Transcriptional and post-transcriptional mechanisms regulating epithelial to mesenchymal transition in breast cancer

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Maren Diepenbruck

aus Gummersbach, Deutschland

Basel, 2015

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von:

Prof. Dr. Gerhard Christofori

Prof. Dr. Mohamed Bentires-Alj

Basel, den 23. Juni 2015

Prof. Dr. Jörg Schibler
Dekan

Summary

The world health organization (WHO) declares cancer as one of the leading causes of mortality worldwide and reported 14 million new cases and 8.2 million cancer-associated deaths in 2012. The term cancer summarizes a broad spectrum of diseases reflecting the common feature of uncontrolled cell proliferation and systemic dissemination of tumor cells. Systemic dissemination of cancer cells requires in principle the invasion of tumor cells into the body's circulation and their outgrowth at a distant site. In breast cancer, which is one of the top five diagnosed cancers among women, as in most cancer types metastatic outgrowth is the leading cause of death.

Epithelial to mesenchymal transition (EMT) is an essential developmental process and comprises the gradual remodeling of epithelial cell architecture and functional capabilities. More precisely, cells lose epithelial cell characteristics like strong cell-cell junctions and an apical-basal cell polarity, which retain cells in a functional epithelial layer. During EMT, cells convert to a low proliferation state and acquire a spindle-like cell shape enabling single cell migration, invasion and increased cell survival. The aberrant activation of EMT promotes breast tumor cell invasion and dissemination, furthermore, its reverse process, mesenchymal to epithelial transition (MET), is believed to support metastatic outgrowth. Hence, we need to better understand the underlying molecular mechanisms controlling the dynamic nature of cell (de)differentiation and its consequences during malignant tumor progression.

In the past years, intensive research has demonstrated that EMT/MET plasticity and its functional implications can be orchestrated by interconnected molecular networks consisting of transcription factors, epigenetic regulators, splicing factors and non-coding RNAs, which can be activated by a plethora of extracellular signals. However, we are just at the beginning to understand the role and regulation of such factors during EMT. Therefore, during my studies I aimed to identify critical players, in particular transcription factors and miRNAs implicated and conserved during normal and cancer-associated cell dedifferentiation and characterized their contribution to cancer progression *in vitro* and *in vivo*.

We established different *in vitro* EMT systems to examine the stepwise morphological transition of epithelial mouse mammary cells by transforming growth

factor β (TGFβ), a potent EMT inducing cytokine. Subsequent global gene expression profiling of various cell dedifferentiation states allowed us to monitor the transcriptomic alterations in a time-resolved manner. In combination with a bioinformatic analysis for DNA-binding motifs, we identified the transcription factor Tead2 as a potential EMT regulator. Tead2 is a transcriptional effector of the Hippo pathway, which tightly controls cell proliferation and organ growth. Upon EMT induction, the nuclear levels of Tead2 increase, which upon direct binding induces a predominantly nuclear localization of its cofactors Yap and Taz. Furthermsore, Tead2 is required during EMT and promotes tumor cell migration, invasion and lung colonization *in vivo*. Genome-wide chromatin immunoprecipitation/next generation sequencing in combination with gene expression profiling revealed the direct transcriptional targets of Tead2 during EMT in epithelial tumor cells. Among other EMT-relevant genes, we identified Zyxin an Actin remodeling and focal adhesion component important for Tead2-induced cell migration and invasion.

Aside from transcriptional control non-coding RNAs can regulate EMT/MET processes. Analyzing global transcriptomic alterations of different cell dedifferentiation states by deep sequencing analysis, we identified a pool of strongly differentially regulated miRNAs. In a combination of screens, we tested their functionality during EMT and mesenchymal tumor cell migration and identified miR-1199-5p as a novel EMT-regulatory miRNA. MiR-1199-5p is transcriptionally downregulated during EMT, and forced expression of miR-1199-5p prevented TGFβ-induced EMT and decreased mesenchymal mammary tumor cell migration and invasion. Furthermore, we report a new double-negative feedback regulation between miR-1199-5p and the EMT transcription factor Zeb1, exemplifying the close interconnections of transcriptional and post-transcriptional networks facilitating epithelial plasticity. In summary, both studies provided new insights into the molecular mechanisms orchestrating EMT and its functional consequences.

Table of contents

Summary

Table of contents

1 General Introduction

1.1 Epithelial to mesenchymal transition

Almost 50 years ago, Elizabeth Hay described the concept of epithelial to mesenchymal transition (EMT) for the first time. She observed dramatic morphological changes of epithelial cells within the chick primitive streak and proposed that these epithelial cells “transformed” into mesenchymal cells and would migrate and form mesodermal and endodermal epithelia during early vertebrate development [1]. Since this process is highly plastic and mesenchymal cells can convert back to an epithelial cell state by a mesenchymal to epithelial transition (MET), the term “transformation” has been replaced by “transition” [2].

1.1.1 Basics of EMT

Epithelial cells are positioned and tightly packed within epithelial layers and are characterized by intercellular junctions, an apical-basal cell polarity and their basal connection to a basement membrane in an epithelial monolayer. These cells display a low migratory potential and sometimes a high proliferation rate. Upon EMT various structural changes occur to remodel epithelial cell architecture and functional capabilities [3, 4] (Figure 1).

(a) Epithelial cell-cell junction disassembly: Epithelial cells are connected by tight and adhesion junctions, desmosomes and gap junctions. Tight junctions mainly function as impermeable barriers within an epithelium. Furthermore, they facilitate mechanical connections between epithelial neighboring cells, which are also reinforced by adhesion junctions and desmosomes. Gap junctions bridge the intercellular space and connect the cytoplasm of epithelial cells. They allow the exchange of small molecules and ions between cells and represent cell communication channels. All four cell-cell junctions are either delocalized and/or degraded during EMT [5]. Epithelial Cadherin (E-cadherin) belongs to the classical type of Cadherins and forms calcium-dependent, trans-membrane adhesion junctions with neighboring epithelial cells. Its cytoplasmic domain builds a multiprotein adhesion complex, which connects E-cadherin with the Actin cytoskeleton via

different Catenins (α, β, δ (p120)) [6]. E-cadherin is required for epithelial cell-cell adhesion and its loss is a hallmark of an EMT [7]. During an EMT, E-cadherin is replaced by neural Cadherin (N-cadherin) an event termed “Cadherin switch” which is important for cell dissemination, migration and invasion [4, 8]. Furthermore, E-cadherin delocalization from the membrane and degradation induces the release and thus nuclear accumulation of β-catenin and p120 where these proteins serve as transcriptional cofactors [9]. Similar to adhesion junctions, epithelial tight junctions are disassembled during EMT and a downregulation of Claudin and Occludin expression and diffuse cytoplasmic localization of ZO-1 (zonula occludens 1) have been described [5]. Desmosomes consist of Plakoglobin, Plakophilin, Desmocollin, Desmoglein and anchor and connect Keratin filaments between neighboring epithelial cells. EMT initiation leads to the disruption of desmosomes and transcriptional downregulation of desmosome proteins [10]

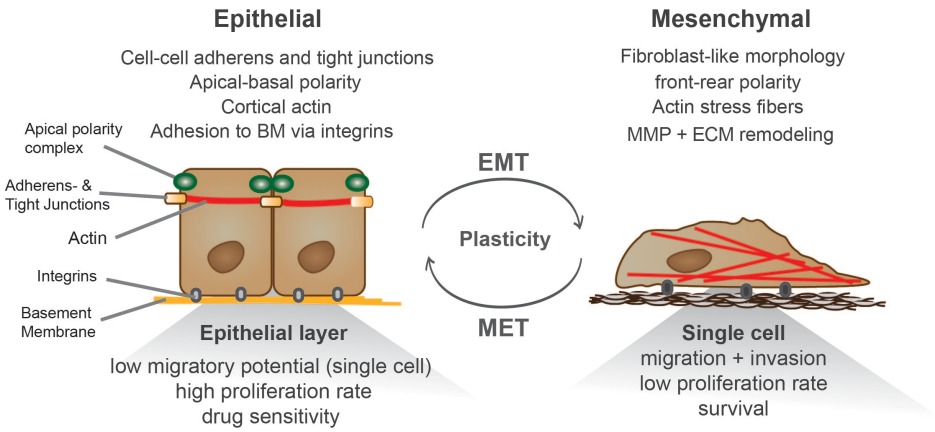


Figure 1: Schematic representation of structural and functional changes during EMT. Epithelial cells within an epithelium are connected via cell-cell junctions and via cell-matrix junctions to the basement membrane via Integrins. Epithelial cells display an apical-basal polarity and cortical actin organization. Upon EMT, cells acquire a fibroblast-like morphology with front-rear polarity and resolve strong cell-cell junctions. Cortical F-Actin is reorganized to Actin stress fibers. Secretion of MMPs and ECM components lead to the remodeling of the surrounding ECM to facilitate single cell migration and invasion, which is further promoted by the dynamic formation of cell-matrix adhesions through Integrins. Additionally, mesenchymal cells switch to a rather low proliferation rate, exhibit increased cell survival and resistance to apoptotic signals.

b) Loss of apical-basal cell polarity: The apical polarity complexes PAR (aPkc-Par3-Par6) and Crumbs (Crumbs-Pals1-Patj) and the basolateral complex Scribble (Scrib-Lgl-Dlg) are localized and connected to tight and adhesion junctions in epithelial cells. The loss of epithelial cell-cell junctions during EMT induces the

dissolution of an apical-basal polarity and promotes a front-rear directional polarity in mesenchymal cells [3].

c) ECM remodeling and cell-matrix interactions: Integrins are type I trans-membrane proteins, they connect the extracellular matrix (ECM) with the cell cytoskeleton and mediate cell adhesion and signal transduction [4]. Integrins form heterodimers composed of a non-covalently bound α and β subunit. In humans 24 types of alpha and 9 types of beta subunits exist and the combination of α/β subunits in heterodimers determines the specific binding to different ECM substrates. Epithelial cells express α6/β4 Integrins, which mediate the interaction with basement membrane proteins such as Laminins. However, the expression of α6/β4 is transcriptionally repressed during EMT and replaced by α5/β1 Integrins promoting dynamic focal adhesion assembly, cell migration and survival by selective binding for instance to Fibronectin [11, 12]. Fibronectin is an ECM glycoprotein highly secreted by mesenchymal cells to facilitate cell migration [13]. Matrix metalloproteinases (MMPs) like MMP2 and MMP9 are endopeptidases secreted by mesenchymal cells to degrade ECM proteins and to facilitate mesenchymal cell invasion [14]. In addition, MMP-mediated ECM degradation liberates/activates a broad spectrum of ECM-trapped growth factors and cytokines, for instance, transforming growth factor β (TGFβ), a potent inducer of EMT [15, 16].

d) Cytoskeleton rearrangements: Epithelial cells display cortical Actin filaments, anchored at cell-cell and cell-matrix junctions and thereby interconnect an epithelial cell layer with an Actin belt. Upon EMT, the Actin cytoskeleton is drastically rearranged to highly dynamic Actin stress fibers, which allows mesenchymal cell contractility and directional movements [2, 4, 9]. In particular, Rho GTPases (RhoA, Rac1, Cdc42) control Actin polymerization and rearrangements in mesenchymal cells. RhoA enforces the formation of Actin stress fibers and Rac1 and Cdc42 mediate the formation of lamellipodia and filopodia, respectively [17]. Apart from Actin cytoskeleton remodeling, the intermediate filament cytoskeleton composition also changes during EMT. Intermediate filaments (IFs) provide cells with mechanical strength and facilitate the trafficking of organelles and other proteins within cells. Upon cell dedifferentiation, epithelial Keratins are replaced by type III Vimentin IFs, which is crucial for mesenchymal cell shape, increased cell motility and focal adhesion dynamics [18].

These structural changes during EMT lead to an elongated cell shape going along with increased migratory and invasive capabilities. Furthermore, mesenchymal cells have been shown to display a lower proliferation rate, increased cell survival, resistance to apoptotic signals and stem cell-like properties [2]. Various extracellular stimuli can induce EMT or MET in a time and context-dependent manner (see 1.2). Networks of transcriptional, post-transcriptional and post-translational mechanisms seem to orchestrate such gradual changes in cell morphology and behavior (see 1.3).

1.1.2 EMT in biological contexts

Based on the biological context, EMT has been subcategorized in three different types: developmental EMT *(type I)*, EMT in wound healing and fibrosis *(type II)* and oncogenic EMT *(type III)* [19]. All three types display similar underlining molecular mechanisms causing cellular remodeling events as described above. However, they differ in their biological consequences [2, 20, 21].

Type I - Developmental EMT: Several rounds of cell dedifferentiation and differentiation are essential during embryogenesis to allow the formation of tissues and organs, thereby, illustrating the plasticity of epithelial cells. A primary EMT occurs during gastrulation, an event forming the three embryonic germ layers (Figure 2). Epithelial cells from the epiblast (a single epithelial layer) undergo EMT and ingress from the primitive streak in the interior of the embryo and generate the mesoderm and endoderm. Remaining epithelial cells in the epiblast form the ectoderm. Mesenchymal cells of the mesoderm undergo MET and form the notochord, somites, the primordium of the urogenital system, the splanchnopleure and the somatopleure. In a second round of EMT/MET processes, cells for instance from the dorsal half of the somite give rise to components of the dermis, muscle and satellite cells. Notably, the generation of the cardiac valves requires three repeats of cell dedifferentiation and subsequent differentiation events. In conclusion, a type I EMT goes hand in hand with its reversal process and aims to generate new tissues with distinct functions during development.

