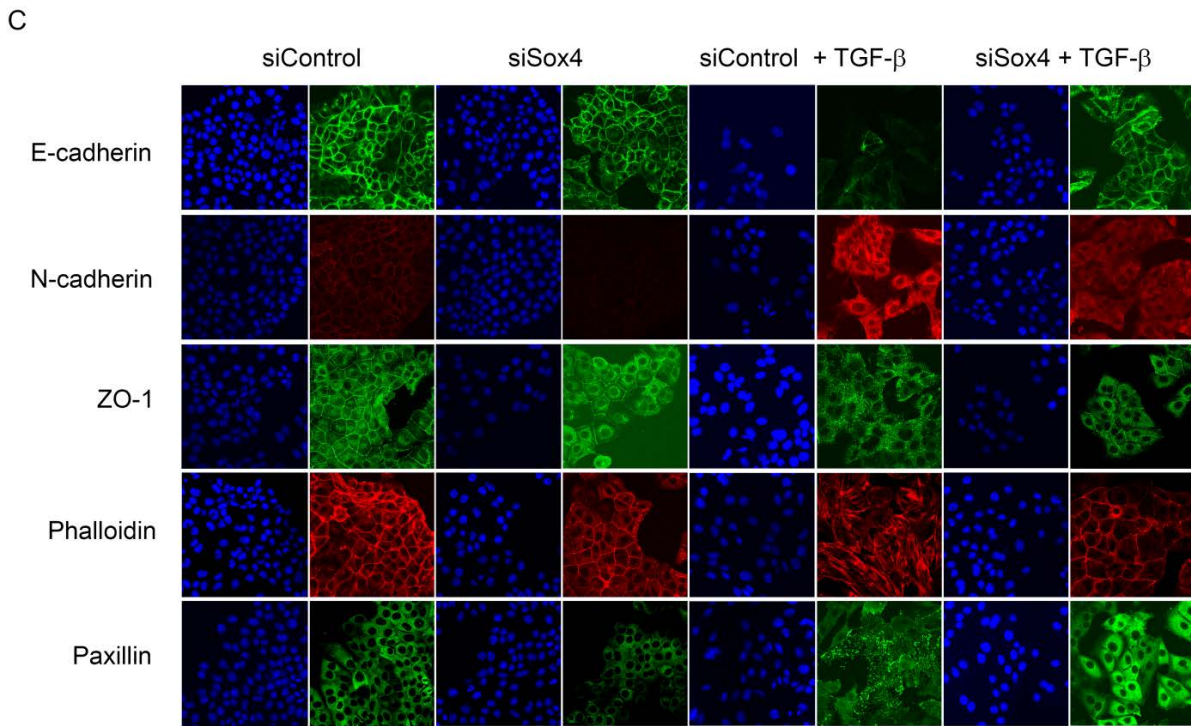
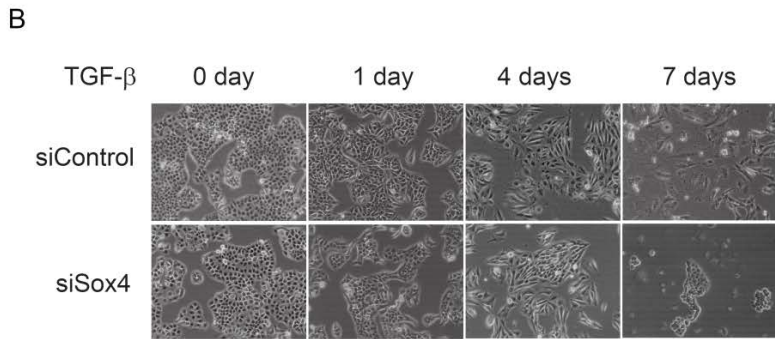
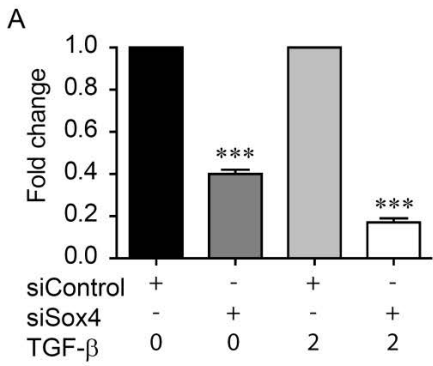
control and E-cadherin-specific shRNA treated MCF7 cells. (D) RT-qPCR for quantification of Sox4 levels upon EMT induction in Py2T cells. Statistical values were calculated by using a paired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

scoring factor in this analysis. Sox4 has AACAA/TG/A binding motif (Figure 1A) and according to MARA analysis, its activity is increased during EMT (Figure 1A). Very interestingly, quantitative RT-PCR experiments revealed that not only the activity, but also Sox4 expression is up-regulated during such TGF- $\beta$ -induced EMT in NMuMG cells (Figure 1B). We further confirmed Sox4 upregulation in a human breast cancer cell line that undergoes EMT upon E-cadherin knockdown (MCF7) (Lehembre et al., 2008) (Figure 1C). In addition, this was also seen in another cell line from PyMT model system of breast cancer that also undergoes EMT in response to TGF- $\beta$  treatment (Waldameier et.al. unpublished data) (Figure 1D). These results suggested a potential oncogenic role for Sox4 during EMT.

### **3.2.3.2 Sox4 depletion prevents EMT**

To directly assess the role of Sox4 in EMT, we used two different siRNAs and pooled them together to knockdown Sox4 in the absence and presence of TGF- $\beta$  in NMuMG cells. This siRNA pool efficiently downregulated Sox4 levels (Figure 2A). Sox4-ablated NMuMG cells were not able to undergo EMT and retained an epithelial phenotype during TGF- $\beta$  treatment (Figure 2B). Immunofluorescence studies revealed that Sox4 depletion in the absence of TGF- $\beta$  led to a retention of all tested epithelial markers similar to control cells while mesenchymal markers like N-cadherin and paxillin were decreased (Figure 2C). On the other hand, in the presence of TGF- $\beta$ , cells were able to maintain E-cadherin and ZO-1 at the membrane while the mesenchymal marker N-cadherin was decreased. Along with these observations, TGF- $\beta$ -mediated cytoskeleton re-modeling was also prevented and stress fibers were reduced. Fibronectin and paxillin staining revealed a significant failure to form focal adhesion after Sox4 reduction (Figure 2C). These observations were further confirmed by immunoblot analysis (Figure 2D). Notably, Sox4 depletion in NMuMG cells treated with TGF- $\beta$  for 15 days (LT) exhibited epithelial cell morphology, demonstrating that Sox4 is not only required for the initiation but also for the maintenance of EMT in NMuMG cells (Figure 2E). We further

Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes



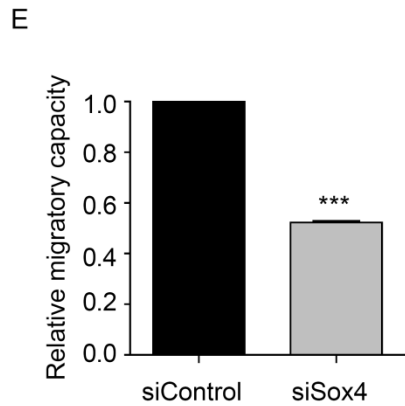
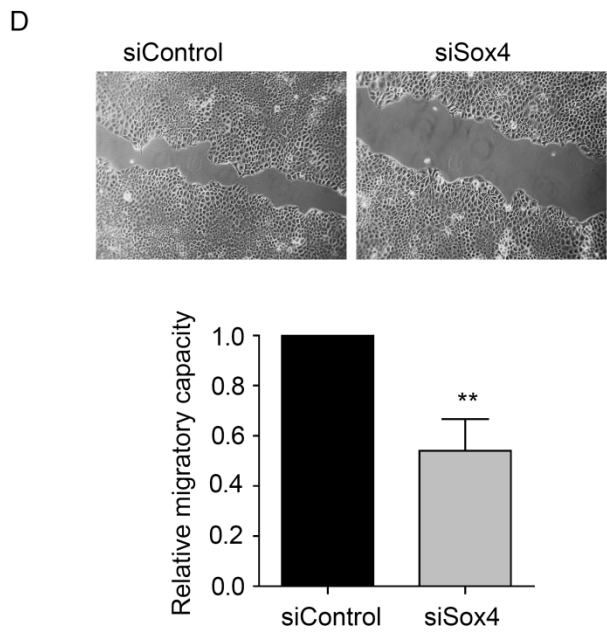
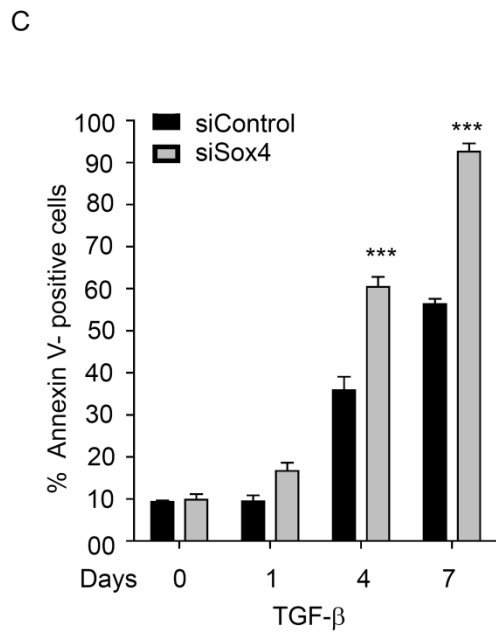
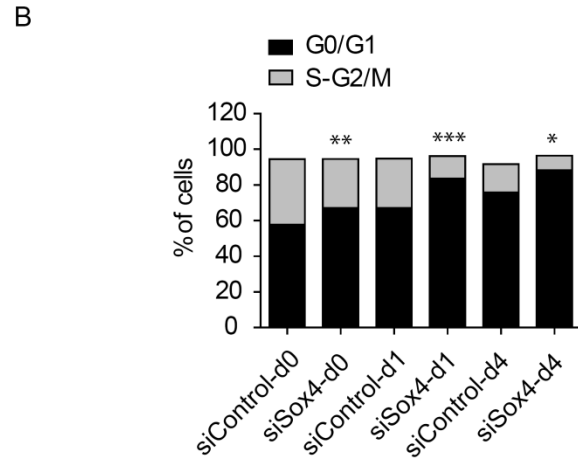
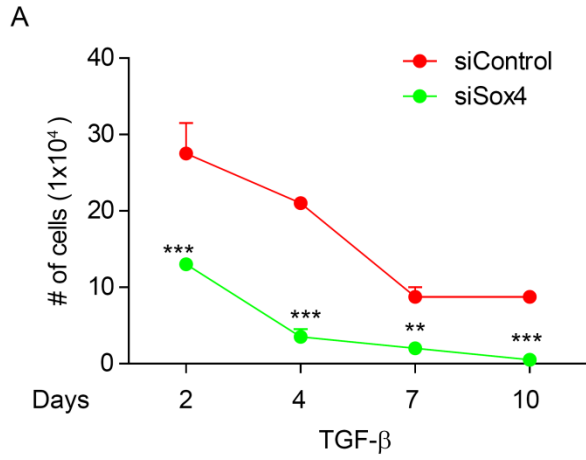
**Figure 2: Depletion of Sox4 blocks epithelial differentiation.** (A) qRT-PCR was performed after siRNA-mediated knockdown of Sox4 in the absence and presence of TGF- $\beta$ . Cells were treated with TGF- $\beta$  for 2 days. (B) Morphology of NMuMG cells transfected with either control or Sox4-specific siRNA during TGF- $\beta$  mediated EMT, as evaluated by phase contrast microscopy. Original magnification was 10X. (C) Immunofluorescence stainings for adherent junction protein, E-cadherin and N-cadherin, tight junction protein ZO-1, the actin cytoskeleton by phalloidin and focal adhesion protein paxillin were performed after Sox4 knockdown in the absence and presence of TGF- $\beta$ . Cells were treated with TGF- $\beta$  for 2 days. Original magnification was 40X. (D) Expression of the epithelial proteins E-cadherin, ZO-1 and  $\beta$ -catenin and the mesenchymal proteins N-cadherin, fibronectin and p120 as detected by immunoblot analysis after siRNA mediated knockdown of Sox4 during the TGF- $\beta$  time-course. Actin was used as a loading control. (E) Morphology of long term TGF- $\beta$  treated NMuMG cells transfected with either control or Sox4-specific siRNA, as evaluated by phase contrast microscopy. Original magnification was 10X. Statistical values were calculated by using a paired, two-tailed t-test. p-value  $\leq$  0.05 indicated with (\*), p-value  $\leq$  0.01 indicated with (\*\*), p-value  $\leq$  0.001 indicated with (\*\*\*)

extended these studies to the Py2T murine breast cancer model. Similar to NMuMG cells, upon a lentiviral mediated stable knockdown of Sox4 Py2T cells failed to undergo EMT at later stages and recovered the epithelial marker E-cadherin (Supplemental figure S1 A) while the mesenchymal marker fibronectin was decreased (Supplemental figure S1 B). We did not observe any change in N-Cadherin expression as well as in morphology after TGF- $\beta$  treatment in Sox4-depleted cells (data not shown). Together, these results strongly argue that Sox4 is required for TGF- $\beta$  induced EMT.

### **3.2.3.3 Sox4 provides a survival advantage during TGF- $\beta$ -induced EMT and supports migration**

Sox4 has been implicated in cell survival in various cancers (Hur et al., 2010; Pramoonjago et al., 2006; Shen et al., 2010). Thus, we next investigated whether Sox4 has a similar function during TGF- $\beta$ -induced EMT in NMuMG cells. Sox4-depleted NMuMG cells showed a significant reduction in cell growth in comparison to control cells and displayed high sensitivity towards TGF- $\beta$ -mediated growth inhibition (Figure 3A). To determine whether this effect was due to alterations in proliferation ability or changes in apoptosis levels, we used propidium iodide staining to perform FACS-based analysis of the cell cycle and further performed Annexin V staining to quantify apoptosis. In comparison to control siRNA-treated cells, a statistically

Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes



**Figure 3: Sox4 supports migration and provides survival advantage to the cells during TGF- $\beta$ -induced EMT.**

(A) Proliferation assays were performed after siRNA-mediated knockdown of Sox4 during a TGF- $\beta$  time-course. Cells were counted using a Neubauer counting chamber. (B) Cell cycle analysis was done after transient knockdown of Sox4 during EMT for 0, 1 and 4 days. Propidium Iodide (PI) was used for the staining. (C) Annexin V staining was performed to quantify cell death in Sox4 knockdown cells during TGF- $\beta$ -induced EMT for 0, 1, 4 and 7 days. (D) Following siRNA-mediated ablation of Sox4 in the absence of TGF- $\beta$ , a wound healing assay was performed after 19 hours of wound creation. (E) Sox4 depleted cells were subjected to a Boyden chamber migration assay for 20 hours using 20% FBS as a chemoattractant. Statistical values were calculated by using an unpaired/paired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

significant G0/G1 block was observed after Sox4 depletion (Figure 3B). In addition, increased apoptosis occurred in Sox4 knockdown cells in the presence of TGF- $\beta$  (Figure 3C). Thus, TGF- $\beta$ -sensitive growth of Sox4-depleted NMuMG cells relies on both increased proliferation and decreased apoptosis during TGF- $\beta$ -induced EMT. These observation led us conclude that Sox4 is required for cell survival and proliferation during TGF- $\beta$ -induced EMT.

We next extended our study in Py2T cells carrying a stable knockdown of Sox4 and performed a similar series of experiments as described above. Unlike NMuMG cells, these cells did not show any apoptotic phenotype (data not shown), although they exhibited a slower growth rate (Supplemental figure S2 A) and G0/G1 arrest (Supplemental figure S2 B). These results suggest that the apoptotic phenotype is most likely prominent in untransformed cells.

Since an increased migration potential is a characteristic feature of the EMT process (Brabletz et al., 2005; Christofori, 2006; Grunert et al., 2003; Huber et al., 2005; Thiery and Sleeman, 2006), we measured the migratory capacity of Sox4-depleted NMuMG cells. Wound healing assays revealed a significant reduction in migration of cells depleted of Sox4 compared to control cells (Figure 3D). Similar results were obtained in trans-well migration assays (Figure 3E) and in Sox4 depleted Py2T cells treated with TGF- $\beta$  for 15 days (Supplemental figure S2 C).

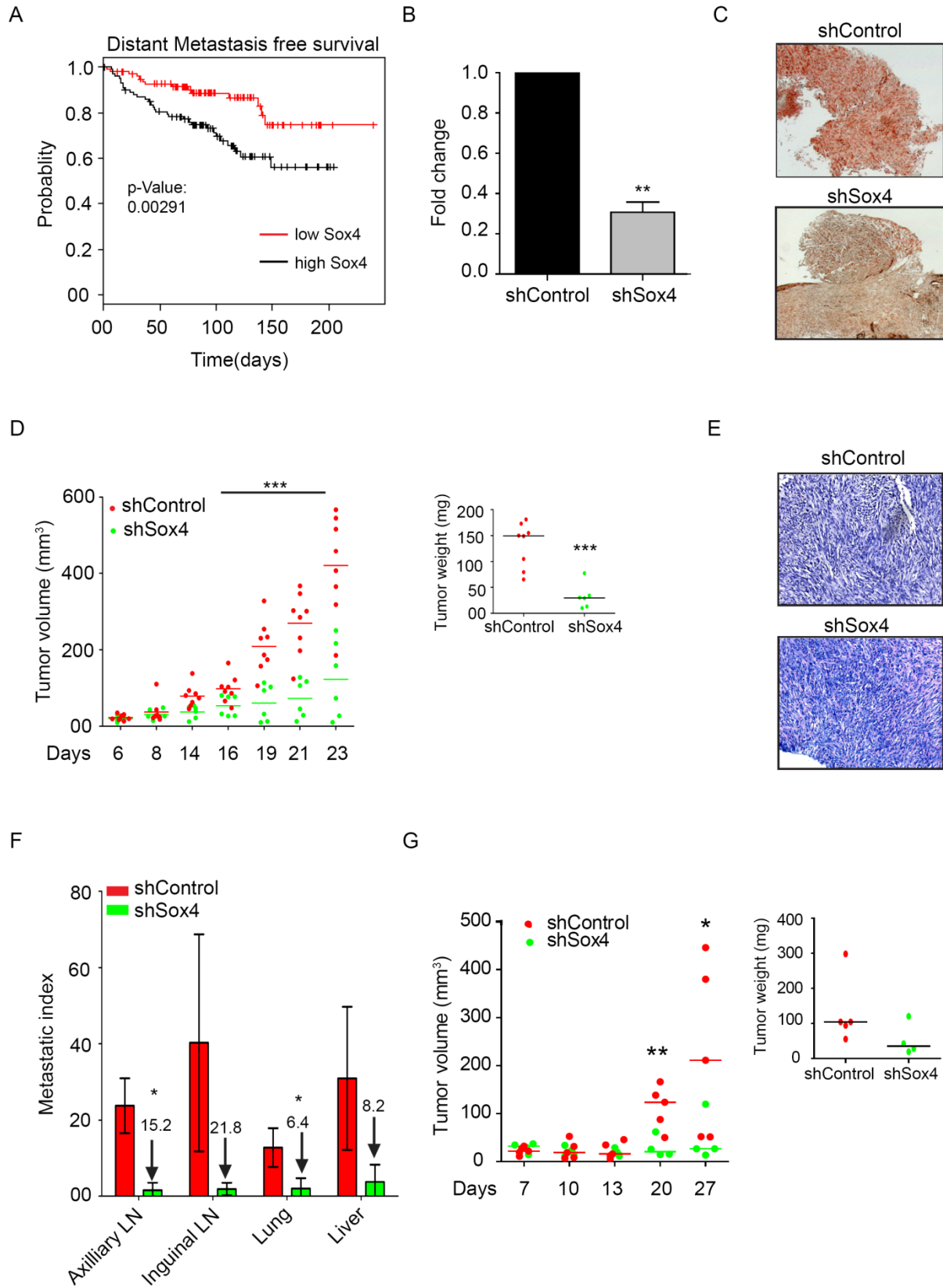
**3.2.3.4 Sox4 is required for TGF- $\beta$ -induced tumorigenesis and metastatic spread**

We next asked whether increased expression of Sox4 correlate with the ability of tumors to grow, progress and metastasize. Analysis of the Schmidt breast cancer database (Schmidt et al., 2008) revealed a correlation between Sox4 expression and metastasis free survival (Figure 4A). This database contains 200 lymph node-negative breast cancer patients that were either treated with modified radical mastectomy or breast-conserving surgeries followed by irradiation and were not exposed to any systemic therapy after surgery. This led us to further investigate whether depletion of Sox4 in Py2T cells could impair the ability of these cells to form tumors and to seed metastasis. In order to address this question, we performed orthotropic mammary fat-pad injections of Sox4-depleted Py2T cells that were also tagged by luciferase expression into nude mice and quantified tumor growth and metastasis in lymph nodes, lungs and liver. To quantitate the knockdown efficiency of these cells before injection, Q-PCR was performed, showing a significant reduction in Sox4 levels (Figure 4B). Immunostaining experiments in tumor sections revealed that an efficient depletion of Sox4 was achieved in the tumors as well (Figure 4C). As shown in Figure 4D, the knockdown of Sox4 in Py2T cells led to a significant reduction in tumor growth, suggesting a crucial role for Sox4 expression in primary tumor growth. H&E staining of tumor sections did not show any morphological differences compared to the control counterpart (Figure 4E). We also performed luciferase assays in the organs known to be target of breast cancer metastasis and found that luciferase levels were decreased in axillary and inguinal lymph nodes as well as in lung and liver upon Sox4 depletion (Supplemental Figure S3). These results argue in favor of the hypothesis that Sox4 is required for breast cancer metastasis. It has recently been shown that tumor size can be correlated with metastasis (Minn et al., 2007). Therefore, we calculated the metastatic index by dividing the average luciferase levels in each organ with average tumor weight (metastatic index), and the results further supported our findings that Sox4 depletion is associated with a decreased metastatic potential (Figure 4F). We further validated our data by performing sub-cutaneous injections of Sox4-depleted Py2T cells and found a similar reduction in tumor growth (Figure 4G).

We next investigated whether knockdown of Sox4 expression in B16-F10 melanomas cells impairs their ability to metastasize. An initial analysis of Sox4 expression upon TGF- $\beta$  treatment for 1, 4, 7 and 10 days in B16-F10 melanoma cells upregulated Sox4 mRNA levels similar to

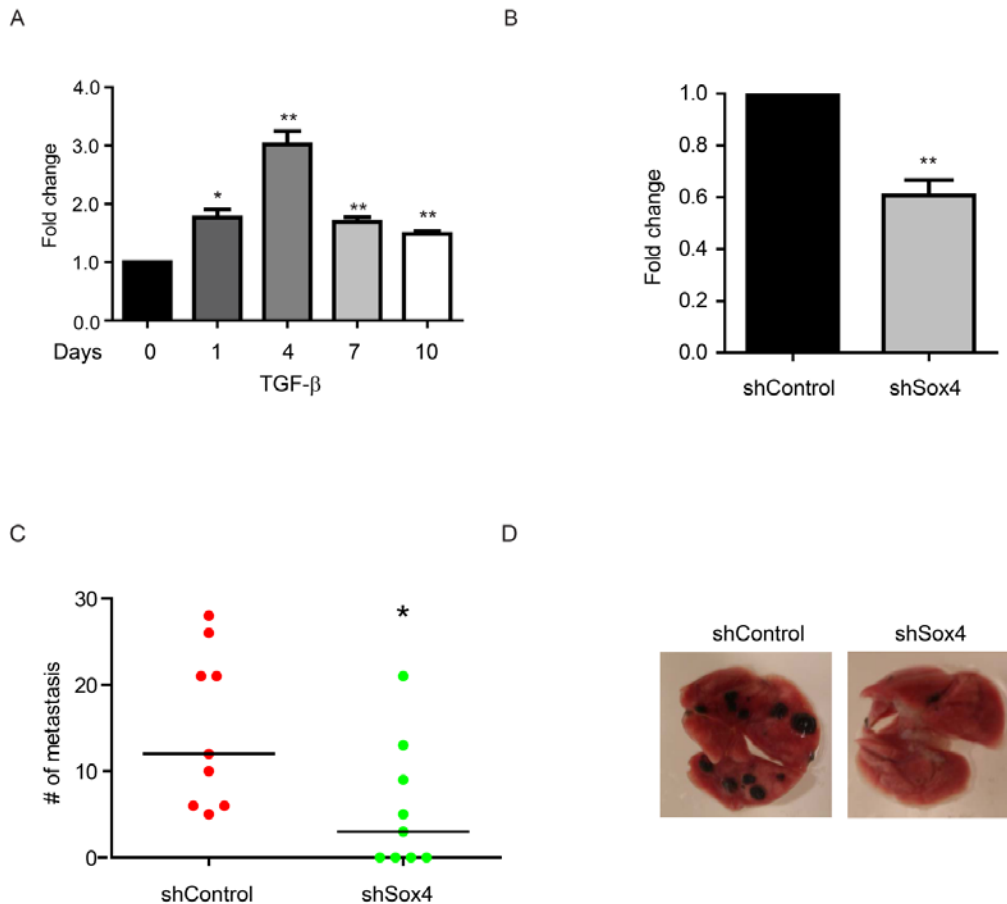


Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes



Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes

**Figure 4: Sox4 is necessary for breast cancer tumorigenesis and metastasis.** (A) Correlation of Sox4 expression in human breast cancer array data set. Kaplan-Meier survival curve for the patient samples classified for having high Sox4 expression and low Sox4 expression to assess the metastasis free survival. (B) RT-qPCR quantification of Sox4 knockdown efficiency in Py2T cells stably transfected with control (shControl) and Sox4-specific (ShSox4) shRNA. (C) Immunohistochemistry for Sox4 was performed on tumor sections to validate the Sox4 knockdown efficiency. (D) Tumor growth was measured following fat-pad injection of shControl and shSox4 cells into 7 weeks old female nude mice. Tumor weight was assessed after sacrificing the nude mice 23 days post-injection. (E) H&E staining on tumor sections derived from mice injected with shControl and shSox4 cells. (F) Metastatic spread was quantified by measuring the relative luciferase levels in lymph nodes, lung and liver of nude mice injected with shControl and shSox4 cells and further by dividing the relative luciferase levels with tumor weight in various organs. (G) Tumor growth was measured following sub-cutaneous injection of shControl and shSox4 cells into the 7 weeks old female nude mice. Tumor weight was assessed after sacrificing the nude mice 27 days of post-injection. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)



**Figure 5: Sox4 is required for metastasis in B16-F10 melanoma cells.** (A) qRT-PCR to determine the expression of Sox4 during TGF- $\beta$ -induced EMT in B16-F10 melanoma cell line. (B) RT-qPCR to assess the knockdown efficiency after stable transfection of shSox4 in B16-F10 melanoma cells. (C) Lung metastasis was measured after the intravenous injection of shControl and shSox4 melanoma cells into the tail-vein of C57 black 6 mice. Mice were sacrificed after 18 days of post-injection. (D) Pictures of lungs after sacrificing mice injected intravenously with shControl and shSox4 melanoma cells. Statistical values were calculated by using an unpaired/paired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

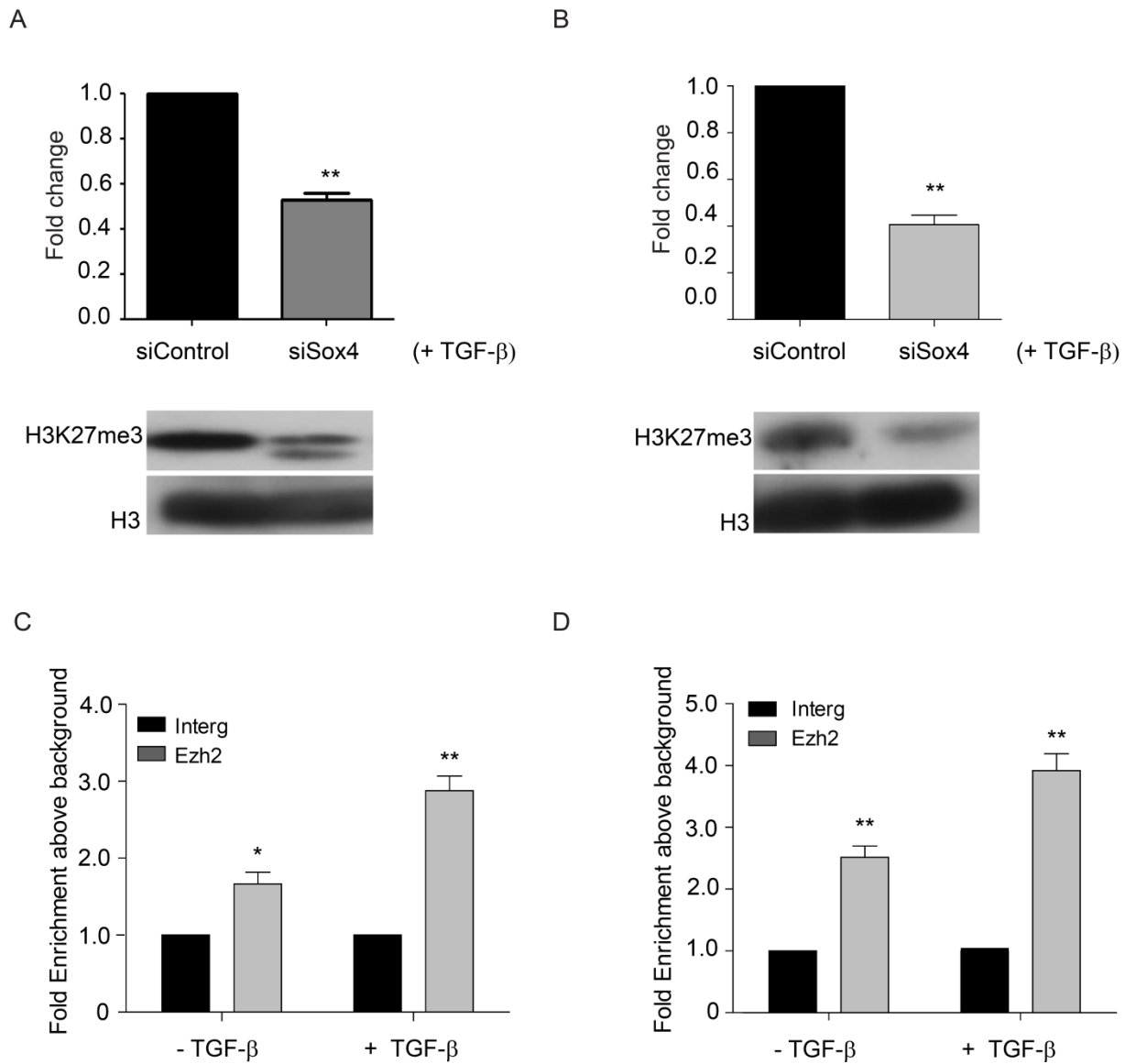
NMuMG cells (Figure 5A). We then performed intravenous injection of stable Sox4 knockdown B16-F10 melanoma cells (shSox4-B16-F10) into the tail-veins of C57BL/6 mice and quantified metastasis to the lung. The results show that Sox4 depleted B16-F10 melanomas cells were significantly impaired in colonizing the lungs (5B, 5C and 5D). Taken together, these observations led us to conclude that Sox4 was an important player in the metastatic process.