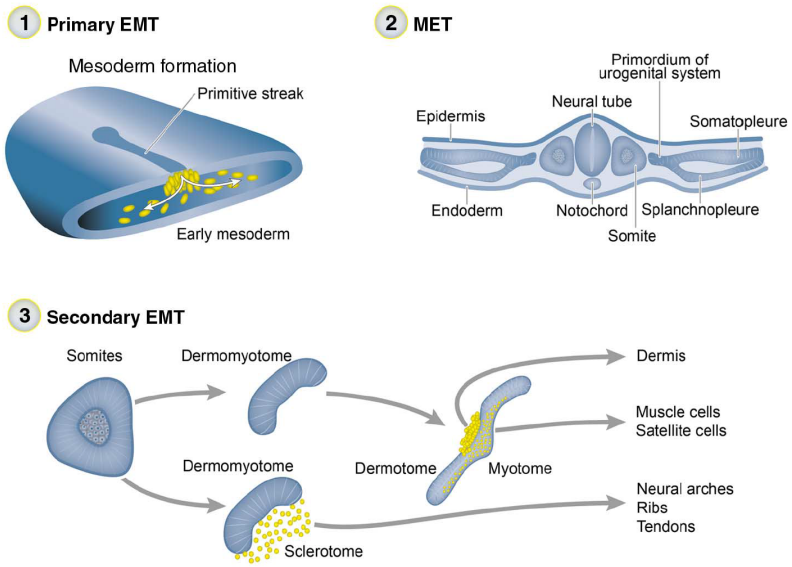


Figure 2: Reversible EMT processes during embryogenesis. (1) Primary EMT allows the migration of cells from the primitive streak to the interior of the embryo and thereby forms the mesoderm and endoderm. **(2)** MET of the mesenchymal mesoderm cells leads to the formation of the notochord, somites, primordium of the urogenital system, splanchnopleure and somatopleure. **(3)** Subsequently, cells for instance from the dermomyotome perform another round of EMT/MET to generate the dermis, muscle and satellite cells (adapted from [20]).

Type II - EMT in wound healing and fibrosis: The second type of EMT displays cell (de)differentiation processes in the adult. EMT is induced upon tissue injury and inflammation in adjacent epithelial cells. The generated mesenchymal cells are able to migrate into the wound, reconstruct and repair the tissue. However, chronic tissue inflammation induces the formation and maintenance of a mesenchymal cell phenotype eventually leading to organ fibrosis.

Type III - Oncogenic EMT: Type III EMT describes the aberrant activation of (partial) dedifferentiation processes in epithelial tumor cells. Oncogenic EMT promotes primary tumor cell dissemination and invasion into the surrounding tissue to eventually seed metastasis in distant organs. The reversion of cell differentiation during tumor progression is thought to be a crucial event to promote cancer cell colonization and the establishment of secondary tumors. Aside from invasive properties, oncogenic EMT seems to endow tumor cells with resistance to anoikis and chemotherapy, evasion of immune surveillance and stem cell-like properties [20].

In summary, epithelial cell plasticity facilitated by EMT and MET processes, is an indispensable feature during physiological processes such as embryogenesis and tissue homeostasis. However, aberrant and uncontrolled activation of EMT correlates with pathological patterns, such as organ fibrosis and promotes malignant tumor progression.

1.1.3 EMT and the metastatic cascade

Genetic and epigenetic instability can induce tumor formation in a multistep process. The malignant form of tumor progression and the cause for more than 90 % of cancer-associated deaths is characterized by a tumor cell`s ability to invade into the surrounding tissue, spread and reestablish secondary tumors at distant organs [22]. However, when metastasis occurs during tumor progression is still debated and has raised two cancer progression models. The linear progression model describes the dissemination of tumor cells from advanced, higher graded tumors and claims metastasis is a rather late event during tumor progression. On the other hand, the parallel progression model suggests an early cell dissemination event and metastases evolve independently of the primary tumor [23]. So far, final proof for one or the other model has not yet been obtained. On a side note, it has been estimated that only 0.01 % of tumor cells entering the systemic circulation are able to form macrometastases, demonstrating that metastases formation is a rather inefficient process [24].

The “invasion metastasis cascade” (Figure 3) describes the sequential processes leading to metastases outgrowth and includes tumor cell migration and invasion into the surrounding tissue. Subsequently, cells brake through the basement membrane and intravasate into the systemic circulations like the lymphatic and blood vascular system. Tumor cells disseminate in the body and have to survive the harsh conditions in the bloodstream and cope with shearing forces, anoikis and immune cells. At a secondary site, cells extravasate, migrate into the organ parenchyma and either enter a dormant cell state and remain there for a long time [25] or start to proliferate, an event termed colonization, and form macrometastases [26]. In the past years, more and more evidence accumulated demonstrating that EMT/MET plasticity of epithelial tumor cells promotes malignant tumor progression. *In vitro* and *in vivo* studies have revealed that EMT of epithelial tumor cells promotes cell migration and invasion into the stroma [2]. Additionally, cells at the invasive tumor front often display a dedifferentiated morphology that is accompanied with a loss of the cell-cell junction protein E-cadherin [27, 28]. Furthermore, circulating tumor cells (CTCs) isolated from the bloodstream of patients with metastatic breast cancers exhibited mesenchymal cell marker expression relative to primary tumor cells [29, 30]. These findings are also consistent with a previous *in vivo* lineage tracing study in a Kras/p53-driven pancreatic tumor mouse model [31]. Here, invading and circulating

pancreatic tumor cells exhibited a rather mesenchymal cell phenotype. Finally, the tumor differentiation status correlated with the clinical outcome, where less differentiated tumors gave a poorer survival prognosis for breast cancer patients [32]. These results indicate that EMT might help tumor cells to overcome the initial steps of the metastatic cascade.

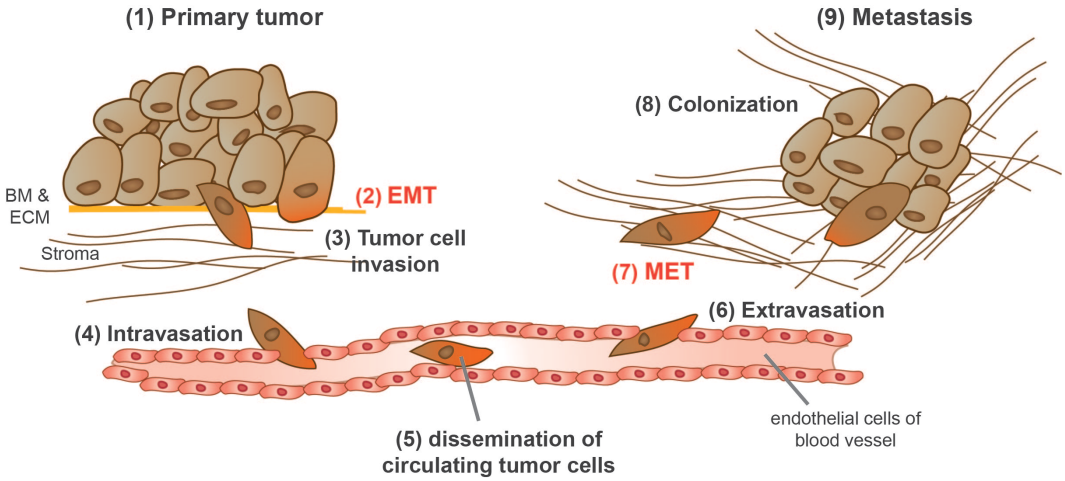


Figure 3: EMT/MET and the metastatic cascade. (1) Hyperproliferative epithelial primary tumor (brown) (2) Some tumor cells undergo EMT (reddish) and gain mesenchymal characteristics and (3) start to invade the surrounding ECM and break through the basement membrane (BM). (4) Tumor cells intravasate into the blood circulation. (5) Circulating tumor cells with mesenchymal characteristics disseminate through the blood system and (6) extravasation at a distant site. (7) Cells invade the parenchyma of the target organ and either enter a dormant state or undergo MET, (8) start proliferating (colonization) and (9) establish macrometastasis.

Tsai and colleagues have demonstrated the importance of EMT/MET dynamics for metastatic outgrowth. In a squamous cell carcinoma mouse model the authors have shown that activation of Twist1, a classical EMT inducer, in the primary tumor induced tumor cell intravasation and dissemination *in vivo*. More importantly, Twist1 inactivation and thereby MET induction at a metastatic site was crucial for tumor cell outgrowth [33]. Similar observations have been reported for other EMT/MET inducing factors, such as Prrx1 [34], Id1 [35] and Snail [36]. Furthermore, such MET events at secondary sites would explain why metastatic lesions often display a differentiated phenotype similar to their primary tumors [37, 38]. Accordingly, Thomas Brabletz has highlighted two potential types of metastasis formation [39]. Within the “Plasticity type I” metastasis model, tumor cells are susceptible for signals from the microenvironment and dynamic EMT/MET processes drive metastasis formation. In detail, initial cell dedifferentiation promotes tumor cell dissemination, intravasation and extravasation into/from the bloodstream. Afterwards, the reversion to a

differentiated cell state promotes tumor cell proliferation and colonization at the metastatic site. Thomas Brabletz has pointed out that metastases can also have a dedifferentiated morphology independent of the morphology of the primary tumor. This phenotype can be explained by the “Genetic type II” metastasis model. Here, tumor cells undergo a permanent, irreversible EMT, exhibit stem-like properties and loose their phenotypic plasticity upon genetic alterations. Such genetic alterations force cancer cell metastatic colonization and outgrowth without the need for an MET [39].

The existence and relevance of EMT during human tumor progression was long debated and its transient nature did and does complicate the search for proof [40]. However, new sophisticated technologies helped and provided evidence that an EMT and maybe more importantly a partial EMT (characterized by the coexistence of epithelial and mesenchymal markers) is indeed implicated in cancer progression [29, 41-43]. Nevertheless, further investigation is needed to pinpoint the function and further dissect the molecular mechanisms of EMT/MET processes during cancer progression.

1.1.4 EMT in breast cancer

Breast cancer is a heterogeneous disease and the most commonly diagnosed cancer in women. The pathological and molecular heterogeneity is reflected in the diverse clinical outcomes demonstrating the importance of breast cancer classification for both patient treatment and prognosis [44].

Histopathologically, invasive breast cancers are divided into different subtypes depending on their morphology. The most common type is the “invasive carcinoma of no special type” and was previously known as “invasive ductal carcinoma”. Less frequently diagnosed are invasive lobular carcinomas or others like metaplastic, tubular, cribriform, mucinous, medullary carcinomas [45]. However, more importantly for breast cancer treatment is their immunopathological classification based on the expression of specific markers: estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2, also known as Neu or ErbB2). Generally spoken, ER- and PR-positive (ER+, PR+) primary tumors have a differentiated morphology, respond well to endocrine therapy and are predictive for a

good clinical prognosis for patients. HER2 gene amplification and overexpression is observed in 20 – 25 % of breast carcinomas. HER2-positive (HER2+) tumors are rather dedifferentiated and aggressive but can be treated by targeted therapies such as HER2 specific tyrosine kinase activity blocking antibodies or small-molecule inhibitors. Tumors negative for all three receptors (ER-/PR-/HER2-) are summarized in the triple-negative breast cancer (TNBC) subtype and display the worst prognosis also because of the lack of successful therapies [46].

In addition to immunopathological classifications invasive breast cancers can be subdivided according to gene expression profiles generated by microarray gene expression analysis [47-49]. This classification distinguishes five subtypes: luminal A, luminal B, basal-like, HER2-enriched and normal breast-like. Latter exhibits a similar expression profile like non-cancerous tissue. The luminal subtypes are often ER+ and resemble a rather differentiated tumor morphology since they express E-cadherin and epithelial Cytokeratins (CK8, CK18). Gene expression profiles of basal-like tumors overlap in part with the TNBC. Later, a sixth subtype, claudin-low, has been added to the molecular classification of breast cancers [50]. Both, basal-like and claudin-low subtypes often assemble metaplastic mammary carcinomas, a histologically mesenchymal, aggressive and chemoresistant tumor [51, 52]. Claudin-low tumors lack expression of luminal, epithelial and proliferation-associated markers and, even more than the basal-like subtype, display an EMT gene signature and stem cell-like features [32]. Of note, hormone and chemotherapy treatment of patients with different breast cancers induced recurrent tumors with an undifferentiated claudin-low gene signature [53]. However, it remains unclear, whether the tumor cells were induced for a claudin-low signature or whether they existed before and have been selectively enriched during cancer treatment. Nevertheless, the study has demonstrated that dedifferentiated tumors with a claudin-low EMT gene signature go hand in hand with chemoresistance and tumor-initiating capacities. Along this line, claudin-low tumors are also predictive for a worse patient survival prognosis than differentiated luminal A tumors. Surprisingly, the prognosis is not worse than for basal-like and other tumor subtypes, which could have been expected if a full EMT cell phenotype goes along with tumor cell invasiveness and metastatic potential [32].

In summary, sub-classifications of human breast cancers clearly helped us to dissect the heterogeneous nature of this cancer type and provided valuable

advantages for breast cancer treatment and prognosis. EMT-like features can be found in certain subtypes of breast cancers, where a dedifferentiated morphology is predictive for poor relapse free survival and overall survival compared to differentiated tumors [32].

1.2 Inducers of EMT

Various extracellular stimuli can activate signaling cascades that mediate EMT progression during development, tissue regeneration and malignant tumor progression. Growth factors, such as fibroblast growth factor (FGF), epidermal growth factor (EGF), insulin-like growth factor (IGF), hepatocyte growth factor (HGF), platelet-derived growth factor (PDGF), vascular endothelial growth factor (VEGF) and transforming growth factor β (TGFβ) can induce EMT in a context-dependent manner. Cells release growth factors into the extracellular microenvironment where they bind to tyrosine kinase receptors (RTKs) located on cells. These receptors transmit and activate MAP kinase and PI3 kinase intracellular signaling pathways inducing cell proliferation and survival. Depending on the growth factor, other signaling pathways such as TGFβ, Wnt, Notch, Hedgehog, JAK-STAT, NF-κB and Hippo signaling can be activated by different stimuli, crosstalk and influence the EMT process. Tissue metabolic and mechanical stresses can also act as powerful inducers of EMT and tumor cell invasion [3].

1.2.1 Canonical and non-canonical TGFβ signaling in EMT

TGFβ signaling in normal tissues regulates cytostasis, cell apoptosis and differentiation and therefore tightly controls tissue development and homeostasis. The cytokine TGFβ belongs to the TGFβ superfamily and consists of three members (TGFβ1-3). TGFβ homodimers bind to TGFβ receptor II (TGFβRII) to eventually form heterotetrameric complexes with TGFβRI. Phosphorylated TGFβRII acts as serine/threonine kinase receptor and trans-phosphorylates serine and threonine residues within the intracellular domain of TGFβRI. TGFβRI recruits, phosphorylates and activates receptor Smads (R-Smads), Smad2 and Smad3, which in turn complex with Smad4, a common Smad (Co-Smad). In the nucleus, the trimeric Smad complex binds to specific DNA motifs, however to achieve high DNA binding and specificity they need to interact with other transcriptional cofactors to activate or repress gene transcription in a context-dependent manner. TGFβ signaling via Smads describes the canonical TGFβ pathway (Figure 4) and controls cell proliferation, apoptosis and differentiation during development and tissue regeneration [3, 54, 55].

TGFβ signaling induces the transcription of Snail, Zeb1 and Twist transcription factors (TFs) [56-58], which can together with Smads repress or activate the transcription of other epithelial or mesenchymal genes, respectively [16]. For instance, Snail interacts with Smad3/4 to repress transcription of the epithelial junction proteins E-cadherin and Occludin [59]. Similarly, Zeb1 and Zeb2 complex with Smad3/4 to repress E-cadherin transcription, thereby, promoting EMT [57]. Of note, Smad3 and Smad4 are crucial mediators of TGFβ-induced EMT, whereas Smad2 can also function as an EMT suppressor [56]. Several groups have demonstrated that TGFβ-induced EMT is a sequential process. Epithelial cells can dedifferentiate into mesenchymal cells through a partial or hybrid cell state, which is characterized by the coexistence of epithelial and mesenchymal markers. This stepwise transition and its (ir)reversibility is tightly controlled by double negative feedback loops of Snail/miR-34 and Zebs/miR-200, depending for instance on the concentration and duration of TGFβ treatment [60-62].

Non-canonical TGFβ signaling via Rho-like GTPases, PI3K and MAPK signaling pathways can contribute to EMT and its functional outputs. Activation of these pathways occurs on the level of TGFβRI and II, which interact and activate certain adaptor proteins. For example, RhoA, Rac and Cdc42 GTPases drive Actin cytoskeleton reorganization and promote the formation of lamellipodia and filopodia [63, 64]. Par6, a cell polarity complex protein, is phosphorylated by TGFβRII and promotes cell-cell tight junction disassembly during EMT [65]. TGFβ signaling can phosphorylate Akt through PI3K activation, which leads to the activation of mammalian target of rapamycin (mTOR) and subsequently S6 kinase 1 (S6K1). Activation of PI3K-Akt has been shown to be required for cell migration and tight junction disassembly [66]. Furthermore, TGFβ activates p38 and c-Jun N-terminal kinase (JNK) via TRAF6/TAK1/MKK signaling. p38 has been shown to be required for TGFβ-induced EMT and apoptosis [67]. JNK induces the expression of the mesenchymal marker Fibronectin and is a critical regulator of cell migration [68]. Finally, TGFβ activates Erk1/2 through the Ras/Raf/Mek1/2 cascade, and Erk1/2 are required for cell-cell adhesion junction disassembly and cell motility [69]. In summary, TGFβ is a potent inducer of EMT and modulates cell dedifferentiation through Smad-dependent and independent pathways.

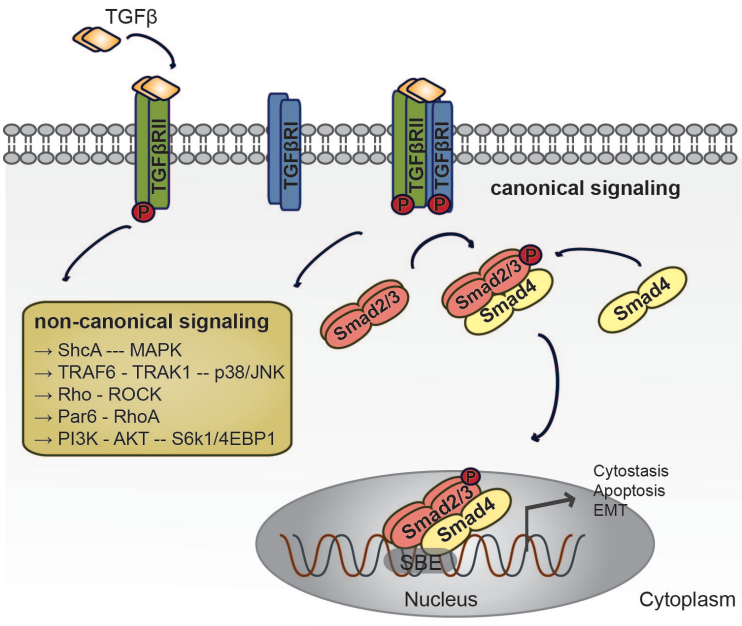


Figure 4: Canonical and non-canonical TGFβ signaling. Binding of TGFβ dimers to TGFβRII induces the assembly of the heterotetrameric receptor complex build by TGFβRII and TGFβRI. Transphosphorylation of TGFβRI phosphorylates Smad2/3, which upon binding to Smad4 translocate to the nucleus and activate or repress gene transcription (canonical signaling). Activated TGFβRs can also transmit the signal to other factors independently of Smads. TGFβ mediates the activation of p38 and JNK through receptor associated TRAF6 and TRAK1. ShcA activates the Ras-Raf-MEK-ERK MAP kinase pathway. RhoA at tight junctions is degraded by the proteasome upon activation of Par6-Smurf1 by TGFβRII. In parallel, RhoA is activated by TGFβ and promotes Actin cytoskeleton rearrangements via ROCK. TGFβ induces PI3K and activates/suppresses S6K1/4EBP1 via AKT.

1.2.2 TGFβ signaling in cancer

In normal tissue and early stages of tumorigenesis TGFβ mainly induces a G1 cell cycle arrest and apoptosis and functions therefore as a tumor suppressor. In advanced carcinomas, tumor cells lose such growth-inhibitory restrictions and accumulate invasive and metastatic properties in response to TGFβ, which can be accompanied by an EMT phenotype [70]. Loss-of-function mutations or deletions of core TGFβ pathway components, such as TGFβRs or Smad4, have been observed in many human cancers like colon, gastric, ovarian, head and neck, esophageal and pulmonary carcinomas [71]. Interestingly, TGFβR mutations are rarely seen in breast cancer. Here, only the tumor suppressive branch of TGFβ signaling (anti-proliferation and pro-apoptosis) seems to be lost while the cancer promotive function is still intact. During tumorigenesis, TGFβ signaling can switch from a tumor suppressor to a tumor promoter, a phenomenon known as the “TGFβ paradox” [70]. In this case, TGFβ promotes tumor growth, cell dedifferentiation, invasion and dissemination and in turn generates a pro-inflammatory and pro-tumorigenic microenvironment by stimulating the release of autocrine mitogens and cytokines. Like a self-activating loop, TGFβ is produced and released by tumor cells and tumor stromal cells [72]. Not surprisingly,

increased TGFβ levels have been found in cancer cells and in the tumor microenvironment of breast cancer patients [73]. Furthermore, high plasma levels of TGFβ in cancer patients are predictive for poor prognosis [71].

1.3 Molecular networks regulating EMT

EMT-inducing stimuli such as various growth factors, other receptor ligands, hypoxia or ECM constitution can induce intracellular signaling cascades, which in turn activate or inactivate various components of interconnected regulatory networks. These networks consist of splicing factors (see 1.3.1), epigenetic regulators (see 1.3.2), TFs (see 1.3.3) and non-coding RNAs (ncRNAs; see 1.3.4). They can orchestrate transcriptional, post-transcriptional and post-translational processes essential for EMT/MET plasticity. Their aberrant regulation can cause oncogenic EMT/MET events and promote malignant tumor progression [74] (Figure 5).

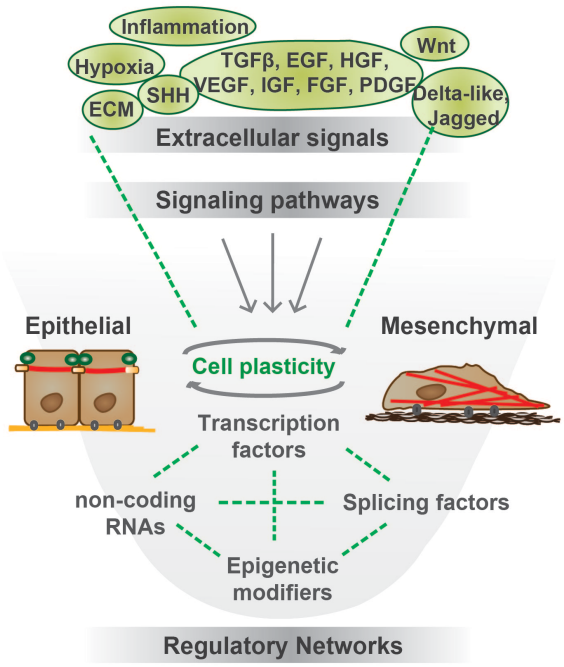


Figure 5: EMT/MET regulatory networks. Various extracellular stimuli activate or inactivate intracellular signaling pathways, which influence the behavior of components of interconnected regulatory networks. Transcription factors, splicing factors, epigenetic modifiers and non-coding RNAs control the dynamic regulation of EMT on the transcriptional, post-transcriptional and post-translational level. Extracellular matrix (ECM), sonic hedgehog (SHH), transforming growth factor β (TGFβ), epidermal growth factor (EGF), hepatocyte growth factor (HGF), vascular endothelial growth factor (VEGF), insulin-like growth factor (IGF), platelet-derived growth factor (PDGF).

1.3.1 Alternative splicing and EMT

Alternative splicing is a post-transcriptional mechanism of gene regulation and facilitates the generation of several (functionally) different RNAs per gene. Genome-wide RNA analysis revealed that approximately 95 % of human multi-exon genes are alternatively spliced. This emphasizes that alternative splicing contributes and expands the RNA and protein diversity encoded by the eukaryotic genome [75, 76]. Shapiro and coworkers have examined the changes in RNA splicing during EMT [77]. Deep sequencing analysis of the transcriptome of human mammary epithelial

(HMLE) versus Twist-induced mesenchmyal HMLE cells has revealed global changes in mRNA splicing patterns during EMT. Several of the EMT-associated alternative splicing events are also conserved in luminal versus basal-like breast cancer cell lines. These results indicate that alternative splicing is tightly regulated during EMT and breast cancer progression. Furthermore, the authors have suggested that an EMT-associated splicing signature could be used as potential biomarkers for breast cancer metastasis [77].

The RNA-binding proteins epithelial splicing factor regulatory protein 1 and 2 (ESRP1 and ESRP2) have been identified as fibroblast growth factor receptor 2 (FGFR2) splicing regulators [78]. Expression of FGFR2 splice variant with exon IIIb (FGFR2-IIIb) is observed in epithelial cells, whereas FGFR2 splice variant with exon IIIc (FGFR2-IIIc) is exclusive in mesenchymal cells. These distinct FGFR2 splice variant expression patterns are essential during development where they control cell proliferation and differentiation [79-81]. Expression of ESRP1 and 2 is strongly downregulated during EMT and are direct targets of EMT TFs such as Snail, Zeb1 and Zeb2 [77, 82, 83]. Loss of ESRP1 or ESRP2 expression induces a switch from FGFR2-IIIb to FGFR2-IIIc isoform in human prostate epithelial cells [78]. Furthermore, stable knockdown of ESRP1 or ESRP1/2 in epithelial mammary cells promotes mesenchmyal cell morphology, accompanied with increased expression of mesenchymal markers and cell motility. However, loss of ESRP2 expression only, does not induce mesenchymal characteristics, suggesting that epithelial cell morphology is predominantly regulated by ESRP1 [84]. Warzecha and colleagues have further identified other known splice isoforms of CD44 [85], p120/Catenin (CTNND1) [86] and Mena (ENAH1) [87] to be regulated by ESRP1/2 and are typical for an epithelial or mesenchymal cell morphology [78] (Figure 6).

A switch in the cell surface marker CD44 from variant isoforms (CD44v) to the standard isoform (CD44s) has been observed to be required during EMT and tumor formation. High expression of CD44s further correlated with high-grade human breast tumors [85]. p120 stabilizes adhesion junctions between epithelial cells [88] and has four different splice isoforms. The shorter isoforms (isoform 3 and 4) lack N-terminal sequences and their expression has been observed in epithelial cells, whereas mesenchymal cells mainly expressed the full-length p120 splice variant. p120 isoforms bind RhoA via a central Armadillo domain site (amino acids 622-628). However, only full-length p120 is able to suppress RhoA activity via the N-terminal

domain and increases tumor cell invasion [86, 89]. Mena is a member of the Ena/VASP family and its isoform Mena+11a is specifically expressed in epithelial pancreatic tumor cell lines [87]. These results suggested that ESRP1/2 regulate a broad spectrum of epithelial gene isoforms on the post-transcriptional level to maintain an epithelial cell morphology and behavior.