### **3.2.3.5 Ezh2 is a transcriptionally regulated by Sox4**

Given the established role for Sox4 as a transcriptional regulator, we next attempted to identify genes that are possibly regulated by Sox4 and through which it may contribute to EMT. We performed a global transcription profiling in control and Sox4 siRNA-treated NMuMG in the absence and presence of TGF- $\beta$  for 2 days in order to reveal the direct targets of Sox4. The differentially expressed genes were analyzed for Sox4 binding motif within 1 Kb of their transcription start site. The resulting gene list of 198 genes was further compared with the EMT time-course expression analysis we performed earlier to reveal that 99 EMT relevant genes were in common. This list was further subjected to gene ontology analysis and we could identify 35 targets that have been implicated in angiogenesis, cell morphogenesis, apoptosis, migration, adhesion, MAPK signaling, cell cycle, cytoskeleton remodeling, EGFR signaling and EMT. In order to strengthen our observation, we assess the direct occupancy of Sox4 at the promoters of these genes. Chromatin Immunoprecipitation (ChIP) assay in NMuMG cells transiently transfected with HA-Sox4 construct using HA-specific antibodies revealed that out of these 35 genes, the promoters of 23 genes are directly bound by Sox4 (including key EMT genes such as Spred1, Edn1, Palld, Cyr61, Ereg, Areg and Yap1) (Supplemental Figure S4 A-I). These Chips experiments further helped us to better define the Sox4 binding motif as A A C A A G/A for the

Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes

target promoters analyzed. It has been shown previously that Sox4 directly regulates Mll, a DNA-binding protein that methylates histone H3 lys4 (H3K4) and positively regulates expression of target genes, in prostate cancer (Scharer et al., 2009). Given emerging evidences of Polycomb group (PcG) proteins, especially the H3K27 methyltransferase Ezh2, in carcinogenesis, we investigated whether Sox4 may have possible regulatory influence on the



**Figure 6: Sox4 directly binds the Ezh2 promoter.** (A-B) RT-qPCR for measurement of Ezh2 expression levels upon Sox4 knockdown in NMuMG (A) and Py2T (B). Immunoblot for the repressive mark H3K27me3 after siRNA- mediated Sox4 ablation in NMuMG (A) and Py2T (B) cells. The cells were treated with TGF- $\beta$  for 2 days. (C-D) Following Chromatin Immunoprecipitation (ChIP) using HA-specific antibody in NMuMG (C) and Py2T (D) cells transiently transfected with HA-Sox4 construct, realtime PCR was performed using Ezh2-promoter specific primers to determine the Sox4 occupancy at this region. Statistical values were calculated by using an unpaired/paired, two-tailed t-test. p-value  $\leq$  0.05 indicated with (\*), p-value  $\leq$  0.01 indicated with (\*\*), p-value  $\leq$  0.001 indicated with (\*\*\*)

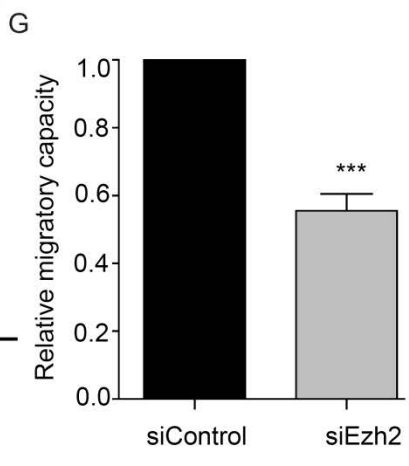
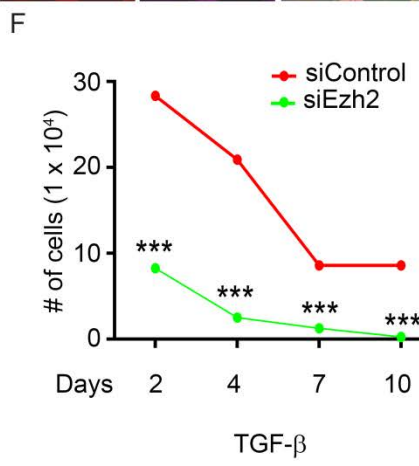
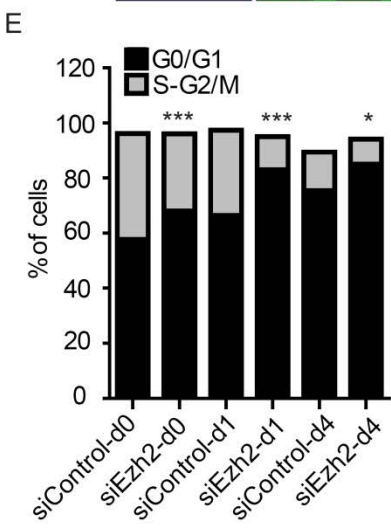
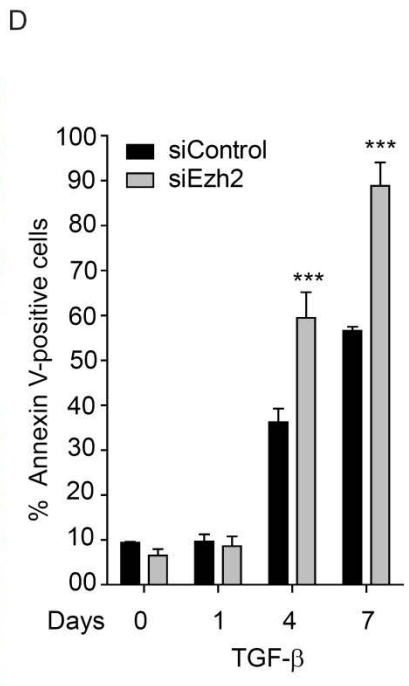
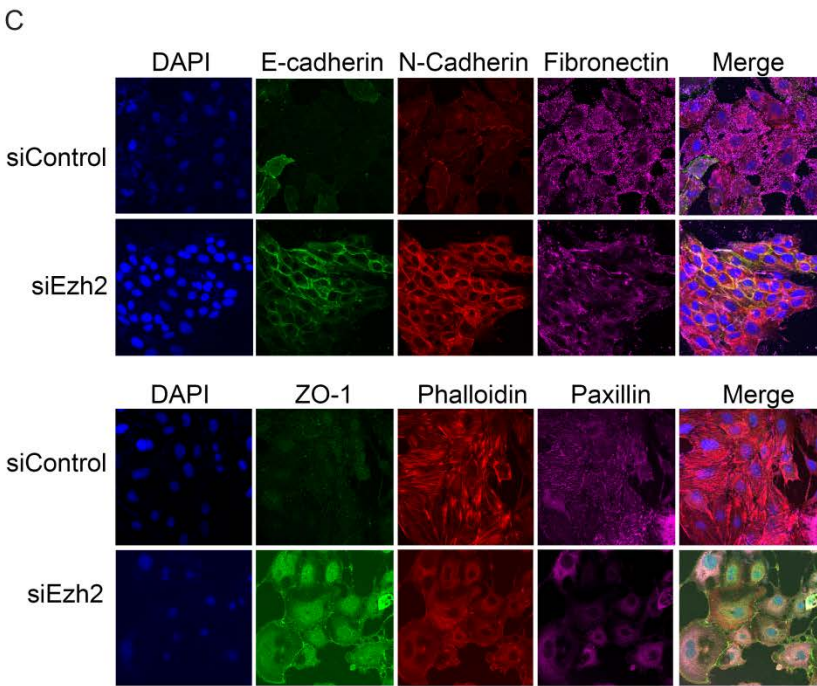
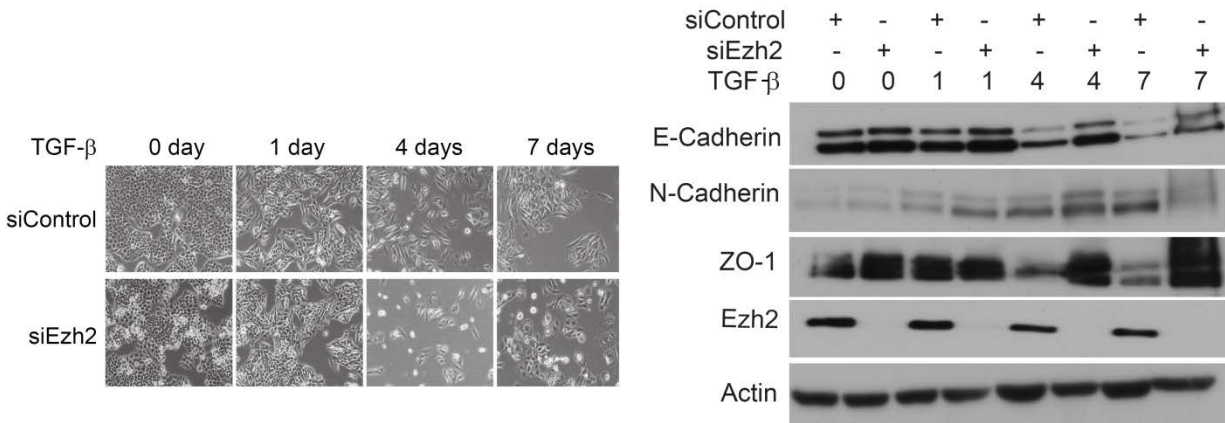
transcriptional regulation of Ezh2. Interestingly, expression array results and further validation by quantitative RT-PCR showed that Ezh2 is significantly downregulated upon Sox4 depletion in presence of TGF- $\beta$  in NMuMG (Fig 6A) and Py2T (Figure 6B). Sox4 knockdown also led to a global reduction in H3K27me3 levels in NMuMG and Py2T cells (Figure 6A and 6B). Interestingly, Chromatin Immunoprecipitation assays revealed that Sox4 directly bound the Ezh2 promoter in NMuMG (Figure 6C) and Py2T cells (Figure 6D). These data suggest that Sox4 directly binds to promoters of key EMT genes, including Ezh2, for their transcriptional activation during TGF- $\beta$  induced EMT.

### 3.2.3.6 Ezh2 knockdown phenocopies Sox4 depletion

We next speculated whether some aspect of EMT regulation by Sox4 are mediated via its regulation of Ezh2 and assessed whether knockdown of Ezh2 phenocopies Sox4 depletion. Interestingly, siRNA-mediated depletion of Ezh2 in NMuMG cells led to a retention of an epithelial phenotype even in the presence of TGF- $\beta$  (Figure 7A). Furthermore, these morphological changes were also accompanied with retention of the epithelial markers like E-cadherin and ZO-1 as well as the loss of the mesenchymal marker N-cadherin (Figure 7B). These results were further supported by immunofluorescence staining for epithelial and mesenchymal markers in Ezh2-depleted NMuMG cells (Figure 7C).

We next investigated whether Ezh2 also exerted a survival effect during EMT similar to that observed for Sox4. Annexin V staining in Ezh2 depleted NMuMG cells during TGF- $\beta$  treatment revealed that apoptosis was significantly increased upon Ezh2 depletion during EMT (Figure

Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes



**Figure 7: Depletion of Ezh2 phenocopies the absence of Sox4 during TGF- $\beta$ -induced EMT.** (A) Phase contrast microscopy was performed to study the morphology of NMuMG cells transfected with siControl and siEzh2 is monitored after TGF- $\beta$  treatment for the indicated period. (B) Expression levels of the epithelial markers, E-cadherin and ZO-1 and mesenchymal marker, N-cadherin were measured by immunoblot analysis in cells treated with either control or Ezh2-specific siRNA during a TGF- $\beta$  time-course. Actin was used as a loading control. (C) Confocal microscopy analysis of NMuMG cells to assess the localization and expression levels of EMT markers after 7 days of TGF- $\beta$  treatment. (D) Apoptosis was measured by Annexin V staining in combination with FACS analysis in control (siControl) and Ezh2-depleted (siEzh2) NMuMG cells treated with TGF- $\beta$  for the days indicated. (E) Propidium iodide staining to quantitate the cell cycle phases after Ezh2 knockdown during TGF- $\beta$ -induced EMT for 0, 1 and 4 days. (F) NMuMG cells transfected with siControl and siEzh2 were treated with TGF- $\beta$  for the days indicated and counted using a Neubauer chamber to assess the proliferation rate during TGF- $\beta$  EMT. (G) Ezh2 depleted cells were subjected to Boyden chamber migration assays using 20% FBS as a chemoattractant. Statistical values were calculated by using an unpaired/paired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

7D). In addition, similar to Sox4 knockdown cells, Ezh2-depleted NMuMG cells were also blocked in the G0/G1 phase of the cell cycle and were not able to proliferate (Figure 7E and 7F). These results suggest that, in line with Sox4 depletion phenotype, decreased and TGF- $\beta$  sensitive growth of Ezh2 depleted NMuMG cells relies on both increased proliferation and decreased apoptosis during TGF- $\beta$ -induced EMT. We also tested whether Ezh2 had an impact on the migration capacity of cell during EMT by performing a trans-well migration assay in Ezh2-depleted NMuMG cells in the absence of TGF- $\beta$ . As expected, these cells were not able to migrate more than the control counterpart (Figure 7G). Together, these observations have led us to conclude that Ezh2 is an downstream effector of Sox4 regulatory network and functions to regulate TGF- $\beta$ -induced EMT in a fashion similar to Sox4.

### 3.2.4 Discussion

Most solid tumors are epithelial in nature and increasing evidences suggest that tumor progression accompanies EMT to enable tumor cells to acquire the capacity to be motile and infiltrate surrounding tissues, thus licensing them to metastasize to distant sites.

Transcription factors play a range of crucial regulatory roles in diverse aspects of cellular function such as cell proliferation, differentiation and survival by regulating transcription of genes underlying the process. Such regulation has to be tightly regulated and is crucial for proper execution of the gene-expression program that defines cellular identity. Misregulation in the expression and function of transcription factors involved in cellular pathways has been linked to various diseases including cancer.

Sox4 is a one of such key transcription factors that has been linked to certain cancers. However, no detailed investigation has been performed to assess its direct relation with EMT. Our study revealed that Sox4 is an inducer of EMT and the absence of Sox4 in breast cancer cells leads to suppression of EMT. Moreover, Sox4 is required for EMT-driven cell migration and provides a survival advantage to cells undergoing EMT by attenuating TGF- $\beta$ -induced apoptosis. More importantly, we also demonstrated that Sox4 is not only required for EMT but also for metastatic spread during breast carcinogenesis. Orthotropic fat-pad injection of Sox4 depleted Py2T cells into nude mice has revealed that metastasis lesions are decreased in different breast cancer metastasis sites including lymph-nodes (axillary and inguinal), lungs and liver. In addition, after surveying the gene expression profiles of human cancer biopsies using the NextBio.com database, we found a significant positive correlation between increased Sox4 expression and metastatic potential of melanoma cancer. Moreover, a positive and significant correlation of increased Sox4 expression with tumorigenicity and advances in tumor stage was also observed. Interestingly, in melanomas it is well known that the resistance towards TGF- $\beta$ -mediated growth inhibition and EMT plays an essential role for their metastatic spread, supporting our findings that Sox4 depletion has a protective function during tumor-progression and metastasis (Poser et al., 2001; Silye et al., 1998; Teicher, 2001). Intravenous injection of Sox4-depleted B16-F10 melanoma cells into the tail-veins of C57/B16 mice revealed that Sox4 also plays a key role in metastasis formation in a melanoma cell model.

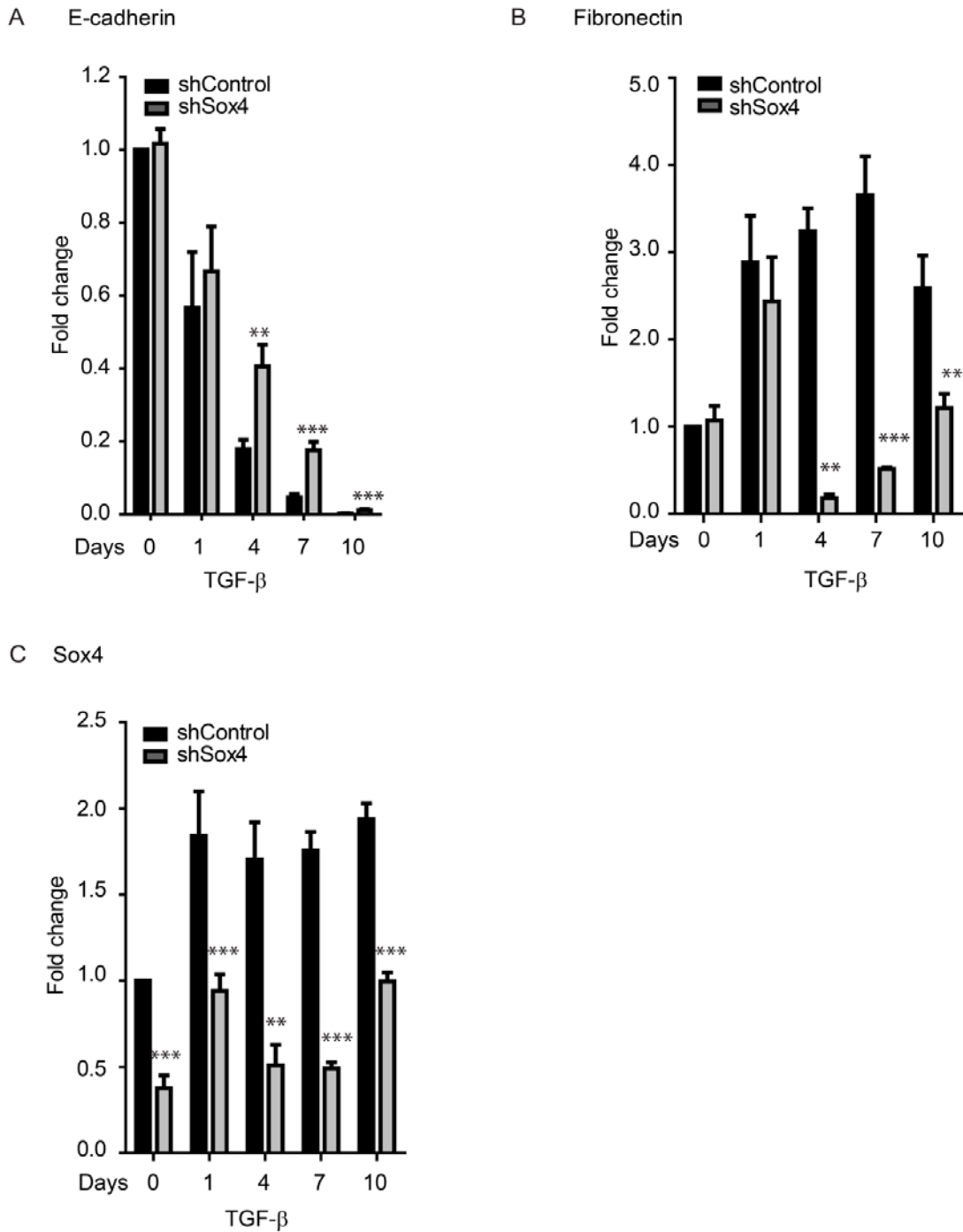
To identify genes that are possibly under direct transcriptional control of Sox4, genomewide expression profiling was performed after Sox4 depletion in the absence and presence of TGF- $\beta$  which identified a number of genes that were deregulated in Sox4 knockdown cells. To reveal the possible direct targets of Sox4, we screened for genes that had a Sox4 binding motif within 1Kb of their transcription start site. In addition, the resulting list was further compared with the



genome-wide expression data of a EMT time-course to identify the EMT relevant genes. This analysis has revealed that Sox4 regulates many key EMT genes, such as Spred1, Edn1, Palld, Cyr61, Ereg, Areg and Yap1 by directly acting on their promoters. In addition, Sox4 also targeted many genes involved in other EMT relevant processes such as angiogenesis, adhesion, migration, morphogenesis, cell cycle regulation and cytoskeleton re-modeling. We further identified the Ezh2 promoter to be bound by Sox4 for transcriptional modulation. Ezh2 is a histone H3K27 methyltransferase that has been shown to repress E-cadherin (Cao et al., 2008) and is therefore implicated in EMT. In line with being a downstream effector of Sox4, Ezh2 contributes to TGF- $\beta$ -induced EMT in a fashion similar to Sox4 with respect to being required for transition from epithelial to mesenchymal identity and cell migration and cell survival events that accompany this process.

Taken together, our findings reveal mechanistic insight into how Sox4 regulates EMT by targeting and controlling the expression of crucial EMT genes, thereby directly contributing to the transcriptional reprogramming underlying this process. Keeping the versatile function of this factor in view, our data are certainly of key importance in the EMT field and will possibly pave the way for a therapeutic value in the long-term.

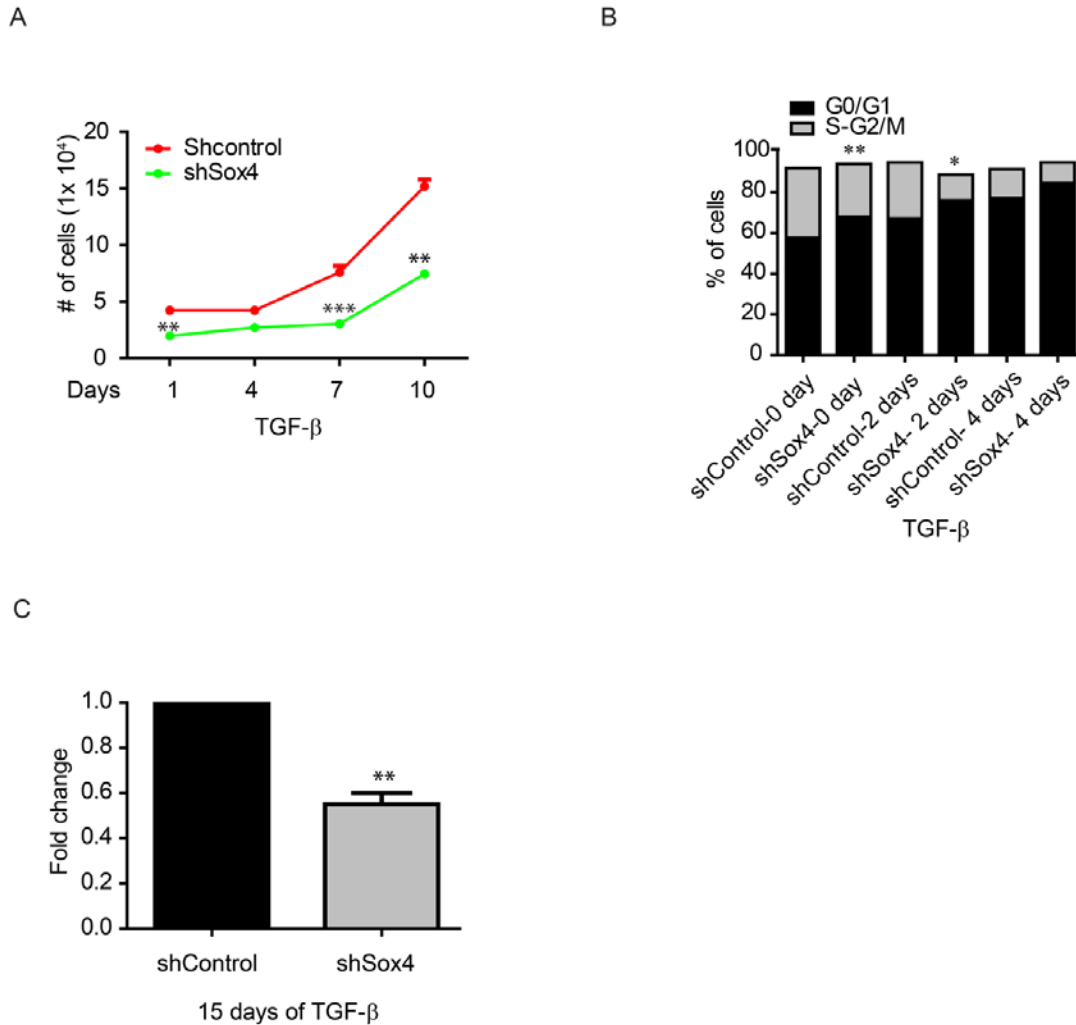
### 3.2.5 Supplemental figures



**Figure S1: Sox4 ablation delays EMT in Py2T cells.** (A-B) RT-qPCR quantification of the epithelial marker E-cadherin (A) and mesenchymal marker fibronectin (B) in Py2T cells stably transfected with control (shControl) and Sox4-specific (ShSox4) shRNA. (C) RT-qPCR to assess the knockdown efficiency after stable transfection of

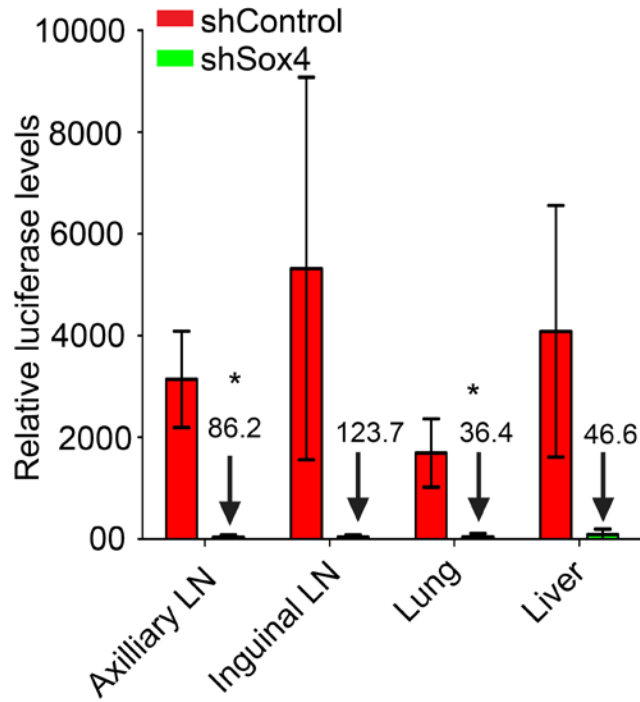
Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes

shSox4 in Py2T cells. Statistical values were calculated by using an unpaired/paired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)



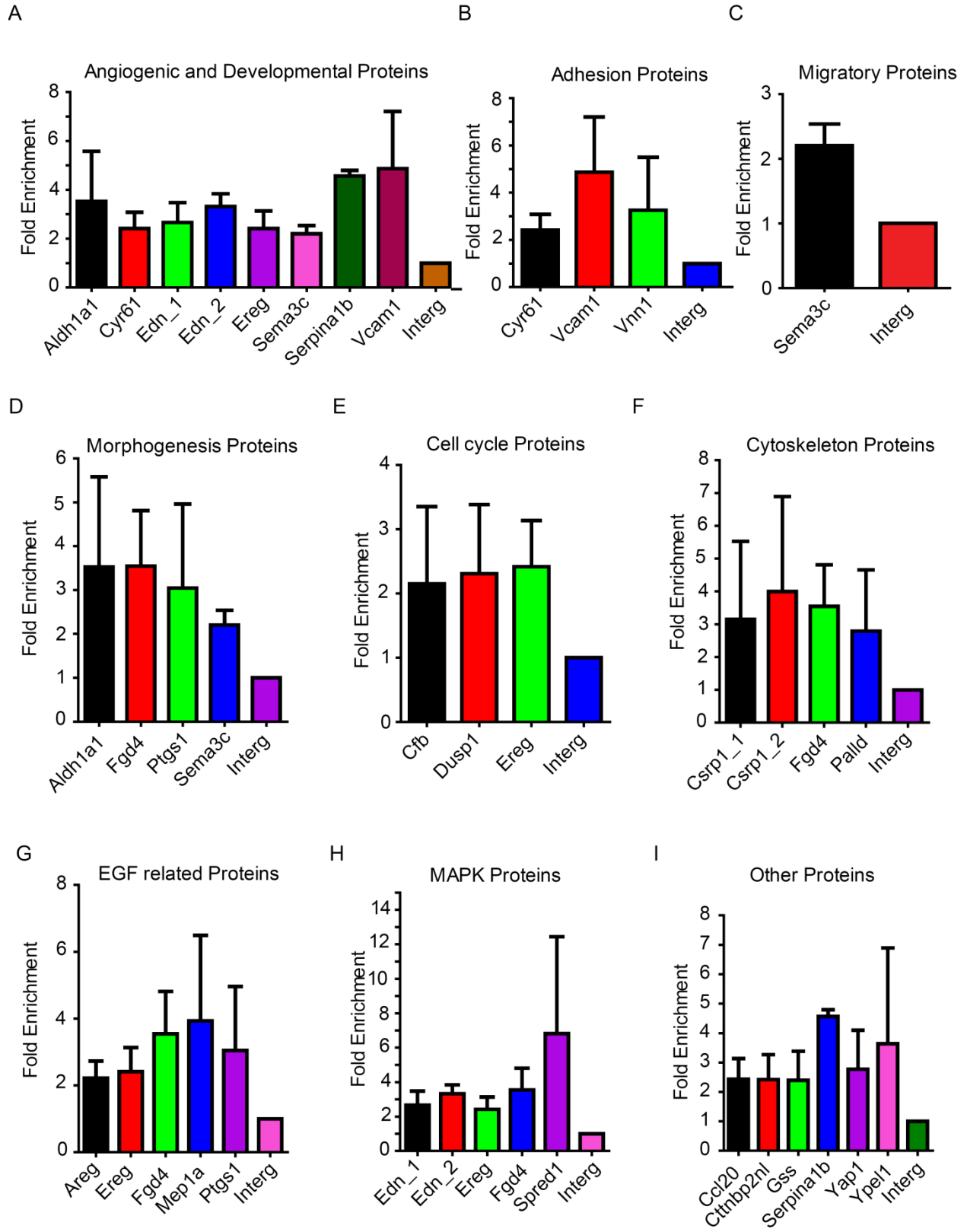
**Figure S2: Sox4 is required for cell proliferation and migration in Py2T cells.** (A) Py2T cells transfected with shControl and shSox4 were treated with TGF- $\beta$  for 2, 4, 7 and 10 days and counted using a Neubauer chamber to assess the proliferation rate during TGF- $\beta$ -induced EMT. (B) Cell cycle analysis was performed by using FACS Diva software after Propidium Iodide staining. (C) Boyden chamber migration assays were performed in Sox4-depleted cells after 15 days of TGF- $\beta$  treatment. 20% FBS was used as chemoattractant and assays were performed for 20 hours. Statistical values were calculated by using an unpaired/paired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes



**Figure S3: Absence of Sox4 impairs breast cancer metastasis.** (A) Metastatic spread was quantified by measuring relative luciferase levels in lymph nodes, lung and liver of nude mice injected with shControl and shSox4. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

Sox4 regulates Epithelial to Mesencymal transition by directly controlling transcription of underlying master genes



**Figure S4: Sox4 directly binds to promoters of crucial EMT genes** (a-i) Chromatin Immunoprecipitation (ChIP) assay is performed using HA-specific antibodies in NMuMG cells transiently transfected with HA-Sox4 construct. Following this, qPCRs were performed to test the occupancy by Sox4 at the promoters of various genes implicated in angiogenesis and development (A), adhesion (B), migration (C), morphogenesis (D), cell cycle (E), cytoskeleton (F), EGF-related (G), Mitogen activated protein kinases related (H) and some other genes (I).

### 3.2.6 Methods and Materials

#### Reagents and antibodies

**Reagents:** TGF- $\beta$  (240-B, R&D systems). DMEM (D5671, Sigma-Aldrich), RPMI-1640 (R0883, Sigma-Aldrich), PBS (D8537, Sigma-Aldrich), trypsin (T4174, Sigma-Aldrich), Opti-MEM (11058, Gibco), FBS (F7524, Sigma-Aldrich), Glutamine (G7513, Sigma-Aldrich), Pencillin/streptomycin (P4333, Sigma-Aldrich), Lipofectamine RNAiMax (11668-019, Invitrogen), Alexa Fluor-488, 568, 633 (Invitrogen), Polybrene (AL-118, Sigma-Aldrich), Puromycin (P7255, Sigma-Aldrich), Fugene HD (12998300, Roche), Trizol (T9424, Sigma-Aldrich), M-MLV reverse transcriptase (M314C 28692233, Promega), SYBR-green PCR MasterMix (Eurogentec) and Bradford reagent (500-0006, Biorad), Protease inhibitor cocktail (P2714, Sigma-Aldrich). **Antibodies: Western Blot:** E-Cadherin (610182, Transduction Laboratories), N-Cadherin (M142, Takara), ZO-1 (617300, Zymed), Fibronectin (F-3648, Sigma-Aldrich),  $\beta$ -Catenin (C 2206, Sigma-Aldrich), p120 (P1870, Sigma-Aldrich), Actin (SC-1616, Sanata Cruz Biotechnology). **Immunofluorescence:** E-Cadherin (13-1900, Zymed), N-Cadherin (610921, Transduction Laboratories), ZO-1 (617300, Zymed), Fibronectin (F-3648, Sigma-Aldrich), Phalloidin (A12380, Invitrogen), Paxillin (13520, Transduction Laboratories), Vimentin (V2258, Sigma). **IHC:** Sox4 (ab52043, Abcam), ChIP: HA (ab9110, Abcam). Apoptosis and Cell cycle: Annexin-V (559934, BD Biosciences) and PI (P4170, Sigma-Aldrich), **Small interfering RNAs:** siControl (Stealth RNAi™ siRNA Negative Controls, 12935-100, Invitrogen), siSox4 (SASI\_Mm01\_00114972 and SASI\_Mm01\_00114974, Sigma-Aldrich) and siEzh2 (SASI\_Mm01\_00061985, SASI\_Mm01\_00061987 & SASI\_Mm01\_00061988, Sigma-Aldrich). **Small hairpin RNA:** shControl (Mission Non-target shRNA control vector, SHC002) and shSox4 (SHCLNG-NM\_009238\_Mouse, TRCN0000012081).

#### Cell lines and cell culture

A subclone of NMuMG cells (NMuMG/E9; hereafter NMuMG) expressing E-cadherin has been previously described (Maeda et al., 2005). MCF7 shControl and MCF7-shEcad have been described before (Lehembre et al., 2008). Py2T cells were derived from Polyoma middle T breast cancer tumor model (unpublished data, Waldmeires et.al). B16-F10 melanoma cells also have been described earlier (Briles and Kornfeld, 1978; Fidler, 1975; Fidler and Bucana, 1977; Fidler et al., 1976; Fidler and Kripke, 1977). NMuMG, MCF7shControl, MCF7-shEcad, Py2T, 293T and PLAT-E cells were cultured in DMEM supplemented with 10% FBS, 2mM glutamine, 100U penicillin and 0.2mg/ml streptomycin. B16-F10 cells were cultured in RPMI-1640 supplemented with 10% FBS, 2mM glutamine, 100U penicillin and 0.2mg/ml streptomycin. All the cells were cultured at 37°C with 5% CO<sub>2</sub> in humid incubator. For TGF- $\beta$  time-course experiments, cells were treated with 2ng/ml TGF- $\beta$  for indicated time point and it was replaced every 2 days. For siRNA transfections, Lipofectamine RNAiMax was used according to the manufacturer's instructions.

### Quantitative RT-PCR

Total RNA was prepared by using a Tri Reagent according to the manufacturer's instructions. RNA was further reverse transcribed with ImProm-II Reverse Transcriptase, and transcripts were quantified by PCR using SYBR-green PCR Mastermix in a real time PCR system (Step One Plus, Applied Biosystems). Human or mouse ribosomal L19 primers were used for normalization. PCR assays were performed in duplicates, and fold induction was calculated against control-treated cell lines using the comparative Ct method ( $\Delta\Delta$  Ct). Following primers were used:

Primer name	Sequences
mRpl19 Forward primer	ctcggtgccggaaaaaca
mRpl19 Reverse primer	tcatccaggtcaccttctca
mSox4 Forward primer	cctcgctctcctcgtcct
mSox4 Reverse primer	tcgtcttcgaactcgtcgt
hRpl19 Forward primer	gatgccggaaaaacaccttg
hRpl19 Reverse primer	cagggcagtgatctccttctg

Sox4 regulates Epithelial to Mesencymal transition by directly  
controlling transcription of underlying master genes

hSox4 Forward primer	ccaaatcttttggggactttt
hSox4 Reverse primer	ctggcccctcaactcctc
mE-cadherin Forward primer	cgaccctgcctctgaatcc
mE-cadherin Reverse primer	tacacgctgggaacatgagc
mN-Cadherin Forward primer	caatgacgtccaccctgttct
mN-Cadherin Reverse primer	ctgccatgactttctacggaga
mFibronectin1 Forward primer	cccagacttatggtggcaatt
mFibronectin1 Reverse primer	atattccgactcgagtctga
mEzh2 Forward primer	caggctggggcatctttatc
mEzh2 Reverse primer	acgaattttgttgcctttc

### **Immunoblot**

Cells were lysed for 1 hour on ice in RIPA-Plus buffer (50mM Tris-HCl, pH8.0), 150mM NaCl, 10% glycerol, 1% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, 2mM CaCl<sub>2</sub>, 1mM dithiothreitol, 1mM sodium fluoride, 0.2mM sodium orthovanadate, 1x protease inhibitor cocktail and further quantified by using Bradford reagent. 50 µg of cleared protein lysates were separated by SDS-PAGE and electroblotted on PDVF membranes, and proteins were visualized with the appropriate primary and secondary antibodies and ECL on superRX films. Depending on the species origin of antibodies, immunoblots were either probed sequentially or on multiple membranes. Adobe Photoshop has been used to excise the relevant portion of the immunoblots from the original scans of X-ray films exposed to chemoluminescence visualization of specific proteins.

### **Immunofluorescence**

siControl, siSox4 and siEzh2 cells were plated on coverslips and treated with TGF-β for mentioned time. The cells were fixed with 4% paraformaldehyde in HBSS and further permeablized with 0.2% Triton for 5 minutes at room temperature. These cells were blocked by using 3.5% goat serum for 15 minutes and incubated with primary antibodies against E-cadherin, N-cadherin, fibronectin, Zo-1, β-catenin, p120 and Ezh2 for 1 hour and then incubated with



fluorochrome-labeled secondary antibody for 1 hour at room temperature. The coverslips were counterstained with DAPI and imaged with a confocal laser-scanning microscope. Data were processed with Adobe Photoshop 7.0 software.

### **Production of lentivirus for in-vivo and in-vitro knockdown studies**

Murine Sox4 shRNAs and control shRNA were purchased from Sigma-Aldrich as described above. For lenti-virus production, 293T cells were transfected with the shRNA expressing lenti-viral vector in combination with the packaging vectors including envelope protein; HDM-pVSV/G, codon-optimized HIV gag-pol; HDM-Hp<sub>g</sub>m2, transactivator of transcription; HDM-Tat1b and pRC-CMV-RaII by Fugene HD. After 48 hours of transfection, viral supernatant was harvested, filtered (0.46  $\mu$ m), supplemented with polybrene (8ng/ml) and used to infect target cells. Infections were performed once a day for two consecutive days. Infected cells were positively selected using Puromycin (5ug/ml).

### **Production of retrovirus for over-expression studies**

HA-Sox4 (kindly provided by Prof. Dr.A-P Tsou) cloned into the retroviral expression vector pBabe. Retroviral particles were produced by transfecting PLAT-E cells with the retroviral expression vectors using Fugene HD. After two days of virus production, retroviral-containing supernatants were harvested, filtered (0.45  $\mu$ m) and added to target cells in presence of polybrene (8 ng/ml). Infections were performed once a day for two days and later on selection was performed by using puromycin. Single clones were picked for the experiments.

### **Wound healing assay**

*In vitro* wound healing assays were done on confluent siControl and siSox4 cells. The media on the confluent cells was replaced with DMEM with 2% fetal bovine serum media and an area of cells was scraped off using a 200ul pipette tips. Light microscopic images were taken at time 0 and at 20 hours. The data was further analyzed by using a ImageJ software.