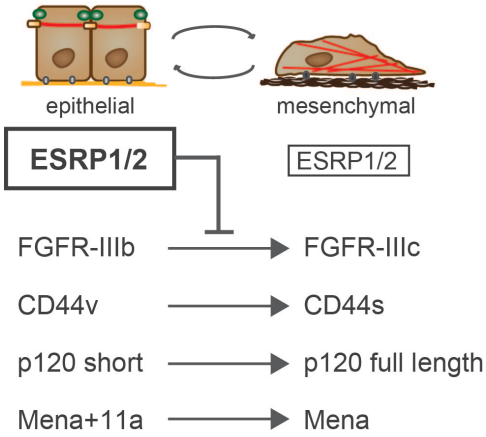


Figure 6: ESRP1/2 regulate EMT/MET plasticity via alternative splicing. Epithelial splicing factor regulatory proteins 1 and 2 (ESRP1/2) are highly expressed in epithelial cells and are transcriptionally downregulated during EMT. ESRP1/2 facilitate alternative splicing of factors specific for epithelial cell morphology: Fibroblast growth factor receptor 2 (FGFR2) with exon IIIb (splice variant with exon IIIc is expressed in mesenchymal cells); CD44 variant isoform (CD44v; CD44 standard (CD44s) isoform is expressed in mesenchymal cells); p120 short isoforms (p120 full length in mesenchymal cells) and Mena with exon 11a (Mena+11).

In addition to ESRP1/2 other RNA-binding splicing factors have been identified to regulate EMT-specific RNA splicing patterns, for instance members of the RBFOX (RNA binding protein FOX1 homologue), CELF (CUGABP Elav-like family), MBNL (muscleblind-like protein) and hnRNP (heterogeneous nuclear ribonucleoprotein) families [77]. RBFOX2, for instance, is moderately upregulated during EMT and regulates mesenchymal tumor cell invasion [90].

In conclusion, alternative splicing is coordinated by RNA-binding splicing factors and adds another layer of dynamic gene expression regulation during EMT. Aside from transcriptional, epigenetic and ncRNA-mediated post-transcriptional control, alternative splicing events regulate cellular plasticity and are associated with tumor progression.

1.3.2 Epigenetic regulation of EMT

Epigenetic mechanisms globally regulate gene expression and cause cell phenotype alterations without affecting the order of nucleotides in DNA sequences. Here, stable but reversible DNA methylation and post-translational histone

modification marks alter and remodel the chromatin architecture. Euchromatin describes open, lightly packed chromatin and genes can be actively transcribed. On the other hand, heterochromatin is rather characterized as closed, tightly packed chromatin and gene expression is repressed. Moreover, epigenetic marks control the On/Off-state of gene transcription in a context-dependent manner [91, 92].

Alterations in DNA methylation and histone modification patterns are commonly observed and lead to aberrant gene expression in cancer [93, 94]. Additionally, several studies have demonstrated global epigenetic chromatin remodeling and reprogramming processes during epithelial cell dedifferentiation. These processes include DNA methylation and histone modifications such as methylation and acetylation marks [95-98].

1.3.2.1 DNA methylation during EMT

DNA methylation implies the covalent attachment of a methyl group to cytosines of CpG di-nucleotide sites within DNA sequences. Such marks usually induce repression of gene transcription and are established by DNA methyltransferases (Dnmts) [99]. The CDH1 gene promoter region has been found to be hypermethylated in breast and prostate carcinoma cell lines, which correlated with a loss of E-cadherin expression. Treatment with 5-aza-2-deoxycytidine, a demethylation agent, partially restored E-cadherin expression [100]. Furthermore, CDH1 promoter hypermethylation correlated with malignant breast cancer progression [101]. A dynamic and reversible regulation of gene promoter hypermethylation during EMT has been reported recently. TGFβ-induced EMT of epithelial MDCK cells induced promoter hypermethylation and transcriptional downregulation of the miR-200 gene cluster - a mechanism also observed in metastatic cancer cell lines. Interestingly, TGFβ-withdrawal reversed such methylation marks and induced a MET [102]. These studies suggest that individual DNA promoter (de)methylation events are a critical epigenetic mechanism for gene silencing during EMT/MET and can drive malignant tumor progression. In contrast, McDonald et al. have reported that genome-wide DNA methylation was unchanged in a TGFβ-induced *in vitro* EMT model. Nonetheless, the authors demonstrated a global reduction in heterochromatin induced by histone modifications, which might have been crucial for cell dedifferentiation [97].

1.3.2.2 Histone modifications during EMT

Histones organize genomic DNA into structural units (nucleosomes) and can be post-translationally modified, for instance by acetylation or methylation of lysine (K) or arginine (R) residues on histone 3 or 4 (H3, H4). Such covalent modifications open or close the chromatin and make it accessible for factors controlling gene transcription. While, generally spoken, histone acetylation usually activates adjacent gene promoters, methylation can activate or repress gene transcription depending on the lysine residue: K4, K36 or K79 – activation of transcription; K9 or K27 – inactivation of transcription [103].

Polycomb group (PcG) complexes

The mammalian PcG complexes consist of two classes of multiprotein complexes, namely polycomb repressor complex 1 (PRC1) and 2 (PRC2). PRC2 consists of three core components: Ezh (enhancer of zeste 1 or 2 (Ezh1 or Ezh2)), Eed (embryonic ectoderm development) and Suz12 (suppressor of zeste 12). PRC2 exhibits histone methyltransferase activity and is initially recruited to target genes to catalyze mono-, di- or trimethylation of H3K27. Subsequent recruitment of PRC1 to H3K27me3 sites is required for the stabilization of such repressive marks [104, 105]. The core PRC1 consists of chromobox-domain (Cbx) proteins, members of the Pcgf family (Pcgf1-6), Ring1 family (Ring1a and Ring1b) and Hph family (Hph1-3) [106].

Both, PRC1 and PRC2, participate in promoting EMT and cancer progression. Bmi1 (B lymphoma Mo-MLV insertion region 1 homolog, also known as Pcgf4) is part of the PRC1 and was the first PcG protein associated with cancer. Bmi1 collaborates with c-Myc and H-Ras to induce cell proliferation and directly represses the Ink4a/Arf locus, which encodes for the tumor suppressors and cell cycle inhibitors p16 and p19Arf. Overexpression of Bmi1 induces neoplastic transformation of mouse embryonic fibroblasts and lymphomas [107-109]. Later, Bmi1 has been connected to EMT. Twist has been shown to induce the expression of Bmi1, which interacts with Ezh2 and Twist1 to repress CDH1 promoter activity and promotes EMT [110]. Additionally, Bmi1 can be post-transcriptionally repressed by miR-200c and miR-203 [111] (Figure 7).

PRC2 can induce EMT by collaborating with factors guiding the complex to certain target genes. Snail has been reported to silence the CDH1 promoter in tumor

cells by recruiting the PRC2 components Ezh2 and Suz12, which mediate H3K27me3 and thus repression of CDH1 expression [112]. Of note, the epithelial cell-specific miR-200 family can post-transcriptionally repress Suz12 expression, which in turn triggers the stabilization of E-cadherin levels [113]. Ezh2 expression is post-transcriptionally controlled by miR-101. Upon EMT induction expression of Snail and Slug represses miR-101 transcription, which in turn leads to the stabilization of Ezh2 levels, EMT and increased migration of oral tongue squamous cell carcinoma (OTSCC) cells [114]. Furthermore, Sox4 directly induces the transcription of Ezh2 during TGFβ-induced EMT of normal mammary epithelial cells, and Sox4 executes its EMT and metastasis promoting function at least in part via Ezh2. Interestingly, many EMT-relevant genes are regulated by H3K27me3 marks during TGFβ-induced EMT and overlap with Ezh2 function during EMT [98]. Snail has been reported to interact with the methyltransferases Suv39h1 (suppressor of variegation 3-9 homolog 1) and G9a to repress CDH1 transcription [115, 116] (Figure 7). Additionally to TFs, long non-coding RNAs (lncRNAs) can act in *cis* and *trans* to recruit PRC2 to specific genomic regions [117-121]. For instance, lncRNA HOTAIR recruits the PRC2 in *trans*, which triggers the generation of H3K27 methylation and subsequent gene silencing of HOX genes in fibroblasts [117].

Interestingly, H3K27me3 marks mediated by polycomb group proteins are often found in combination with activating histone modifications, such as H3K4m3. In this case, nearby gene promoters are transcriptionally repressed, but can be easily activated when the repressive mark is removed [122, 123]. It is believed that such bivalent marks could allow a dynamic gene regulation and might be important for epithelial-mesenchymal cell plasticity. Chaffer and colleagues have recently reported a bivalent Zeb1 promoter configuration in basal $CD44^{low}$ non-cancer stem cells (CSCs). Upon TGFβ stimulation, the repressive K27 mark was removed, the Zeb1 promoter was activated and cells rapidly entered a $CD44^{high}$ stem cell-like state [124]. However, additional research is needed to test whether poised gene promoters are fundamental for rapid and reversible EMT cell plasticity.

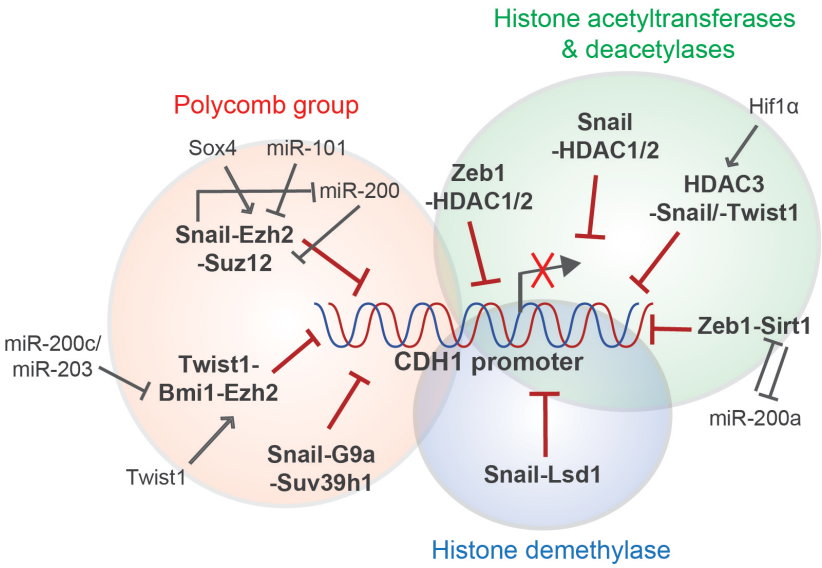


Figure 7: Histone modifiers collaborate with EMT TFs to silence the CDH1 promoter during EMT. Members of the polycomb group complex (red), histone acetyltransferases and deacetylases (green) and histone demethylases (blue) regulate E-cadherin expression during EMT. Their (post)-transcriptional regulation by TFs or miRNAs is depicted.

Histone (de)acetylation

Histone acetylation results in euchromatin and active gene transcription. Histone acetyl transferases (HATs) mediate the covalent binding of an acetyl group, for instance to lysine histone residues. Histone deacetylases (HDACs) revert such modifications and induce transcriptional repression of genes [125]. Recruitment of the HDAC1/2-containing NuRD (Mi2/nucleosome remodeling and deacetylase) repressor complex by Snail results in CDH1 gene silencing [126] (Figure 7). Twist1 together with the NuRD complex also represses E-cadherin expression and loss of Twist1 or complex components reduced the metastatic behavior of mouse mammary cells [127]. HDAC3 is essential during hypoxia-induced EMT and metastasis. It is a direct target of Hif1α and represses E-cadherin transcription in collaboration with Snail and Twist1. Additionally, it promotes expression of N-cadherin and Vimentin [128]. The deacetylase Sirt1 is another example for epigenetic CDH1 gene silencing. EGF-induced EMT in epithelial prostate cells upregulates Sirt1 and is recruited by Zeb1 to the proximal CDH1 promoter. H3 deacetylation suppressed CDH1 transcription [129]. Interestingly, in a TGFβ-induced mammary EMT model, Sirt1 is embedded in a double negative feedback loop with miR-200a [130]. Similarly to Sirt1, Zeb1 can also recruit HDAC1/2 to the CDH1 promoter, mediate H3 and H4 deacetylation and induced its transcriptional repression in pancreatic tumor cells

[131] (Figure 7). Surprisingly, only a class I HDAC inhibitor (mocetinostat) consistently restored E-cadherin expression and Zeb1-induced drug resistance to gemcitabine of pancreatic cells *in vivo* [132].

Histone methylation

Histone methylation or demethylation is catalyzed by lysine methyltransferases or demethylases, respectively. Depending on the lysine residue and cellular context, histone methylation marks can either activate (H3K4, H3K36 and H3K79 marks) or repress gene transcription (H3K9, H3K27 and H4K20 marks) [125]. Lysine specific demethylase 1 (Lsd1) has been the first histone demethylase identified, it removes methyl groups of H3K4m2/3 sites to reduce gene transcription [133]. Lsd1 interacts directly with the SNAG domain of Snail and both collaborate to reduce the H3K4m2 mark on epithelial gene promoters, such as CDH1, CLDN7 and KRT8 (encoding for Claudin 7 and Cytokeratin 8, respectively) in human mammary epithelial cells [134]) (Figure 7). Furthermore, Lsd1 is highly expressed in estrogen-negative breast tumors, which display a rather dedifferentiated gene signature [135]. Enzymatic inhibition of Lsd1-Snail interaction triggers the re-expression of epithelial genes and suppressed cancer cell invasion [136]. Of note, Lsd1 can also remove methyl groups of repressive H3K9m3 marks and reactivate gene transcription [137]. McDonald et al. have shown a reversible reduction in heterochromatin and an increase in euchromatin during TGFβ-induced EMT, measured by H3K9me2 and H3K4me3 marks. The same group has further reported that the changes in chromatin reprogramming were in part dependent on Lsd1 [97]. These results, underline the importance of Lsd1 activity on different histone marks during epithelial cell dedifferentiation processes.

Taken together, epigenetic modifications, such as DNA methylation and different histone modifications, contribute to epithelial cell plasticity and malignant tumor progression by regulating global chromatin architecture and the expression of individual genes.

1.3.3 Transcriptional control of EMT

A plethora of extracellular stimuli can activate cellular signaling pathways (Wnt, NFkB, Notch, Ras-ERK1/2, Hif1/2, TGFβ), which are known to induce EMT (Figure 5). The majority of them converge on the level of EMT-regulatory TFs, which can initiate and/or maintain the overall transcriptomic changes during EMT in a tissue- and developmental-dependent manner [138].

1.3.3.1 EMT transcription factors – drivers of a cell dedifferentiation program

The TF families Snail (zinc finger proteins Snail and Slug), Zeb (zinc finger and homeodomain proteins Zeb1 and Zeb2) and Twist (basic helix-loop-helix proteins E12, E47, Twist1, Twist2 and ID) play a central role during developmental and oncogenic EMT and are considered master EMT TFs [139].

EMT TFs are potent inducers of the epithelial cell dedifferentiation process and act as transcriptional suppressors and activators. They coordinate the expression of cell architecture proteins, other transcriptional remodelers and even control the expression of each other to drive EMT [139]. The gene promoter of the cell adhesion junction protein E-cadherin is a common target of all EMT TFs. In cooperation with epigenetic modifiers, Snail [112, 115, 116, 126, 134], Zeb [129] and Twist [127] directly silence the CDH1 gene promoter and inhibit E-cadherin expression (see 1.3.2; Figure 7) – a crucial event during EMT [16]. In addition to E-cadherin, the EMT TFs further suppress the expression of other epithelial genes encoding for cell-cell adhesion (Occludin, ZO-1, Claudins, Desmoplakin, Plakoglobin) and polarity complex proteins (Crumbs3, Pals). At the same time, they promote the transcription of mesenchymal genes encoding for N-cadherin, MMPs or ECM proteins to induce and maintain EMT, cell motility, invasion and survival [3] (Figure 8).

Various cellular signaling pathways can trigger the transcription/activation of EMT TFs [139]. Snail and Slug can be activated by TGFβ, Wnt, Notch, PI3K, NFκB signaling. Zeb TFs are activated by TGFβ, Wnt and, like Twist1, by MAPK signaling. Importantly, post-translational modifications like phosphorylation and sumoylation determine their cellular localization, activity and degradation (Figure 8). Moreover,

miRNAs such as members of the miR-200 and miR-34 families post-transcriptionally bind and negatively regulate Zeb and Snail cellular levels [39].

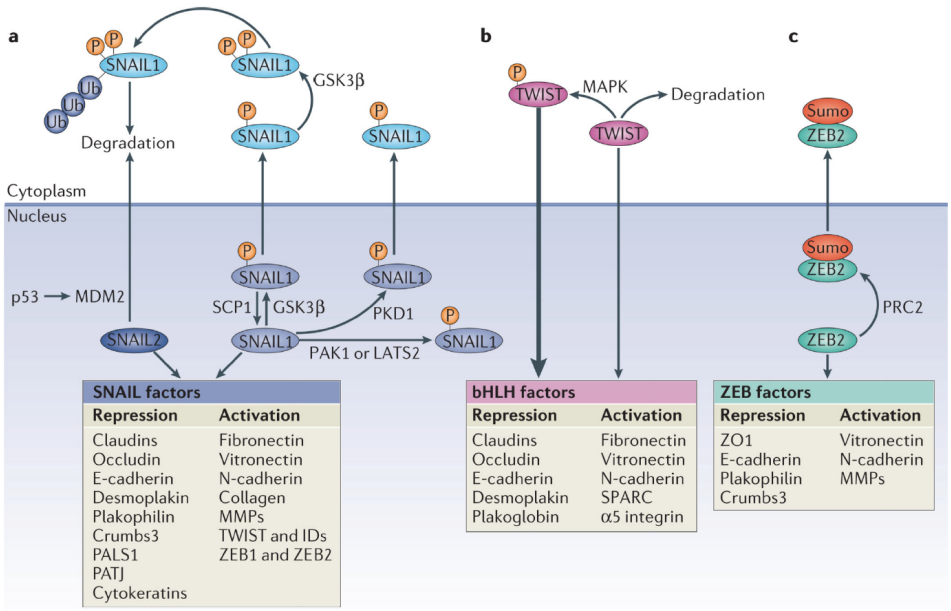


Figure 8: EMT TFs – regulation and downstream targets during EMT. **(a)** Snail can be phosphorylated by the glycogen synthase kinase 3β (GSK3β) and protein kinase D1 (PKD1), which induces the nuclear export and transcriptional inactivation of Snail. A second phosphorylation by GSK3β leads to its ubiquitylation and proteasomal degradation. Phosphorylation of Snail by p21 activated kinase 1 (PAK1), large tumor suppressor 2 (Lats2) or dephosphorylation by small C-terminal domain phosphatase 1 (SCP1) triggers Snail nuclear localization. Slug (also called Snail2) cellular localization is controlled by p53. Both TFs activate or repress gene transcription of various factors involved in EMT (purple box). **(b)** Phosphorylation of Twist by p38, JUN N-terminal kinase (JNK) and ERK induces Twist nuclear translocation, where it functions as transcriptional activator and repressor (red box). **(c)** Zeb2 controls the transcription of various structural proteins involved in EMT (green box). Zeb2 can be sumoylated by the polycomb group repressor complex 2 (PRC2), which leads to its cytoplasmic localization (adapted from [3]).

Knockout of Snail, Zeb or Twist family members in mice are embryonic lethal since their loss induces defects, for instance in gastrulation, neural tube closure or mesoderm differentiation [140-142]. In addition, abnormal activation of the EMT TFs has been observed in various human cancers. For instance, aberrant stimulation of signaling pathways (TGFβ, Wnt, Notch) or failure of their post-transcriptional or –translational control can cause their reactivation in cancers and often correlate with a poor clinical outcome [139, 143]. Yang and coworkers have demonstrated for the first time that suppression of Twist1 expression in orthotopically transplanted mouse mammary cells inhibited lung metastasis formation in mice. Conversely, forced expression of Twist 1 in epithelial MDCK or HMEC cells induced a mesenchymal cell morphology accompanied with a downregulation of E-cadherin and increased cell migration *in vitro.* Elevated levels of Twist1 have been found in different human

metastatic breast cancer cell lines and in lobular breast cancer tissues, a breast cancer type that usually lacks E-cadherin expression [110]. Furthermore, Twist1 expression correlates with high-grade ductal carcinoma [144], Slug expression is increased in primary breast cancer tissues and both correlate with poor clinical outcome [145]. Twist1 and Snail expression are associated with E-cadherin repression, metastasis of breast carcinoma [146, 147] and a poor prognosis [148]. Forced expression of Zeb1 promotes metastasis in a mouse xenograft model for colorectal cancer [149]. High expression of Zeb1 is further linked to reduced E-cadherin levels and increased aggressiveness of human breast carcinoma [148]. These results reveal that EMT TFs control epithelial cancer cell differentiation, cell invasiveness and their expression correlate with malignant tumor progression.

Aside from Snail, Zeb, Twist EMT TFs other TFs have been shown to control or contribute to cell EMT/MET plasticity like Prrx1 [34], Sox4 [98], Klf4 [150], FoxC2 [151], p53 [152] or Tead2 [153] and the list keeps growing.

1.3.3.2 The Tead transcription factors

Tead family, structure and function

The TEA (transcriptional enhancer factors) domain TF family consists of four evolutionary conserved family members in mammals (Tead1-4) and belongs to the helix-turn-helix class of TFs. Tead2 gene expression can be detected already at the 2-cell stage and it is the only family member expressed within the first seven days during mouse embryogenesis. At later developmental stages and in the adult, the other Tead family members are abundantly expressed in a tissue-specific but with overlapping patterns [154-156]. It has been shown that Teads are required for cardiogenesis [157], myogenesis [158], neural crest formation [159], notochord [160] and trophectoderm [161] formation. Furthermore, it should be noted that Tead1 and 2 double-knockout mice exhibited more severe growth defects and developmental abnormalities on embryonic day 8.5 than the single knockouts, suggesting redundant and distinct functions for Tead1 and Tead2 [160]. Tead homologs are also found in invertebrates. For instance, Scalloped (Sd) is the Tead homolog in *Drosophila melanogaster* and is required for wing development [162, 163]. Egl-44 regulates the differentiation of touch-sensitive cells and egg-laying motor neurons in *Caenorhabditis elegans* [164, 165].

All four mammalian Tead TFs have the same protein structure and share a high degree of sequence homology. The N-terminal DNA-binding domain is 99 % identical between the Teads [166] and binds to a DNA motif (5`-GGAATG-3`) found in the SV40 enhancer elements and in (proximal) promoter regions of target genes. Tead TFs lack a transactivation domain; therefore, their transcriptional activity depends on the interaction with transcription cofactors via their C-terminal cofactor-binding domain [155, 167, 168]. Here, several cofactors have been reported in mammals: **(a)** Yap (Yes-associated protein) [169] and Taz (transcriptional coactivator with PDZ-binding motif) [170], both downstream effectors of the Hippo pathway **(b)** Vestigial-like proteins 1-4 [171-173] and **(c)** p160 family of nuclear receptor coactivators [174].

Teads, transcriptional effectors of the Hippo pathway

Yap and Taz are cofactors of Tead TFs. Tead TFs interact with nuclear Yap and Taz in mammals to regulate genes that are involved in cell proliferation, survival and anti-apoptosis [175]. In *Drosophila*, Sd binds to the Yap homologue Yorki (Yki) to facilitate cell growth [176]. Yap and Taz in mammals and Yki in *Drosophila* are conserved downstream targets of the Hippo pathway, which can control the subcellular localization of the Tead/Sd cofactors [175].

The core Hippo pathway members, Warts (Wts), Salvador (Sav) and Hippo (Hpo) were originally discovered by genetic mosaic screens in *Drosophila* and mutations in these factors lead to tissue overgrowth [177-183]. Homologues of these factors have been found in mammals: Mst1 and Mst2 (Hpo), Sav1 (Sav), Lats1 and Lats2 (Wts). Of note, expression of human Hippo pathway homologues are able to rescue the proliferation phenotype of mutant pathway members in *Drosophila* [184], underlining the high degree of functional conservation. The core Hippo pathway in mammals consists of a serine kinase cascade (Figure 9A). The STE20 kinases Mst1 and Mst2 interact with the adaptor protein Sav1 and directly phosphorylate and activate Lats1, Lats2 and Mob1 (*Drosophila* homolog of Mats). Lats1/2 kinases complex with Mob1 and phosphorylate Yap and Taz. The amino acid residues serine 127 and serine 89 of Yap and Taz, respectively, generate a binding site for interaction with 14-3-3 proteins and triggers their cytoplasmic localization and transcriptional cofactor inactivation. Furthermore, phosphorylation of serine 381 of Yap and serine 311 of Taz by Lats1/2 induce their proteasomal degradation [175, 185, 186].

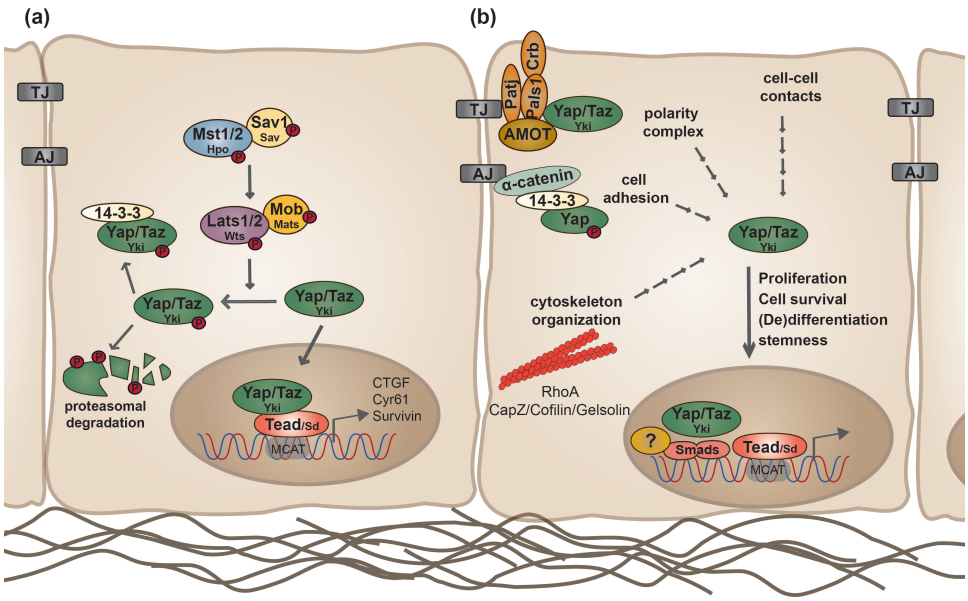


Figure 9: Regulation of Yap/Taz cellular localization. (a) The core mammalian Hippo pathway consists of a kinase cascade, which controls Yap/Taz (Yki; Yorki in *Drosophila*) nuclear and cytoplasmic localization. Phosphorylation of Mst1/2 (Hpo; Hippo) and Sav1 (Sav; Salvador) leads to the phosphorylation and activation of Lats1/2 (Wts; Warts) and Mob (Mats; Mob as tumor suppressor), which in turn phosphorylate Yap/Taz. Activation of the Hippo pathway leads either to the cytoplasmic retention of Yap/Taz bound to 14-3-3 proteins or to their proteasomal degradation upon additional phosphorylation events. Inactivation of the Hippo pathway kinase cascade induces the translocation of Yap/Taz into the nucleus. They bind for instance to Tead (Sd; Scalloped) TFs to regulate genes involved in cell proliferation and survival. **(b)** Yap/Taz localization are further controlled by cell polarity complexes, cell-cell junctions and the cytoskeleton organization. In the nucleus, Yap/Taz bind to different TFs and control cell proliferation, survival, (de)differentiation and stemness. Mammalian STE20-like protein kinase 1/2 (Mst1/2), Salvador homologue 1 (Sav1), larger tumor suppressor 1/2 (Lats1/2), Mob kinase activator (Mob), Yes-associated protein (Yap), transcriptional coactivator with PDZ binding motif (Taz or WWTR1), TEA domain TF family (Tead), Angiomotin family of proteins (AMOT).