### **Migration assay**

Cell migration was assessed in siControl and siSox4 cells in a transwell migration assay (pore size: 8  $\mu\text{m}$ ; Falcon BD).  $10^4$  cells were seeded in 2% FBS/DMEM (Sigma) in the upper chamber and the lower chamber was filled with 20% FBS/DMEM. After 20 hours of incubation at 37°C, cells in the upper chamber were carefully removed with a cotton swap and the cells that had traversed the membrane were fixed in 4% paraformaldehyde/PBS, stained with DAPI. Pictures of the membrane were taken at a 10x magnification using a fluorescent microscope (Nikon Diaphot 300). Quantification was done using the software ImageJ.

### **Apoptosis assay (Annexin assay)**

Following supernatant collection to get the floating cells for the assay, cells were trypsinized and mixed with supernatant. Furtheron, cells were washed twice with cold PBS and resuspended in 1X Annexin V binding buffer (10mM Hepes (pH 7.4), 0.14 M NaCl, 0.25 mM CaCl<sub>2</sub>) at a concentration of  $1 \times 10^6$  cells/ml. 5  $\mu\text{l}$  of Cy5 Annexin V was added to the 100  $\mu\text{l}$  of cells ( $1 \times 10^5$ ) and incubated for 15 min on ice in the dark. Stained cells were filtered with a 40 $\mu\text{m}$  mesh and analyzed on a FACSCanto II using DIVA software.

### **Cell growth curve**

$1 \times 10^4$  cells were seeded in each well of 24-well plate and cell numbers were assessed for mentioned days by using a Neubauer counting chamber.

### **PI Staining**

siControl, siSox4 and siEzh2 cells were trypsinized and fixed in 70% ice-cold Ethanol for overnight. Washed twice with PBS and resuspended in sodium citrate buffer with 5 $\mu\text{g/ml}$  PI for overnight. Stained cells were analyzed by FACSCanto II using DIVA software.

### **Fat-pad injection, Subcutaneous injection and Intravenous injection**

7 weeks old female nude mice were injected orthotopically into the fat-pad with  $1 \times 10^6$  shControl and shSox4 Py2T cells in 100 $\mu\text{l}$  of PBS. After 23 days of treatment, mice were sacrificed and tumor, lymph-nodes, lungs and liver were isolated, weighted and prepared as described in histological analysis.

shControl and shSox4 B16-F10 melanoma cells were harvested by trypsinization and resuspended in PBS to a final concentration of  $1 \times 10^5$  cells/200 $\mu$ l. 8 weeks old C57/Bl6 mice were injected intravenously through tail-vein with 200  $\mu$ l of each cell line, and waited for 18 days before sacrificing the mice. After sacrificing the mice, lung metastasis were counted.

shControl and shSox4 cells were harvested by trypsinization and resuspended in PBS to a final concentration of  $1 \times 10^6$  cells/100 $\mu$ l. 8 weeks old C57/Bl6 mice were injected into the right flank with each cell line, and tumors were allowed to form for 27 days. Mice were sacrificed when the tumors reached a diameter of 1.5 cm and lymph-nodes, lungs and liver were isolated and processed for histological analysis.

### **Histological analysis**

The preparation of immunohistochemical analysis was performed as described previously (Perl et al., 1998). Briefly, for the paraffin section preparation, tumors, lungs and livers from above mouse experiment were removed and fixed in HBS-Ca<sup>2+</sup> (HEPES-buffered saline, 1 mM CaCl<sub>2</sub>) containing 4% paraformaldehyde for overnight and processed further for H&E staining. For OCT frozen sections, organs were removed and fixed in HBS-Ca<sup>2+</sup> (HEPES-buffered saline, 1 mM CaCl<sub>2</sub>) containing 4% paraformaldehyde for 2 h at 4 °C. Tissues were incubated overnight at 4 °C in HBS-Ca<sup>2+</sup>/20% sucrose, embedded in O.C.T. compound (Tissue Tek), and frozen in liquid nitrogen. H & E Staining was performed as described before (Perl et al., 1998) and evaluated on an AxioVert microscope.

### **Immunohistochemistry**

shControl and shSox4 Py2T tumor sections deparaffined in xylene and rehydrated. Sox4 antigenic recovery was carried out with a pressure cooker for 15 min in 10mM sodium citrate buffer with 0.05% Tween 20, pH 6.0. Slides were washed with TBS with 0.025% Triton X-100 and blocked with 10% goat serum with 1% BSA in TBS for 1 hour at room temperature. After blocking, the sections were incubated with Sox4 antibody in TBS with 1% BSA for overnight. Sections were washed again in TBS with 0.025% Triton X-100 and incubated with 0.3 %H<sub>2</sub>O<sub>2</sub> to block the endogenous peroxidase. Sections were washed again with TBS and bound antibodies

were detected using the ABC vector-horseradish peroxidase kit according to the manufacturer's instructions.

### **Microarray processing and data analysis**

RNA was isolated from NMuMG cells transfected with control siRNA or Sox4 siRNA and treated with TGF- $\beta$  for 0 and 2 days using Tri Reagent (Sigma-Aldrich). RNA quality and quantity was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies). The manufacturer's protocols for the GeneChip platform by Affimetrix were followed. Methods included synthesis of the first- and second-strand cDNA followed by synthesis of cRNA by *in vitro* transcription, subsequent synthesis of single-stranded cDNA, biotin labeling and fragmentation of cDNA and hybridization with the microarray slide (GeneChip® Mouse Gene 1.0 ST array), posthybridization washings and detection of the hybridized cDNAs using a streptavidin-coupled fluorescent dye. Hybridized Affimetrix GeneChips were scanned using an Affimetrix GeneChip 3000 scanner. Image generation and feature extraction were performed using Affimetrix GCOS Software and quality control was performed using Affimetrix Expression Console Software. Raw microarray data were normalized with Robust Multi-Array (RMA) and analyzed using Partek® Genomics Suite Software (Partek Inc.). One-way analysis of variance (ANOVA) and asymptotic analysis were used to identify significantly differentially expressed genes. The gene ontology (GO) tool from Partek® Genomics Suite Software as well as the David gene ontology software were used for further analysis.

### **Chromatin Immunoprecipitation**

ChIP experiments were performed as previously described (Weber et al, 2007). In brief, crosslinked chromatin was sonicated to achieve an average fragment size of 500 bp. Starting with 100  $\mu$ g of chromatin and 5  $\mu$ g of anti-HA antibody, 1  $\mu$ l of ChIP material and 1  $\mu$ l of input material were used for quantitative real-time PCR using specific primers covering the motif of Sox4 in the promoter of target genes. Primers covering an intergenic region are used as a control. The efficiencies of PCR amplification were normalized for between the primer pairs. Following primers were used for ChIP –PCR.

Sox4 regulates Epithelial to Mesencymal transition by directly  
controlling transcription of underlying master genes

Name	Sequences
Aldh1a1 FP	ggtttgctggtagccatggt
Aldh1a1 RP	cattcccttttctgctttg
Cyr61 FP	ggcaagtggagaagggtga
Cyr61 RP	gcacctcgagagaaggacac
Edn_1 FP	ttccttcttgctaactctg
Edn1_1 RP	agatcacctactttccacca
Edn_2 FP	gtttgcattgagttccattg
Edn_2 RP	gtcacgaacagcagagagaaga
Sema3c FP	agtctcaaagtgttgggacat
Sema3c RP	ttatgctgacctttccacactg
Serpina1b FP	attctctgcaaaggaaacaaa
Serpina1b RP	ggggctcagagtaggttatgtg
Vcam1 FP	tctgcatcaacgtcctttca
Vcam 1 RP	cccattatcatgagtcactctttt
Vnn1 FP	gttcaagtgacagctgagtgct
Vnn1 RP	ccctggggttttctttaaattc
Fgd4 FP	cttcagaatgagcctgtttcaa
Fgd4 RP	tgcatgcatgaaaactacacac
Ptgs1 FP	aaaacaactcccctcaccttt
Ptgs1 Rp	gggcagtgagtgggatgtaa
Cfb FP	tccttggacggagatacagg
Cfb RP	ggaagagacaacagggtgga
Dusp1 FP	cagcttctgttcagtggagatg
Dusp1 RP	ttgctgtgtagctctggctagt
Ereg FP	gcatttgagacaggcacaga
Ereg RP	ccctcagcttccaatgtgat
Csrp1_1 FP	gttcaaggccattctggtctac
Csrp1_1 RP	gctatgggtgggagtgttagag

Sox4 regulates Epithelial to Mesencymal transition by directly  
controlling transcription of underlying master genes

Csrp1_2 FP	tgacttctctttgcacacctatg
Csrp1_2 RP	ctcacagttaggtctccgcttt
Palld FP	gaagaactgaccacatggctaa
Palld RP	aattacctcccagccttttctc
Areg FP	cattatgcagctgctttgga
Areg RP	tttcgcttatggtggaacc
Mep1a1 FP	gagtcaccacgagacaagca
Mep1a1 RP	ggggctttgttacacaggaa
Spred1 FP	actcatgccagttccattcttt
Spred1 RP	acgctcaagtccccgttact
Ccl20 FP	gttgagactggtgtttccaca
Ccl20 RP	ccagtcctagaggggaaagatt
Ctnnbp2nl FP	aattaagccaccacaaggtgtt
Ctnnbp2nl RP	gaggaagtaccctcctctggtc
Gss FP	taggttgccagaggatgagg
Gss RP	tgtgagatggggacactcaa
Yap1 FP	agcttcaaaaaccccgttct
Yap1 RP	ggctaaagcagcacaggaac
Ypel1 FP	gctgggctacgggtccaaaaca
Ypel RP	gcgccccatccggcccaggct
Ezh2 FP	gtcacacgccttcctttcagt
Ezh2 RP	gcctgagccaagtttgaatagt

### Statistical analysis

Statistical analysis and graphs were generated using the GraphPad Prism software (GraphPad Software Inc, San Diego, CA). All statistical analysis was done by unpaired, two-sided t-test. Normality testing was performed using the Kolmogorov-Smirnov test with Dallal-Wilkinson-Lillie for p-values.

### **3.3 Polycomb-dependent mechanisms regulate Epithelial to Mesenchymal transition**

Neha Tiwari\*, Vijay K. Tiwari\*, Piotr Balwierz, Phil Arnold, Erik Van Nimwegen, Dirk Schübeler and Gerhard Christofori

(\*Equal contribution)

*Manuscript in preparation*

#### **3.3.1 Abstract**

Epithelial-to-mesenchymal transition (EMT) represents a key example of cell plasticity that underlies crucial steps during development, wound healing and carcinogenesis. Despite emerging evidences that genes involved in EMT are altered in tumors via both genetic and epigenetic changes, the later remains largely unknown. In the current study, we comprehensively investigated the role of two prominent epigenetic modifications – Histone 3 Lysine 27 trimethylation (H3K27me3) and DNA methylation during TGF- $\beta$ -induced EMT of mammary epithelial cells. A genome-wide ChIP-seq analysis for the H3K27me3 mark during six consecutive stages of EMT progression uncovered a number of key EMT genes that are transcriptionally regulated by the Polycomb-mediated H3K27me3 during this process. Genes such as *Mcam*, *Pdgfrb* and *Itga5* are enriched with H3K27me3 and repressed in epithelial cells and treatment with TGF- $\beta$  leads to a progressive loss of this mark that accompanies their transcriptional activation and acquisition of EMT. Another set of genes such as *Cdh1*, *Ocln* and *Cdx2*, gain this mark and become repressed during TGF- $\beta$ -induced EMT. We further show that the coordinated activities of both *Ezh1* and *Ezh2* are responsible for H3K27me3-mediated repression in these cells as co-depletion of these two enzymes not only removes the H3K27me3 mark from the target promoters and de-represses them but also prevents EMT. A genomewide analysis of promoter methylation using Methylated DNA Immunoprecipitation (MeDIP) in combination with tiling arrays revealed no major changes during EMT. Taken together, our data provides evidence that Polycomb machinery-mediated epigenetic reprogramming underlies transcriptional changes driving EMT.

#### **3.3.2 Introduction**

TGF- $\beta$  belongs to most potent inducers of the epithelial-mesenchymal transition (EMT). Cancer cells which undergo EMT are able to detach from the primary tumor by breaking their cell-cell contact and follow the chemoattractive path through extracellular matrix to invade into the surrounding tissue and form metastatic lesions at distant sites (Grunert et al., 2003; Polyak and Weinberg, 2009; Thiery and Sleeman, 2006). Interestingly, besides promoting invasiveness, TGF- $\beta$ -induced EMT was shown to induce the transition of transformed and immortalized human mammary epithelial cells into mesenchymal cancer cells with stem cell traits, thus linking EMT to tumor cell plasticity (Cao et al., 2008; Shipitsin et al., 2007). It is established that such TGF- $\beta$ -mediated cellular transformation involves activation of downstream signaling cascades and subsequent transcriptional reprogramming. In addition to the remarkable variety of transcription factors and co-regulators, the role of chromatin-mediated transcription control in transcriptional regulation has been increasingly appreciated. However, there is no solid evidence that EMT involves epigenetic reprogramming events to initiate transcriptional changes.

The Polycomb Group (PcG) and trithorax Group (trxG) of proteins are involved in defining cellular memory and prevent changes in cell type-specific transcription programs to maintain cell identity (Bantignies and Cavalli, 2006; Cao et al., 2005; Cao et al., 2002; Cao and Zhang, 2004a; Cao and Zhang, 2004b; Jacobs and van Lohuizen, 1999; Kuzmichev et al., 2004; Negishi et al., 2007; Vire et al., 2006). Among Polycomb complexes, Polycomb repressive complex 2 (PRC), comprised of EZH2, EED, SUZ12, RbAp46/48 and AEBP2 (E(z), Esc, Su(z)12, and RbAp48 in *Drosophila*) of which EZH2 bears the enzymatic activity that trimethylates Histone H3 at Lys 27 (H3K27me3) (Cao and Zhang, 2004a; Cao and Zhang, 2004b; Satijn et al., 2001; Sparmann and van Lohuizen, 2006a). Such PRC2 activity is known to associate with transcriptional repression by several mechanisms. PRC2 has been shown to recruit DNA methyltransferases (DNMT) to the chromatin for repression (Vire et al., 2006). PRC2 binding has also been reported to attract Histone deacetylases (HDAC) to deacetylate the histone tails for gene repression (Wang et al., 2004b). Furthermore, it has also been proposed that Polycomb complexes lead to condensed chromatin structure and thus physically hinder the transcription machinery (Wang et al., 2004b) PRC2 is also known to facilitate binding of another Polycomb complex, PRC1, which then contributes to the maintenance of gene repression (Cao et al., 2005; Wang et al., 2004a; Wang et



al., 2004b). Thus, PRC2 is thought to be involved in the initiation and PRC1 in the maintenance of gene repression.

Ezh1, unlike Ezh2, is a part of a non-canonical PRC2 and mediates methylation of H3K27 similar to Ezh2 and functions in the maintenance of embryonic stem cell pluripotency and plasticity (Shen et al., 2008). Ezh1 was shown to colocalize with the H3K27 trimethylation mark in Ezh2 *-/-* cells and to preferentially preserve this mark at development-related genes. Co-depletion of both Ezh1 and Ezh2 abolished residual methylation on H3K27 and derepressed H2K27-trimethylated target genes (Shen et al., 2008). Compared to Ezh2, Ezh1 has very weak methyltransferase activity but efficiently represses transcription and compacts chromatin (Margueron et al., 2008). It is also been demonstrated that Ezh1 is expressed in non-proliferative cells while Ezh2 is expressed in proliferative cells which again argues for their evolution for different roles.

In the current study, we investigated whether the Polycomb machinery is involved in transcriptional reprogramming during EMT. Since it has been shown earlier that E-cadherin can be repressed by Ezh2 in the prostate cancer (Cao et al., 2008) and that low Ezh2 expression levels are correlated with metastasis-free survival in breast cancer (Kleer et al., 2003), we investigated whether Ezh2-catalyzed H3K27me<sub>3</sub> is involved in the transcriptional modulation during TGF- $\beta$ -induced EMT in breast carcinogenesis. Towards this, we used the normal mammary epithelial cells, NMuMG, which undergo progressive EMT upon TGF- $\beta$  treatment (Supplemental figure 1a). We performed Chromatin Immunoprecipitation using H3K27me<sub>3</sub>-specific antibody at six different morphological stages during TGF- $\beta$ -induced EMT followed by next generation sequencing (ChIP-seq). A comparison of identified H2K27me<sub>3</sub> enriched genes with global transcription profiling of the corresponding stages revealed a number of key EMT genes that may possibly be transcriptionally regulated by the PRC2 machinery. By carrying out loss of function studies, we demonstrate that double-knockdown of Ezh1 and Ezh2 not only prevents TGF- $\beta$ -induced EMT but also blocks EMT-driven cell migration. Such depletion further suppresses TGF- $\beta$ -induced apoptosis, arguing for a H3K27me<sub>3</sub>-mediated regulation of survival-associated genes. Gene ontology studies on gene regulated by Ezh1 and Ezh2 further supported our data for a Polycomb-mediated mechanisms regulating transcription of genes underlying EMT-related migration, adhesion and survival. Moreover, comprehensive MeDIP-chip analysis

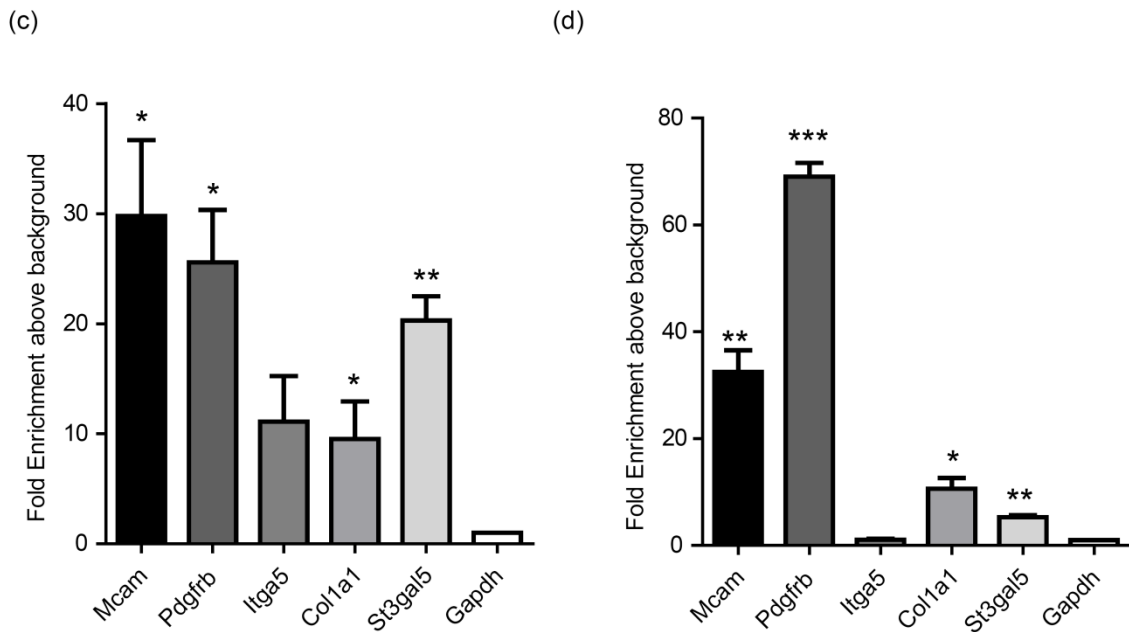
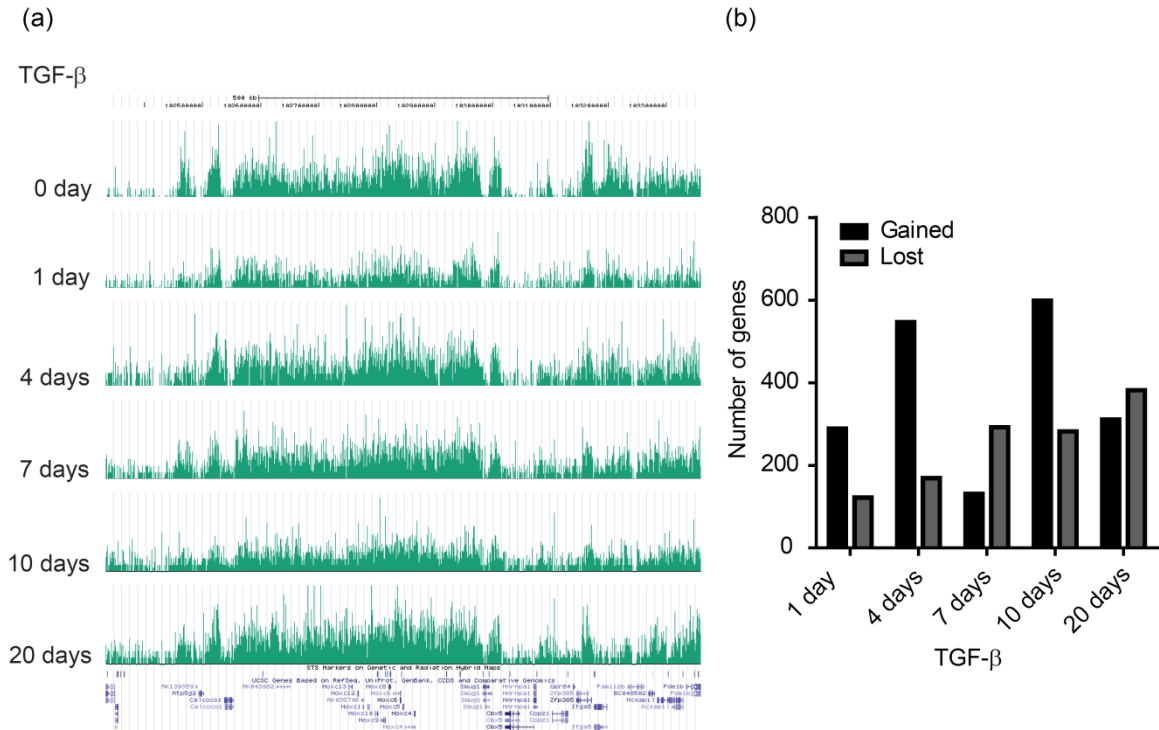
at three different stages of TGF- $\beta$ -induced EMT revealed no significant changes in DNA methylation. In summary, our data reveals new insights into the epigenetic reprogramming of key EMT genes by Polycomb-mediated mechanisms.

### **3.3.3 Results**

#### **3.3.3.1 Genome-wide analysis of H3K27me3 mark reveals widespread epigenetic reprogramming during TGF- $\beta$ -induced EMT**

To assess the role of Polycomb group of proteins during EMT, we employed NMuMG cells which are untransformed normal murine mammary gland cell line and undergo progressive EMT in response to TGF- $\beta$  (Lehembre et al., 2008; Miettinen et al., 1994; Piek et al., 1999a), acquiring the complete mesenchymal transition by the end of 10 days (Supplemental figure S1a). To illustrate the dynamics of the PRC2-mediated repressive histone modification H3K27me3 during EMT, we performed Chromatin immunoprecipitation (ChIP) using a H3K27me3-specific antibody upon TGF- $\beta$  treatment in NMuMG cells for 1, 4, 7, 10 and 20 days and subjected the precipitated DNA to next generation sequencing (ChIP-seq). This high resolution profiling showed that H3K27me3 is enriched in broad domains for the Hox locus that mostly remains unchanged during EMT (Figure 1a). We next attempted to reveal H3K27me3 peaks associated with genes using a sliding window approach (see experimental Procedures). This analysis identified a total of 2034 genes (Figure 1b). To correlate H3K27me3 enrichment with transcription levels of associated genes, we compared this list with the genome wide expression analysis performed in the non-treated NMuMG cells as well as those treated with TGF- $\beta$  for 1, 4, 7, 10 and 20 days. Using this comparison, we sorted out 90 genes which are transcriptionally upregulated during EMT and lose the H3K27me3 mark and 181 genes which are transcriptionally downregulated during EMT and gain the H3K27me3 mark (Supplemental figure 1b, 1c, 1d and 1e and table 1). We further validated the ChIP-Seq data using single gene controls (Figure 1c). To further strengthen our data, we used the Py2T cells (Waldameier et.al, unpublished data) which are derived for MMTV-PyMT transgenic mice tumors and undergo EMT upon TGF- $\beta$  treatment in a similar fashion to NMuMG cells. All tested promoters were

similarly enriched for H3K27me3 in Py2T cells (Figure 1d). In summary, genomewide location analysis identifies targets of the PRC2-associated mark H3K27me3 during EMT.



**Figure 1:** Genome-wide reprogramming of the Polycomb-associated mark H3K27me3 during TGF- $\beta$ -induced EMT. (a) Following ChIP using H3K27me3-specific antibody in non-treated as well as TGF- $\beta$  treated (1, 4, 7, 10 and 20 days) NMuMG cells, next-generation sequencing was carried out. Genome browser view of H3K27me3 ChIP enrichment at the HoxD cluster for several stages of TGF- $\beta$ -induced EMT. (b) Dynamics of H3K27me3 upon TGF- $\beta$  treatment for 1, 4, 7, 10 and 20 days in NMuMG cells. The numbers of genes have been counted at each stage in comparison with untreated NMuMG cells and having a Z-Score above 3.5. (c) & (d) ChIP-qPCR validation of H3K27me3 enrichment at various identified targets in NMuMG (c) and Py2T (d) cells. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

### 3.3.3.2 Polycomb targets are key EMT genes

To gain insight into the categories of genes represented within the promoter class, we performed a gene ontology analysis revealing that many of the H3K27me3 targets genes have already been implicated in various cancers and in EMT. The EMT relevant genes include E-cadherin (Cdh1) (Cavallaro, 2004; Wheelock et al., 2008), melanoma cell adhesion molecule (Mcam) (Zabouo et al., 2009), platelet-derived growth factor receptor beta (Pdgfrb) (van Zijl et al., 2009), Snail2 (Snai2) (Herranz et al., 2008; Peinado et al., 2007), Integrin alpha 5 (Itga5) (Kim et al., 2010), Occludin (Ocln) (Ikenouchi et al., 2003), Fibroblast growth factor receptor 2 (FGFR2) (Kato and Kato, 2009) and Inhibitor of differentiation 1 (Id1) (Tobin et al., 2011). In addition, genes related to extracellular matrix, cell adhesion, cytoskeleton re-modeling, apoptosis, cell-cycle, cell signaling pathways and DNA replication were also over-represented (Table 2, Table 3 and Supplemental Figure S2).

### 3.3.3.3 Ezh1 and Ezh2 both contribute to the H3K27me3-mediated repression

It is long known that Ezh2, a catalytic subunit of the PcG complex, is a main player in setting the H3K27me3 mark at the promoters. There is emerging evidence that Ezh1, a member of the non-canonical PcG complex also mediates transcriptional repression via H3K27me3 (Ezhkova et al., 2011; Margueron et al., 2008; Shen et al., 2008; Stojic et al., 2011). To dissect the contribution of the two enzymes of the PRC complex in our system, we depleted Ezh1 and Ezh2 expression individually as well as together by using pools of two or three siRNAs, respectively. RT-qPCRs

Polycomb-dependent mechanisms regulate Epithelial to  
Mesenchymal transition

(a)

Downregulated						
1110032A04Rik	Cdh1	Fbln1	Kif12	Oxr1	Sgsh	Zfp704
5033414D02Rik	Cdk6	Fbxo2	Krt5	Pard6b	Sh3bgr3	
Abcb6	Cdx2	Fdps	Krt7	Pbx4	Slamf9	
Abcd3	Cebpb	Fgfr2	Krt79	Pdik1	Slbp	
Adcy9	Cfb	Fhdc1	Krtcap2	Pdzk1ip1	Slc17a2	
Agxt2l2	Chaf1b	Folr1	Lasp1	Phgdh	Slc25a35	
Aldh3a1	Chchd6	Fosl1	Lcn2	Pim3	Slc25a39	
Ank	Cldn9	Foxa2	Lgals3bp	Plekha7	Slc31a2	
Anxa9	Clic3	Fzd5	Lgals4	Plk4	Slc35a5	
Apoa1bp	Cltb	Gaa	Lgals6	Pola2	Slc7a11	
Arhgef10l	Cnm1	Gatc	Lig1	Pon3	Slc9a3r1	
Arl4a	Cox15	Gcnt1	Limk2	Ppt2	Slc9a3r2	
Arl6ip6	Cpn1	Gdf15	Lonp1	Pqlc3	Smpd3	
Atp6v0e2	Cpne2	Gmip	Lpin1	Prkca	Snap29	
Atp9a	Cutc	Gnai1	Lsm4	Prkcz	Sorbs1	
Avpi1	Cyb561	Gnb1l	Ltr	Psrc1	Spc25	
Bcl6	Cyp51	Gpd1	Macrocl1	Rad23a	Spire2	
Birc5	Cys1	Gpt	Mea1	Rapgef3	Stra13	
Brp16	Dsg2	Grhr	Mgat3	Rasip1	Tapbl3	
Cables2	Dusp6	Grtp1	Mlh1	Recql4	Tgfb3	
Calm3	Egln3	Gstm3	Mlph	Rfc1	Tmbim1	
Camk1g	Entpd7	Gstm5	Mthfd2	Rnf167	Tmco3	
Car2	Eps8l2	Hal	Nacc2	S100a8	Tmem51	
Cav2	Eps8l3	Hells	Ncapd2	Scn8a	Tnxb	
Ccdc88c	Erlin1	Helq	Nedd4l	Scnn1a	Trpt1	
Ccna2	Esrp2	Hoxb13	Nfat5	Sec14l1	Tspan15	
Ccne2	Ets2	Hpx	Notch1	Selenbp1	Ubxn8	
Cd274	Fam134a	Hvcn1	Nt5c3	Sema3c	Vps28	
Cdc20	Fam83h	Id1	Ocl1	Sema3e	Vps53	
Cdc6	Fas	Igfb7	Ocln	Sema4a	Zcchc7	

(b)

Upregulated		
Acta2	Gabbr1	Pcnx
Actg2	Glipr2	Pde4dip
Acvr1	Gnaz	Pdgfrb
Adarb1	Gpx8	Pkia
Aebp1	H2-DMa	Pkp1
Apob48r	Hk2	Plek
Asgr1	Ifitm3	Prrc1
AU021092	Il11	Pstpip1
Bmp7	Il24	Ptp4a3
Cacnb3	Il6ra	Rgs16
Capp	Itga5	Rgs3
Ccdc109b	Kcnab3	Sdc3
Cd1d1	Kremen1	Sema4f
Chn1	Krt17	Serpinf1
Chna1	Lce1g	Sh3bp2
Clstn1	Lcp1	Snai2
Cmah	Lgr6	Sparcl1
Cnn1	Lpin3	St3gal5
Col1a1	Ltbp1	St6galnac4
Col4a3	Mcam	St6galnac6
Col4a4	Meox1	Stmn4
Col7a1	Mpdz	Tcf7
Cxcl5	Mras	Tcn2
Cyp11a1	Myl7	Tfpi
Dbh	Myo1b	Tnfsf13b
Dll4	Ncoa1	Tpm3
Fam117a	Nfatc1	Trim54
Fgd1	Nfatc2	U46068
Fkbp1b	Npr2	Wnt7a
Foxf2	Osbpl5	Wnt9a

**Table 1:** Genes that are transcriptionally downregulated and gain the H3K27me3 mark during EMT (a) and vice versa (b).

Polycomb-dependent mechanisms regulate Epithelial to  
Mesenchymal transition

Upregulated genes_Lost mark	
Top GeneGo Pathway Maps	
name	p-Value
Development_Regulation of epithelial-to-mesenchymal transition (EMT)	1.5799E-05
Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	0.000223001
Neurophysiological process_Dopamine D2 receptor transactivation of PDGFR in CNS	0.000288001
Development_Transactivation of PDGFR in non-neuronal cells by Dopamine D2 receptor	0.000489235
Cell adhesion_Tight junctions	0.000763587
Apoptosis and survival_APRIL and BAFF signaling	0.00096713
Cell adhesion_Chemokines and adhesion	0.001556797
Immune response_NF-AT signaling and leukocyte interactions	0.001567452
Development_TGF-beta-dependent induction of EMT via RhoA, PI3K and ILK.	0.001567452
Immune response_Histamine signaling in dendritic cells	0.001995675
Top GeneGo Process Networks	
name	p-Value
Muscle contraction	1.40761E-05
Development_EMT_Regulation of epithelial-to-mesenchymal transition	2.3359E-05
Signal transduction_WNT signaling	0.000119396
Development_Skeletal muscle development	0.000209977
Cytoskeleton_Actin filaments	0.000710113
Development_Ossification and bone remodeling	0.002136825
Cell adhesion_Integrin priming	0.002407517
Development_Neurogenesis_Synaptogenesis	0.004220694
Cytoskeleton_Regulation of cytoskeleton rearrangement	0.004575658
Cell adhesion_Amyloid proteins	0.005489456

Downregulated genes_Gained mark	
Top GeneGo Pathway Maps	
name	p-Value
Signal transduction_cAMP signaling	1.43305E-07
Regulation of lipid metabolism_Regulation of lipid metabolism by niacin and isoprenaline	8.44346E-06
Development_NOTCH-induced EMT	4.60186E-05
ENaC regulation in airways (normal and CF)	0.000244319
Cell cycle_Role of APC in cell cycle regulation	0.000382927
Regulation of CFTR activity (norm and CF)	0.000409088
Transport_ACM3 in salivary glands	0.001000374
Development_Notch Signaling Pathway	0.001198934
Cell adhesion_Histamine H1 receptor signaling in the interruption of cell barrier integrity	0.001423589
Development_Thrombopoietin-regulated cell processes	0.001423589
Top GeneGo Process Networks	
name	p-Value
Cell adhesion_Amyloid proteins	1.38193E-05
Signal transduction_WNT signaling	3.75641E-05
Development_Hedgehog signaling	5.54331E-05
Cell cycle_Core	0.000210764
Development_Blood vessel morphogenesis	0.000359325
Cell cycle_S phase	0.001194119
Cardiac development_Wnt_beta-catenin, Notch, VEGF, IP3 and integrin signaling	0.001246825
Signal Transduction_Cholecystokinin signaling	0.003825099
Neurophysiological process_Corticoliberin signaling	0.004963407
Corticoliberin signaling	0.004963407

**Table 2:** Gene-Ontology analysis revealed crucial EMT relevant pathways under regulation by Polycomb. GeneGo software was used to categorize the genes into functional groups. Gene ontology on the genes that lose the H3K27me3 mark and are transcriptionally up-regulated (a) during EMT and vice-versa (b).