The regulation of Yap and Taz by the Hippo pathway has been shown to facilitate cell-contact inhibition, a mechanism fundamental in organ size control and maintenance [187]. In sparsely growing cells, the Hippo pathway kinase cascade is inactivated. This leads to hypophosphorylation of Yap and Taz and their nuclear localization, where they bind to Tead TFs to stimulate cell proliferation [188]. Conversely, the upstream pathway members are activated in densely growing cells. Lats1/2 phosphorylates Yap and Taz, which leads to their cytoplasmic localization and cell growth arrest [187]. Consequences of Hippo pathway deregulations are obvious. Inactivation of kinases Mst1/2 in hepatocytes induce cell proliferation, liver overgrowth and tumor formation *in vivo* [189]. Lats1-deficient mice develop ovarian and soft tissue tumors [190]. Conditional activation and subsequent inactivation of Yap induces reversible liver overgrowth in mice [191] and expansion of small intestine progenitor cells and their dedifferentiation [192]. Such organ overgrowth

defects have also been reported in *Drosophila* upon changing Yki-Sd transcriptional activities [176, 187, 191, 193].

Increasing evidence indicates that deregulation of Hippo pathway components are also implicated in human cancers. Downregulation of Lats1/2 and Mst1/2 expression upon promoter hypermethylation has been reported in breast cancer, soft tissue sarcomas, colorectal cancer, astrocytomas and acute leukemia [194-197]. Mutations in Mob1 and Sav1 have been observed in renal cancer cell lines and in skin melanoma [180, 198]. The upstream hippo pathway members negatively control cell proliferation, their functions are mainly inhibited in human cancers and are therefore classified as tumor suppressors. Yap and Taz, on the other hand, are mainly acting as oncogenes. Yap gene amplification and/or increased expression has been seen in many human cancers [199]. Increased expression of Yap and Taz evokes their nuclear localization, which correlated with breast cancer tumor grade [200-202]. Taz is overexpressed in 20 % of human breast cancers [203] and higher levels of Taz are also found in colorectal cancer, glioblastoma, melanoma and non-small cell lung cancer [199]. Such overgrowth defects triggered by Yap and Taz can be induced via Tead transcription factors [188]. Therefore, cancer therapy aims to disrupt the Yap-Tead or Taz-Tead interface to interfere with oncogenic cell proliferation. For instance, Verteporfin has been shown to inhibit Tead-Yap interaction and reduced liver overgrowth upon Yap overexpression or inactivation of Neurofibromin 2 (Nf2, also called Merlin; an activator of the Hippo kinase cascade) in mice [204]. Another compound, Dobutamine, has been identified in a drug screen to induce cytoplasmic localization of Yap and by this suppresses Yap-Tead transcriptional activity in the nucleus [205]. Recently, Statins, used for treatment of hypercholesterolemia, have been shown to inhibit the nuclear localization of Yap and Taz [206]. Ideally, such and future drugs will be tested in cancer patients with amplified or overexpressed Yap and Taz genes.

Tead/Yap/Taz in cell (de)differentiation

Several studies in human untransformed MCF10A cells have demonstrated that forced expression of Yap and Taz not solely induced cell proliferation, but also promoted EMT, cell migration, invasion and anchorage-independent growth via Tead transcription factors. The Yap/Taz/Tead complex induces tumorigenic cell properties and controls target genes like CTGF, CYR61, KI67, AXL, c-MYC and SURVIVIN [170,

187, 191, 207-211]. Expression of Yap-Tead further promotes tumor formation and breast cancer metastasis *in vivo* [188, 212]. More and more evidence accumulated and reveal that the Yap/Taz/Tead complex is also important for stem cell maintenance. In undifferentiated embryonic, neural and hematopoietic cells the expression of Tead2 and Yap is enriched [213]. Additionally, Yap is activated in induced pluripotent stem (iPS) cells [214]. Similar reports in tissue-specific stem cells have shown that high levels of Yap and Taz induced stem cell expansion and blocked their differentiation [192, 215-217]. These results suggest that the Yap/Taz/Tead complex controls cell proliferation and (de)differentiation, processes enabling tissue development and homeostasis and can further force tumorigenesis if deregulated. In the past years, several research groups have investigated which stimuli can influence endogenous Yap/Taz localization and transcriptional activity aside from the control of the Hippo pathway. Interestingly, some of the factors or mechanisms are implicated in the EMT process:

Apical epithelial polarity complex: The Crumbs polarity complex consists of Crumbs, Pals and Patj, is localized at the apical plasma membrane in epithelial cells and becomes delocalized upon EMT. It has been shown that this complex interacts with phosphorylated Yap and Taz and mediates their cytoplasmic localization in densely growing mammary epithelial cells [218] (Figure 9B). Additionally, Yap and Taz can also interact with Smad2/3/4, which in turn is required for nuclear localization of Smads upon TGFβ stimulation [219]. In confluent cultures, Smads are activated by TGFβ, however, fail to translocate to the nucleus. Subsequent Hippo pathway inactivation (loss of Lats1 or 2) or disruption of the Crumbs complex (loss Crb3 or Pals1) releases Yap/Taz from the membrane and induces their nuclear co-localization with Smads upon TGFβ treatment [218]. These findings suggest that the Hippo pathway and epithelial cell polarity control Yap/Taz cellular localization and thereby interfere with canonical TGFβ signaling.

Tight and adhesion junctions: α-catenin links the Actin cytoskeleton to Cadherin adhesion junctions in epithelial cells. It has been observed that α-catenin negatively controls the activity of Yap by binding to the phosphorylated cytoplasmic Yap/14-3-3 complex and localizes it to cell adhesion junctions [217, 220] (Figure 9B). Loss of α-catenin, β-catenin or disruption of adhesion junctions by calcium depletion induces the nuclear localization of Yap [221, 222] and hyperproliferation of human keratinocytes [217].

Angiomotin (Amot) protein family members are important for tight junction assembly and epithelial cell polarity [223]. Amot1 and 2 can directly interact with Yap and Taz, localize them to tight junctions and thereby inhibit Yap/Taz transcriptional activity [224-226]. Downregulation of Amot2 in epithelial MDCK cells leads to Yap/Taz nuclear localization and transcriptional activation [224]. Furthermore, loss of Amot2 in epithelial MCF10A cells promotes nuclear localization of Yap and EMT thereby phenocoping an overexpression of Yap in MCF10A cells [226].

Actin cytoskeleton organization: Mechanical cues coming from cell stretching/crumpling and ECM elasticity influence cell shape. Cells can sense and adapt to such cues and stimulate proliferation, differentiation, apoptosis and migration [227]. Interestingly, all these physical stimuli affect Yap and Taz cellular localization and transcriptional activity [228-231]. For instance, cells plated on a stiff matrix display a nuclear localization of Yap and Taz. Conversely, Yap and Taz are excluded from the nucleus in cells grown on a soft matrix [228]. Additionally, round, compact cells display cytoplasmic Yap and Taz, whereas spreaded, stretched cells exhibit nuclear Yap and Taz [228, 229]. These findings suggest that mechanical cues coming from cell substrate stiffness and cell shape directly control Yap/Taz localization and activity. Interestingly, confluent, stretched epithelial cells display partial nuclear Yap and Taz. Here, mechanical signals partially uncouple Yap/Taz regulation from the Hippo pathway and overcome cell contact-mediated proliferation inhibition [229].

Rearrangement of the Actin cytoskeleton can translate and adapt cell shape and behavior to the above-mentioned cellular mechanical cues [227]. Consequently, interfering with Actin cytoskeleton polymerization, more precisely with the Actin polymerization inhibitors CapZ, Cofilin and Gelsolin, seems to disrupt the regulation of Yap/Taz cellular localization and Yap/Taz/Tead transcriptional activity [229-231]. Surprisingly, the impact of Actin polymerization on Yki regulation seems to be conserved in *Drosophila*, since loss of Actin capping proteins in the wing disc promotes Yki transcriptional activity and tissue overgrowth [232, 233]. Additionally, Rho GTPases have been reported to trigger Yap dephosphorylation and nuclear localization. Aside Rac1 and Cdc42, RhoA strongly increased Yap transcriptional activity, however the mechanism behind this regulation is so far unknown [228, 231, 234]. These findings suggest that the Actin cytoskeleton is a major regulator of

Yap/Taz/Tead activity (Figure 9B). However, the exact mechanism how F-Actin and Rho GTPases regulate Yap/Taz localization has to be investigated.

Collectively, loss of epithelial polarity and cell junctions, Actin stress fiber formation and ECM remodeling are crucial events during EMT and facilitate cell migration and tumor progression [4]. One can speculate whether such events reinforce Yap/Taz/Tead transcriptional activity during EMT in addition to other EMT inducers (e.g. TGFβ signaling). In the past years, a broad collection of various stimuli and pathways has been discovered, which all influence Yap/Taz/Tead activity [235]. However, the physiological context of Yap/Taz/Tead activity during tissue development, homeostasis and disease has to be investigated in the future.

1.3.4. Post-transcriptional control of EMT by ncRNAs

More than 80 % of the human genome can be transcribed into RNA, however less than 2 % of the genome contains protein-coding capacity [236]. In the past decades, more and more evidence accumulated demonstrating that several non-coding RNA (ncRNA) species are powerful regulators of gene expression and therefore heavily influence most cellular processes [237, 238].

1.3.4.1 miRNAs

MicroRNAs (miRNAs) are a class of ncRNAs with an approximate length of 22 nucleotides. They are ubiquitously expressed, conserved across species and control approximately 60 % of all human protein coding genes on the post-transcriptional level [239, 240].

Canonical miRNA biogenesis and function

In 1993, the group of Victor Ambros has identified the first miRNA in *Caenorhabditis elegans* [241] and demonstrated an essential function for ncRNA lin-4 during larval development. Lin-4 reduced the protein levels of lin-14, which possesses several lin-4 complementary sequence elements within its mRNA 3` untranslated region (UTR). Therefore, the authors have suggested an antisense RNA-RNA interaction of lin-4 and lin-14, where lin-4 negatively controls lin-14

translation. Since 1993, more than 2500 mature human miRNAs have been identified and annotated (miRBase version 21) [242]. Beside a few reports demonstrating an activation of mRNA translation, miRNAs are mainly described as post-transcriptional repressors of RNAs [243-246].

MiRNAs are mainly located within intronic regions of coding or non-coding host genes or in intergenic regions, latter exhibiting their own gene promoter [240, 247]. Few miRNAs have been reported to be located in exonic regions of genes [248]. In the canonical miRNA biogenesis pathway (Figure 10), miRNAs are transcribed by RNA polymerase II (Pol II) either as monocistronic or polycistronic primary transcript (pri-miRNA) cluster. Pri-miRNAs contain a local hairpin structure, which exhibit the mature miRNA sequence. The nuclear RNase III Drosha binds to the double stranded RNA-binding protein DiGeorge syndrome chromosomal region 8 (DGCR8) to form the microprocessor complex, which endonucleolytically cleaves the pri-miRNA to release an ~ 65 nucleotide long miRNA precursor molecule (pre-miRNA). Several non-canonical miRNA biogenesis pathways, leading to the production of pre-miRNAs in a Drosha/DGCR8- or Dicer-independent manner, have been reported and reviewed [249]. Following Drosha/DGCR8 processing, Exportin 5 binds to pre-miRNAs to actively transport them to the cytoplasm via the cofactor RAN-GTP. Here, the appropriate length and 3` overhangs of the miRNA precursors are essential for a successful nuclear export by Exportin 5. Dicer, a second RNase III enzyme in the miRNA biogenesis pathway, binds to the RNA-binding cofactor TRBP and cleaves the pre-miRNA close to the terminal loop to generate a double stranded miRNA:miRNA* complex. Subsequently, only one strand (guide strand) of the miRNA:miRNA* complex is loaded into the miRNA-induced silencing complex (miRISC). The mature miRNA binds via imperfect or perfect base pairing to the 3` or 5` UTRs of the target RNA and induces its translational repression or deadenylation and degradation, respectively [250, 251]. More precisely, the nucleotides 2-7 within the 5` end of a miRNA are termed as seed sequence and bind the target RNA. The group of David Bartel has extensively studied miRNA:RNA interactions and their contextual features. They have presented four types of canonical miRNA:RNA recognition sites, which have different efficacies on target repression: 6-mer site (miRNA positions 2-7), 7-mer-A1 site (miRNA positions 1-7), 7-mer-m8 site (miRNA position 2-8) and 8-mer site (miRNA positions 1-8) (order of recognition sites displays increased target repression potential). Additional recognition sites or base pairing of

nucleotides 12-17 in the mature miRNA can further enhance 7mer-8 site-mediated target repression [252].