Polycomb-dependent mechanisms regulate Epithelial to  
Mesenchymal transition

(a) Upregulated genes\_Lost mark

Extracellular Matrix	Cell Adhesion	Cytoskeleton Organization	Neuronal Differentiation
Col4a4	Aebp1	Capg	Bmp7
Col1a1	Clstn1	Cnn1	Sema4f
Col4a3	Col4a3	Lcp1	Serpinf1
Col7a1	Col7a1	Myo1b	Wnt7a
Ltbp1	Itga5	Plek	
Sparc1	Mcam	Pstpip1	
Wnt9a	Mpdz	Sdc3	
Wnta7a	Pkp1	Tpm3	
	Plek	Trim54	
	Pstpip1		

Angiogenesis	Cell Projection	Apoptosis/Survival	Cancer
Acv11	Clstn1	Bmp7	Col4a4
Col1a1	Fgd1	Col4a3	Pdgfrb
Dbh	Lcp1	Dbh	Tcf7
Dll4	Mpdz	Snai2	Tpm3
	Myl7	Tcf7	Wnt7a
	Pstip1	Tnfs13b	Wnt9a
	Tpm3		

(b) Downregulated genes\_Gained mark

Cell Cycle	DNA Replication	Cytoskeleton Organization	Morphogenesis
Avpi1	Ccne2	Bcl6	Car2
Birc5	Cdc6	Birc5	Cdh1
Cables2	Chaf1b	Fhdc1	Cdx2
Calm3	Krt7	Lasp1	Cebpb
Ccna2	Lig1	Lpin1	Fgfr2
Ccne2	Pla2	Mlh1	Foxa2
Cdc20	Rfc1	Prkcz	Hells
Cdc6		Psrc1	Hoxb13
Cdk6		Sorbs1	Id1
Chaf1b		Spc25	Notch1
Hells		Tnxb	Phgdh
Krt7			Plk4
Lig1			Sema3c
Mlh1			Tgfr3
Ncapd2			
Pard6b			
Phgdh			
Pim3			
Psrc1			
Smpd3			
Spc25			

Angiogenesis	Cell-Cell junction	Apoptosis/Survival	Cancer
Cdx2	Cdh1	Bcl6	Birc5
Fgfr2	Cldn9	Birc5	Ccne2
Fzd5	Dsg2	Cdh1	Cdh1
Id1	Lasp1	Cebpb	Cdk6
Notch1	Ocln	Fas	Egnl3
Sema3c	Pard6b	Hells	Fas
Tgfr3	Prkcz	Mlh1	Fgfr2
	Snap29	Notch1	Fzd5
	Sorbs1	Prkca	Mlh1
			Prkca

**Table 3:** Function analysis of regulated genes. David software was used to group genes into functional categories. (a) Genes that lose the H3K27me3 during EMT have a role in extra-cellular matrix organization, cell adhesion, cytoskeleton organization, neuronal differentiation, angiogenesis, cell projection formation, cell proliferation, apoptosis/survival and cancers. (b) GenesA that gain the H3K27me3 mark are necessary for cytoskeleton

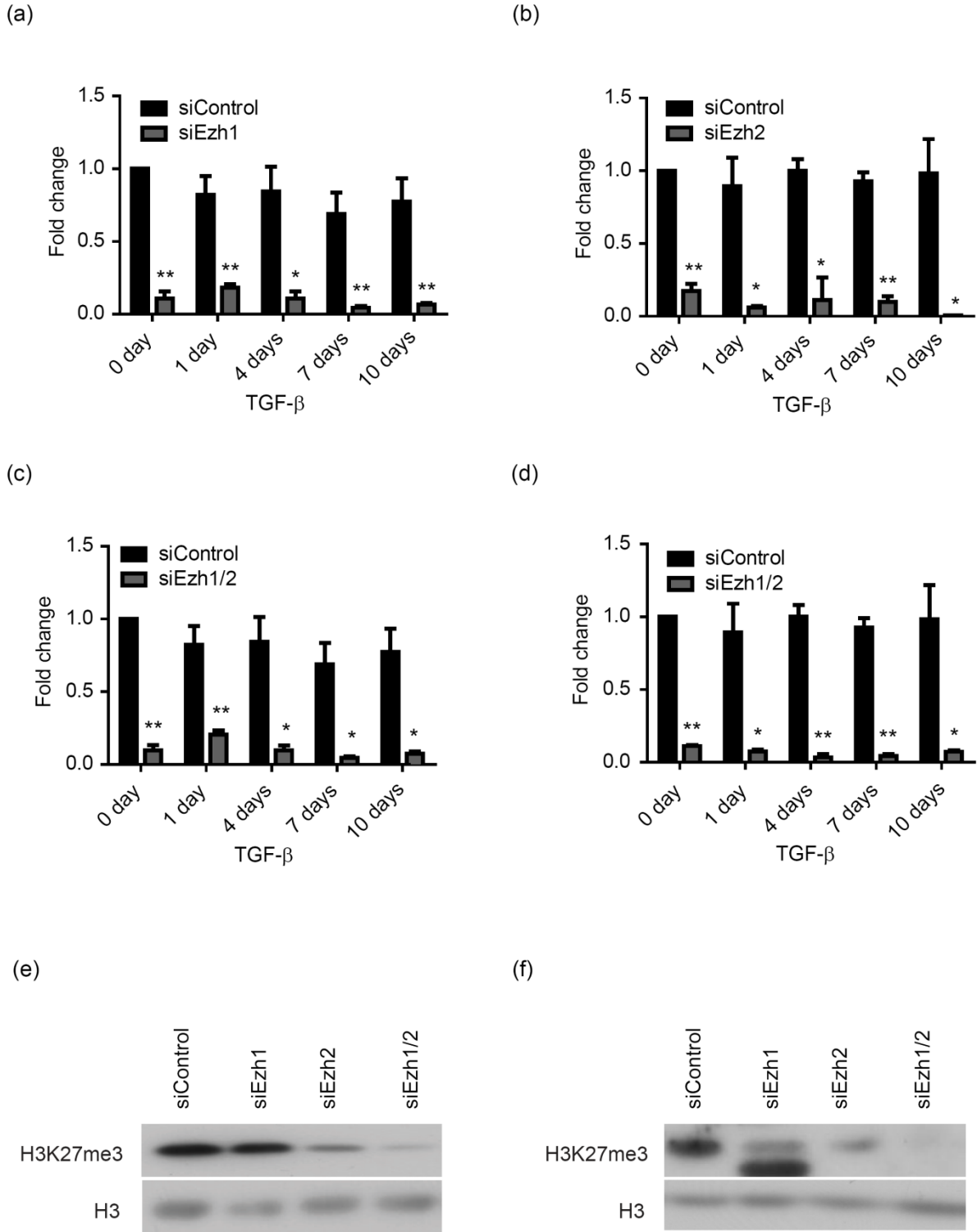
organization, angiogenesis, apoptosis/survival and cancer. Furthermore, they are also required for cell-cycle, replication, and morphogenesis of organs and cell-cell integrity.

confirmed efficient depletion of Ezh1 and Ezh2 expression (Figure 2a-d). Immunoblot analysis suggested that Ezh2, to the large extent, contributed to the overall H3K27me3 levels in NMuMG cells (Figure 2e). To assess whether reduction of Ezh1 and Ezh2 (Ezh1/2) provide an additive effect, we tested transfected cells with pooled Ezh1 and Ezh2-specific siRNAs. Immunoblot analysis revealed that double-knockdown provides a more efficient reduction in global H3K27me3 levels (Figure 2e). We further determined the H3K27me3 levels in Py2T cells upon depletion of Ezh1, Ezh2 as well as double-knockdown of Ezh1 and Ezh2. This experiment revealed a reduction in global H3K27me3 levels even after Ezh1 knockdown and the depletion efficiency is higher after double knockdown of Ezh1 and Ezh2 (Figure 2f). Together, these results suggest that Ezh1 and Ezh2 together contribute to the H3K27me3 levels in NMuMG and Py2T cells.

#### **3.3.3.4 Co-depletion of Ezh1 and Ezh2 blocks epithelial differentiation**

We next employed loss of function studies for studying the impact of Ezh1/2 on EMT. Interestingly, Ezh1/2-ablated cells retained their epithelial phenotype and were not able to undergo EMT upon TGF- $\beta$  treatment (Figure 3a). This further accompanied the retention of the epithelial marker E-cadherin at the same time decrease in the levels of mesenchymal markers such as N-cadherin and fibronectin (Figure 3b; Supplemental figure S3a-c). Immunofluorescence studies showed that Ezh1/2 knockdown cells were able to maintain the adherent junction protein E-cadherin and tight junction protein ZO-1 at their membranes, while the mesenchymal marker N-cadherin remained unchanged to the control cells after 7 days of TGF- $\beta$  treatment (Figure 3c). Moreover, TGF- $\beta$ -mediated cytoskeleton re-modeling was also prevented and stress fibers were reduced in cells depleted of Ezh1/2 compared to control cells. Fibronectin and paxillin staining showed almost no focal adhesions after Ezh1/2 reduction (Figure 3c). Similar to NMuMG cells, a lentiviral-mediated stable knockdown of Ezh1/2 in Py2T cells resulted in failure to undergo EMT and also retained epithelial markers such as E-cadherin and showed a decrease in mesenchymal markers such as N-cadherin and fibronectin (Supplemental figure S3d-f). Together, these results strongly argue that Ezh1/2 is required for EMT.





**Figure 2:** Ezh1 and Ezh2 together contribute to the overall levels of H3K27me3 in NMuMG cells. (a-d) Realtime RT-PCR was carried out to test the knockdown efficiency following siRNA-mediated single knockdowns of Ezh1 (a) or Ezh2 (b) in NMuMG cells during TGF- $\beta$ -mediated EMT. Ezh1(c) and Ezh2 (d) transcript levels were further

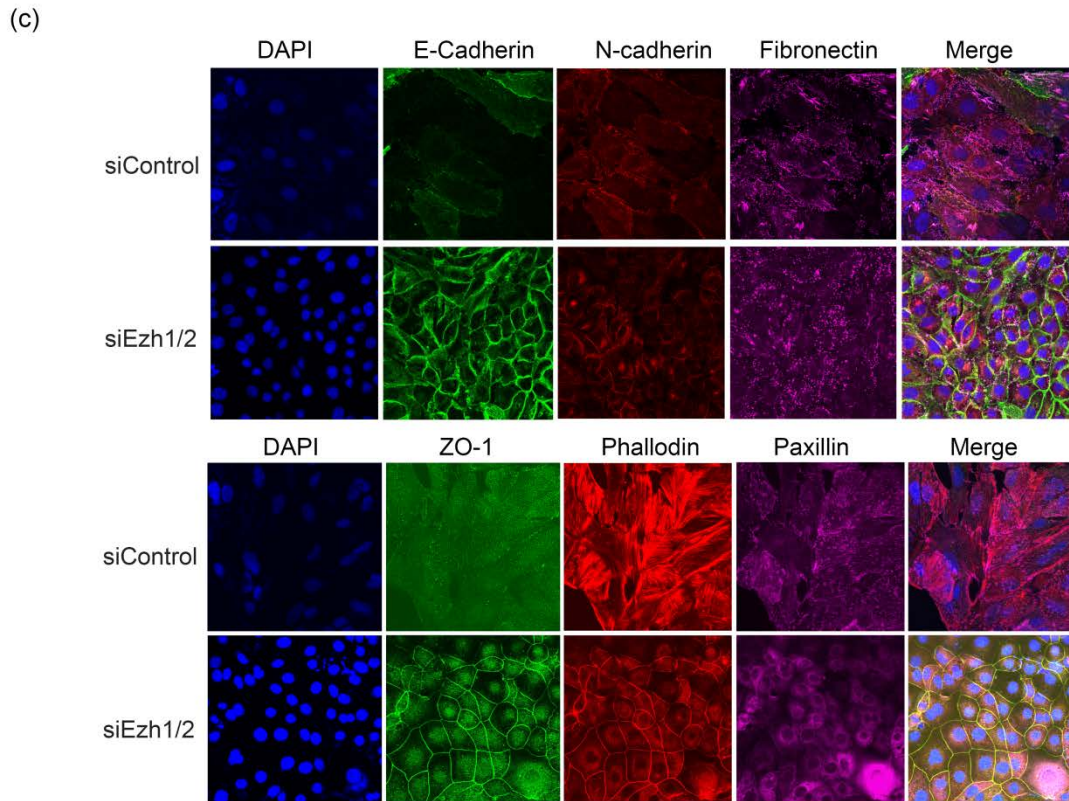
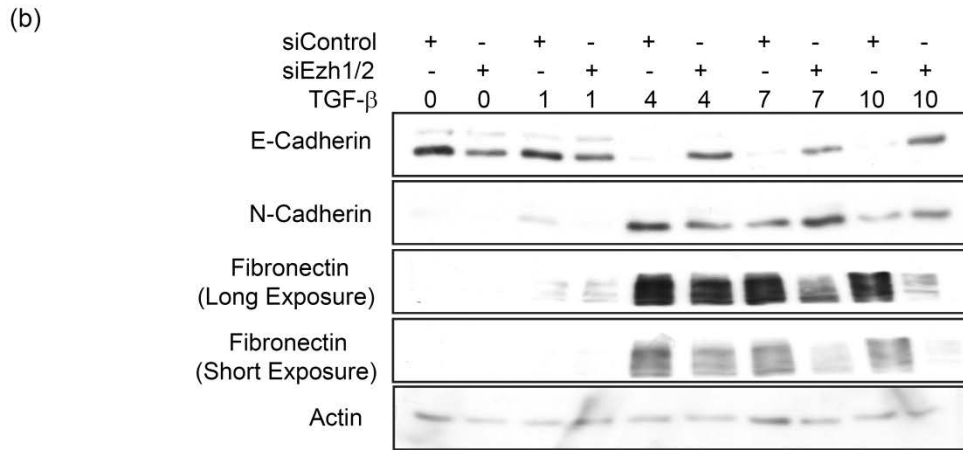
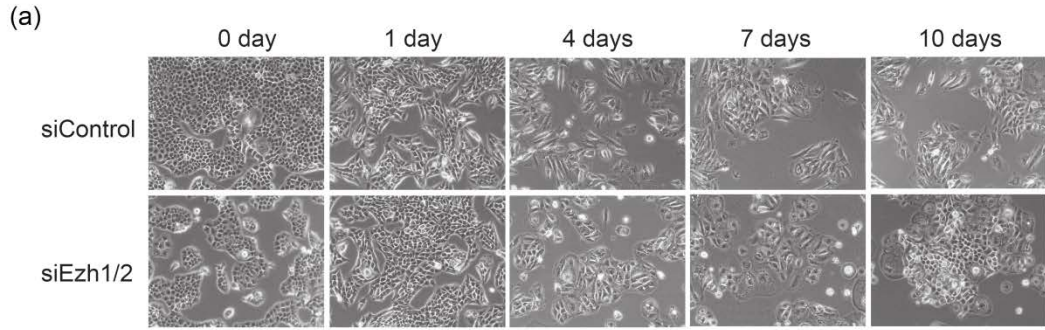
measured upon depletion of Ezh1 and Ezh2 together. Cells were treated with TGF- $\beta$  for 0, 1, 4, 7 and 10 days. (e-f) Total histones were isolated from Ezh1 and Ezh2 depleted or Ezh1/Ezh2 co-depleted NMuMG (e) and Py2T (f) cells. Total H3K27me3 levels were detected using specific antibodies by immunoblotting. Depletion of both Ezh1 and Ezh2 leads to a highly efficient reduction in H3K27me3. Total H3 was used as loading control. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)

### **3.3.3.5 Ezh1 and Ezh2 ablation together prevent cell migration but does not provides a survival advantage to the cells during EMT**

Cell migration and invasion are critical parameters in the metastatic dissemination of cancer cells and the formation of metastasis, the major cause of death in cancer patients (Brabletz et al., 2005; Christofori, 2006; Grunert et al., 2003; Huber et al., 2005; Thiery and Sleeman, 2006; Yilmaz and Christofori, 2010). Since migration has an important role in forming metastasis via EMT, we assessed the migratory capacity of Ezh1/2 knockdown cells. Transwell migration assays revealed a significantly lower chemo-tactic migration of Ezh1/2 knockdown cells compared to control cells in the absence of TGF- $\beta$  (Figure 4a). These findings were further confirmed in Py2T cells (Supplemental Figure S3a).

Ezh1 was shown to be expressed in non-proliferative cells while Ezh2 is known to be expressed in proliferative cells. Ezh2 depletion was shown to result in slower proliferation of prostate cancer (Varambally et al., 2002). Thus, we next investigated whether Ezh1/2 depletion can affect the proliferation and survival during TGF- $\beta$ -induced EMT in NMuMG cells. Surprisingly, loss of Ezh1/2 led to an increase in proliferation of NMuMG cells upon TGF- $\beta$  treatment (Figure 4b). No obvious changes were detectable upon Ezh1/2 knockdown in the absence of TGF- $\beta$ . To determine whether this effect was due to alterations in proliferation or changes in apoptosis rates, we analyzed cell-cycle as well as the rates of apoptosis. Compared to control siRNA-treated cells, no difference was observed in cell-cycle profile upon Ezh1/2 depletion in the absence as well as in the presence of TGF- $\beta$  for 2 and 4 days (Figure 4c). On the other hand, Annexin V staining showed a clear reduction in the fraction of apoptotic cells, arguing for a pro-survival

Polycomb-dependent mechanisms regulate Epithelial to Mesenchymal transition

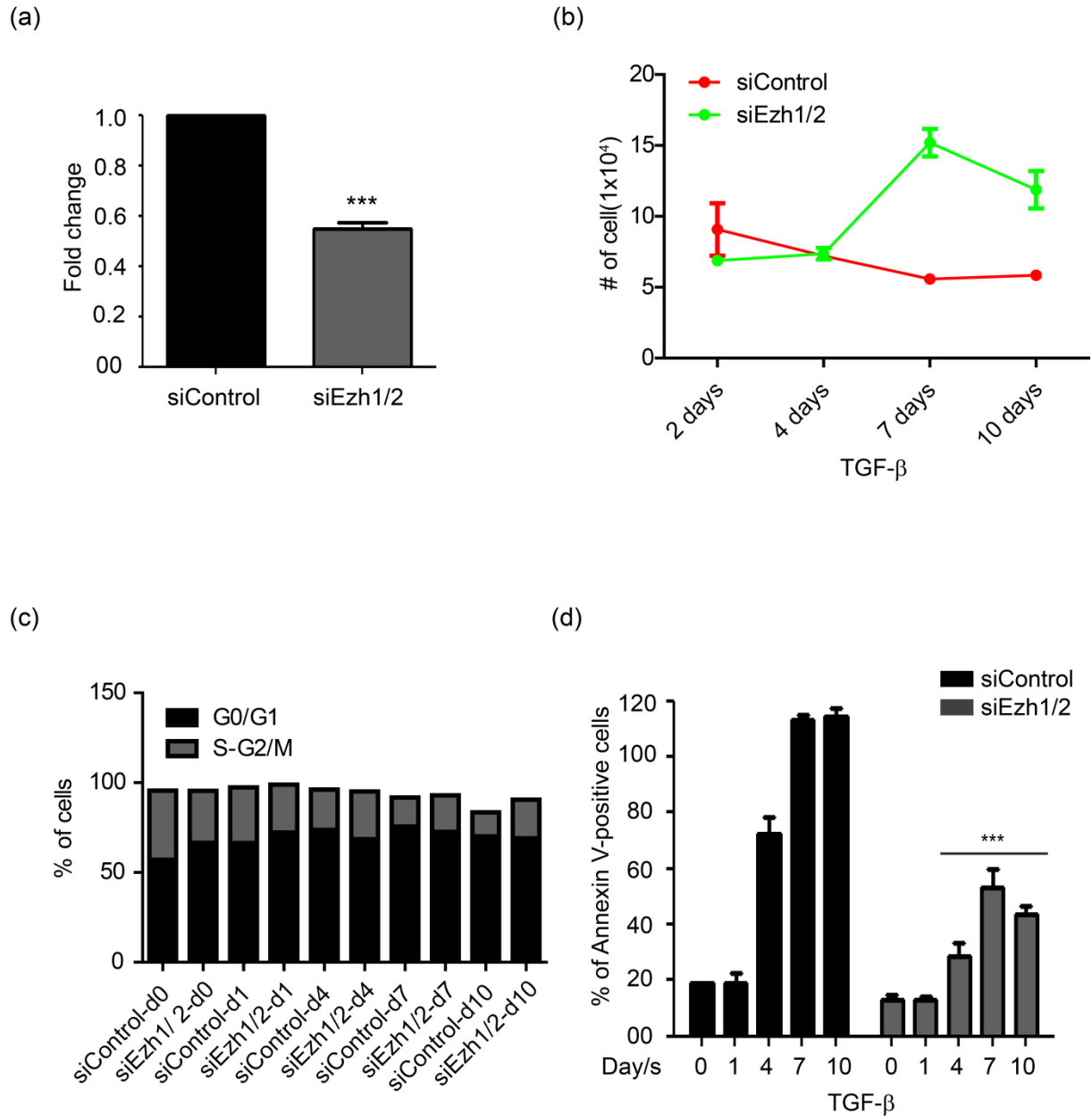


**Figure 3:** Ezh1/2 depletion prevents epithelial differentiation. (a) Cells were treated with TGF- $\beta$  for 0, 1, 4, 7 and 10 days and evaluated by phase contrast microscopy. Cells retained epithelial morphology upon Ezh1/2 depletion, while control cells progressed to EMT during TGF- $\beta$  treatment. Original magnification was 10x. (b) Immunoblot analysis for the epithelial marker E-cadherin and the mesenchymal markers N-cadherin and fibronectin during TGF- $\beta$ -induced EMT in NMuMG cells after depletion of Ezh1/2. Actin was used as a loading control. Cells were treated with TGF- $\beta$  for the indicated time-periods. (c) Immunofluorescence studies showing localization and expression levels of the adherent junction proteins E-cadherin and N-cadherin and the tight junction protein ZO-1 after 7 days of TGF- $\beta$  treatment in NMuMG cells depleted of Ezh1/2. Focal adhesions were visualized by fibronectin and paxillin staining and cytoskeleton-remodeling by phalloidin staining. Original magnification was 40x.

effect due to attenuation of TGF- $\beta$ -induced apoptosis (Figure 4d). We further extended our studies in Py2T cells and found that unlike NMuMG cells, these cells do not undergo apoptosis during TGF- $\beta$  treatment suggesting that apoptotic phenotype is biased toward untransformed cells. However, we did observe a significant increase in proliferation after Ezh1/2 knockdown in Py2T cells (Supplemental Figure S4b). Furthermore, we could not notice any significant differences in the cell cycle profiles upon Ezh1/2 depletion during the TGF- $\beta$  time-course (Supplemental Figure S4c).

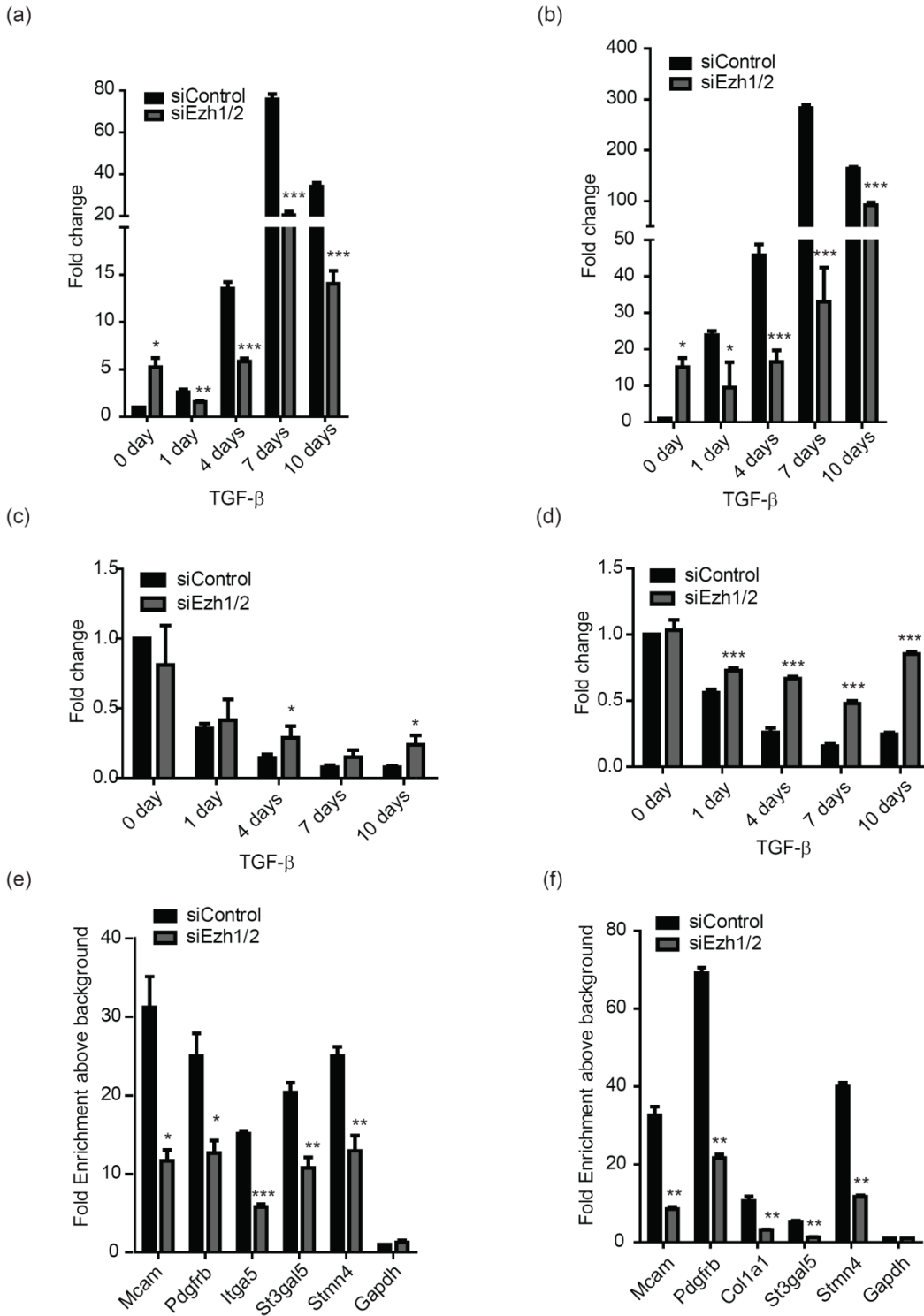
### 3.3.3.6 Dual role of Ezh1/2 during Polycomb-mediated regulation of EMT

Our ChIP-seq analysis suggested a Polycomb-mediated regulation of many key EMT genes. We next attempted to assay whether the transcription of these genes was affected upon Ezh1/2 depletion. The promoters of *Mcam* and *Pdgfr $\beta$*  are highly enriched with H3K27me3 mark in untreated NMuMG cells while *Cdx2* and *Ocln* gain this mark during TGF- $\beta$ -induced EMT. Ezh1/2 depletion led to the loss of H3K27me3 at the promoter of *Mcam* and *Pdgfrb* (Fig 5a and 5b) with a concomitant increase in their expression (Figure 5a-b). Along with this function, once the mark is removed from the genes which are already methylated, Ezh1/2 starts blocking the EMT process and leads to reduction in their expression levels. A similar loss of H3K27me3 mark was observed at the promoters of *Cdx2* and *Ocln* (Fig 5e) that resulted in their de-repression upon depletion of Ezh1/2 in TGF- $\beta$ -treated cells. (Figure 5c-d). In addition, Ezh1/2 reduction



**Figure 4:** Ezh1/2 depletion prevents TGF- $\beta$ -induced apoptosis and migration. (a) Ezh1/2 depleted NMuMG cells were subjected to a Boyden chamber migration assay for 20 hours. 20% FBS was used as a chemoattractant. (b) Proliferation assays were performed after siRNA-mediated knockdown of Ezh1/2 during TGF- $\beta$  time-course. Cells were counted using a Neubauer counting chamber. (c) Cell cycle analysis was done after transient knockdown of Ezh1/2 during EMT for 0, 1 and 4 days. Propidium Iodide (PI) was used for the staining. (d) Annexin V staining was performed to quantify cell death in Ezh1/2 knockdown NMuMG cells during TGF- $\beta$ -induced EMT for 0, 1, 4, 7 and 10 days. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.001$  indicated with (\*\*\*).

Polycomb-dependent mechanisms regulate Epithelial to Mesenchymal transition



**Figure 5:** Depletion of Ezh1/2 results in transcriptional deregulation of H3K27me3-target genes (a-d) Realtime RT-PCR was performed to assess the expression levels of Mcam (a) and Pdgfrb (b), which are already enriched with

H3K27me3 in untreated NMuMG cells and Cdx2 (c) and Ocln (d), which gain this mark upon TGF- $\beta$  treatment for 1, 4, 7 and 10 days. (e-f) Following chromatin immunoprecipitation using H3K27me3-specific antibody, in NMuMG (e) and Py2T (f) cells after Ezh1/2 depletion, qPCRs was performed using promoter-specific primers. Depletion of Ezh1/2 leads to a significant reduction in H3K27me3 enrichment at target gene promoters. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)).

removed H3K27me3 marks from the targeted promoters in NMuMG and Py2T cells and activated the genes (Fig. 5e and 5f).

### 3.3.3.7 MeDIP analysis shows no significant difference in methylation pattern

Since, it has been shown before that premalignant cells can acquire *de novo* DNA methylation at biologically relevant sites early in the carcinogenic process in a deterministic manner (Dumont et al., 2008; Ruike et al., 2010), we investigated whether similar modes also operated in our model system of TGF- $\beta$ -induced EMT in NMuMG cells.

We profiled DNA methylation using the MeDIP technique (Weber et al., 2005) in NMuMG cells untreated as well as treated with TGF- $\beta$  for 10 and 20 days followed by detection using custom tiling arrays that cover 10% of the mouse genome including all well annotated promoters, several large multi-gene loci and the complete chromosome 19 (Nimblegen HD 2.1 arrays). Comparison of promoter methylation levels revealed no promoters with significant differences at day 10 and day 20 of TGF- $\beta$  treated cells compared to control cells (Supplemental Figure S5a-c). However, given previous evidences of silencing of Cdh1 by DNA methylation, we performed bisulfite sequencing of DNA derived from control cells as well as those treated with TGF- $\beta$  for 10 and 20 days. Analysis of the Cdh1 promoter also indicated only a very marginal increase in promoter CpG methylation in day 20 samples (data not shown). Overall, these data suggest that TGF- $\beta$ -induced EMT does not involve any reprogramming of promoter DNA methylation.

### 3.3.4 Discussion

Epithelial-to-mesenchymal transition (EMT) underlies crucial steps in many developmental and disease events including cancer. Both genetic and epigenetic pathways have been implicated in EMT gene regulation. Polycomb Group of proteins have been implicated in cancers of various subtypes (Mills, 2010; Sauvageau and Sauvageau, 2010; Sparmann and van Lohuizen, 2006b), but no comprehensive studies have been performed in models of EMT to reveal their role in transcriptional regulation of EMT genes. In the current study, by performing ChIP-seq analysis for H3K27me3 for several stages of EMT progression, we reveal targets of epigenetic reprogramming during this process. These genes include a number of key EMT genes normally repressed in epithelial cells, such as *Mcam*, *Pdgrfb* and *Itga5*. These cells lose this mark and get activated upon treatment with TGF- $\beta$  that follows EMT. On the other hand, another set of genes that define epithelial identity, such as *Cdh1*, *Ocln* and *Cdx2*, gain this mark and become repressed during TGF- $\beta$ -induced EMT. We functionally show that both *Ezh1* and *Ezh2* contribute to the H3K27me3 levels in these cells as upon knockdown of these enzymes, both global as well as promoter-specific H3K27me3 levels drop significantly, which further accompanies transcriptional activation of associated genes. These changes further result in the blockage of EMT. Our data argue for a model where, in normal epithelial cells, the Polycomb machinery is involved in repressing transcription of genes, the activation of which would otherwise result in loss of epithelial identity. Upon induction of EMT with TGF- $\beta$  leading to EMT, genomewide remodeling in H3K27me3 levels occurs and the function of Polycomb switches to repressing epithelial cell-type-specific genes and mesenchymal genes get de-repressed. These data provide convincing evidence that Polycomb-mediated transcriptional regulatory mechanisms underlie phenotypic changes during EMT.

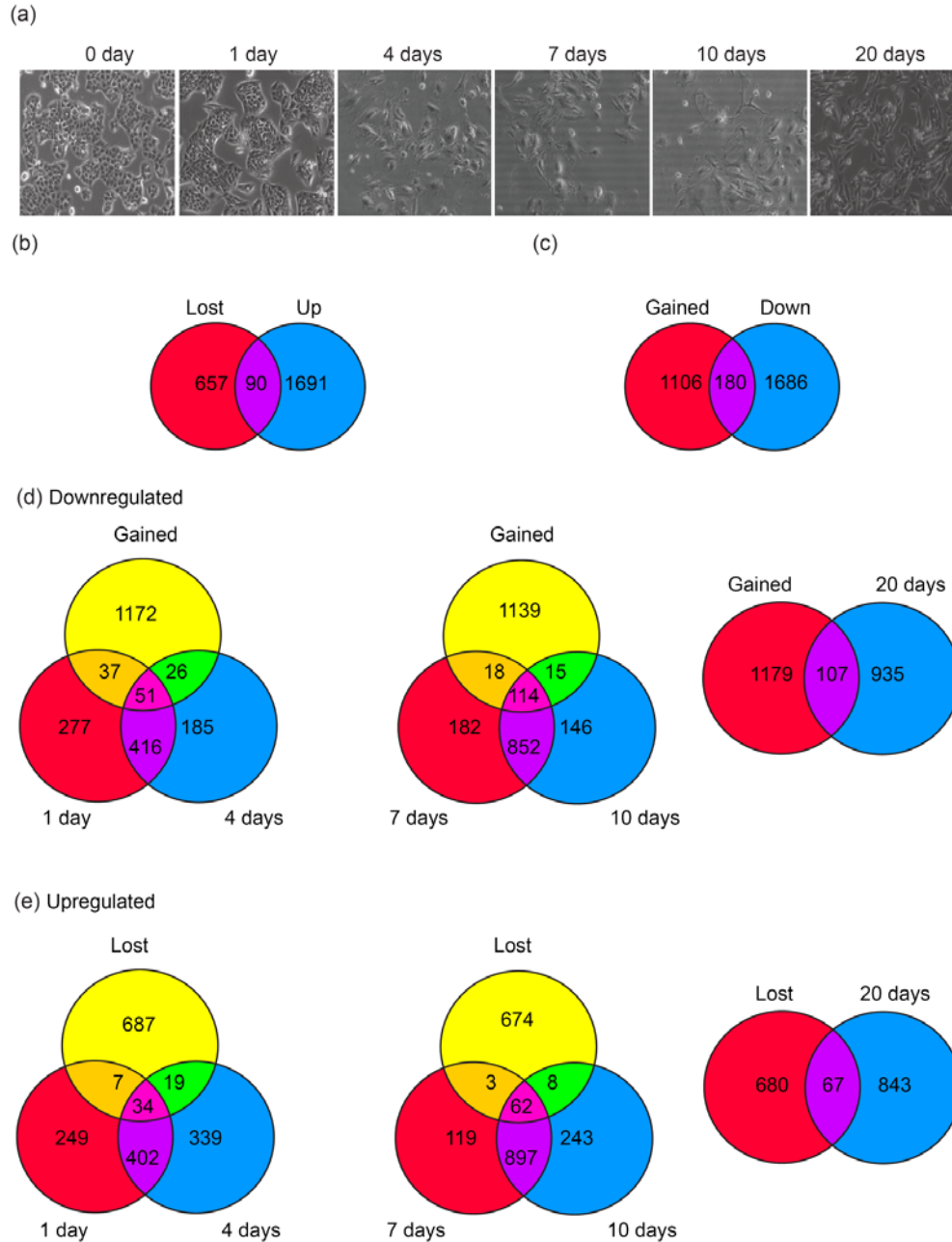
*Ezh2* was shown to directly repress the transcription of *KLF2* in cancer cells (Taniguchi et al., 2011). Expression of a non-coding RNA, *HOTAIR*, was implicated in reprogramming of PRC2 function in metastatic progression including in breast cancer (Kogo et al., 2011). Another study showed that *Ezh2* may further regulate cancer cell growth and invasiveness by regulating miRNAs (Cao et al., 2011). An important marker of epithelial cell identity, *Cdh1*, was shown to be silenced by the Polycomb machinery by a Snail-dependent recruitment of *Ezh2* to the *Cdh1* promoter during EMT (Cao et al., 2008). In line with these observations, our data identified *Cdh1* among the targets that gain H3K27me3 mark during TGF- $\beta$ -induced EMT. *Cdh1* was also



shown to be silenced by DNA methylation in a variety of human cancers (Esteller, 2007; Esteller, 2008b; Graff et al., 1995; Yoshiura et al., 1995). Genomewide analysis of DNA methylation did not reveal any promoters that exhibit reprogramming of DNA methylation during EMT. However, bisulfite sequencing discovered a mild, but detectable, increase in promoter CpG methylation at *Cdh1* promoters at day 20 of TGF- $\beta$ -induced EMT.

These observations demarcate the dynamics of the Polycomb-associated mark and its role in an experimental model of EMT. Such reprogramming seems crucial for proper EMT progression, as knockdown of this mark had substantial effects on cell migration and proliferation. It has been recently shown that EMT accompanies a global reduction in the heterochromatin mark H3 Lys9 dimethylation (H3K9Me<sub>2</sub>), an increase in the euchromatin mark H3 Lys4 trimethylation (H3K4Me<sub>3</sub>) and an increase in the transcriptional mark H3 Lys36 trimethylation (H3K36Me<sub>3</sub>) (McDonald et al., 2011b). Thus it is very likely that other chromatin modifications not examined so far are also reprogrammed during EMT. However, given a number of previous studies that linked Polycomb-mediated mechanisms with carcinogenesis (Mills, 2010; Sauvageau and Sauvageau, 2010; Sparmann and van Lohuizen, 2006b), our study is of utmost relevance to the field where we outline genomewide targets of the Polycomb machinery for transcriptional modulation during several stages of EMT. Our findings may reflect a general mechanism for cell-fate transitions including other cases of EMT such as development. It will be further interesting to explore epigenetic events in the reverse process, mesenchymal-to-epithelial transition (MET), such as those involved in normal development or disease as well as in vitro induced pluripotent stem cells (iPSCs) generation. It is important to note that the epigenetic reprogramming of the Polycomb-associated H3K27me<sub>3</sub> mark during TGF- $\beta$ -induced EMT very closely resembles stem cell differentiation (Mohn et al 2008). However, in contrast to stem cell differentiation, DNA methylation patterns remain unchanged during EMT. Taken together, our data reveal strong evidence that Polycomb machinery-mediated epigenetic mechanisms underlie transcriptional changes that are crucial for epithelial to mesenchymal transition. It will be further exciting to investigate whether similar epigenetic reprogramming occurs in other cases of EMT and in experimental systems where in cell-fate transitions are observed.

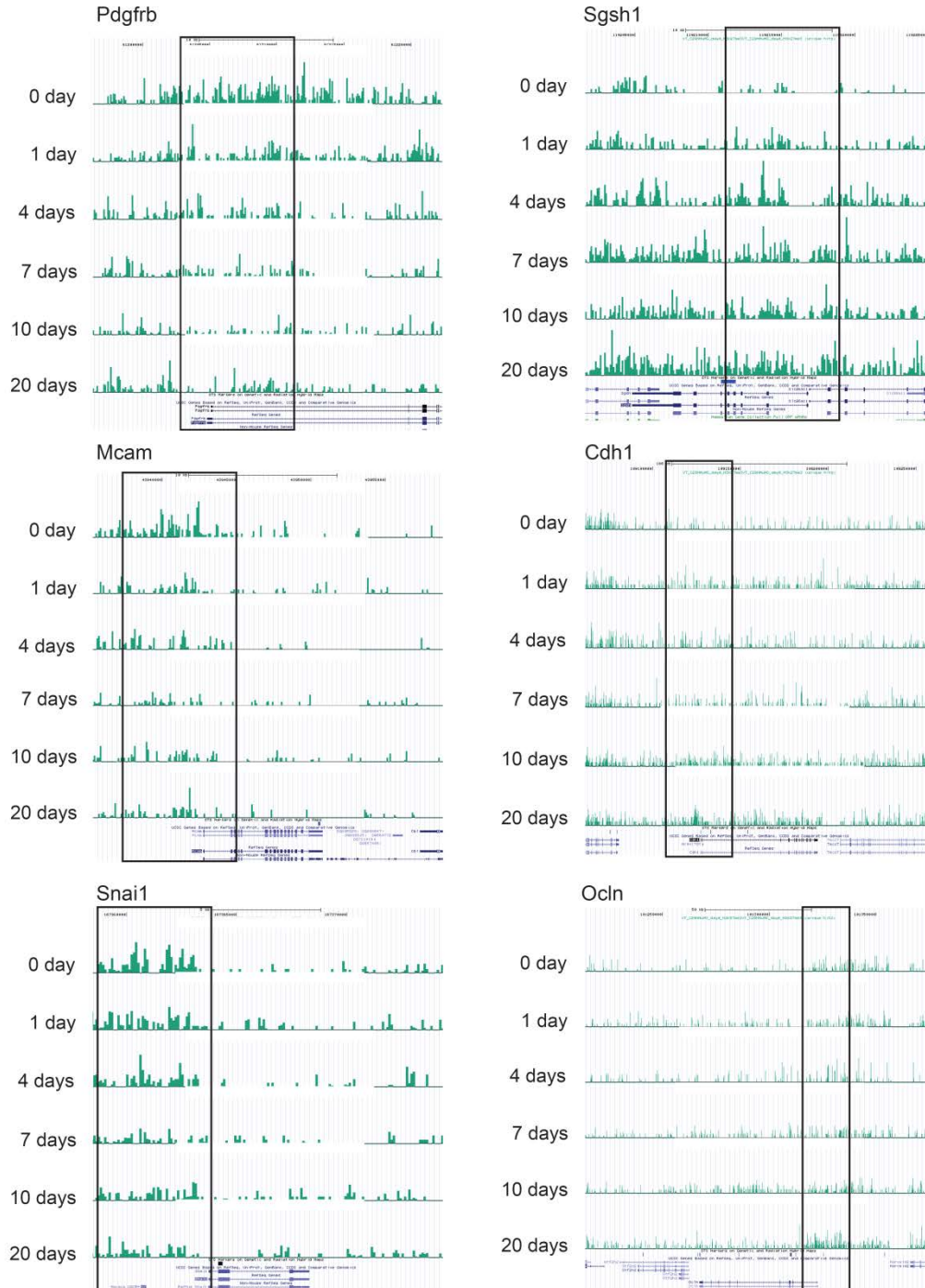
### **3.3.5 Supplemental data**



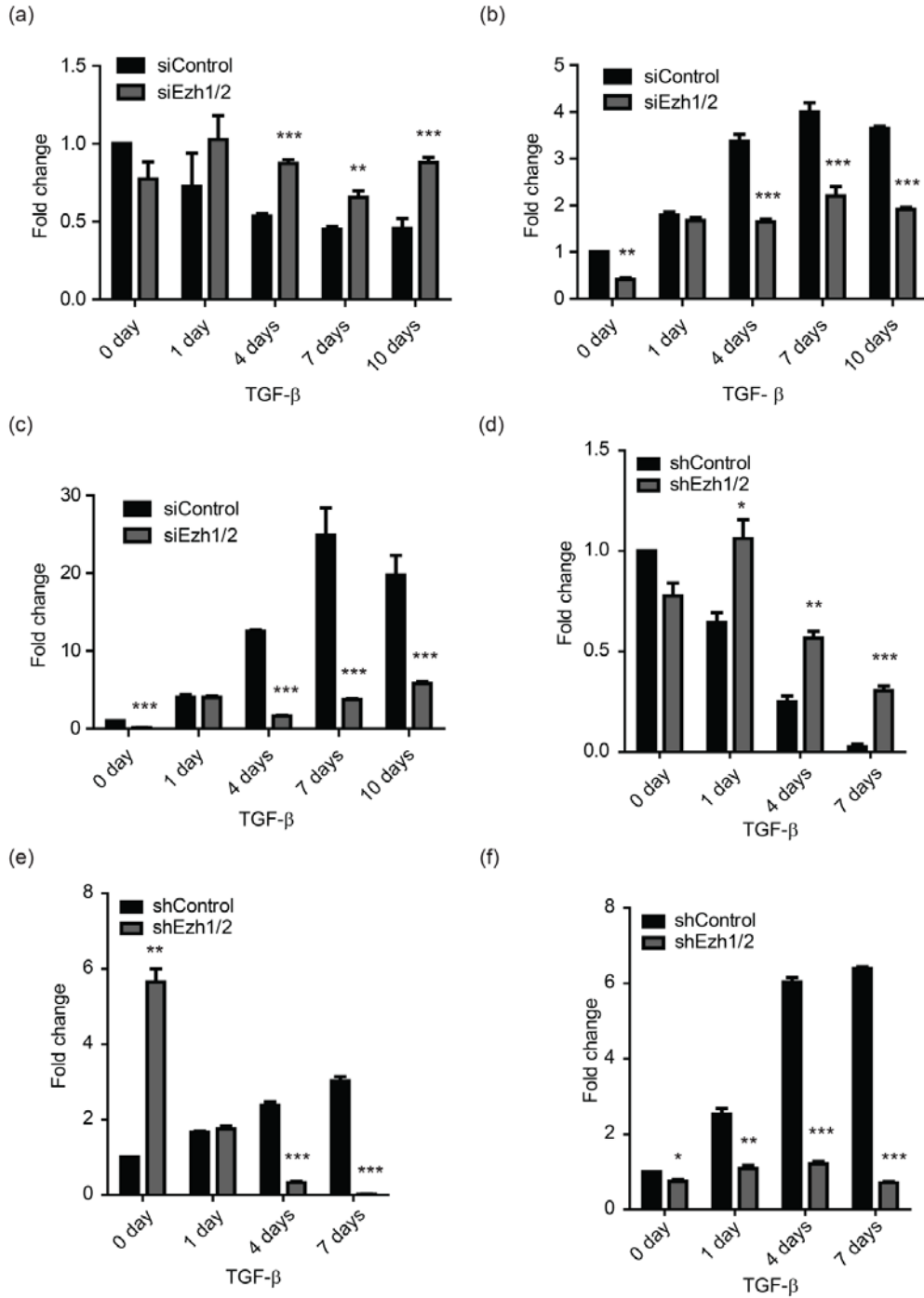
**Figure S1:** Dynamics of PcG mark H3K27me3 during EMT in NMuMG cells. (a) Morphological changes are observed after TGF- $\beta$  treatment of NMuMG cells for 0, 1, 4, 7, 10 and 20 days by using a phase-contrast microscopy, reflecting epithelial to mesenchymal transition. Original magnification was 10x. (b-c) Venn-Diagram showing the overlap between genes that lose the H3K27me3 mark during TGF- $\beta$ -induced EMT and those that are transcriptionally induced upon TGF- $\beta$  treatment (b) and vice-versa (c). (d-e) Systemic comparison of genes gaining

Polycomb-dependent mechanisms regulate Epithelial to  
Mesenchymal transition

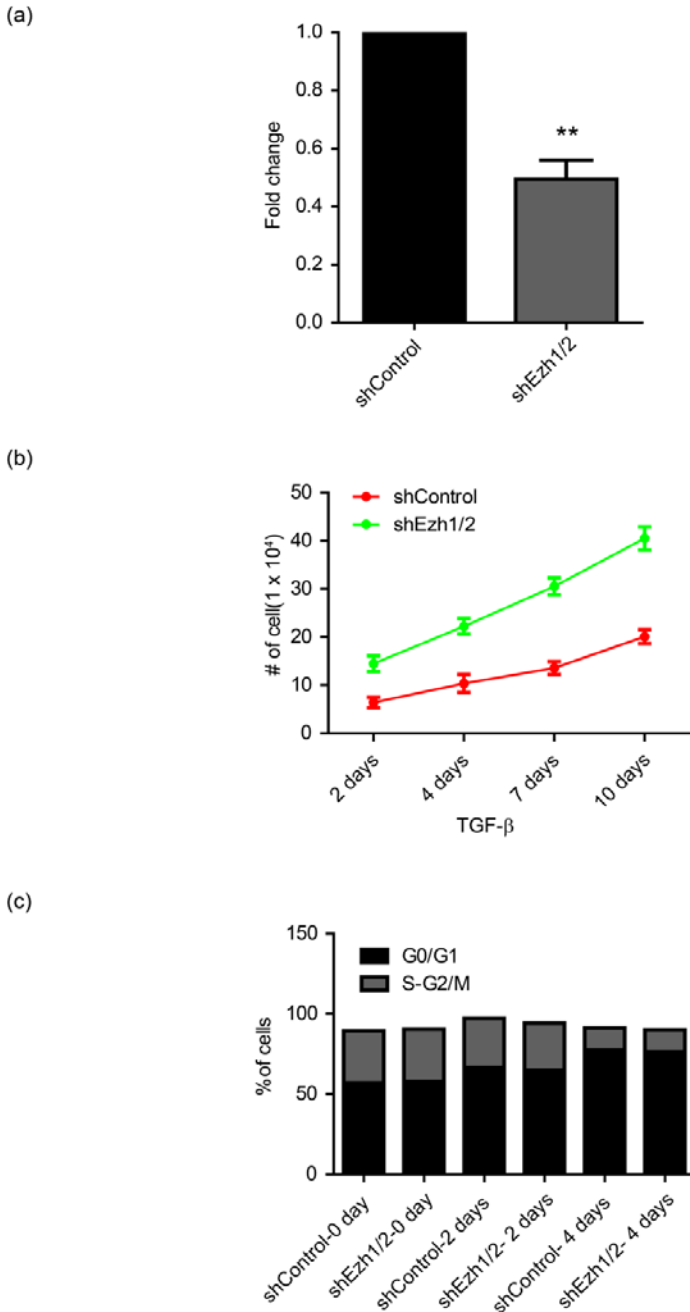
(c) or losing (d) H3K27me3 mark during EMT with genes that are transcriptionally downregulated or upregulated during TGF- $\beta$ -induced EMT.



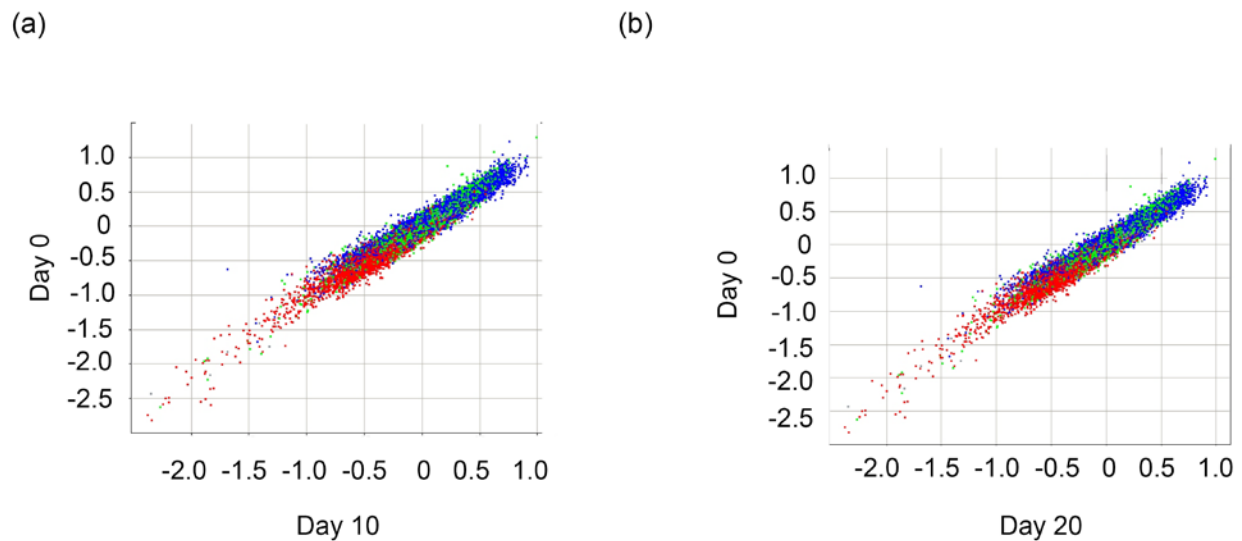
**Figure S2:** H3K27me3 is enriched at the promoters of key EMT genes. Genome browser view of H3K27me3 ChIP enrichment at the *Mcam*, *Pdgfrb*, *Snai1*, *Sgsh1*, *Cdh1* and *Ocln* gene loci during TGF- $\beta$ -induced EMT in NMuMG cells. Promoter regions are marked by the rectangular boxes.



**Figure S3:** Ezh1/2 depletion prevents EMT. (a-f) Realtime RT-PCR was carried out to quantify the expression levels of EMT markers E-cadherin (a and d), N-cadherin (b and e) and fibronectin (c and f) in NMuMG (a, b and c) and Py2T cells (d, e and f) during TGF- $\beta$  mediated EMT. NMuMG cells were treated with TGF- $\beta$  for 0, 1, 4, 7 and 10 days while Py2T cells were treated for 0, 1, 4 and 7 days. Py2T cells were stably transfected with Ezh1/2 shRNA. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)



**Figure S4:** Ezh1/2 are required for cell migration but not for survival. (a) Boyden-chamber migration assays were performed to assess the migratory capacity of stably transfected Ezh1/2 depleted Py2T cells treated with TGF- $\beta$  for 15 days. The assays were carried out for 20 hours and 20% FBS was used as a chemo-attractant. (b) Proliferation assays were done to assess the proliferation rate after Ezh1/2 ablation in Py2T cells. Neubauer counting chambers were used for the counting. (c) Cell cycle analysis was done in the absence as well as in the presence of TGF- $\beta$  (2 and 4 days) after Ezh1/2 reduction in Py2T cells. PI staining was used to distinguish the cell cycle phases. Statistical values were calculated by using an unpaired, two-tailed t-test. p-value  $\leq 0.05$  indicated with (\*), p-value  $\leq 0.01$  indicated with (\*\*), p-value  $\leq 0.001$  indicated with (\*\*\*)



**Figure S5:** DNA methylation pattern remain unchanged during TGF- $\beta$ -mediated EMT in NMuMG cells. (a-b) Scatter plot comparing averaged DNA methylation values from replicate microarrays for all mouse promoters from nontreated NMuMG cells versus NMuMG cells treated with TGF- $\beta$  for 10 days (a) and 20 days (b).

#### 4.1.6 Methods and Materials

##### Reagents and antibodies

**Reagents:** TGF- $\beta$  (240-B, R&D systems). DMEM (D5671, Sigma-Aldrich), PBS (D8537, Sigma-Aldrich), trypsin (T4174, Sigma-Aldrich), Opti-MEM (11058, Gibco), FBS (F7524,

Sigma-Aldrich), Glutamine (G7513, Sigma-Aldrich), Pencillin/streptomycin (P4333, Sigma-Aldrich), Lipofectamine RNAiMax (11668-019, Invitrogen), Alexa Fluor-488, 568, 633 (Invitrogen), Polybrene (AL-118, Sigma-Aldrich), Puromycin (P7255, Sigma-Aldrich), Fugene HD (12998300, Roche), Trizol (T9424, Sigma-Aldrich), M-MLV reverse transcriptase (M314C 28692233, Promega), SYBR-green PCR MasterMix (Eurogentec) and Bradford reagent (500-0006, Biorad), Protease inhibitor cocktail (P2714, Sigma-Aldrich). **Antibodies: Western Blot:** E-Cadherin (610182, Transduction Laboratories), N-Cadherin (M142, Takara), ZO-1 (617300, Zymed), Fibronectin (F-3648, Sigma-Aldrich), Actin (SC-1616, Sanata Cruz Biotechnology). **Immunofluorescence:** E-Cadherin (13-1900, Zymed), N-Cadherin (610921, Transduction Laboratories), ZO-1 (617300, Zymed), Phalloidin (A12380, Invitrogen) and Paxillin (13520, Transduction Laboratories). **Chromatin Immunoprecipitation:** H3K27me3 (abcam) **Apoptosis and Cell cycle:** Annexin-V (559934, BD Biosciences) and PI (P4170, Sigma-Aldrich). **Small interfering RNAs:** siControl (Stealth RNAi™ siRNA Negative Controls, 12935-100, Invitrogen), siEzh1 (SASI\_Mm01\_00114972 and SASI\_Mm01\_00114974, Sigma-Aldrich) and siEzh2 (SASI\_Mm01\_00061987 & SASI\_Mm01\_00061988, Sigma-Aldrich). **Small hairpin RNA:** shControl (Mission Non-target shRNA control vector, SHC002) and shEzh1/2 (SHCLNG-NM\_007971\_Mouse, TRCN0000039041).

### **Cell lines and cell culture**

A subclone of NMuMG cells (NMuMG/E9; hereafter NMuMG) expressing E-cadherin has been previously described (Jechlinger et al., 2002). MCF7 shControl and MCF7-shEcad have been described before (Lehembre et al., 2008). Py2T cells were derived from Polyoma middle T breast cancer tumor model (unpublished data, Waldmeier et.al). NMuMG, MCF7shControl, MCF7-shEcad, Py2T, 293T and PLAT-E cells were cultured in DMEM supplemented with 10% FBS, 2mM glutamine, 100U penicillin and 0.2mg/ml streptomycin. All the cells were cultured at 37°C with 5% CO<sub>2</sub> in humid incubator. For TGF- $\beta$  time-course experiments, cells were treated with 2ng/ml TGF- $\beta$  for indicated time point and it was replaced every 2 days. For siRNA transfections, Lipofectamine RNAiMax was used according to the manufacturer's instructions.

### **Microarray processing and data analysis**

RNA was isolated from NMuMG cells transfected with NMuMG cells treated with TGF- $\beta$  for 0, 1, 4, 7 and 10 days by RNAeasy Mini kit (Sigma-Aldrich). Non-treated cells are used as a control. RNA quality and quantity was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies). The manufacturer's protocols for the GeneChip platform by Affimetrix were followed. Methods included synthesis of the first- and second-strand cDNA followed by synthesis of cRNA by *in vitro* transcription, subsequent synthesis of single-stranded cDNA, biotin labeling and fragmentation of cDNA and hybridization with the microarray slide (GeneChip® Mouse Gene 430 2.0 array), post-hybridization washings and detection of the hybridized cDNAs using a streptavidin-coupled fluorescent dye. Hybridized Affimetrix GeneChips were scanned using an Affimetrix GeneChip 3000 scanner. Image generation and feature extraction were performed using Affimetrix GCOS Software and quality control was performed using Affimetrix Expression Console Software. Raw microarray data were normalized with Robust Multi-Array (RMA) and analyzed using Partek® Genomics Suite Software (Partek Inc.). One-way analysis of variance (ANOVA) and asymptotic analysis were used to identify significantly differentially expressed genes. The gene ontology (GO) tools from Metacore Software as well as the David gene ontology software were used for further analysis

### **Chromatin Immunoprecipitation**

ChIP experiments were performed as previously described (Weber et al, 2007). In brief, crosslinked chromatin was sonicated to achieve an average fragment size of 500 bp. Starting with 100  $\mu$ g of chromatin and 5  $\mu$ g of anti-H3K27me3 antibody, 1  $\mu$ l of ChIP material and 1  $\mu$ l of input material were used for quantitative real-time PCR using specific primers covering the 1000 bp promoter region from the transcription start site. Primers covering an intergenic region are used as a control. The efficiencies of PCR amplification were normalized for between the primer pairs. Following primers were used for ChIP –PCR. Relative positions from the transcription start site are also mentioned.



Name	Sequences
Mcam FP	ggtccccgctagtagtgacaaa
Mcam RP	ggttgaaggagcaatgacagggtg
Pdgfrb FP	gaaaacagacacacgcgtccac
Pdgfrb RP	caccacacactttgggggaaag
Itga5 FP	cccagaggtgattccttctca
Itga5 RP	cctccccctccttccagatgta
Col1a1 FP	tggactccttcccttcttcc
Col1a1 RP	atcttgatggagagctgggagga
St3gal5 FP	ccacctactctcggctggagt
St3gal5 RP	cgtcacgaggataagggagacca
Stmn4 FP	tcatcttactcccagccttc
Stmn4 RP	gcttggcaattggacagtctct
Gapdh FP	ctctgctcctcctgttcc
Gapdh RP	tcctagaccgtagcagtg

### Next Generation Sequencing and Analysis

The ChIP libraries were prepared with the Illumina ChIP-Seq DNA Sample Prep Kit (Cat# IP-102-1001) according to Illumina's instructions and sequenced on the Genome Analyzer 2 following the manufacturer's protocols.

#### Genomic coordinates

The July 2007 *M. musculus* genome assembly (NCBI37/mm9) provided by NCBI (<http://www.ncbi.nlm.nih.gov/genome/guide/mouse/>) and the Mouse Genome Sequencing Consortium ([http://www.sanger.ac.uk/Projects/M\\_musculus/](http://www.sanger.ac.uk/Projects/M_musculus/)) was used as a basis for all analyses. Annotation of known RefSeq transcripts was obtained from UCSC (<http://hgdownload.cse.ucsc.edu/goldenPath/mm9/database/refGene.txt.gz> from Oct 18, 2009).

#### Read filtering, alignment and weighting

Low-complexity reads were filtered out based on their dinucleotide entropy (removing <1% of the reads). Alignments to the mouse genome were performed by the software bowtie (version

0.9.9.1) (Niessen et al., 2008) with parameters -v 2 -a -m 100, tracking up to 100 best alignment positions per query and allowing at most two mismatches. To track genomically untemplated hits (e.g., exon-exon junctions or missing parts in the current assembly), the reads were also mapped to an annotation database containing known mouse sequences (miRNA from <ftp://ftp.sanger.ac.uk/pub/mirbase/sequences/13.0>, rRNA,snRNA, snoRNA and RefSeq mRNA from GenBank <http://www.ncbi.nlm.nih.gov/sites/entrez>, downloaded on July 16, 2009, tRNA from <http://lowelab.ucsc.edu/GtRNAdb/> and piRNA from NCBI (accessions DQ539889 to DQ569912). In that case, all best hits with at most two mismatches were tracked. Each alignment was weighted by the inverse of the number of hits. In the cases where a read had more hits to an individual sequence from the annotation database than to the whole genome, the former number of hits was selected to ensure that the total weight of a read did not exceed one. All quantifications were based on weighted alignments.

### **Peak finding**

Genomic regions of increased ChIP-seq read alignment densities were identified using macs (version 1.3.7.1) (Shi and Massague, 2003), using a pool of read alignments from all biological replicates and cellular stages (weights rounded to integers) as input, parameters --mfold=8 --gsize=2700000000 --tsize=36 and default values for all other parameters. IP enrichments (see below) of resulting peak candidates were calculated and peak candidates with enrichments lower than 2-fold above background (combining biological replicates) were removed.

### **Calculation of peak enrichments in genomic regions**

Enrichment of peaks in genomic regions (see above for definition) were calculated as the ratio of observed over expected number of peaks in a region, where the observed number is the count of all peaks overlapping a region by more than half of their length, and the expected number is the fraction of genomic bases in that region type, multiplied with the total number of peaks.

### **Calculation of IP enrichments**

IP enrichments of a genomic region (TSS windows or peak region) were calculated as  $e = \log_2\left(\frac{n_{fg}/N_{fg} * \min(N_{fg}, N_{bg}) + p}{n_{bg}/N_{bg} * \min(N_{fg}, N_{bg}) + p}\right)$ , where  $n_{fg}$  and  $n_{bg}$  are the summed weights of overlapping foreground and background (input chromatin) read alignments, respectively.  $N_{fg}$  and  $N_{bg}$  are the total number of aligned reads in foreground and

background samples, and  $p$  is a pseudocount constant ( $p=16$ ) used to regularize enrichments based on low counts that would otherwise be dominated by sampling noise.

### MeDIP-chip data analysis

Nimblegen array intensity files were read and  $\log_2$  enrichments ( $\log_2$  bound/input ratios) for each individual probe were calculated using the R package Ringo. Probe-level as well as promoter level  $\log_2$  enrichments showed good reproducibility between the two replicates.

### Quantitative RT-PCR

Total RNA was prepared by using a Tri Reagent according to the manufacturer's instructions. RNA was further reverse transcribed with ImProm-II Reverse Transcriptase, and transcripts were quantified by PCR using SYBR-green PCR Mastermix in a real time PCR system (Step One Plus, Applied Biosystems). Human or mouse ribosomal L19 primers were used for normalization. PCR assays were performed in duplicates, and fold induction was calculated against control-treated cell lines using the comparative Ct method ( $\Delta\Delta$  Ct). Following primers were used:

Primer name	Sequences
mRpl19 Forward primer	ctcgttgccggaaaaaca
mRpl19 Reverse primer	tcatccaggtcaccttctca
mEzh2 Forward primer	caggctggggcatctttatc
mEzh2 Reverse primer	acgaatthttgtgcctttc
mE-cadherin Forward primer	cgaccctgcctctgaatcc
mE-cadherin Reverse primer	tacacgctgggaaacatgagc
mN-Cadherin Forward primer	caatgacgtccaccctgttct
mN-Cadherin Reverse primer	ctgccatgactttctacggaga
mFibronectin1 Forward primer	cccagacttatggtggcaatt
mFibronectin1 Reverse primer	atattccgactcgagtctga
mMcam Forward primer	actggtgtgcgtcttctgttcg
mMcam Reverse primer	gcttttctctctctggcacacc

mPdgrfb Forward primer	acctgcagagacctcaaaaggtg
mPdgrfb Reverse primer	ctgatcttctcccagaaagtcaca
mCdx2 Forward primer	catcaccatcaggaggaaaagtga

### **Immunoblot**

Cells were lysed for 1 hour on ice in RIPA-Plus buffer (50mM Tris-HCl, pH8.0), 150mM NaCl, 10% glycerol, 1% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, 2mM CaCl<sub>2</sub>, 1mM dithiothreitol, 1mM sodium fluoride, 0.2mM sodium orthovanadate, 1x protease inhibitor cocktail and further quantified by using Bradford reagent. 50 µg of cleared protein lysates were separated by SDS-PAGE and electroblotted on PDVF membranes, and proteins were visualized with the appropriate primary and secondary antibodies and ECL on superRX films. Depending on the species origin of antibodies, immunoblots were either probed sequentially or on multiple membranes. Adobe Photoshop has been used to excise the relevant portion of the immunoblots from the original scans of X-ray films exposed to chemoluminescence visualization of specific proteins.

### **Immunofluorescence**

siControl and siEzh1/2 cells were plated on coverslips and treated with TGF-β for mentioned time. The cells were fixed with 4% paraformaldehyde in HBSS and further permeabilized with 0.2% Triton for 5 minutes at room temperature. These cells were blocked by using 3.5% goat serum for 15 minutes and incubated with primary antibodies against E-cadherin, N-cadherin, fibronectin, ZO-1, vimentin, paxillin and phalloidin for 1 hour and then incubated with fluoro-chrome-labeled secondary antibody for 1 hour at room temperature. The coverslips were counterstained with DAPI and imaged with a confocal laser-scanning microscope. Data were processed with Adobe Photoshop 7.0 software.

### **Production of lentivirus for in-vivo and in-vitro knockdown studies**

Murine Ezh1/2 shRNAs and control shRNA were purchased from Sigma-Aldrich as described above. For lenti-virus production, 293T cells were transfected with the shRNA expressing lenti-

viral vector in combination with the packaging vectors including envelope protein; HDM-pVSV/G, codon-optimized HIV gag-pol; HDM-Hp<sub>g</sub>m2, transactivator of transcription; HDM-Tat1b and pRC-CMV-RaII by Fugene HD. After 48 hours of transfection, viral supernatant was harvested, filtered (0.46 μm), supplemented with polybrene (8ng/ml) and used to infect target cells. Infections were performed once a day for two consecutive days. Infected cells were positively selected using Puromycin (5ug/ml).

### **Migration assay**

Cell migration was assessed by transwell migration assay (pore size: 8 μm; Falcon BD). 10<sup>4</sup> cells were seeded in 2% FBS/DMEM (Sigma) in the upper chamber and the lower chamber was filled with 20% FBS/DMEM. After 20 hours of incubation at 37°C, cells in the upper chamber were carefully removed with a cotton swap and the cells that had traversed the membrane were fixed in 4% paraformaldehyde/PBS, stained with DAPI. Pictures of the membrane were taken at a 10x magnification using a fluorescent microscope (Nikon Diaphot 300). Quantification was done using the software ImageJ.

### **Apoptosis assay (Annexin assay)**

Cells were washed twice with cold PBS and resuspended in 1X Annexin V binding buffer at a concentration of 1 x 10<sup>6</sup> cells/ml. 5 μl of Cy5 Annexin V was added to the 100 μl of cells (1 x 10<sup>5</sup>) and incubated for 15 min on ice in the dark. Stained cells were filtered with a 40um mesh and analyzed on a FACSCanto II using DIVA software.

### **Cell growth curve**

1x10<sup>4</sup> cells were seeded in each well of 24-well plate and cell numbers were assessed for mentioned days by using a Neubauer counting chamber.

### **PI Staining**

Cells were trypsinized and fixed in 70% ice-cold Ethanol for overnight. Washed twice with PBS and resuspended in sodium citrate buffer with 5μg/ml PI for overnight. Stained cells were analyzed by FACSCanto II using DIVA software.

### **Statistical analysis**

Statistical analysis and graphs were generated using the GraphPad Prism software (GraphPad Software Inc, San Diego, CA). All statistical analysis was done by unpaired, two-sided t-test. Normality testing was performed using the Kolmogorov-Smirnov test with Dallal-Wilkinson-Lillie for p-values. All the data are shown as mean  $\pm$  SD and are representative of at least three independent experiments. p-value  $\leq$  0.05 indicated with (\*), p-value  $\leq$  0.01 indicated with (\*\*), p-value  $\leq$  0.001 indicated with (\*\*\*) in whole paper.