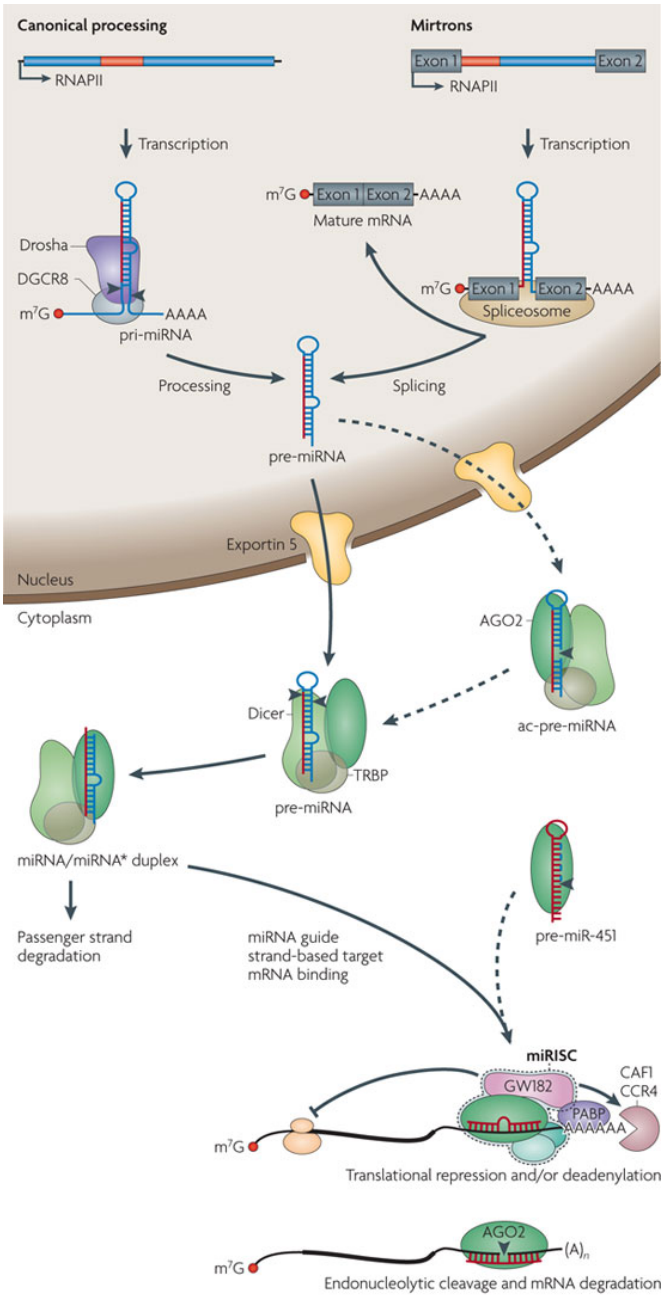


Figure 10: miRNA biogenesis. MiRNAs are transcribed from intergenic or intronic genomic regions by RNA polymerase II (RNAPII). During canonical miRNA processing, the microprocessor complex consisting of DiGeorge syndrome chromosomal region 8 (DGCR8) and nuclear RNase III Drosha, binds and cleaves the pri-miRNA to release a ~ 65 nucleotide pre-miRNA, which is transported to the cytoplasm via Exportin 5. The RNase III Dicer bound to the RNA-binding cofactor TRBP cleaves the pre-miRNA to generate the ~ 20 nucleotide double stranded miRNA:miRNA* complex. The guide strand (red) of the miRNA duplex is loaded into the miRNA-induced silencing complex (miRISC) and binds to a complementary RNA sequence in the target RNA and induces its translational repression or deadenylation and degradation (adapted from [261]).

An important aspect of miRNA function is that one miRNA can control the transcripts of several RNAs (pleitropy) and one RNA can be targeted by different miRNAs [239, 246, 253, 254]. Furthermore, miRNAs are often positioned within gene regulatory networks such as feedback or feedforward loops with TFs [255, 256]. Therefore, miRNAs are thought to function as molecular switches in cells to allow fine-tuning of essential cellular processes such as proliferation, (de)differentiation

and survival during normal development and homeostasis. They further endow these processes with the required robustness [257-259] and, consequently, deregulation of miRNAs is implicated in diseases such as cancer [260].

miRNAs in EMT

In order to identify new EMT-regulatory miRNAs, several research groups have profiled global miRNA expression in various cellular EMT models mainly using microarray analysis. They have either extracted and characterized individual differentially expressed miRNAs during EMT or generated a miRNA signature specific for an epithelial or mesenchymal cell phenotype. Since there is an imperfect miRNA expression profile overlap comparing different studies, one can conclude that miRNA expression and regulation is highly cell context dependent [254, 262-267].

miR-200 family: the guardian of an epithelial morphology

Strikingly, one miRNA family has appeared in most of the previously mentioned miRNA expression profiling studies. Members of the miR-200 family are associated with an epithelial cell morphology and their expression is decreased upon dedifferentiation across various cell types and EMT-inducing stimuli [254, 262-264]. The miR-200 family consists of five family members and is organized as two gene clusters on chromosome 1 (miR-200a/b, miR-429) and chromosome 12 (miR-200c, miR-141) in the human genome [265]. Ectopic expression of individual family members blocked TGFβ-induced EMT in normal murine mammary gland cells (NMuMG) [268] or induced MET in mesenchymal canine kidney cells (MDCK) cells *in vitro* [254]. In both studies, forced expression of miR-200s led to a stabilization of E-cadherin adhesion junctions and transcriptional downregulation of the EMT TFs Zeb1 and Zeb2. Here, miR-200 family members directly repress the transcripts of Zeb1/2 on the post-transcriptional level by binding to specific recognition sites in their 3` UTRs [269]. Conversely, TGFβ-induced Zeb1 binds to specific motifs in the miR-200 promoters to negatively control their transcription during EMT. This reciprocal regulation of Zeb1/2 and miR-200 family members not solely determines cell morphology but also controls cell migration and invasion *in vitro*, for instance, of the highly metastatic human MDA-MB-231 breast cancer cells [269, 270] (Figure 11).

Recently, a genome-wide analysis of direct mRNA targets of miR-200a and miR-200b in MDA-MB-231 cells has been published [271]. MiR-200a/miR-141 and miR-200b/c/miR-429 form two seed classes within the miR-200 family where the miRNA seed regions differ in only one nucleotide. In this way, the authors have covered both distinct sets of miR-200 mRNA targets in their analysis. Using Argonaute high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP) and subsequent barcoded cDNA analysis, Bracken et al. have identified around 2000 transcripts directly bound by miR-200a and miR-200b, which are mainly associated with cytoskeleton remodeling and genes implicated in EMT, TGFβ-, Wnt- and Rho-ROCK-signaling pathways. This study has demonstrated that the miR-200 family can target multiple genes implicated in cell architecture and transcriptomic remodeling crucial during EMT and positioned the miR-200 family as a central regulator of epithelial cell plasticity [270-272].

miRNAs control cellular architecture during EMT

The epithelial cell-cell adhesion junction protein E-cadherin facilitates normal epithelial tissue formation and its loss induces EMT *in vitro* and promotes tumor cell dissemination and local invasion in carcinomas [6, 28, 273, 274]. Several studies have shown that E-cadherin is directly targeted by several miRNAs during EMT and malignant tumor progression [275-278] (Figure 11).

MiR-9 is induced by MYC/MYCN and reduces E-cadherin levels post-transcriptionally, thus promoted breast cancer cell motility and invasiveness *in vitro*. Furthermore, loss of E-cadherin activated β-catenin signaling and VEGFA expression, which further triggered tumor angiogenesis and metastasis of the orthotopically transplanted human breast cancer cell line SUM149 in immune-deficient mice [275].

The same group has identified miR-10b to be strongly expressed in metastatic breast cancer cell lines such as MDA-MB-231 compared to non-metastatic SUM149 cells. Forced expression of miR-10b in SUM149 cells induced cell migration/invasion and lung metastasis in immune deficient mice [279]. In addition, E-cadherin is directly regulated by miR-10b and their expression is inversely correlated during human breast cancer development [276, 280].

MiR-23a is induced by canonical TGFβ signaling and directly targets E-cadherin in human lung adenocarcinoma epithelial cells (A549). Ectopic expression of miR-23a induced an EMT-like morphology of A549 cells [277]. Additionally, miR-92a belongs to the oncogenic miR-17-92 gene cluster and targets E-cadherin mRNA 3`UTR. Their expression levels are inversely correlated in esophageal squamous cell lines (ESC) and ESC carcinoma (ESCC) tissues and high levels of miR-92a in ESCC is associated with lymph node metastasis and a poor survival rate of ESCC patients [278].

While E-cadherin is relocalized and downregulated during EMT, N-cadherin is often upregulated (Cadherin switch) to promote cell motility and migration [4]. MiR-194 negatively controls N-cadherin expression and reduces liver metastasis of transplanted human liver cancer cells in immune deficient mice [281].

The small GTPase RhoA controls actin cytoskeleton organization and epithelial cell polarity during EMT. MiR-155 is transcriptionally upregulated by canonical TGFβ signaling in normal mouse epithelial cells and directly targets and reduces RhoA mRNA levels. Moreover, inhibition of miR-155 expression prevents TGFβ-induced EMT [262].

The mesenchymal intermediate filament Vimentin is a post-transcriptional target of miR-506. Overexpression of miR-506 blocks TGFβ-induced EMT in MCF10A cells and suppresses the migratory/invasive properties of MDA-MB-231 cells *in vitro* [282]. Additional miRNAs, such as miR-30a [283], miR-17-3p [284], miR-124 and miR-203 [285], have been reported to negatively regulate Vimentin expression.

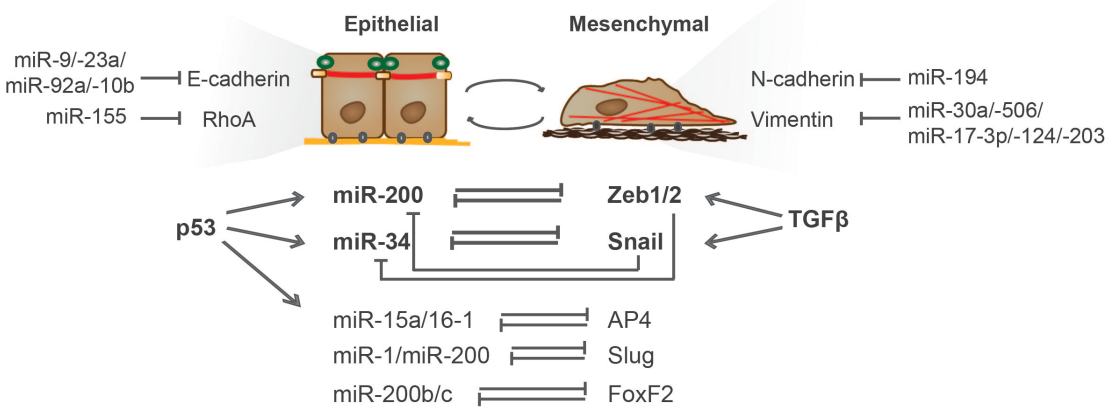


Figure 11: MiRNAs control cell architecture and TFs during EMT. *Top*: Different miRNAs directly control the expression of epithelial and mesenchymal cell architecture proteins on the post-transcriptional level. *Bottom*: Double negative feedback loops between miRNAs and TFs regulate EMT plasticity. p53 induces the

expression of miR-200 family members, miR-34 and miR-15a/16-1 gene cluster in epithelial cells. TGFβ triggers the expression of Zeb1/2 and Snail during EMT initiation.

miRNAs targeting EMT-regulatory TFs

The TF families Zeb, Snail and Twist are expressed by different EMT-inducing stimuli and are considered EMT drivers (see 1.3.3.1). Importantly, these TFs are highly controlled on the post-transcriptional level by miRNAs.

The key EMT TFs Zeb1 and Zeb2 are tightly controlled during EMT by the miR-200 family as mentioned above [254, 265, 268, 269]. Similar to the regulation of miR-200s, miR-203 is also downregulated during TGFβ-induced EMT in MDCK cells and in mesenchymal breast cancer cells [254]. Ectopic expression of miR-203 directly represses Zeb1 mRNA levels and, consequently, stabilizes E-cadherin junctions, however the effect is not as strong as for the miR-200 family. MiR-130b expression has been reported to be downregulated during p53 mutant-induced EMT in endometrial cancer cells. Restoration of miR-130b in mesenchymal endometrial cancer cells impairs cancer cell invasion by directly controlling the cellular levels of Zeb1 [286]. Snail is directly regulated by miR-30 in murine hepatocytes and TGFβ-induced EMT in AML12 cells leads to a downregulation of miR-30b/c/d/e along with an upregulation of Snail. Overexpression of miR-30b triggers the repression of Snail protein levels and blocks TGFβ-induced EMT and AML12 cell migration [287]. The other TF family member Slug is post-transcriptionally regulated by miR-1 and miR-200b. Both miRNAs are repressed upon TGFβ-induced EMT in a murine epithelial prostate cancer cell line. Forced expression of miR-1 and miR-200b prevents Slug expression, induces MET and inhibits invasion of human prostate cancer cells [288]. MiR-580 targets the 3` UTR of Twist1 mRNA and represses its expression in human mammary epithelial MCF10A cells [289]. These miRNAs are a few selected examples of a plethora of reports about post-transcriptional regulators controlling master EMT TFs.