## 5.0 References

- Aaboe, M., Birkenkamp-Demtroder, K., Wiuf, C., Sorensen, F. B., Thykjaer, T., Sauter, G., Jensen, K. M., Dyrskjot, L., and Orntoft, T. (2006). SOX4 expression in bladder carcinoma: clinical aspects and in vitro functional characterization. *Cancer Res* 66, 3434-3442.
- Aberle, H., Schwartz, H., and Kemler, R. (1996). Cadherin-catenin complex: protein interactions and their implications for cadherin function. *J Cell Biochem* 61, 514-523.
- Adams, J. C. (2001). Cell-matrix contact structures. *Cell Mol Life Sci* 58, 371-392.
- Adams, J. M., and Cory, S. (2007). The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* 26, 1324-1337.
- Ahn, S. G., Kim, H. S., Jeong, S. W., Kim, B. E., Rhim, H., Shim, J. Y., Kim, J. W., Lee, J. H., and Kim, I. K. (2002). Sox-4 is a positive regulator of Hep3B and HepG2 cells' apoptosis induced by prostaglandin (PG)A(2) and delta(12)-PGJ(2). *Exp Mol Med* 34, 243-249.
- Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J., and Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 100, 3983-3988.
- Alcorn, J. F., Guala, A. S., van der Velden, J., McElhinney, B., Irvin, C. G., Davis, R. J., and Janssen-Heininger, Y. M. (2008). Jun N-terminal kinase 1 regulates epithelial-to-mesenchymal transition induced by TGF-beta1. *Journal of cell science* 121, 1036-1045.
- Anastasiadis, P. Z. (2007). p120-ctn: A nexus for contextual signaling via Rho GTPases. *Biochim Biophys Acta* 1773, 34-46.
- Anest, V., Hanson, J. L., Cogswell, P. C., Steinbrecher, K. A., Strahl, B. D., and Baldwin, A. S. (2003). A nucleosomal function for I kappa B kinase-alpha in NF-kappa B-dependent gene expression. *Nature* 423, 659-663.
- Ansieau, S., Bastid, J., Doreau, A., Morel, A. P., Bouchet, B. P., Thomas, C., Fauvet, F., Puisieux, I., Doglioni, C., Piccinin, S., *et al.* (2008). Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer Cell* 14, 79-89.
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merckenschlager, M., and Fisher, A. G. (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8, 532-538.
- Baek, S. H. When signaling kinases meet histones and histone modifiers in the nucleus. *Mol Cell* 42, 274-284.
- Baeriswyl, V., and Christofori, G. (2009). The angiogenic switch in carcinogenesis. *Semin Cancer Biol* 19, 329-337.
- Bajenoff, M., Egen, J. G., Koo, L. Y., Laugier, J. P., Brau, F., Glaichenhaus, N., and Germain, R. N. (2006). Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity* 25, 989-1001.
- Bakin, A. V., Rinehart, C., Tomlinson, A. K., and Arteaga, C. L. (2002). p38 mitogen-activated protein kinase is required for TGFbeta-mediated fibroblastic transdifferentiation and cell migration. *J Cell Sci* 115, 3193-3206.
- Balwierz, P. J., Carninci, P., Daub, C. O., Kawai, J., Hayashizaki, Y., Van Belle, W., Beisel, C., and van Nimwegen, E. (2009). Methods for analyzing deep sequencing expression data: constructing the human and mouse promoterome with deepCAGE data. *Genome Biol* 10, R79.
- Bantignies, F., and Cavalli, G. (2006). Cellular memory and dynamic regulation of polycomb group proteins. *Curr Opin Cell Biol* 18, 275-283.
- Bardoni, B., and Mandel, J. L. (2002). Advances in understanding of fragile X pathogenesis and FMRP function, and in identification of X linked mental retardation genes. *Curr Opin Genet Dev* 12, 284-293.
- Barrios-Rodiles, M., Brown, K. R., Ozdamar, B., Bose, R., Liu, Z., Donovan, R. S., Shinjo, F., Liu, Y., Dembowy, J., Taylor, I. W., *et al.* (2005). High-throughput mapping of a dynamic signaling network in mammalian cells. *Science* 307, 1621-1625.
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling

- of histone methylations in the human genome. *Cell* 129, 823-837.
- Battle, 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.
- Bauerschmitz, G. J., Ranki, T., Kangasniemi, L., Ribacka, C., Eriksson, M., Porten, M., Herrmann, I., Ristimaki, A., Virkkunen, P., Tarkkanen, M., *et al.* (2008). Tissue-specific promoters active in CD44+CD24-/low breast cancer cells. *Cancer Res* 68, 5533-5539.
- Bell, O., Conrad, T., Kind, J., Wirbelauer, C., Akhtar, A., and Schubeler, D. (2008). Transcription-coupled methylation of histone H3 at lysine 36 regulates dosage compensation by enhancing recruitment of the MSL complex in *Drosophila melanogaster*. *Mol Cell Biol* 28, 3401-3409.
- Bell, O., Wirbelauer, C., Hild, M., Scharf, A. N., Schwaiger, M., MacAlpine, D. M., Zilbermann, F., van Leeuwen, F., Bell, S. P., Imhof, A., *et al.* (2007). Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in *Drosophila*. *EMBO J* 26, 4974-4984.
- Ben-Saadon, R., Zaaroor, D., Ziv, T., and Ciechanover, A. (2006). The polycomb protein Ring1B generates self atypical mixed ubiquitin chains required for its in vitro histone H2A ligase activity. *Mol Cell* 24, 701-711.
- Berdasco, M., and Esteller, M. (2010). Aberrant epigenetic landscape in cancer: how cellular identity goes awry. *Dev Cell* 19, 698-711.
- Berezovska, O. P., Glinskii, A. B., Yang, Z., Li, X. M., Hoffman, R. M., and Glinsky, G. V. (2006). Essential role for activation of the Polycomb group (PcG) protein chromatin silencing pathway in metastatic prostate cancer. *Cell Cycle* 5, 1886-1901.
- Bergers, G., and Benjamin, L. E. (2003). Tumorigenesis and the angiogenic switch. *Nat Rev Cancer* 3, 401-410.
- Bergh, J., Norberg, T., Sjogren, S., Lindgren, A., and Holmberg, L. (1995). Complete sequencing of the p53 gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. *Nat Med* 1, 1029-1034.
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.
- Berrier, A. L., Mastrangelo, A. M., Downward, J., Ginsberg, M., and LaFlamme, S. E. (2000). Activated R-ras, Rac1, PI 3-kinase and PKCepsilon can each restore cell spreading inhibited by isolated integrin beta1 cytoplasmic domains. *J Cell Biol* 151, 1549-1560.
- Bershadsky, A., Kozlov, M., and Geiger, B. (2006). Adhesion-mediated mechanosensitivity: a time to experiment, and a time to theorize. *Curr Opin Cell Biol* 18, 472-481.
- Berx, G., and van Roy, F. (2009). Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harb Perspect Biol* 1, a003129.
- Bhattaram, P., Penzo-Mendez, A., Sock, E., Colmenares, C., Kaneko, K. J., Vassilev, A., Depamphilis, M. L., Wegner, M., and Lefebvre, V. (2010). Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat Commun* 1, 9.
- Bhowmick, N. A., Neilson, E. G., and Moses, H. L. (2004). Stromal fibroblasts in cancer initiation and progression. *Nature* 432, 332-337.
- Bhutani, N., Burns, D. M., and Blau, H. M. (2011). DNA demethylation dynamics. *Cell* 146, 866-872.
- Blasco, M. A. (2005). Telomeres and human disease: ageing, cancer and beyond. *Nat Rev Genet* 6, 611-622.
- Blobe, G. C., Liu, X., Fang, S. J., How, T., and Lodish, H. F. (2001). A novel mechanism for regulating transforming growth factor beta (TGF-beta) signaling. Functional modulation of type III TGF-beta receptor expression through interaction with the PDZ domain protein, GIPC. *J Biol Chem* 276, 39608-39617.
- Bogenrieder, T., and Herlyn, M. (2003). Axis of evil: molecular mechanisms of cancer metastasis. *Oncogene* 22, 6524-6536.
- Bolos, V., Peinado, H., Perez-Moreno, M. A., Fraga, M. F., Esteller, M., and Cano, A. (2003).



- The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *J Cell Sci* *116*, 499-511.
- Bonnet, D., and Dick, J. E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* *3*, 730-737.
- Boominathan, L. (2010). The tumor suppressors p53, p63, and p73 are regulators of microRNA processing complex. *PLoS One* *5*, e10615.
- Bos, J. L. (2005). Linking Rap to cell adhesion. *Curr Opin Cell Biol* *17*, 123-128.
- Bos, J. L., de Bruyn, K., Enserink, J., Kuiperij, B., Rangarajan, S., Rehmann, H., Riedl, J., de Rooij, J., van Mansfeld, F., and Zwartkuis, F. (2003). The role of Rap1 in integrin-mediated cell adhesion. *Biochem Soc Trans* *31*, 83-86.
- Boutet, A., De Frutos, C. A., Maxwell, P. H., Mayol, M. J., Romero, J., and Nieto, M. A. (2006). Snail activation disrupts tissue homeostasis and induces fibrosis in the adult kidney. *EMBO J* *25*, 5603-5613.
- 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.
- Boyer, L. A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L. A., Lee, T. I., Levine, S. S., Wernig, M., Tajonar, A., Ray, M. K., *et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* *441*, 349-353.
- Brabletz, T., Jung, A., Spaderna, S., Hlubek, F., and Kirchner, T. (2005). Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nat Rev Cancer* *5*, 744-749.
- Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* *20*, 1123-1136.
- Briles, E. B., and Kornfeld, S. (1978). Isolation and metastatic properties of detachment variants of B16 melanoma cells. *J Natl Cancer Inst* *60*, 1217-1222.
- Brown, R. L., Reinke, L. M., Damerow, M. S., Perez, D., Chodosh, L. A., Yang, J., and Cheng, C. (2011). CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. *J Clin Invest* *121*, 1064-1074.
- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* *84*, 843-851.
- Bryan, T. M., Englezou, A., Gupta, J., Bacchetti, S., and Reddel, R. R. (1995). Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J* *14*, 4240-4248.
- Burk, U., Schubert, J., Wellner, U., Schmalhofer, O., Vincan, E., Spaderna, S., and Brabletz, T. (2008). A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep* *9*, 582-589.
- Burkhardt, D. L., and Sage, J. (2008). Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat Rev Cancer* *8*, 671-682.
- Burridge, K., and Chrzanowska-Wodnicka, M. (1996). Focal adhesions, contractility, and signaling. *Annu Rev Cell Dev Biol* *12*, 463-518.
- Bussing, I., Slack, F. J., and Grosshans, H. (2008). let-7 microRNAs in development, stem cells and cancer. *Trends Mol Med* *14*, 400-409.
- Calin, G. A., and Croce, C. M. (2006). MicroRNA signatures in human cancers. *Nat Rev Cancer* *6*, 857-866.
- Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A* *101*, 2999-3004.
- 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.
- Cao, Q., Yu, J., Dhanasekaran, S. M., Kim, J. H., Mani, R. S., Tomlins, S. A., Mehra, R.,

- Laxman, B., Cao, X., Kleer, C. G., *et al.* (2008). Repression of E-cadherin by the polycomb group protein EZH2 in cancer. *Oncogene* *27*, 7274-7284.
- Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol Cell* *20*, 845-854.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* *298*, 1039-1043.
- Cao, R., and Zhang, Y. (2004a). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* *14*, 155-164.
- Cao, R., and Zhang, Y. (2004b). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol Cell* *15*, 57-67.
- Carette, G., Di Padova, M., Micales, B., Lyons, G. E., and Sartorelli, V. (2004). The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev* *18*, 2627-2638.
- Cariati, M., Naderi, A., Brown, J. P., Smalley, M. J., Pinder, S. E., Caldas, C., and Purushotham, A. D. (2008). Alpha-6 integrin is necessary for the tumorigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. *Int J Cancer* *122*, 298-304.
- Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W. J., Anderson, S., Yates, J., Washburn, M. P., and Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* *123*, 581-592.
- Cavallaro, U. (2004). N-cadherin as an invasion promoter: a novel target for antitumor therapy? *Curr Opin Investig Drugs* *5*, 1274-1278.
- Cavallaro, U., and Christofori, G. (2004). Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat Rev Cancer* *4*, 118-132.
- Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* *10*, 295-304.
- Cerutti, H., and Casas-Mollano, J. A. (2009). Histone H3 phosphorylation: universal code or lineage specific dialects? *Epigenetics* *4*, 71-75.
- Cha, T. L., Zhou, B. P., Xia, W., Wu, Y., Yang, C. C., Chen, C. T., Ping, B., Otte, A. P., and Hung, M. C. (2005). Akt-mediated phosphorylation of EZH2 suppresses methylation of lysine 27 in histone H3. *Science* *310*, 306-310.
- Chamberlain, S. J., Yee, D., and Magnuson, T. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* *26*, 1496-1505.
- Chang, C. J., Chao, C. H., Xia, W., Yang, J. Y., Xiong, Y., Li, C. W., Yu, W. H., Rehman, S. K., Hsu, J. L., Lee, H. H., *et al.* (2011). p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nat Cell Biol* *13*, 317-323.
- Chang, L., Jones, Y., Ellisman, M. H., Goldstein, L. S., and Karin, M. (2003). JNK1 is required for maintenance of neuronal microtubules and controls phosphorylation of microtubule-associated proteins. *Developmental cell* *4*, 521-533.
- Chang, T. C., Yu, D., Lee, Y. S., Wentzel, E. A., Arking, D. E., West, K. M., Dang, C. V., Thomas-Tikhonenko, A., and Mendell, J. T. (2008). Widespread microRNA repression by Myc contributes to tumorigenesis. *Nat Genet* *40*, 43-50.
- Chen, C. R., Kang, Y., Siegel, P. M., and Massague, J. (2002). E2F4/5 and p107 as Smad cofactors linking the TGFbeta receptor to c-myc repression. *Cell* *110*, 19-32.
- Chen, C. S., Mrksich, M., Huang, S., Whitesides, G. M., and Ingber, D. E. (1997). Geometric control of cell life and death. *Science* *276*, 1425-1428.
- Chen, H., Tu, S. W., and Hsieh, J. T. (2005). Down-regulation of human DAB2IP gene expression mediated by polycomb Ezh2 complex and histone deacetylase in prostate cancer. *J Biol Chem* *280*, 22437-22444.
- Chen, X., Johns, D. C., Geiman, D. E., Marban, E., Dang, D. T., Hamlin, G., Sun, R., and Yang, V. W. (2001). Kruppel-like factor 4 (gut-enriched Kruppel-like factor) inhibits cell proliferation by blocking G1/S progression of the cell cycle. *J Biol Chem* *276*, 30423-30428.

- Chen, Z. F., and Behringer, R. R. (1995). twist is required in head mesenchyme for cranial neural tube morphogenesis. *Genes Dev* 9, 686-699.
- Cheng, C. W., Wu, P. E., Yu, J. C., Huang, C. S., Yue, C. T., Wu, C. W., and Shen, C. Y. (2001). Mechanisms of inactivation of E-cadherin in breast carcinoma: modification of the two-hit hypothesis of tumor suppressor gene. *Oncogene* 20, 3814-3823.
- Cheng, N., Chytil, A., Shyr, Y., Joly, A., and Moses, H. L. (2008). Transforming growth factor-beta signaling-deficient fibroblasts enhance hepatocyte growth factor signaling in mammary carcinoma cells to promote scattering and invasion. *Mol Cancer Res* 6, 1521-1533.
- Cheung, M., Abu-Elmagd, M., Clevers, H., and Scotting, P. J. (2000a). Roles of Sox4 in central nervous system development. *Brain Res Mol Brain Res* 79, 180-191.
- Cheung, P., Allis, C. D., and Sassone-Corsi, P. (2000b). Signaling to chromatin through histone modifications. *Cell* 103, 263-271.
- Cheung, P., Tanner, K. G., Cheung, W. L., Sassone-Corsi, P., Denu, J. M., and Allis, C. D. (2000c). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol Cell* 5, 905-915.
- Cho, W. C. (2010a). MicroRNAs in cancer - from research to therapy. *Biochim Biophys Acta* 1805, 209-217.
- Cho, W. C. (2010b). MicroRNAs: potential biomarkers for cancer diagnosis, prognosis and targets for therapy. *Int J Biochem Cell Biol* 42, 1273-1281.
- Christofori, G. (2006). New signals from the invasive front. *Nature* 441, 444-450.
- Chrzanowska-Wodnicka, M., and Burridge, K. (1996). Rho-stimulated contractility drives the formation of stress fibers and focal adhesions. *J Cell Biol* 133, 1403-1415.
- Clayton, A. L., Rose, S., Barratt, M. J., and Mahadevan, L. C. (2000). Phosphoacetylation of histone H3 on c-fos- and c-jun-associated nucleosomes upon gene activation. *EMBO J* 19, 3714-3726.
- Collen, A., Hanemaaijer, R., Lupu, F., Quax, P. H., van Lent, N., Grimbergen, J., Peters, E., Koolwijk, P., and van Hinsbergh, V. W. (2003). Membrane-type matrix metalloproteinase-mediated angiogenesis in a fibrin-collagen matrix. *Blood* 101, 1810-1817.
- Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, D., and van Roy, F. (2001). The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol Cell* 7, 1267-1278.
- Condeelis, J., and Segall, J. E. (2003). Intravital imaging of cell movement in tumours. *Nat Rev Cancer* 3, 921-930.
- Conery, A. R., Cao, Y., Thompson, E. A., Townsend, C. M., Jr., Ko, T. C., and Luo, K. (2004). Akt interacts directly with Smad3 to regulate the sensitivity to TGF-beta induced apoptosis. *Nat Cell Biol* 6, 366-372.
- Cui, H., Hu, B., Li, T., Ma, J., Alam, G., Gunning, W. T., and Ding, H. F. (2007). Bmi-1 is essential for the tumorigenicity of neuroblastoma cells. *Am J Pathol* 170, 1370-1378.
- Curto, M., Cole, B. K., Lallemand, D., Liu, C. H., and McClatchey, A. I. (2007). Contact-dependent inhibition of EGFR signaling by Nf2/Merlin. *J Cell Biol* 177, 893-903.
- Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* 111, 185-196.
- Dang, D. T., Bachman, K. E., Mahatan, C. S., Dang, L. H., Giardiello, F. M., and Yang, V. W. (2000). Decreased expression of the gut-enriched Kruppel-like factor gene in intestinal adenomas of multiple intestinal neoplasia mice and in colonic adenomas of familial adenomatous polyposis patients. *FEBS Lett* 476, 203-207.
- Das, C., Lucia, M. S., Hansen, K. C., and Tyler, J. K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* 459, 113-117.
- Davidson, L. A., and Keller, R. E. (1999). Neural tube closure in *Xenopus laevis* involves medial migration, directed protrusive activity, cell intercalation and convergent extension. *Development* 126, 4547-4556.
- Davies, M., Robinson, M., Smith, E., Huntley, S., Prime, S., and Paterson, I. (2005). Induction of an epithelial to mesenchymal transition in

- human immortal and malignant keratinocytes by TGF-beta1 involves MAPK, Smad and AP-1 signalling pathways. *J Cell Biochem* 95, 918-931.
- de Grouw, E. P., Raaijmakers, M. H., Boezeman, J. B., van der Reijden, B. A., van de Locht, L. T., de Witte, T. J., Jansen, J. H., and Raymakers, R. A. (2006). Preferential expression of a high number of ATP binding cassette transporters in both normal and leukemic CD34+CD38- cells. *Leukemia* 20, 750-754.
- De Smaele, E., Zazzeroni, F., Papa, S., Nguyen, D. U., Jin, R., Jones, J., Cong, R., and Franzoso, G. (2001). Induction of gadd45beta by NF-kappaB downregulates pro-apoptotic JNK signalling. *Nature* 414, 308-313.
- De Souza, C. P., Osmani, A. H., Wu, L. P., Spotts, J. L., and Osmani, S. A. (2000). Mitotic histone H3 phosphorylation by the NIMA kinase in *Aspergillus nidulans*. *Cell* 102, 293-302.
- Dean, M., Fojo, T., and Bates, S. (2005). Tumour stem cells and drug resistance. *Nat Rev Cancer* 5, 275-284.
- DeBerardinis, R. J., Lum, J. J., Hatzivassiliou, G., and Thompson, C. B. (2008). The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab* 7, 11-20.
- Derijard, B., Hibi, M., Wu, I. H., Barrett, T., Su, B., Deng, T., Karin, M., and Davis, R. J. (1994). JNK1: a protein kinase stimulated by UV light and Ha-Ras that binds and phosphorylates the c-Jun activation domain. *Cell* 76, 1025-1037.
- Derynck, R., Akhurst, R. J., and Balmain, A. (2001). TGF-beta signaling in tumor suppression and cancer progression. *Nat Genet* 29, 117-129.
- Deshpande, A., Sicinski, P., and Hinds, P. W. (2005). Cyclins and cdks in development and cancer: a perspective. *Oncogene* 24, 2909-2915.
- Desmedt, C., Piette, F., Loi, S., Wang, Y., Lallemand, F., Haibe-Kains, B., Viale, G., Delorenzi, M., Zhang, Y., d'Assignies, M. S., *et al.* (2007). Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res* 13, 3207-3214.
- Dion, M. F., Altschuler, S. J., Wu, L. F., and Rando, O. J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. *Proc Natl Acad Sci U S A* 102, 5501-5506.
- Downs, J. A., Lowndes, N. F., and Jackson, S. P. (2000). A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature* 408, 1001-1004.
- Dumont, N., Wilson, M. B., Crawford, Y. G., Reynolds, P. A., Sigaroudinia, M., and Tlsty, T. D. (2008). Sustained induction of epithelial to mesenchymal transition activates DNA methylation of genes silenced in basal-like breast cancers. *Proc Natl Acad Sci U S A* 105, 14867-14872.
- Dyxhoorn, D. M., Wu, Y., Xie, H., Yu, F., Lal, A., Petrocca, F., Martinvalet, D., Song, E., Lim, B., and Lieberman, J. (2009). miR-200 enhances mouse breast cancer cell colonization to form distant metastases. *PLoS One* 4, e7181.
- Dyson, M. H., Thomson, S., Inagaki, M., Goto, H., Arthur, S. J., Nightingale, K., Iborra, F. J., and Mahadevan, L. C. (2005). MAP kinase-mediated phosphorylation of distinct pools of histone H3 at S10 or S28 via mitogen- and stress-activated kinase 1/2. *J Cell Sci* 118, 2247-2259.
- Eden, S., Hashimshony, T., Keshet, I., Cedar, H., and Thorne, A. W. (1998). DNA methylation models histone acetylation. *Nature* 394, 842.
- Egeblad, M., and Werb, Z. (2002). New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2, 161-174.
- El Fahime, E., Torrente, Y., Caron, N. J., Bresolin, M. D., and Tremblay, J. P. (2000). In vivo migration of transplanted myoblasts requires matrix metalloproteinase activity. *Exp Cell Res* 258, 279-287.
- Eramo, A., Lotti, F., Sette, G., Pillozzi, E., Biffoni, M., Di Virgilio, A., Conticello, C., Rucio, L., Peschle, C., and De Maria, R. (2008). Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ* 15, 504-514.
- Erhardt, S., Su, I. H., Schneider, R., Barton, S., Bannister, A. J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovskiy, A., and Surani, M. A. (2003). Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. *Development* 130, 4235-4248.

- Ernst, T., Chase, A. J., Score, J., Hidalgo-Curtis, C. E., Bryant, C., Jones, A. V., Waghorn, K., Zoi, K., Ross, F. M., Reiter, A., *et al.* (2010). Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet* 42, 722-726.
- Esquela-Kerscher, A., and Slack, F. J. (2006). Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* 6, 259-269.
- Esteller, M. (2007). Epigenetic gene silencing in cancer: the DNA hypermethylome. *Hum Mol Genet* 16 *Spec No 1*, R50-59.
- Esteller, M. (2008a). Epigenetics in cancer. *N Engl J Med* 358, 1148-1159.
- Esteller, M. (2008b). Epigenetics in cancer. *N Engl J Med* 358, 1148-1159.
- Esteve, P. O., Chin, H. G., Smallwood, A., Feehery, G. R., Gangisetty, O., Karpf, A. R., Carey, M. F., and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev* 20, 3089-3103.
- Ezhkova, E., Lien, W. H., Stokes, N., Pasolli, H. A., Silva, J. M., and Fuchs, E. (2011). EZH1 and EZH2 cogovern histone H3K27 trimethylation and are essential for hair follicle homeostasis and wound repair. *Genes Dev* 25, 485-498.
- Ezhkova, E., Pasolli, H. A., Parker, J. S., Stokes, N., Su, I. H., Hannon, G., Tarakhovsky, A., and Fuchs, E. (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. *Cell* 136, 1122-1135.
- Feldman, N., Gerson, A., Fang, J., Li, E., Zhang, Y., Shinkai, Y., Cedar, H., and Bergman, Y. (2006). G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* 8, 188-194.
- Feldmann, G., Fendrich, V., McGovern, K., Bedja, D., Bisht, S., Alvarez, H., Koorstra, J. B., Habbe, N., Karikari, C., Mullendore, M., *et al.* (2008). An orally bioavailable small-molecule inhibitor of Hedgehog signaling inhibits tumor initiation and metastasis in pancreatic cancer. *Mol Cancer Ther* 7, 2725-2735.
- Feron, O. (2009). Pyruvate into lactate and back: from the Warburg effect to symbiotic energy fuel exchange in cancer cells. *Radiother Oncol* 92, 329-333.
- Fidler, I. J. (1975). Biological behavior of malignant melanoma cells correlated to their survival in vivo. *Cancer Res* 35, 218-224.
- Fidler, I. J. (2003). The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat Rev Cancer* 3, 453-458.
- Fidler, I. J., and Balch, C. M. (1987). The biology of cancer metastasis and implications for therapy. *Curr Probl Surg* 24, 129-209.
- Fidler, I. J., and Bucana, C. (1977). Mechanism of tumor cell resistance to lysis by syngeneic lymphocytes. *Cancer Res* 37, 3945-3956.
- Fidler, I. J., Darnell, J. H., and Budmen, M. B. (1976). Tumoricidal properties of mouse macrophages activated with mediators from rat lymphocytes stimulated with concanavalin A. *Cancer Res* 36, 3608-3615.
- Fidler, I. J., and Kripke, M. L. (1977). Metastasis results from preexisting variant cells within a malignant tumor. *Science* 197, 893-895.
- Fidler, I. J., and Radinsky, R. (1996). Search for genes that suppress cancer metastasis. *J Natl Cancer Inst* 88, 1700-1703.
- Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H., and Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* 438, 1116-1122.
- Foster, K. W., Frost, A. R., McKie-Bell, P., Lin, C. Y., Engler, J. A., Grizzle, W. E., and Ruppert, J. M. (2000). Increase of GSK3 messenger RNA and protein expression during progression of breast cancer. *Cancer Res* 60, 6488-6495.
- Foster, K. W., Liu, Z., Nail, C. D., Li, X., Fitzgerald, T. J., Bailey, S. K., Frost, A. R., Louro, I. D., Townes, T. M., Paterson, A. J., *et al.* (2005). Induction of KLF4 in basal keratinocytes blocks the proliferation-differentiation switch and initiates squamous epithelial dysplasia. *Oncogene* 24, 1491-1500.
- Fournier, A. K., Campbell, L. E., Castagnino, P., Liu, W. F., Chung, B. M., Weaver, V. M., Chen, C. S., and Assoian, R. K. (2008). Rac-dependent cyclin D1 gene expression regulated by cadherin- and integrin-mediated adhesion. *J Cell Sci* 121, 226-233.
- Francis, N. J., Follmer, N. E., Simon, M. D., Aghia, G., and Butler, J. D. (2009). Polycomb proteins remain bound to chromatin and DNA

- during DNA replication in vitro. *Cell* *137*, 110-122.
- Friedl, P. (2004). Prespecification and plasticity: shifting mechanisms of cell migration. *Curr Opin Cell Biol* *16*, 14-23.
- Friedl, P., and Brocker, E. B. (2000). The biology of cell locomotion within three-dimensional extracellular matrix. *Cell Mol Life Sci* *57*, 41-64.
- Friedl, P., and Wolf, K. (2003). Tumour-cell invasion and migration: diversity and escape mechanisms. *Nat Rev Cancer* *3*, 362-374.
- Friedl, P., and Wolf, K. (2010). Plasticity of cell migration: a multiscale tuning model. *J Cell Biol* *188*, 11-19.
- Fuchs, D., Daniel, V., Sadeghi, M., Opelz, G., and Naujokat, C. (2010). Salinomycin overcomes ABC transporter-mediated multidrug and apoptosis resistance in human leukemia stem cell-like KG-1a cells. *Biochem Biophys Res Commun* *394*, 1098-1104.
- Fuchs, D., Heinold, A., Opelz, G., Daniel, V., and Naujokat, C. (2009). Salinomycin induces apoptosis and overcomes apoptosis resistance in human cancer cells. *Biochem Biophys Res Commun* *390*, 743-749.
- Gallin, W. J., Edelman, G. M., and Cunningham, B. A. (1983). Characterization of L-CAM, a major cell adhesion molecule from embryonic liver cells. *Proc Natl Acad Sci U S A* *80*, 1038-1042.
- Gao, F., Foat, B. C., and Bussemaker, H. J. (2004). Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC Bioinformatics* *5*, 31.
- Ginestier, C., Liu, S., Diebel, M. E., Korkaya, H., Luo, M., Brown, M., Wicinski, J., Cabaud, O., Charafe-Jauffret, E., Birnbaum, D., *et al.* (2010). CXCR1 blockade selectively targets human breast cancer stem cells in vitro and in xenografts. *J Clin Invest* *120*, 485-497.
- Glinsky, G. V. (2007). Stem cell origin of death-from-cancer phenotypes of human prostate and breast cancers. *Stem Cell Rev* *3*, 79-93.
- Goll, M. G., and Bestor, T. H. (2005). Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* *74*, 481-514.
- Gomis, R. R., Alarcon, C., He, W., Wang, Q., Seoane, J., Lash, A., and Massague, J. (2006a). A FoxO-Smad synexpression group in human keratinocytes. *Proc Natl Acad Sci U S A* *103*, 12747-12752.
- Gomis, R. R., Alarcon, C., Nadal, C., Van Poznak, C., and Massague, J. (2006b). C/EBPbeta at the core of the TGFbeta cytostatic response and its evasion in metastatic breast cancer cells. *Cancer Cell* *10*, 203-214.
- Gopalkrishnan, R. V., Kang, D. C., and Fisher, P. B. (2001). Molecular markers and determinants of prostate cancer metastasis. *J Cell Physiol* *189*, 245-256.
- Graff, J. R., Herman, J. G., Lapidus, R. G., Chopra, H., Xu, R., Jarrard, D. F., Isaacs, W. B., Pitha, P. M., Davidson, N. E., and Baylin, S. B. (1995). E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. *Cancer Res* *55*, 5195-5199.
- Gregory, P. A., Bert, A. G., Paterson, E. L., Barry, S. C., Tsykin, A., Farshid, G., Vadas, M. A., Khew-Goodall, Y., and Goodall, G. J. (2008a). The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol* *10*, 593-601.
- Gregory, P. A., Bracken, C. P., Bert, A. G., and Goodall, G. J. (2008b). MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle* *7*, 3112-3118.
- Grille, S. J., Bellacosa, A., Upson, J., Klein-Szanto, A. J., van Roy, F., Lee-Kwon, W., Donowitz, M., Tschlis, P. N., and Larue, L. (2003). The protein kinase Akt induces epithelial mesenchymal transition and promotes enhanced motility and invasiveness of squamous cell carcinoma lines. *Cancer Res* *63*, 2172-2178.
- Grimaud, C., Bantignies, F., Pal-Bhadra, M., Ghana, P., Bhadra, U., and Cavalli, G. (2006). RNAi components are required for nuclear clustering of Polycomb group response elements. *Cell* *124*, 957-971.
- Grunert, S., Jechlinger, M., and Beug, H. (2003). Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. *Nat Rev Mol Cell Biol* *4*, 657-665.
- Gupta, P. B., Chaffer, C. L., and Weinberg, R. A. (2009a). Cancer stem cells: mirage or reality? *Nat Med* *15*, 1010-1012.
- Gupta, P. B., Onder, T. T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R. A., and Lander,

- E. S. (2009b). Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* *138*, 645-659.
- Hajra, K. M., Chen, D. Y., and Fearon, E. R. (2002). The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Res* *62*, 1613-1618.
- Hanahan, D., and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell* *100*, 57-70.
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* *144*, 646-674.
- Hansen, K. H., Bracken, A. P., Pasini, D., Dietrich, N., Gehani, S. S., Monrad, A., Rappsilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol* *10*, 1291-1300.
- Hay, E.D. (1968). Organization and fine structure of epithelium and mesenchyme in the developing chick embryo. In *Epithelial-Mesenchymal Interactions* (ed. R. Fleischmayer & R. E. Billingham), pp. 31-55. Baltimore, U.S.A.: Williams & Wilkins.
- Hazan, R. B., Qiao, R., Keren, R., Badano, I., and Suyama, K. (2004). Cadherin switch in tumor progression. *Ann N Y Acad Sci* *1014*, 155-163.
- Hegerfeldt, Y., Tusch, M., Brocker, E. B., and Friedl, P. (2002). Collective cell movement in primary melanoma explants: plasticity of cell-cell interaction, beta1-integrin function, and migration strategies. *Cancer Res* *62*, 2125-2130.
- Hemavathy, K., Guru, S. C., Harris, J., Chen, J. D., and Ip, Y. T. (2000). Human Slug is a repressor that localizes to sites of active transcription. *Mol Cell Biol* *20*, 5087-5095.
- Heppner, G. H., and Miller, B. E. (1983). Tumor heterogeneity: biological implications and therapeutic consequences. *Cancer Metastasis Rev* *2*, 5-23.
- 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.
- Hezel, A. F., and Bardeesy, N. (2008). LKB1; linking cell structure and tumor suppression. *Oncogene* *27*, 6908-6919.
- Hiraoka, N., Allen, E., Apel, I. J., Gyetko, M. R., and Weiss, S. J. (1998). Matrix metalloproteinases regulate neovascularization by acting as pericellular fibrinolysins. *Cell* *95*, 365-377.
- Hirsch, H. A., Iliopoulos, D., Tschlis, P. N., and Struhl, K. (2009). Metformin selectively targets cancer stem cells, and acts together with chemotherapy to block tumor growth and prolong remission. *Cancer Res* *69*, 7507-7511.
- Hong, C. S., and Saint-Jeannet, J. P. (2005). Sox proteins and neural crest development. *Semin Cell Dev Biol* *16*, 694-703.
- Hood, J. D., and Cheresch, D. A. (2002). Role of integrins in cell invasion and migration. *Nat Rev Cancer* *2*, 91-100.
- Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F., *et al.* (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* *102*, 279-291.
- Hsu, P. P., and Sabatini, D. M. (2008). Cancer cell metabolism: Warburg and beyond. *Cell* *134*, 703-707.
- Hu, M. C., Qiu, W. R., and Wang, Y. P. (1997). JNK1, JNK2 and JNK3 are p53 N-terminal serine 34 kinases. *Oncogene* *15*, 2277-2287.
- Huang, C., Rajfur, Z., Borchers, C., Schaller, M. D., and Jacobson, K. (2003). JNK phosphorylates paxillin and regulates cell migration. *Nature* *424*, 219-223.
- Huber, M. A., Kraut, N., and Beug, H. (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr Opin Cell Biol* *17*, 548-558.
- Hur, W., Rhim, H., Jung, C. K., Kim, J. D., Bae, S. H., Jang, J. W., Yang, J. M., Oh, S. T., Kim, D. G., Wang, H. J., *et al.* (2010). SOX4 overexpression regulates the p53-mediated apoptosis in hepatocellular carcinoma: clinical implication and functional analysis in vitro. *Carcinogenesis* *31*, 1298-1307.
- Ikenouchi, J., Matsuda, M., Furuse, M., and Tsukita, S. (2003). Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. *J Cell Sci* *116*, 1959-1967.

- Irina, O., and Friedl, P. (2009). Mechanisms of collective cell migration at a glance. *J Cell Sci* *122*, 3203-3208.
- Ingber, D. E., Prusty, D., Sun, Z., Betensky, H., and Wang, N. (1995). Cell shape, cytoskeletal mechanics, and cell cycle control in angiogenesis. *J Biomech* *28*, 1471-1484.
- Ireton, R. C., Davis, M. A., van Hengel, J., Mariner, D. J., Barnes, K., Thoreson, M. A., Anastasiadis, P. Z., Matrisian, L., Bundy, L. M., Sealy, L., *et al.* (2002). A novel role for p120 catenin in E-cadherin function. *J Cell Biol* *159*, 465-476.
- Itoh, F., Asao, H., Sugamura, K., Heldin, C. H., ten Dijke, P., and Itoh, S. (2001). Promoting bone morphogenetic protein signaling through negative regulation of inhibitory Smads. *EMBO J* *20*, 4132-4142.
- Jacobs, J. J., Kieboom, K., Marino, S., DePinho, R. A., and van Lohuizen, M. (1999a). The oncogene and Polycomb-group gene *bmi-1* regulates cell proliferation and senescence through the *ink4a* locus. *Nature* *397*, 164-168.
- Jacobs, J. J., Scheijen, B., Voncken, J. W., Kieboom, K., Berns, A., and van Lohuizen, M. (1999b). *Bmi-1* collaborates with *c-Myc* in tumorigenesis by inhibiting *c-Myc*-induced apoptosis via *INK4a/ARF*. *Genes Dev* *13*, 2678-2690.
- Jacobs, J. J., and van Lohuizen, M. (1999). Cellular memory of transcriptional states by Polycomb-group proteins. *Semin Cell Dev Biol* *10*, 227-235.
- Jacques, T. S., Relvas, J. B., Nishimura, S., Pytela, R., Edwards, G. M., Streuli, C. H., and French-Constant, C. (1998). Neural precursor cell chain migration and division are regulated through different beta1 integrins. *Development* *125*, 3167-3177.
- Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* *33 Suppl*, 245-254.
- Jechlinger, M., Grunert, S., and Beug, H. (2002). Mechanisms in epithelial plasticity and metastasis: insights from 3D cultures and expression profiling. *J Mammary Gland Biol Neoplasia* *7*, 415-432.
- Jin, L., Hope, K. J., Zhai, Q., Smadja-Joffe, F., and Dick, J. E. (2006). Targeting of CD44 eradicates human acute myeloid leukemic stem cells. *Nat Med* *12*, 1167-1174.
- Johnson, C. D., Esquela-Kerschner, A., Stefani, G., Byrom, M., Kelnar, K., Ovcharenko, D., Wilson, M., Wang, X., Shelton, J., Shingara, J., *et al.* (2007). The *let-7* microRNA represses cell proliferation pathways in human cells. *Cancer Res* *67*, 7713-7722.
- Junttila, M. R., and Evan, G. I. (2009). p53--a Jack of all trades but master of none. *Nat Rev Cancer* *9*, 821-829.
- Kalluri, R. (2009). EMT: when epithelial cells decide to become mesenchymal-like cells. *J Clin Invest* *119*, 1417-1419.
- Kalluri, R., and Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *J Clin Invest* *119*, 1420-1428.
- Kang, Y., Chen, C. R., and Massague, J. (2003). A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. *Mol Cell* *11*, 915-926.
- Kang, Y., and Massague, J. (2004). Epithelial-mesenchymal transitions: twist in development and metastasis. *Cell* *118*, 277-279.
- Katoh, Y., and Katoh, M. (2009). FGFR2-related pathogenesis and FGFR2-targeted therapeutics (Review). *Int J Mol Med* *23*, 307-311.
- Katz, J. P., Perreault, N., Goldstein, B. G., Actman, L., McNally, S. R., Silberg, D. G., Furth, E. E., and Kaestner, K. H. (2005). Loss of *Klf4* in mice causes altered proliferation and differentiation and precancerous changes in the adult stomach. *Gastroenterology* *128*, 935-945.
- Kauffman, E. C., Robinson, V. L., Stadler, W. M., Sokoloff, M. H., and Rinker-Schaeffer, C. W. (2003). Metastasis suppression: the evolving role of metastasis suppressor genes for regulating cancer cell growth at the secondary site. *J Urol* *169*, 1122-1133.
- Keirsebilck, A., Bonne, S., Staes, K., van Hengel, J., Nollet, F., Reynolds, A., and van Roy, F. (1998). Molecular cloning of the human p120ctn catenin gene (*CTNND1*): expression of multiple alternatively spliced isoforms. *Genomics* *50*, 129-146.
- Kennedy, K. M., and Dewhirst, M. W. (2010). Tumor metabolism of lactate: the influence and



- therapeutic potential for MCT and CD147 regulation. *Future Oncol* 6, 127-148.
- Kim, J. B., Ko, E., Han, W., Shin, I., Park, S. Y., and Noh, D. Y. (2008a). Berberine diminishes the side population and ABCG2 transporter expression in MCF-7 breast cancer cells. *Planta Med* 74, 1693-1700.
- Kim, J. B., Lee, K. M., Ko, E., Han, W., Lee, J. E., Shin, I., Bae, J. Y., Kim, S., and Noh, D. Y. (2008b). Berberine inhibits growth of the breast cancer cell lines MCF-7 and MDA-MB-231. *Planta Med* 74, 39-42.
- Kim, R., Emi, M., and Tanabe, K. (2007). Cancer immunoediting from immune surveillance to immune escape. *Immunology* 121, 1-14.
- Kim, S., Kang, H. Y., Nam, E. H., Choi, M. S., Zhao, X. F., Hong, C. S., Lee, J. W., Lee, J. H., and Park, Y. K. (2010). TMPRSS4 induces invasion and epithelial-mesenchymal transition through upregulation of integrin alpha5 and its signaling pathways. *Carcinogenesis* 31, 597-606.
- Kind, J., Vaquerizas, J. M., Gebhardt, P., Gentzel, M., Luscombe, N. M., Bertone, P., and Akhtar, A. (2008). Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell* 133, 813-828.
- Klebes, A., Sustar, A., Kechris, K., Li, H., Schubiger, G., and Kornberg, T. B. (2005). Regulation of cellular plasticity in *Drosophila* imaginal disc cells by the Polycomb group, trithorax group and lama genes. *Development* 132, 3753-3765.
- Kleer, C. G., Cao, Q., Varambally, S., Shen, R., Ota, I., Tomlins, S. A., Ghosh, D., Sewalt, R. G., Otte, A. P., Hayes, D. F., *et al.* (2003). EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc Natl Acad Sci U S A* 100, 11606-11611.
- Klinowska, T. C., Soriano, J. V., Edwards, G. M., Oliver, J. M., Valentijn, A. J., Montesano, R., and Streuli, C. H. (1999). Laminin and beta1 integrins are crucial for normal mammary gland development in the mouse. *Dev Biol* 215, 13-32.
- Klose, R. J., Kallin, E. M., and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet* 7, 715-727.
- Klymkowsky, M. W. (2005). beta-catenin and its regulatory network. *Hum Pathol* 36, 225-227.
- Kokudo, T., Suzuki, Y., Yoshimatsu, Y., Yamazaki, T., Watabe, T., and Miyazono, K. (2008). Snail is required for TGFbeta-induced endothelial-mesenchymal transition of embryonic stem cell-derived endothelial cells. *J Cell Sci* 121, 3317-3324.
- Kondo, M., Cubillo, E., Tobiume, K., Shirakihara, T., Fukuda, N., Suzuki, H., Shimizu, K., Takehara, K., Cano, A., Saitoh, M., and Miyazono, K. (2004). A role for Id in the regulation of TGF-beta-induced epithelial-mesenchymal transdifferentiation. *Cell Death Differ* 11, 1092-1101.
- Korpai, M., Lee, E. S., Hu, G., and Kang, Y. (2008). The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *J Biol Chem* 283, 14910-14914.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.
- Kowanetz, M., Valcourt, U., Bergstrom, R., Heldin, C. H., and Moustakas, A. (2004). Id2 and Id3 define the potency of cell proliferation and differentiation responses to transforming growth factor beta and bone morphogenetic protein. *Mol Cell Biol* 24, 4241-4254.
- Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., *et al.* (2003a). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell* 11, 721-729.
- Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D. P., Beattie, B. K., Emili, A., Boone, C., *et al.* (2003b). Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* 23, 4207-4218.
- Ku, M., Koche, R. P., Rheinbay, E., Mendenhall, E. M., Endoh, M., Mikkelsen, T. S., Presser, A., Nusbaum, C., Xie, X., Chi, A. S., *et al.* (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* 4, e1000242.

- Kurdistani, S. K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. *Cell* *117*, 721-733.
- Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. *Mol Cell* *14*, 183-193.
- Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* *410*, 116-120.
- Lamouille, S., and Derynck, R. (2007). Cell size and invasion in TGF-beta-induced epithelial to mesenchymal transition is regulated by activation of the mTOR pathway. *J Cell Biol* *178*, 437-451.
- Lanzuolo, C., Roure, V., Dekker, J., Bantignies, F., and Orlando, V. (2007). Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol* *9*, 1167-1174.
- Laping, N. J., Grygielko, E., Mathur, A., Butter, S., Bomberger, J., Tweed, C., Martin, W., Fornwald, J., Lehr, R., Harling, J., *et al.* (2002). Inhibition of transforming growth factor (TGF)-beta1-induced extracellular matrix with a novel inhibitor of the TGF-beta type I receptor kinase activity: SB-431542. *Mol Pharmacol* *62*, 58-64.
- Lapuk, A., Marr, H., Jakkula, L., Pedro, H., Bhattacharya, S., Purdom, E., Hu, Z., Simpson, K., Pachter, L., Durinck, S., *et al.* (2010). Exon-level microarray analyses identify alternative splicing programs in breast cancer. *Mol Cancer Res* *8*, 961-974.
- Lauffenburger, D. A., and Horwitz, A. F. (1996). Cell migration: a physically integrated molecular process. *Cell* *84*, 359-369.
- Le, M. T., Teh, C., Shyh-Chang, N., Xie, H., Zhou, B., Korzh, V., Lodish, H. F., and Lim, B. (2009). MicroRNA-125b is a novel negative regulator of p53. *Genes Dev* *23*, 862-876.
- Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* *72*, 73-84.
- Lee, N., Maurange, C., Ringrose, L., and Paro, R. (2005). Suppression of Polycomb group proteins by JNK signalling induces transdetermination in *Drosophila* imaginal discs. *Nature* *438*, 234-237.
- Lee, T. I., Jenner, R. G., Boyer, L. A., Guenther, M. G., Levine, S. S., Kumar, R. M., Chevalier, B., Johnstone, S. E., Cole, M. F., Isono, K., *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* *125*, 301-313.
- Lehembre, F., Yilmaz, M., Wicki, A., Schomber, T., Strittmatter, K., Ziegler, D., Kren, A., Went, P., Derksen, P. W., Berns, A., *et al.* (2008). NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin. *EMBO J* *27*, 2603-2615.
- Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* *13*, 1192-1200.
- Lele, T. P., Thodeti, C. K., and Ingber, D. E. (2006). Force meets chemistry: analysis of mechanochemical conversion in focal adhesions using fluorescence recovery after photobleaching. *J Cell Biochem* *97*, 1175-1183.
- Levine, S. S., King, I. F., and Kingston, R. E. (2004). Division of labor in polycomb group repression. *Trends Biochem Sci* *29*, 478-485.
- Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* *120*, 15-20.
- Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., Wicha, M., Clarke, M. F., and Simeone, D. M. (2007). Identification of pancreatic cancer stem cells. *Cancer Res* *67*, 1030-1037.
- Liao, J. H., Chen, J. S., Chai, M. Q., Zhao, S., and Song, J. G. (2001). The involvement of p38 MAPK in transforming growth factor beta1-induced apoptosis in murine hepatocytes. *Cell Res* *11*, 89-94.
- Liao, Y. L., Sun, Y. M., Chau, G. Y., Chau, Y. P., Lai, T. C., Wang, J. L., Horng, J. T., Hsiao, M., and Tsou, A. P. (2008). Identification of SOX4 target genes using phylogenetic footprinting-based prediction from expression microarrays suggests that overexpression of SOX4 potentiates metastasis in hepatocellular carcinoma. *Oncogene*.

- Lienert, F., Wirbelauer, C., Som, I., Dean, A., Mohn, F., and Schubeler, D. (2011). Identification of genetic elements that autonomously determine DNA methylation states. *Nat Genet.*
- Lioubinski, O., Muller, M., Wegner, M., and Sander, M. (2003). Expression of Sox transcription factors in the developing mouse pancreas. *Dev Dyn* 227, 402-408.
- Liu, P., Ramachandran, S., Ali Seyed, M., Scharer, C. D., Laycock, N., Dalton, W. B., Williams, H., Karanam, S., Datta, M. W., Jaye, D. L., and Moreno, C. S. (2006a). Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. *Cancer Res* 66, 4011-4019.
- Liu, W. F., Nelson, C. M., Pirone, D. M., and Chen, C. S. (2006b). E-cadherin engagement stimulates proliferation via Rac1. *J Cell Biol* 173, 431-441.
- Lo, H. W., Hsu, S. C., Xia, W., Cao, X., Shih, J. Y., Wei, Y., Abbruzzese, J. L., Hortobagyi, G. N., and Hung, M. C. (2007). Epidermal growth factor receptor cooperates with signal transducer and activator of transcription 3 to induce epithelial-mesenchymal transition in cancer cells via up-regulation of TWIST gene expression. *Cancer Res* 67, 9066-9076.
- Lo, W. S., Trievel, R. C., Rojas, J. R., Duggan, L., Hsu, J. Y., Allis, C. D., Marmorstein, R., and Berger, S. L. (2000). Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14. *Mol Cell* 5, 917-926.
- Long, J., Zuo, D., and Park, M. (2005). Pc2-mediated sumoylation of Smad-interacting protein 1 attenuates transcriptional repression of E-cadherin. *J Biol Chem* 280, 35477-35489.
- Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., *et al.* (2005). MicroRNA expression profiles classify human cancers. *Nature* 435, 834-838.
- Lu, Z., Ghosh, S., Wang, Z., and Hunter, T. (2003). Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion. *Cancer Cell* 4, 499-515.
- Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- Ma, L., Young, J., Prabhala, H., Pan, E., Mestdagh, P., Muth, D., Teruya-Feldstein, J., Reinhardt, F., Onder, T. T., Valastyan, S., *et al.* (2010). miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat Cell Biol* 12, 247-256.
- Maeda, M., Johnson, K. R., and Wheelock, M. J. (2005). Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition. *J Cell Sci* 118, 873-887.
- 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.
- Manni, I., Artuso, S., Careccia, S., Rizzo, M. G., Baserga, R., Piaggio, G., and Sacchi, A. (2009). The microRNA miR-92 increases proliferation of myeloid cells and by targeting p63 modulates the abundance of its isoforms. *FASEB J* 23, 3957-3966.
- Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., Dynlacht, B. D., and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* 32, 503-518.
- Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* 6, 838-849.
- Martinez, A. M., Colomb, S., Dejardin, J., Bantignies, F., and Cavalli, G. (2006). Polycomb group-dependent Cyclin A repression in *Drosophila*. *Genes Dev* 20, 501-513.
- Massague, J. (2008). TGFbeta in Cancer. *Cell* 134, 215-230.
- Massari, M. E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol* 20, 429-440.
- Matise, L. A., Pickup, M. W., and Moses, H. L. (2009). TGF-beta helps cells fly solo. *Nat Cell Biol* 11, 1281-1284.
- Matsumoto, Y., Tanaka, K., Harimaya, K., Nakatani, F., Matsuda, S., and Iwamoto, Y.

- (2001). Small GTP-binding protein, Rho, both increased and decreased cellular motility, activation of matrix metalloproteinase 2 and invasion of human osteosarcoma cells. *Jpn J Cancer Res* 92, 429-438.
- McDonald, O. G., Wu, H., Timp, W., Doi, A., and Feinberg, A. P. (2011a). Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol* 18, 867-874.
- McDonald, O. G., Wu, H., Timp, W., Doi, A., and Feinberg, A. P. (2011b). Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol* 18, 867-874.
- Medici, D., Hay, E. D., and Olsen, B. R. (2008). Snail and Slug promote epithelial-mesenchymal transition through beta-catenin-T-cell factor-4-dependent expression of transforming growth factor-beta3. *Mol Biol Cell* 19, 4875-4887.
- Mempel, T. R., Junt, T., and von Andrian, U. H. (2006). Rulers over randomness: stroma cells guide lymphocyte migration in lymph nodes. *Immunity* 25, 867-869.
- Miettinen, P. J., Ebner, R., Lopez, A. R., and Derynck, R. (1994). TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *J Cell Biol* 127, 2021-2036.
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.
- Miller, M. J., Wei, S. H., Cahalan, M. D., and Parker, I. (2003). Autonomous T cell trafficking examined in vivo with intravital two-photon microscopy. *Proc Natl Acad Sci U S A* 100, 2604-2609.
- Miller, M. J., Wei, S. H., Parker, I., and Cahalan, M. D. (2002). Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. *Science* 296, 1869-1873.
- Mills, A. A. (2010). Throwing the cancer switch: reciprocal roles of polycomb and trithorax proteins. *Nat Rev Cancer* 10, 669-682.
- Min, J., Zaslavsky, A., Fedele, G., McLaughlin, S. K., Reczek, E. E., De Raedt, T., Guney, I., Strohlic, D. E., Macconail, L. E., Beroukhim, R., *et al.* (2010). An oncogene-tumor suppressor cascade drives metastatic prostate cancer by coordinately activating Ras and nuclear factor-kappaB. *Nat Med* 16, 286-294.
- Minn, A. J., Gupta, G. P., Padua, D., Bos, P., Nguyen, D. X., Nuyten, D., Kreike, B., Zhang, Y., Wang, Y., Ishwaran, H., *et al.* (2007). Lung metastasis genes couple breast tumor size and metastatic spread. *Proc Natl Acad Sci U S A* 104, 6740-6745.
- Mo, Y. Y., and Reynolds, A. B. (1996). Identification of murine p120 isoforms and heterogeneous expression of p120cas isoforms in human tumor cell lines. *Cancer Res* 56, 2633-2640.
- Morali, O. G., Delmas, V., Moore, R., Jeanney, C., Thiery, J. P., and Larue, L. (2001). IGF-II induces rapid beta-catenin relocation to the nucleus during epithelium to mesenchyme transition. *Oncogene* 20, 4942-4950.
- Morel, A. P., Lievre, M., Thomas, C., Hinkal, G., Ansieau, S., and Puisieux, A. (2008). Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS One* 3, e2888.
- Moreno-Bueno, G., Cubillo, E., Sarrío, D., Peinado, H., Rodríguez-Pinilla, S. M., Villa, S., Bolos, V., Jorda, M., Fabra, A., Portillo, F., *et al.* (2006). Genetic profiling of epithelial cells expressing E-cadherin repressors reveals a distinct role for Snail, Slug, and E47 factors in epithelial-mesenchymal transition. *Cancer Res* 66, 9543-9556.
- Mougiakakos, D., Choudhury, A., Lladser, A., Kiessling, R., and Johansson, C. C. (2010). Regulatory T cells in cancer. *Adv Cancer Res* 107, 57-117.
- Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., and Simon, J. A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* 111, 197-208.
- Nabeshima, K., Inoue, T., Shimao, Y., Okada, Y., Itoh, Y., Seiki, M., and Koono, M. (2000). Front-cell-specific expression of membrane-type 1 matrix metalloproteinase and gelatinase A during cohort migration of colon carcinoma cells induced by hepatocyte growth factor/scatter factor. *Cancer Res* 60, 3364-3369.

- Negishi, M., Saraya, A., Miyagi, S., Nagao, K., Inagaki, Y., Nishikawa, M., Tajima, S., Koseki, H., Tsuda, H., Takasaki, Y., *et al.* (2007). Bmi1 cooperates with Dnmt1-associated protein 1 in gene silencing. *Biochem Biophys Res Commun* *353*, 992-998.
- Nelson, W. J., and Nusse, R. (2004). Convergence of Wnt, beta-catenin, and cadherin pathways. *Science* *303*, 1483-1487.
- Niessen, K., Fu, Y., Chang, L., Hoodless, P. A., McFadden, D., and Karsan, A. (2008). Slug is a direct Notch target required for initiation of cardiac cushion cellularization. *J Cell Biol* *182*, 315-325.
- Nieto, M. A. (2010). The Ins and Outs of the Epithelial to Mesenchymal Transition in Health and Disease. *Annu Rev Cell Dev Biol*.
- Nowak, S. J., and Corces, V. G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet* *20*, 214-220.
- O'Carroll, D., Erhardt, S., Pagani, M., Barton, S. C., Surani, M. A., and Jenuwein, T. (2001). The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* *21*, 4330-4336.
- Ohm, J. E., McGarvey, K. M., Yu, X., Cheng, L., Schuebel, K. E., Cope, L., Mohammad, H. P., Chen, W., Daniel, V. C., Yu, W., *et al.* (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat Genet* *39*, 237-242.
- Ohnishi, S., Ohnami, S., Laub, F., Aoki, K., Suzuki, K., Kanai, Y., Haga, K., Asaka, M., Ramirez, F., and Yoshida, T. (2003). Downregulation and growth inhibitory effect of epithelial-type Kruppel-like transcription factor KLF4, but not KLF5, in bladder cancer. *Biochem Biophys Res Commun* *308*, 251-256.
- Okada, T., Lopez-Lago, M., and Giancotti, F. G. (2005). Merlin/NF-2 mediates contact inhibition of growth by suppressing recruitment of Rac to the plasma membrane. *J Cell Biol* *171*, 361-371.
- Oltean, S., Sorg, B. S., Albrecht, T., Bonano, V. I., Brazas, R. M., Dewhirst, M. W., and Garcia-Blanco, M. A. (2006). Alternative inclusion of fibroblast growth factor receptor 2 exon IIIc in Dunning prostate tumors reveals unexpected epithelial mesenchymal plasticity. *Proc Natl Acad Sci U S A* *103*, 14116-14121.
- Ostrand-Rosenberg, S., and Sinha, P. (2009). Myeloid-derived suppressor cells: linking inflammation and cancer. *J Immunol* *182*, 4499-4506.
- Ozben, T. (2006). Mechanisms and strategies to overcome multiple drug resistance in cancer. *FEBS Lett* *580*, 2903-2909.
- Ozdamar, B., Bose, R., Barrios-Rodiles, M., Wang, H. R., Zhang, Y., and Wrana, J. L. (2005). Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity. *Science* *307*, 1603-1609.
- Pajares, M. J., Ezponda, T., Catena, R., Calvo, A., Pio, R., and Montuenga, L. M. (2007). Alternative splicing: an emerging topic in molecular and clinical oncology. *Lancet Oncol* *8*, 349-357.
- Pals, S. T., de Gorter, D. J., and Spaargaren, M. (2007). Lymphoma dissemination: the other face of lymphocyte homing. *Blood* *110*, 3102-3111.
- Papagiannakopoulos, T., Shapiro, A., and Kosik, K. S. (2008). MicroRNA-21 targets a network of key tumor-suppressive pathways in glioblastoma cells. *Cancer Res* *68*, 8164-8172.
- Pardali, K., and Moustakas, A. (2007). Actions of TGF-beta as tumor suppressor and prometastatic factor in human cancer. *Biochim Biophys Acta* *1775*, 21-62.
- Park, S. M., Gaur, A. B., Lengyel, E., and Peter, M. E. (2008). The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev* *22*, 894-907.
- Parker, K. K., and Ingber, D. E. (2007). Extracellular matrix, mechanotransduction and structural hierarchies in heart tissue engineering. *Philos Trans R Soc Lond B Biol Sci* *362*, 1267-1279.
- Partanen, J. I., Nieminen, A. I., and Klefstrom, J. (2009). 3D view to tumor suppression: Lkb1, polarity and the arrest of oncogenic c-Myc. *Cell Cycle* *8*, 716-724.
- Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol* *27*, 3769-3779.

- Paterson, E. L., Kolesnikoff, N., Gregory, P. A., Bert, A. G., Khew-Goodall, Y., and Goodall, G. J. (2008). The microRNA-200 family regulates epithelial to mesenchymal transition. *ScientificWorldJournal* 8, 901-904.
- 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.
- Perez-Cadahia, B., Drobnic, B., and Davie, J. R. (2009). H3 phosphorylation: dual role in mitosis and interphase. *Biochem Cell Biol* 87, 695-709.
- Perez-Moreno, M., Jamora, C., and Fuchs, E. (2003). Sticky business: orchestrating cellular signals at adherens junctions. *Cell* 112, 535-548.
- Perez-Soler, R., Neamati, N., Zou, Y., Schneider, E., Doyle, L. A., Andreeff, M., Priebe, W., and Ling, Y. H. (1997). Annamycin circumvents resistance mediated by the multidrug resistance-associated protein (MRP) in breast MCF-7 and small-cell lung UMCC-1 cancer cell lines selected for resistance to etoposide. *Int J Cancer* 71, 35-41.
- Perl, A. K., Wilgenbus, P., Dahl, U., Semb, H., and Christofori, G. (1998). A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature* 392, 190-193.
- Perrais, M., Chen, X., Perez-Moreno, M., and Gumbiner, B. M. (2007). E-cadherin homophilic ligation inhibits cell growth and epidermal growth factor receptor signaling independently of other cell interactions. *Mol Biol Cell* 18, 2013-2025.
- Peters, A. H., and Schubeler, D. (2005). Methylation of histones: playing memory with DNA. *Curr Opin Cell Biol* 17, 230-238.
- Peyrieras, N., Hyafil, F., Louvard, D., Ploegh, H. L., and Jacob, F. (1983). Uvomorulin: a nonintegral membrane protein of early mouse embryo. *Proc Natl Acad Sci U S A* 80, 6274-6277.
- Piek, E., Moustakas, A., Kurisaki, A., Heldin, C. H., and ten Dijke, P. (1999a). TGF-(beta) type I receptor/ALK-5 and Smad proteins mediate epithelial to mesenchymal transdifferentiation in NMuMG breast epithelial cells. *J Cell Sci* 112 ( Pt 24), 4557-4568.
- Piek, E., Moustakas, A., Kurisaki, A., Heldin, C. H., and ten Dijke, P. (1999b). TGF-(beta) type I receptor/ALK-5 and Smad proteins mediate epithelial to mesenchymal transdifferentiation in NMuMG breast epithelial cells. *Journal of cell science* 112 ( Pt 24), 4557-4568.
- Pino, M. S., Balsamo, M., Di Modugno, F., Mottolese, M., Alessio, M., Melucci, E., Milella, M., McConkey, D. J., Philippar, U., Gertler, F. B., *et al.* (2008). Human Mena+11a isoform serves as a marker of epithelial phenotype and sensitivity to epidermal growth factor receptor inhibition in human pancreatic cancer cell lines. *Clin Cancer Res* 14, 4943-4950.
- Pitts, W. C., Rojas, V. A., Gaffey, M. J., Rouse, R. V., Esteban, J., Frierson, H. F., Kempson, R. L., and Weiss, L. M. (1991). Carcinomas with metaplasia and sarcomas of the breast. *Am J Clin Pathol* 95, 623-632.
- Platt, V. M., and Szoka, F. C., Jr. (2008). Anticancer therapeutics: targeting macromolecules and nanocarriers to hyaluronan or CD44, a hyaluronan receptor. *Mol Pharm* 5, 474-486.
- Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolzheimer, E., *et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 122, 517-527.
- Polyak, K., and Weinberg, R. A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer* 9, 265-273.
- Poser, I., Dominguez, D., de Herreros, A. G., Varnai, A., Buettner, R., and Bosserhoff, A. K. (2001). Loss of E-cadherin expression in melanoma cells involves up-regulation of the transcriptional repressor Snail. *J Biol Chem* 276, 24661-24666.
- Pramoonjago, P., Baras, A. S., and Moskaluk, C. A. (2006). Knockdown of Sox4 expression by RNAi induces apoptosis in ACC3 cells. *Oncogene* 25, 5626-5639.
- Prigent, C., and Dimitrov, S. (2003). Phosphorylation of serine 10 in histone H3, what for? *J Cell Sci* 116, 3677-3685.

- Prince, M. E., Sivanandan, R., Kaczorowski, A., Wolf, G. T., Kaplan, M. J., Dalerba, P., Weissman, I. L., Clarke, M. F., and Ailles, L. E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc Natl Acad Sci U S A* *104*, 973-978.
- Radpour, R., Berekati, Z., Kohler, C., Lv, Q., Burki, N., Diesch, C., Bitzer, J., Zheng, H., Schmid, S., and Zhong, X. Y. (2011). Hypermethylation of tumor suppressor genes involved in critical regulatory pathways for developing a blood-based test in breast cancer. *PLoS One* *6*, e16080.
- Razin, A., and Cedar, H. (1977). Distribution of 5-methylcytosine in chromatin. *Proc Natl Acad Sci U S A* *74*, 2725-2728.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* *406*, 593-599.
- Remy, I., Montmarquette, A., and Michnick, S. W. (2004). PKB/Akt modulates TGF-beta signalling through a direct interaction with Smad3. *Nat Cell Biol* *6*, 358-365.
- Restivo, A., Piacentini, G., Placidi, S., Saffirio, C., and Marino, B. (2006). Cardiac outflow tract: a review of some embryogenetic aspects of the conotruncal region of the heart. *Anat Rec A Discov Mol Cell Evol Biol* *288*, 936-943.
- Rhodes, D. R., Sanda, M. G., Otte, A. P., Chinnaiyan, A. M., and Rubin, M. A. (2003). Multiplex biomarker approach for determining risk of prostate-specific antigen-defined recurrence of prostate cancer. *J Natl Cancer Inst* *95*, 661-668.
- Riccioni, R., Dupuis, M. L., Bernabei, M., Petrucci, E., Pasquini, L., Mariani, G., Cianfriglia, M., and Testa, U. (2010). The cancer stem cell selective inhibitor salinomycin is a p-glycoprotein inhibitor. *Blood Cells Mol Dis* *45*, 86-92.
- Ringrose, L., and Paro, R. (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* *134*, 223-232.
- Riveline, D., Zamir, E., Balaban, N. Q., Schwarz, U. S., Ishizaki, T., Narumiya, S., Kam, Z., Geiger, B., and Bershadsky, A. D. (2001). Focal contacts as mechanosensors: externally applied local mechanical force induces growth of focal contacts by an mDia1-dependent and ROCK-independent mechanism. *J Cell Biol* *153*, 1175-1186.
- Robinson, P. J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R. G., and Rhodes, D. (2008). 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction. *J Mol Biol* *381*, 816-825.
- Robinson, P. J., and Rhodes, D. (2006). Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* *16*, 336-343.
- Rogakou, E. P., Boon, C., Redon, C., and Bonner, W. M. (1999). Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J Cell Biol* *146*, 905-916.
- Rowland, B. D., Bernards, R., and Peeper, D. S. (2005). The KLF4 tumour suppressor is a transcriptional repressor of p53 that acts as a context-dependent oncogene. *Nat Cell Biol* *7*, 1074-1082.
- Ruike, Y., Imanaka, Y., Sato, F., Shimizu, K., and Tsujimoto, G. (2010). Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. *BMC Genomics* *11*, 137.
- Sampson, V. B., Rong, N. H., Han, J., Yang, Q., Aris, V., Soteropoulos, P., Petrelli, N. J., Dunn, S. P., and Krueger, L. J. (2007). MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. *Cancer Res* *67*, 9762-9770.
- Sanchez-Tillo, E., Lazaro, A., Torrent, R., Cuatrecasas, M., Vaquero, E. C., Castells, A., Engel, P., and Postigo, A. (2010). ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. *Oncogene* *29*, 3490-3500.
- Santibanez, J. F. (2006a). JNK mediates TGF-beta1-induced epithelial mesenchymal transdifferentiation of mouse transformed keratinocytes. *FEBS letters* *580*, 5385-5391.
- Santibanez, J. F. (2006b). JNK mediates TGF-beta1-induced epithelial mesenchymal

- transdifferentiation of mouse transformed keratinocytes. *FEBS Lett* 580, 5385-5391.
- Sassone-Corsi, P., Mizzen, C. A., Cheung, P., Crosio, C., Monaco, L., Jacquot, S., Hanauer, A., and Allis, C. D. (1999). Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. *Science* 285, 886-891.
- Satijn, D. P., Hamer, K. M., den Blaauwen, J., and Otte, A. P. (2001). The polycomb group protein EED interacts with YY1, and both proteins induce neural tissue in *Xenopus* embryos. *Mol Cell Biol* 21, 1360-1369.
- Saunders, A., Core, L. J., and Lis, J. T. (2006). Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* 7, 557-567.
- Sauvageau, M., and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell* 7, 299-313.
- Savagner, P., Valles, A. M., Jouanneau, J., Yamada, K. M., and Thiery, J. P. (1994). Alternative splicing in fibroblast growth factor receptor 2 is associated with induced epithelial-mesenchymal transition in rat bladder carcinoma cells. *Mol Biol Cell* 5, 851-862.
- Savagner, P., Yamada, K. M., and Thiery, J. P. (1997). The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. *J Cell Biol* 137, 1403-1419.
- Scharer, C. D., McCabe, C. D., Ali-Seyed, M., Berger, M. F., Bulyk, M. L., and Moreno, C. S. (2009). Genome-wide promoter analysis of the SOX4 transcriptional network in prostate cancer cells. *Cancer Res* 69, 709-717.
- Scheel, A. H., Beyer, U., Agami, R., and Dobbstein, M. (2009). Immunofluorescence-based screening identifies germ cell associated microRNA 302 as an antagonist to p63 expression. *Cell Cycle* 8, 1426-1432.
- Schilham, M. W., Moerer, P., Cumano, A., and Clevers, H. C. (1997). Sox-4 facilitates thymocyte differentiation. *Eur J Immunol* 27, 1292-1295.
- Schilham, M. W., Oosterwegel, M. A., Moerer, P., Ya, J., de Boer, P. A., van de Wetering, M., Verbeek, S., Lamers, W. H., Kruisbeek, A. M., Cumano, A., and Clevers, H. (1996). Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. *Nature* 380, 711-714.
- Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B. E., *et al.* (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat Genet* 39, 232-236.
- Schmidt, M., Bohm, D., von Torne, C., Steiner, E., Puhl, A., Pilch, H., Lehr, H. A., Hengstler, J. G., Kolbl, H., and Gehrman, M. (2008). The humoral immune system has a key prognostic impact in node-negative breast cancer. *Cancer Res* 68, 5405-5413.
- Schubeler, D., MacAlpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D. E., O'Neill, L. P., Turner, B. M., Delrow, J., *et al.* (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev* 18, 1263-1271.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* 128, 735-745.
- Schwartz, Y. B., and Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. *Nat Rev Genet* 8, 9-22.
- Seftor, E. A., Meltzer, P. S., Kirschmann, D. A., Pe'er, J., Maniotis, A. J., Trent, J. M., Folberg, R., and Hendrix, M. J. (2002). Molecular determinants of human uveal melanoma invasion and metastasis. *Clin Exp Metastasis* 19, 233-246.
- Segre, J. A., Bauer, C., and Fuchs, E. (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nat Genet* 22, 356-360.
- Semenza, G. L. (2008). Tumor metabolism: cancer cells give and take lactate. *J Clin Invest* 118, 3835-3837.
- Semenza, G. L. (2010). HIF-1: upstream and downstream of cancer metabolism. *Curr Opin Genet Dev* 20, 51-56.
- Seoane, J., Le, H. V., Shen, L., Anderson, S. A., and Massague, J. (2004). Integration of Smad and forkhead pathways in the control of



- neuroepithelial and glioblastoma cell proliferation. *Cell* 117, 211-223.
- Shapiro, I. M., Cheng, A. W., Flytzanis, N. C., Balsamo, M., Condeelis, J. S., Oktay, M. H., Burge, C. B., and Gertler, F. B. (2011). An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genet* 7, e1002218.
- Shaw, R. J. (2009). Tumor suppression by LKB1: SIK-ness prevents metastasis. *Sci Signal* 2, pe55.
- Shay, J. W., and Wright, W. E. (2000). Hayflick, his limit, and cellular ageing. *Nat Rev Mol Cell Biol* 1, 72-76.
- Shen, R., Pan, S., Qi, S., Lin, X., and Cheng, S. (2010). Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 in gastric cancer. *Biochem Biophys Res Commun* 394, 1047-1052.
- Shen, X., Liu, Y., Hsu, Y. J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G. C., and Orkin, S. H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* 32, 491-502.
- Shenouda, S. K., and Alahari, S. K. (2009). MicroRNA function in cancer: oncogene or a tumor suppressor? *Cancer Metastasis Rev* 28, 369-378.
- Sherr, C. J., and McCormick, F. (2002). The RB and p53 pathways in cancer. *Cancer Cell* 2, 103-112.
- Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., and Casero, R. A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* 119, 941-953.
- Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* 113, 685-700.
- Shi, Y., Sawada, J., Sui, G., Affar el, B., Whetstine, J. R., Lan, F., Ogawa, H., Luke, M. P., and Nakatani, Y. (2003). Coordinated histone modifications mediated by a CtBP co-repressor complex. *Nature* 422, 735-738.
- Shie, J. L., Chen, Z. Y., O'Brien, M. J., Pestell, R. G., Lee, M. E., and Tseng, C. C. (2000). Role of gut-enriched Kruppel-like factor in colonic cell growth and differentiation. *Am J Physiol Gastrointest Liver Physiol* 279, G806-814.
- Shields, J. D., Kourtis, I. C., Tomei, A. A., Roberts, J. M., and Swartz, M. A. (2010). Induction of lymphoidlike stroma and immune escape by tumors that express the chemokine CCL21. *Science* 328, 749-752.
- Shields, J. M., Christy, R. J., and Yang, V. W. (1996). Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. *J Biol Chem* 271, 20009-20017.
- Shin, M. S., Fredrickson, T. N., Hartley, J. W., Suzuki, T., Akagi, K., and Morse, H. C., 3rd (2004). High-throughput retroviral tagging for identification of genes involved in initiation and progression of mouse splenic marginal zone lymphomas. *Cancer Res* 64, 4419-4427.
- Shipitsin, M., Campbell, L. L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., *et al.* (2007). Molecular definition of breast tumor heterogeneity. *Cancer Cell* 11, 259-273.
- Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844-847.
- Siegel, P. M., and Massague, J. (2003). Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer* 3, 807-821.
- Silye, R., Karayiannakis, A. J., Syrigos, K. N., Poole, S., van Noorden, S., Batchelor, W., Regele, H., Sega, W., Boesmueller, H., Krausz, T., and Pignatelli, M. (1998). E-cadherin/catenin complex in benign and malignant melanocytic lesions. *J Pathol* 186, 350-355.
- Simian, M., Hirai, Y., Navre, M., Werb, Z., Lochter, A., and Bissell, M. J. (2001). The interplay of matrix metalloproteinases, morphogens and growth factors is necessary for branching of mammary epithelial cells. *Development* 128, 3117-3131.
- Singh, R. K., Gutman, M., Bucana, C. D., Sanchez, R., Llansa, N., and Fidler, I. J. (1995). Interferons alpha and beta down-regulate the expression of basic fibroblast growth factor in human carcinomas. *Proc Natl Acad Sci U S A* 92, 4562-4566.

- Sinner, D., Kordich, J. J., Spence, J. R., Opoka, R., Rankin, S., Lin, S. C., Jonatan, D., Zorn, A. M., and Wells, J. M. (2007). Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. *Mol Cell Biol* 27, 7802-7815.
- Song, L. B., Li, J., Liao, W. T., Feng, Y., Yu, C. P., Hu, L. J., Kong, Q. L., Xu, L. H., Zhang, X., Liu, W. L., *et al.* (2009). The polycomb group protein Bmi-1 represses the tumor suppressor PTEN and induces epithelial-mesenchymal transition in human nasopharyngeal epithelial cells. *J Clin Invest* 119, 3626-3636.
- Sood, A. K., Seftor, E. A., Fletcher, M. S., Gardner, L. M., Heidger, P. M., Buller, R. E., Seftor, R. E., and Hendrix, M. J. (2001). Molecular determinants of ovarian cancer plasticity. *Am J Pathol* 158, 1279-1288.
- Spaderna, S., Schmalhofer, O., Wahlbuhl, M., Dimmler, A., Bauer, K., Sultan, A., Hlubek, F., Jung, A., Strand, D., Eger, A., *et al.* (2008). The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer. *Cancer Res* 68, 537-544.
- Sparmann, A., and van Lohuizen, M. (2006a). Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* 6, 846-856.
- Sparmann, A., and van Lohuizen, M. (2006b). Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer* 6, 846-856.
- Sperr, W. R., Florian, S., Hauswirth, A. W., and Valent, P. (2005). CD 33 as a target of therapy in acute myeloid leukemia: current status and future perspectives. *Leuk Lymphoma* 46, 1115-1120.
- Spivakov, M., and Fisher, A. G. (2007). Epigenetic signatures of stem-cell identity. *Nat Rev Genet* 8, 263-271.
- St Croix, B., Sheehan, C., Rak, J. W., Florenes, V. A., Slingerland, J. M., and Kerbel, R. S. (1998). E-Cadherin-dependent growth suppression is mediated by the cyclin-dependent kinase inhibitor p27(KIP1). *J Cell Biol* 142, 557-571.
- Stehbens, S. J., Paterson, A. D., Crampton, M. S., Shewan, A. M., Ferguson, C., Akhmanova, A., Parton, R. G., and Yap, A. S. (2006). Dynamic microtubules regulate the local concentration of E-cadherin at cell-cell contacts. *J Cell Sci* 119, 1801-1811.
- Stojic, L., Jasencakova, Z., Prezioso, C., Stutzer, A., Bodega, B., Pasini, D., Klingberg, R., Mozzetta, C., Margueron, R., Puri, P. L., *et al.* (2011). Chromatin regulated interchange between polycomb repressive complex 2 (PRC2)-Ezh2 and PRC2-Ezh1 complexes controls myogenin activation in skeletal muscle cells. *Epigenetics Chromatin* 4, 16.
- Strutz, F., Zeisberg, M., Ziyadeh, F. N., Yang, C. Q., Kalluri, R., Muller, G. A., and Neilson, E. G. (2002). Role of basic fibroblast growth factor-2 in epithelial-mesenchymal transformation. *Kidney Int* 61, 1714-1728.
- Su, I. H., Basavaraj, A., Krutchinsky, A. N., Hobert, O., Ullrich, A., Chait, B. T., and Tarakhovskiy, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol* 4, 124-131.
- Suganuma, T., and Workman, J. L. (2011). Signals and combinatorial functions of histone modifications. *Annu Rev Biochem* 80, 473-499.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.
- Talmadge, J. E., and Fidler, I. J. (2010). AACR centennial series: the biology of cancer metastasis: historical perspective. *Cancer Res* 70, 5649-5669.
- Tan, J., Yang, X., Zhuang, L., Jiang, X., Chen, W., Lee, P. L., Karuturi, R. K., Tan, P. B., Liu, E. T., and Yu, Q. (2007). Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev* 21, 1050-1063.
- Taunton, J., Hassig, C. A., and Schreiber, S. L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* 272, 408-411.
- Tavazoie, S. F., Alarcon, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P. D., Gerald, W. L., and Massague, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* 451, 147-152.
- Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007). How chromatin-

- binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* *14*, 1025-1040.
- Teicher, B. A. (2001). Malignant cells, directors of the malignant process: role of transforming growth factor-beta. *Cancer Metastasis Rev* *20*, 133-143.
- Teng, M. W., Swann, J. B., Koebel, C. M., Schreiber, R. D., and Smyth, M. J. (2008). Immune-mediated dormancy: an equilibrium with cancer. *J Leukoc Biol* *84*, 988-993.
- Terranova, R., Yokobayashi, S., Stadler, M. B., Otte, A. P., van Lohuizen, M., Orkin, S. H., and Peters, A. H. (2008). Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. *Dev Cell* *15*, 668-679.
- Thiery, J. P. (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* *2*, 442-454.
- Thiery, J. P., and Morgan, M. (2004). Breast cancer progression with a Twist. *Nat Med* *10*, 777-778.
- Thiery, J. P., and Sleeman, J. P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol* *7*, 131-142.
- Thomas, D. A., and Massague, J. (2005). TGF-beta directly targets cytotoxic T cell functions during tumor evasion of immune surveillance. *Cancer Cell* *8*, 369-380.
- Thomson, S., Mahadevan, L. C., and Clayton, A. L. (1999). MAP kinase-mediated signalling to nucleosomes and immediate-early gene induction. *Semin Cell Dev Biol* *10*, 205-214.
- Thuault, S., Valcourt, U., Petersen, M., Manfioletti, G., Heldin, C. H., and Moustakas, A. (2006). Transforming growth factor-beta employs HMGA2 to elicit epithelial-mesenchymal transition. *J Cell Biol* *174*, 175-183.
- Tiwari, V. K., McGarvey, K. M., Licchesi, J. D., Ohm, J. E., Herman, J. G., Schubeler, D., and Baylin, S. B. (2008). PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol* *6*, 2911-2927.
- Tobin, N. P., Sims, A. H., Lundgren, K. L., Lehn, S., and Landberg, G. (2011). Cyclin D1, Id1 and EMT in Breast cancer. *BMC Cancer* *11*, 417.
- Ton-That, H., Kaestner, K. H., Shields, J. M., Mahatanakoon, C. S., and Yang, V. W. (1997). Expression of the gut-enriched Kruppel-like factor gene during development and intestinal tumorigenesis. *FEBS Lett* *419*, 239-243.
- Tournier, C., Hess, P., Yang, D. D., Xu, J., Turner, T. K., Nimnual, A., Bar-Sagi, D., Jones, S. N., Flavell, R. A., and Davis, R. J. (2000). Requirement of JNK for stress-induced activation of the cytochrome c-mediated death pathway. *Science* *288*, 870-874.
- Trang, P., Weidhaas, J. B., and Slack, F. J. (2008). MicroRNAs as potential cancer therapeutics. *Oncogene* *27 Suppl 2*, S52-57.
- Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994). The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J* *13*, 3822-3831.
- Tse, J. C., and Kalluri, R. (2007). Mechanisms of metastasis: epithelial-to-mesenchymal transition and contribution of tumor microenvironment. *J Cell Biochem* *101*, 816-829.
- Tsimberidou, A. M., Giles, F. J., Estey, E., O'Brien, S., Keating, M. J., and Kantarjian, H. M. (2006). The role of gemtuzumab ozogamicin in acute leukaemia therapy. *Br J Haematol* *132*, 398-409.
- Tsukita, S., Furuse, M., and Itoh, M. (1999). Structural and signalling molecules come together at tight junctions. *Curr Opin Cell Biol* *11*, 628-633.
- Ulloa, L., and Tabibzadeh, S. (2001). Lefty inhibits receptor-regulated Smad phosphorylation induced by the activated transforming growth factor-beta receptor. *J Biol Chem* *276*, 21397-21404.
- Vaissiere, T., Sawan, C., and Herceg, Z. (2008). Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutat Res* *659*, 40-48.
- Valastyan, S., Chang, A., Benaich, N., Reinhardt, F., and Weinberg, R. A. (2010). Concurrent suppression of integrin alpha5, radixin, and RhoA phenocopies the effects of miR-31 on metastasis. *Cancer Res* *70*, 5147-5154.

- van de Wetering, M., Oosterwegel, M., van Norren, K., and Clevers, H. (1993). Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes. *EMBO J* *12*, 3847-3854.
- van der Velden, J. L., Alcorn, J. F., Guala, A. S., Badura, E. C., and Janssen-Heininger, Y. M. (2010). c-Jun N-Terminal Kinase 1 Promotes TGF-1-induced Epithelial to Mesenchymal Transition via Control of Linker Phosphorylation and Transcriptional Activity of Smad3. *Am J Respir Cell Mol Biol*.
- van Leeuwen, F., Gafken, P. R., and Gottschling, D. E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. *Cell* *109*, 745-756.
- van Nimwegen, E. (2007). Finding regulatory elements and regulatory motifs: a general probabilistic framework. *BMC Bioinformatics* *8 Suppl 6*, S4.
- van Zijl, F., Mair, M., Csiszar, A., Schneller, D., Zulehner, G., Huber, H., Eferl, R., Beug, H., Dolznig, H., and Mikulits, W. (2009). Hepatic tumor-stroma crosstalk guides epithelial to mesenchymal transition at the tumor edge. *Oncogene* *28*, 4022-4033.
- Vander Heiden, M. G., Cantley, L. C., and Thompson, C. B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* *324*, 1029-1033.
- Vandewalle, C., Comijn, J., De Craene, B., Vermassen, P., Bruyneel, E., Andersen, H., Tulchinsky, E., Van Roy, F., and Berx, G. (2005). SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic Acids Res* *33*, 6566-6578.
- Varambally, S., Cao, Q., Mani, R. S., Shankar, S., Wang, X., Ateeq, B., Laxman, B., Cao, X., Jing, X., Ramnarayanan, K., *et al.* (2008). Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. *Science* *322*, 1695-1699.
- Varambally, S., Dhanasekaran, S. M., Zhou, M., Barrette, T. R., Kumar-Sinha, C., Sanda, M. G., Ghosh, D., Pienta, K. J., Sewalt, R. G., Otte, A. P., *et al.* (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* *419*, 624-629.
- Vazquez-Martin, A., Oliveras-Ferraros, C., Cufi, S., Martin-Castillo, B., and Menendez, J. A. (2010). Metformin and energy metabolism in breast cancer: from insulin physiology to tumour-initiating stem cells. *Curr Mol Med* *10*, 674-691.
- Vazquez-Martin, A., Oliveras-Ferraros, C., Del Barco, S., Martin-Castillo, B., and Menendez, J. A. (2011). The anti-diabetic drug metformin suppresses self-renewal and proliferation of trastuzumab-resistant tumor-initiating breast cancer stem cells. *Breast Cancer Res Treat* *126*, 355-364.
- Velden, J. L., Alcorn, J. F., Guala, A. S., Badura, E. C., and Janssen-Heininger, Y. M. (2011). c-Jun N-terminal kinase 1 promotes transforming growth factor-beta1-induced epithelial-to-mesenchymal transition via control of linker phosphorylation and transcriptional activity of Smad3. *Am J Respir Cell Mol Biol* *44*, 571-581.
- Venables, J. P., Klinck, R., Bramard, A., Inkel, L., Dufresne-Martin, G., Koh, C., Gervais-Bird, J., Lapointe, E., Froehlich, U., Durand, M., *et al.* (2008). Identification of alternative splicing markers for breast cancer. *Cancer Res* *68*, 9525-9531.
- Vesuna, F., Lisok, A., Kimble, B., and Raman, V. (2009). Twist modulates breast cancer stem cells by transcriptional regulation of CD24 expression. *Neoplasia* *11*, 1318-1328.
- Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane-Robinson, C., Allis, C. D., and Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. *EMBO J* *15*, 2508-2518.
- Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M., *et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* *439*, 871-874.
- Vlaminckx, K., Vakaet, L., Jr., Mareel, M., Fiers, W., and van Roy, F. (1991). Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. *Cell* *66*, 107-119.
- Vogel, M. J., Guelen, L., de Wit, E., Peric-Hupkes, D., Loden, M., Talhout, W., Feenstra, M., Abbas, B., Classen, A. K., and van Steensel, B. (2006). Human heterochromatin proteins

- form large domains containing KRAB-ZNF genes. *Genome Res* 16, 1493-1504.
- Wade, P. A. (2001). Methyl CpG-binding proteins and transcriptional repression. *Bioessays* 23, 1131-1137.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., and Zhang, Y. (2004a). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873-878.
- Wang, L., Brown, J. L., Cao, R., Zhang, Y., Kassis, J. A., and Jones, R. S. (2004b). Hierarchical recruitment of polycomb group silencing complexes. *Mol Cell* 14, 637-646.
- Wang, S., and Olson, E. N. (2009). AngiomiRs--key regulators of angiogenesis. *Curr Opin Genet Dev* 19, 205-211.
- Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., and Zhao, K. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40, 897-903.
- Warzecha, C. C., Jiang, P., Amirikian, K., Dittmar, K. A., Lu, H., Shen, S., Guo, W., Xing, Y., and Carstens, R. P. (2010). An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. *EMBO J* 29, 3286-3300.
- Warzecha, C. C., Sato, T. K., Nabet, B., Hogenesch, J. B., and Carstens, R. P. (2009). ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. *Mol Cell* 33, 591-601.
- Wei, D., Gong, W., Kanai, M., Schlunk, C., Wang, L., Yao, J. C., Wu, T. T., Huang, S., and Xie, K. (2005). Drastic down-regulation of Kruppel-like factor 4 expression is critical in human gastric cancer development and progression. *Cancer Res* 65, 2746-2754.
- Weidhaas, J. B., Babar, I., Nallur, S. M., Trang, P., Roush, S., Boehm, M., Gillespie, E., and Slack, F. J. (2007). MicroRNAs as potential agents to alter resistance to cytotoxic anticancer therapy. *Cancer Res* 67, 11111-11116.
- Wellner, U., Schubert, J., Burk, U. C., Schmalhofer, O., Zhu, F., Sonntag, A., Waldvogel, B., Vannier, C., Darling, D., zur Hausen, A., *et al.* (2009). The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat Cell Biol* 11, 1487-1495.
- Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318-324.
- Wheelock, M. J., Shintani, Y., Maeda, M., Fukumoto, Y., and Johnson, K. R. (2008). Cadherin switching. *J Cell Sci* 121, 727-735.
- Whitcomb, S. J., Basu, A., Allis, C. D., and Bernstein, E. (2007). Polycomb Group proteins: an evolutionary perspective. *Trends Genet* 23, 494-502.
- Whiteman, E. L., Liu, C. J., Fearon, E. R., and Margolis, B. (2008). The transcription factor snail represses Crumbs3 expression and disrupts apico-basal polarity complexes. *Oncogene* 27, 3875-3879.
- Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D. J., Campan, M., Young, J., Jacobs, I., and Laird, P. W. (2007). Epigenetic stem cell signature in cancer. *Nat Genet* 39, 157-158.
- Wildenberg, G. A., Dohn, M. R., Carnahan, R. H., Davis, M. A., Lobdell, N. A., Settleman, J., and Reynolds, A. B. (2006). p120-catenin and p190RhoGAP regulate cell-cell adhesion by coordinating antagonism between Rac and Rho. *Cell* 127, 1027-1039.
- Wilkes, M. C., Murphy, S. J., Garamszegi, N., and Leof, E. B. (2003). Cell-type-specific activation of PAK2 by transforming growth factor beta independent of Smad2 and Smad3. *Mol Cell Biol* 23, 8878-8889.
- Willis, S. N., and Adams, J. M. (2005). Life in the balance: how BH3-only proteins induce apoptosis. *Curr Opin Cell Biol* 17, 617-625.
- Wilson, M. E., Yang, K. Y., Kalousova, A., Lau, J., Kosaka, Y., Lynn, F. C., Wang, J., Mrejen, C., Episkopou, V., Clevers, H. C., and German, M. S. (2005). The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas. *Diabetes* 54, 3402-3409.
- Wirbelauer, C., Bell, O., and Schubeler, D. (2005). Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone

- modifications show a promoter-proximal bias. *Genes Dev* *19*, 1761-1766.
- Wolffe, A. P., and Hayes, J. J. (1999). Chromatin disruption and modification. *Nucleic Acids Res* *27*, 711-720.
- Wysocka, J., Swigut, T., Xiao, H., Milne, T. A., Kwon, S. Y., Landry, J., Kauer, M., Tackett, A. J., Chait, B. T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* *442*, 86-90.
- Xia, N., Thodeti, C. K., Hunt, T. P., Xu, Q., Ho, M., Whitesides, G. M., Westervelt, R., and Ingber, D. E. (2008). Directional control of cell motility through focal adhesion positioning and spatial control of Rac activation. *FASEB J* *22*, 1649-1659.
- Xue, F., Takahara, T., Yata, Y., Xia, Q., Nonome, K., Shinno, E., Kanayama, M., Takahara, S., and Sugiyama, T. (2008). Blockade of Rho/Rho-associated coiled coil-forming kinase signaling can prevent progression of hepatocellular carcinoma in matrix metalloproteinase-dependent manner. *Hepatology* *38*, 810-817.
- Yamamoto, Y., Verma, U. N., Prajapati, S., Kwak, Y. T., and Gaynor, R. B. (2003). Histone H3 phosphorylation by IKK-alpha is critical for cytokine-induced gene expression. *Nature* *423*, 655-659.
- Yamashita, M., Fatyol, K., Jin, C., Wang, X., Liu, Z., and Zhang, Y. E. (2008). TRAF6 mediates Smad-independent activation of JNK and p38 by TGF-beta. *Mol Cell* *31*, 918-924.
- Yanagisawa, M., Huveldt, D., Kreinest, P., Lohse, C. M., Cheville, J. C., Parker, A. S., Copland, J. A., and Anastasiadis, P. Z. (2008). A p120 catenin isoform switch affects Rho activity, induces tumor cell invasion, and predicts metastatic disease. *J Biol Chem* *283*, 18344-18354.
- Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R. A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* *117*, 927-939.
- Yang, J., and Weinberg, R. A. (2008). Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell* *14*, 818-829.
- Yang, L., Lin, C., and Liu, Z. R. (2006). P68 RNA helicase mediates PDGF-induced epithelial mesenchymal transition by displacing Axin from beta-catenin. *Cell* *127*, 139-155.
- Yang, L., Pang, Y., and Moses, H. L. (2010a). TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol* *31*, 220-227.
- Yang, M. H., Hsu, D. S., Wang, H. W., Wang, H. J., Lan, H. Y., Yang, W. H., Huang, C. H., Kao, S. Y., Tzeng, C. H., Tai, S. K., *et al.* (2010b). Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition. *Nat Cell Biol* *12*, 982-992.
- Yang, X., Lay, F., Han, H., and Jones, P. A. (2010c). Targeting DNA methylation for epigenetic therapy. *Trends Pharmacol Sci* *31*, 536-546.
- Yang, X., Pursell, B., Lu, S., Chang, T. K., and Mercurio, A. M. (2009). Regulation of beta 4-integrin expression by epigenetic modifications in the mammary gland and during the epithelial-to-mesenchymal transition. *J Cell Sci* *122*, 2473-2480.
- Yilmaz, M., and Christofori, G. (2010). Mechanisms of motility in metastasizing cells. *Mol Cancer Res* *8*, 629-642.
- Yilmaz, M., Christofori, G., and Lehenbre, F. (2007). Distinct mechanisms of tumor invasion and metastasis. *Trends Mol Med* *13*, 535-541.
- Yori, J. L., Johnson, E., Zhou, G., Jain, M. K., and Keri, R. A. (2010). Kruppel-like factor 4 inhibits epithelial-to-mesenchymal transition through regulation of E-cadherin gene expression. *The Journal of biological chemistry* *285*, 16854-16863.
- Yoshida, C., and Takeichi, M. (1982). Teratocarcinoma cell adhesion: identification of a cell-surface protein involved in calcium-dependent cell aggregation. *Cell* *28*, 217-224.
- Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T., and Hirohashi, S. (1995). Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proc Natl Acad Sci U S A* *92*, 7416-7419.
- Yu, J., Cao, Q., Mehra, R., Laxman, B., Tomlins, S. A., Creighton, C. J., Dhanasekaran,

- S. M., Shen, R., Chen, G., Morris, D. S., *et al.* (2007). Integrative genomics analysis reveals silencing of beta-adrenergic signaling by polycomb in prostate cancer. *Cancer Cell* *12*, 419-431.
- Yu, L., Hebert, M. C., and Zhang, Y. E. (2002). TGF-beta receptor-activated p38 MAP kinase mediates Smad-independent TGF-beta responses. *EMBO J* *21*, 3749-3759.
- Zabouo, G., Imbert, A. M., Jacquemier, J., Finetti, P., Moreau, T., Esterni, B., Birnbaum, D., Bertucci, F., and Chabannon, C. (2009). CD146 expression is associated with a poor prognosis in human breast tumors and with enhanced motility in breast cancer cell lines. *Breast Cancer Res* *11*, R1.
- Zamir, E., and Geiger, B. (2001). Molecular complexity and dynamics of cell-matrix adhesions. *J Cell Sci* *114*, 3583-3590.
- Zavadil, J., Bitzer, M., Liang, D., Yang, Y. C., Massimi, A., Kneitz, S., Piek, E., and Bottinger, E. P. (2001). Genetic programs of epithelial cell plasticity directed by transforming growth factor-beta. *Proc Natl Acad Sci U S A* *98*, 6686-6691.
- Zavadil, J., and Bottinger, E. P. (2005). TGF-beta and epithelial-to-mesenchymal transitions. *Oncogene* *24*, 5764-5774.
- Zhang, S., Schafer-Hales, K., Khuri, F. R., Zhou, W., Vertino, P. M., and Marcus, A. I. (2008). The tumor suppressor LKB1 regulates lung cancer cell polarity by mediating cdc42 recruitment and activity. *Cancer Res* *68*, 740-748.
- Zhang, W., Deng, H., Bao, X., Lerach, S., Girton, J., Johansen, J., and Johansen, K. M. (2006). The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in *Drosophila*. *Development* *133*, 229-235.
- Zhang, X., Liu, S., Hu, T., He, Y., and Sun, S. (2009). Up-regulated microRNA-143 transcribed by nuclear factor kappa B enhances hepatocarcinoma metastasis by repressing fibronectin expression. *Hepatology* *50*, 490-499.
- Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev* *15*, 2343-2360.
- Zhang, Y. E. (2009). Non-Smad pathways in TGF-beta signaling. *Cell Res* *19*, 128-139.
- Zhang, Z., Vuori, K., Wang, H., Reed, J. C., and Ruoslahti, E. (1996). Integrin activation by R-ras. *Cell* *85*, 61-69.
- Zhao, W., Hisamuddin, I. M., Nandan, M. O., Babbin, B. A., Lamb, N. E., and Yang, V. W. (2004). Identification of Kruppel-like factor 4 as a potential tumor suppressor gene in colorectal cancer. *Oncogene* *23*, 395-402.
- Zippo, A., De Robertis, A., Serafini, R., and Oliviero, S. (2007). PIM1-dependent phosphorylation of histone H3 at serine 10 is required for MYC-dependent transcriptional activation and oncogenic transformation. *Nat Cell Biol* *9*, 932-944.

## 5.0 Curriculum vitae

### 5.1 Personal information

**Name:** Neha Tiwari  
**Date of Birth:** 09 September 1982  
**Place of Birth:** Faizabad, Uttar Pradesh, India  
**Nationality:** Indian  
**Status:** Married, One child  
**Address:** Rosentalstrasse 28  
CH-4058 Basel

### 5.2 Experience and Education

#### 5.2.1 Experience

Institute of Biochemistry and Genetics, Basel, Switzerland  
University of Basel

***PhD Student*** **February 2008**

Department of Genetics and Pathology, Uppsala, Sweden  
Uppsala University

***Guest Researcher*** **August 2007 to December 2007**

The Sidney Kimmel Comprehensive Cancer Centre, Baltimore, USA  
Johns Hopkins University School of Medicine,

***Research Technologist*** **July 2006 to August 2007**

Department of Biochemistry, Faizabad, India  
Dr.R.M.L. Avadh University

***Technical Assistant*** **November 2004 to September 2005**

#### 5.2.2 Education

Department of Biochemistry, Faizabad, India



Dr. R.M.L. Avadh University

*MSc (Masters of Science)*

**2002-2004**

K.S. Saket PG College,

Faizabad, India

Dr. R.M.L. Avadh University

*BSc (Bachelors of Science)*

**1999-2002**

Central School,

New Delhi, India

Central Board of Secondary Education

*Pre-degree*

**1997-1999**

Central School,

New Delhi, India

Central Board of Secondary Education

*SSC (High School)*

**1995-1997**

### **5.3 Workshops, Courses and Conferences**

#### **5.3.1 Workshops**

- "2009 EMBO Molecular Medicine workshop-Invasive Growth: a Genetic Programme for Stem cells and Cancer"- **Congress Center, Torino, Italy**
- SGV Meeting for mouse Experiments, **University of Basel, Switzerland**
- Workshop on "Parallel Sequencing", **D-BSSE, Basel, Switzerland**
- LTK module 1 for animal experiment, **Friedrich Miescher Institute, Basel, Switzerland**
- Workshop on "preparation of cDNA library from single cells"- **Cambridge University, UK**
- Workshop on "Introduction to Cell Culture"- **Johns Hopkins University, USA**
- Workshop on "Introduction to Protein Expression"-**Johns Hopkins University, USA**
- Workshop on "Advanced lab techniques", Cytogenetics Laboratory, **Banaras Hindu**

**University, India**

- Workshop on “National Young Scientist Workshop on Environmental Microbiology and Parasitology”- **Allahabad University, India**

**5.3.2 Courses**

- English: Academic Writing workshop at **University of Basel, Basel, Switzerland, Jun 2011**
- R Course at **Friedrich Miescher Institute, Basel, Switzerland, Feb 2011**
- Genetic approaches in biomedical research at **University of Basel, Basel, Switzerland, Jan 2010**
- Translation control and Post-translation protein modification at **University of Basel, Basel, Switzerland, Jan 2010**
- Cell interactions in Development and Disease at **University of Basel, Basel, Switzerland, Jan 2009**
- Structure, Processing and Function of RNA at **University of Basel, Basel, Switzerland, Mar 2009**
- Bio-informatics course at **Friedrich Miescher Institute, Basel, Switzerland, May 2009**
- Experimentelle Krebsforschung I: Biologie und Molekularbiologie at **University of Basel, Basel, Switzerland, Aug 2008**
- Transcription, Regulation and Gene expression in Eukaryotes at **University of Basel, Basel, Switzerland, Jun 2008**
- Research Compliance Training Courses (includes Human Subjects Research, Animal Care and Use, Conflict of Interest and Commitment), at **Johns Hopkins University, Baltimore, USA, Dec 2006**
- HIPAA Compliance Training Course at **Johns Hopkins University, Baltimore, USA, Dec 2006**
- Blood Borne Pathogen Course at **Johns Hopkins University School of Medicine, Baltimore, USA, Dec 2006**

### 5.3.2 Conferences

- "1st International SystemsX.ch Conference on Systems Biology ", **Congress Center, Basel, Switzerland, Oct 2011**
- "Biovalley life Science day" and "Miptec", **Congress Center, Basel, Switzerland, Sep 2011**
- EuCC (European Cancer Center) Symposium, **Restaurant Seegarten, Basel, Switzerland, May 2011**
- "Biovalley life Science day" and "Miptec", **Congress Center, Basel, Switzerland, Sep 2010**
- “Swiss Stem Cell Network Meeting”, **Biocenter, Basel, Switzerland, Feb 2010**
- “Multiple Sclerosis Meeting”, **ZLF, Basel, Switzerland, Nov 2009**
- "Biovalley life Science day" and "Miptec", **Congress Center, Basel, Switzerland, Oct 2009**
- "MicroEnviMet symposium"- **Univeristy of Torino, Torino, Italy, Sep 2009**
- “D-BSSE System Biology symposium, **Basel, Switzerland, Apr 2009**
- TUMIC (An integrated concept of tumor metastasis) symposium- **Univeristy of Basel, Basel, Switzerland, Jan 2009**
- "Biovalley life Science day" and "Miptec", **Congress Center, Basel, Switzerland, Oct 2008**
- EuCC (European Cancer Center) Symposium, **Hotel Dorint, Basel, Switzerland, May 2008**

### 5.4 **Achievements**

- Poster presentation on Biovalley life Science day 2011. The title of the poster was “Sox4: a key modulator of EMT”.  
**Neha Tiwari**, Erik van Nimwegen and Gerhard Christofori
- Poster presentation on Biovalley life Science day 2009. The title of the poster was “KLF4 and SOX4 as regulator of EMT”.  
**Neha Tiwari**, Mahmut yilmaz, Lorenz Waldeimer, Dorothea Maass, Chantal

Achermann, Lukas Manhart, Phil Arnold, Eric Van Nimwegen and Gerhard Christofori  
**(Best poster award)**

- Poster presentation in EMBO Molecular Medicine Meeting 2009. The title of the poster was "KLF4 and SOX4 as regulator of EMT".

**Neha Tiwari**, Mahmut yilmaz, Lorenz Waldeimer, Dorothea Maass, Chantal Achermann, Lukas Manhart, Phil Arnold, Eric Van Nimwegen and Gerhard Christofori

- Poster presentation in IMPaCT (Innovation Minds in Prostate Cancer) Meeting 2007. The title of the poster was "Consequences of the inhibition of CDK5 in Prostate cancer Cells". Barry D Nelkin. Christopher J. Strock, **Neha Tiwari** and Micheal J. Ochs.

## 5.5 Publications

- Sprouty 4 expression prevents angiogenesis in Rip-Tag mouse model  
**Tiwari N**, Jaggi F, Antoniadis H, Christofori G (In Process).
- Polycomb-dependent Mechanisms Regulate Epithelial to Mesenchymal Transition  
**Tiwari N** #, Tiwari VK#, Arnold P, Balweirz P, Nimwegen E, Schübeler D, Christofori (In Process)  
 #-Authors contributed equally
- KLF4 directly regulates transcription of genes crucial for Epithelial to Mesenchymal Transition  
**Tiwari N**, Arnold P, Nimwegen E, Christofori G (In Process).
- Sox4 regulates epithelial to mesenchymal transition by directly controlling transcription of underlying master genes  
**Tiwari N**, Arnold P, Mikhail P, Schaller N, Christofori G (Submitted).
- Transcription factor Dlx2 protects from TGF $\beta$ -induced cell-cycle arrest and apoptosis  
 Yilmaz M#, Maaß D#, **Tiwari N**, Waldmeier L, Schmidt P, Lehembre F, Christofori G  
 EMBO J. 2011  
 #-Authors contributed equally
- A single-nucleotide deletion in the POMP 5' UTR causes a transcriptional switch and altered epidermal proteasome distribution in KLICK genodermatosis.  
 Dahlqvist J, Klar J, **Tiwari N**, Schuster J, Törmä H, Badhai J, Pujol R, van Steensel MA,

Brinkhuizen T, Gijzen L, Chaves A, Tadini G, Vahlquist A, Dahl N.  
Am J Hum Genet. 2010

## 6.0 Acknowledgements

The present thesis is the result of work carried out at Institute of Biomedicine, University of Basel. Support from a lot of people has directly or indirectly helped me to reach this stage of my life and I wish to express my gratitude towards them.

First and foremost I would like to express my sincere thanks to Prof. Dr. Gerhard Christofori, Director of the Department and my PhD supervisor, who provided me with the opportunity to work under his guidance. I appreciate all of his efforts, including his time spent with me and for his stimulating conceptual inputs that made my Ph.D. experience productive and stimulating. He has enlightened me with his wide knowledge and deep intuitions about where it should go and what is necessary to get there. The joy and enthusiasm he has for his research is contagious and this was and will always be a source of motivation for me. His very kind personality combined with an everlasting support for me made it very easy through difficult times.

I would also like to show my gratitude to the members of the Christofori's group who contributed immensely to my personal and professional life. The group has been a source of many friendships as well as good advice and collaborations. I will miss you all a lot! I wish you a life full of happiness and pray that you all become stars in the sky of science. I convey my special thanks to Anna and Ernesta for their efforts during the final preparation of this thesis.

I also owe my deepest gratitude to Roland, Angelika, Helena, Ursula, Isabel, Petra and Pascal, without them this thesis would not have been possible. I would also like to say thanks to Joerg whose support and care for me will be memorable.

I would like to thank my parents-in-law, parents and other members of the family for their love and encouragement, whose support and continuous trust in my capabilities was inspiration to my progress. I owe most to my other half, my husband Vijay, for being THE BEST, for his love, care, guidance, support, and encouragement. You are not only my husband, but my best friend and soul mate. I am very much indebted to my sweetheart and my soul, Tejas, for his unconditional love and giving me the feelings that cannot be described in words. Since you

arrived I realized what I had missed most in my life so far. Mamma loves you Tejas and you have the sweetest smile in the whole world!