One important mechanism that might be (partially) responsible for the reversibility of cell differentiation and dedifferentiation are double-negative feedback loops between miRNAs and key EMT TFs, functioning as molecular switches (Figure 11). Zeb1/2 and Snail TFs are induced upon TGFβ-signaling [139]. Both of them can directly target the promoter regions of miR-200 and miR-34 family members, inhibiting their transcription and inducing cell dedifferentiation, survival and stemness.

MiR-200 and miR-34 are embedded in a double-negative feedback loop with Zeb and Snail. Wildtype p53 can directly activate transcription of miR-200 and miR-34 family members, which in turn directly repress Zeb and Snail and subsequently initiate epithelial cell differentiation [143]. p53 has been reported to control another double-negative feedback loop during EMT/MET plasticity in colorectal cancer. Upon DNA damage, p53 induces the downregulation of the TF AP4 via miR-15a and miR-16-1. Both miRNAs target the 3` UTR of AP4 and induce MET, decrease cell migration, invasion and metastasis of colorectal cancer cells in a xenograft model. Conversely, AP4 negatively controls the transcripts of miR-15a/16-1 by binding to two alternative transcriptional start sites within an intronic region [290].

Slug has been reported to regulate and to be regulated by miR-1 and miR-200 during EMT-mediated progression of prostate adenocarcinoma [288]. Similarly, miR-200b/c and Forkhead-box F2 (Foxf2) control their transcript levels in a reciprocal feedback loop to regulate EMT, invasiveness and metastasis of lung cancer cells [291].

Accumulating evidence reporting double-negative feedback loops between miRNAs and TFs suggest their importance during reversible (de)differentiation processes during embryonic development as well as in malignant tumor progression and are currently and in the future an area of active research.

Regulation of miRNA expression during EMT

Different extracellular signals can stimulate the activity of TFs, which initiate and maintain the overall transcriptomic changes during EMT [138]. Further, TFs can directly regulate the expression of individual miRNAs or miRNA clusters and either activate or repress their transcription. In the context of EMT, an important group of TFs like Zeb1/2, Snail, Slug and p53 controlling miRNA expression has been pointed out earlier [39, 288].

Epigenetic modifications can also control and organize miRNA expression during EMT and has been recently reported for the miR-200 family. The transcriptional start sites of the miR-200a/b/429 and miR-200c/141 gene clusters are located within canonical CpG islands and, therefore, exhibit potential sites for hypermethylation-mediated gene silencing. It has been demonstrated that upon TGFβ-induced EMT of MDCK cells the promoter regions of the two miRNA clusters get hypermethylated

accompanied with a decrease in miR-200s expression. Interestingly, TGFβ withdrawal induces MET of MDCK cells and a reversion of the DNA methylation marks. Hypermethylation of miR-200 gene clusters is also observed in different human cancer cell lines and treatment with 5′-aza-2′-deoxycytidine, a Dnmt inhibitor, restores the expression of miR-200 family members in the cells [102]. These results indicate a dynamic epigenetic regulation of miR-200 expression during EMT/MET.

A genome wide analysis of miRNA gene sequences has revealed that about half of the genes are associated with CpG islands [292]. Bisulfite genomic sequencing analysis has demonstrated that also the miR-34b/c cluster among others is strongly hypermethylated in metastatic tumor cells. Reintroduction of miR-34b/c represses cell migration *in vitro* and metastasis formation in a mouse xenograft model [293]. These findings further demonstrate the importance of epigenetic miRNA expressional regulation during EMT and tumor progression.

miRNAs can exert tumor suppressive and oncogenic functions

In 2002, Calin and coworkers have reported the first tumor suppressor miRNA implicated in chronic B-cell lymphocytic leukemia (B-CLL). They have found that the miR-15a/16-1 gene cluster on chromosome 13q14 is frequently deleted or downregulated in 68 % of B-CLL patients [294]. Subsequent miRNA expression analyses in tumor versus normal tissue has provided evidence that miRNAs are deregulated in cancer cells [295, 296] and can contribute to human cancer initiation and progression [260, 297]. Deregulation of miRNA expression has been shown to have different causes: (a) genomic miRNA gene amplifications, translocations, deletions [298]; (b) aberrant upstream TF activation [299-302]; (c) altered epigenetic modifications like promoter methylation [293] or histone acetylation at genomic loci of miRNA coding genes [303]; (d) deregulation of miRNA biogenesis pathway components (e.g. Dicer1) inducing impaired miRNA processing and a global reduction in mature miRNAs [304]. Furthermore, alterations in miRNA target site sequence (point mutations in the seed match, translocations or 3` UTR shortening of the mRNA sequence) [305-307] or aberrant expressed competitive endogenous RNAs (ceRNAs) [308, 309] can lead to an escape from miRNA repressive control in cancer.

Up to date, several studies have reported a deregulation of miRNAs in breast cancer analyzed by microarray or deep sequencing analysis [296, 310-313]. Together with the large amount of gene expression, genomic and clinical data from The Cancer Genome Atlas (TCGA), deregulated miRNAs have been further analyzed and linked to different breast cancer subtypes [313, 314].

One of the first human miRNAs discovered to affect breast cancer metastasis has been miR-335 [315]. MiR-335 is found to be lost in metastatic breast cancer cells caused by genetic gene deletion [316]. Loss of miR-335 induces expression of Sox4 and Tenascin-C, which promotes cancer cell migration, invasion and metastasis [315]. Other identified “metastaMirs” (metastasis-regulating miRNAs) in breast cancer are for instance: let-7 [317], miR-10b [279], miR-9 [275], miR-21 [318] and they mainly control tumor cell invasiveness.

A direct link coupling breast cancer metastasis and EMT/MET cell plasticity has been reported for a few miRNAs. The miR-200 family has been identified as EMT suppressor by directly targeting the TFs Zeb1 and Zeb2 to maintain an epithelial cell morphology and inhibit cell migration and invasion of breast cancer cells [254, 265, 268, 269]. MiR-200 family members are transcriptionally downregulated in tumors enriched in an EMT signature such as mesenchymal endometrial carcinomas [267] and in the claudin-low breast cancer subtype [319]. Doxycycline-induced expression of the miR-200c/141 cluster in transplanted Claudin-low tumor cells induces a MET gene signature and decreases lung metastasis *in vivo* [320]. Additionally, forced expression of miR-200b in a mutant K-ras/p53 lung cancer model blocks TGFβ-induced EMT, invasion and metastasis *in vivo* [321]. In these studies, the miR-200 family has functioned as early suppressors of the tumor invasion metastasis cascade by controlling EMT/MET processes.

However, another study has reported a pro metastatic role for miR-200s in breast cancer, as high expression of miR-200 family members is associated with decreased distant relapse-free survival in breast cancer patients [322]. Forced expression of miR-200s in mesenchymal, weakly metastatic mouse tumor cells (4TO7) has induced an epithelial phenotype *in vitro*, but enhanced lung colonization of 4TO7 cells injected intravenously in mice. These findings are consistent with a previous report [323] and support the notion that at later stages in the metastatic cascade miR-200 family members can induce a MET process favoring tumor cell colonization and lethal macrometastases formation at a secondary site. In addition, miR-200 family

members can promote breast cancer metastatic outgrowth by regulating other genes than the ones controlling EMT/MET cell plasticity. For instance, miR-200s target the COPII vesicle component Sec23a, leading to reduced secretion of the metastasis suppressing factors TINAGL1 and IGFBP4 [322], thereby, generating a pro-metastatic niche for tumor cells. In summary, miR-200 family members seem to function as tumor suppressors and tumor promoters, which may depend on their activation in early or late events during the metastatic cascade. Clearly more investigations are required to dissect the function of miR-200s during breast cancer progression.

Other miRNAs such as miR-29a and miR-103/107 have been described as crucial modulators of EMT in breast cancer progression. MiR-29a is upregulated in mesenchymal, metastatic mouse mammary cells compared to their epithelial counterparts. Overexpression of miR-29a in epithelial cells induces EMT and metastasis *in vivo* and its high expression is observed in metastatic human breast cancers [266]. MiR-103/107 expression is also associated with breast cancer metastasis and a poor prognosis for breast cancer patients [324]. MiR-103/107 directly repress Dicer1 mRNA and induce EMT, migration/invasion and metastasis in human mammary epithelial cells *in vivo*. Inhibition of miR-103/107 blocked these cellular effects. Generally, a loss of the miRNA biogenesis component Dicer1 has been reported to promote tumorigenesis by decreasing global mature miRNA expression [325, 326], a phenomenon observed in human breast and other cancers [295, 314].

miRNAs in cancer therapy

In the past, several studies have clearly demonstrated that miRNAs can function as oncogenes or tumor suppressors and, therefore, have a central role in the initiation and progression of various cancer types [327]. To interfere with the deregulation of specific miRNAs, two promising approaches have been established for miRNA-based therapies in cancer:

(a) Sequence-specific inhibition of miRNAs to block their oncogenic function (targeting therapy). Antisense oligonucleotides bind with perfect sequence complementarity to endogenous mature miRNAs to inhibit their function. Ma and colleagues have demonstrated that systemic delivery of an anti-miR-10b

oligonucleotide reduced lung metastasis in a mouse orthotopic transplantation breast cancer model [328].

(b) Delivery of small double stranded RNAs to mimic mature miRNAs and restore their tumor suppressive function, which is reduced or lost in cancer cells (replacement therapy). MRX34 is the first miRNA mimic that has entered phase I clinical trials for miRNA-based cancer therapy. MRX34 replaces the function of miR-34 in patients with primary liver cancer or liver metastasis [329].

Still, major challenges for miRNA-based therapies remain. For instance, enhance oligonucleotide stability and their effective delivery into target cells. For this purpose, oligonucleotides can be encapsulated in liposomes or polymer-based nanoparticles to protect them from serum nucleases. Additional chemical oligonucleotide modifications (2`-O-methyl (2`-O-Me), 2`-O-methyoxyethyl (2`-O-MOE) and locked nucleic acids (LNAs), 3` end cholesterol groups)) have also helped to increase their stability, delivery through cell membranes and their tissue distribution [330]. However, “unknowns” for miRNA-based therapy still remain, for instance, the levels of miRNA expression in cancer cells achieved by the replacement therapy is hard to estimate. Potential off-target effects evoked by introducing DNA/RNA inhibitors or mimics into cells are unexplored. Future investigations have to address these “unknowns” to utilize miRNA-based therapies in cancer treatment.

1.3.4.2 piRNAs

PIWI interacting RNAs (piRNAs) are a class of ncRNAs of 24-30 nucleotides in length, expressed as single stranded RNAs, produced in a dicer-independent manner and are involved in epigenetic regulation (DNA methylation) and transposon silencing. Two studies have revealed that piRNA expression can be deregulated in cancer cell lines and primary tumors [331, 332]. PiRNA-Hep1 (PiT-Hep1) expression was upregulated in hepatocellular carcinoma tumors compared to normal liver tissue and loss and gain of function experiments of piR-Hep1 in hepatocellular carcinoma cells induced a decrease or increase in cell migration and invasion, respectively [333]. These results indicate that piRNAs can be involved in cancer development, whether they regulate EMT in this context, has to be investigated in the future.

1.3.4.3 lncRNAs

Long non-coding RNAs (lncRNAs) are a group of transcripts with a length of > 200 nucleotides. They mediate gene expression by multiple (post)-transcriptional mechanisms and are involved in many biological processes and in tumorigenesis [334].

LncRNAs such as HULC [335], Malat1 [336] and HOTAIR [337] are aberrantly expressed in different human cancer types and affect tumor initiation and progression [335, 337-340]. Recently, the lncRNA lncRNA-ATB has been identified and characterized as post-transcriptional regulator of oncogenic EMT [341]. LncRNA-ATB is highly expressed in hepatoma cells undergoing TGFβ-induced EMT and functions as natural competitive endogenous RNA (ceRNA) for Zeb1 and Zeb2. LncRNA-ATB exhibits several miR-200 binding sites and expression of lncRNA-ATB reduces the levels of “free” miR-200 family members, therefore, stabilizing Zeb1/2 mRNAs during EMT. Forced expression of lncRNA-ATB induces EMT, invasion of hepatoma cells *in vitro* and metastasis *in vivo*. At the same time, lncRNA-ATB binds IL11 mRNA, stabilizes its expression and activates STAT3 signaling, which promotes tumor cell colonization at a secondary site [341]. In addition, an EMT lncRNA expression profile of normal mouse mammary cells (NMuMG) induced to undergo EMT upon TGFβ treatment has been recently published and the authors identified lncRNA-HIT as a new EMT-regulatory lncRNA. LncRNA-HIT expression has been further associated with invasive human breast carcinomas [342], however, the mechanisms by which lncRNA-HIT regulates EMT are still unknown.

1.3.4.4 circRNAs

Circular RNAs (circRNAs) have already been discovered in 1991 [343], however, novel bioinformatic approaches and deep sequencing technologies have identified these RNAs as new class of highly abundant, evolutionary conserved ncRNAs [344, 345]. The group of Gregory J. Goodall has recently demonstrated that hundreds of circRNAs are regulated during TGFβ-induced EMT of human mammary epithelial cells (HMLE). The RNA binding protein and alternative splicing factor Quaking (QKI) binds to sites flanking circRNA-forming exons and promotes their generation during EMT [346]. Two circRNAs, Cdr1 and Sry, have been described to function as miRNA decoys, thereby, controlling miRNA inhibitory functions [347, 348]. Since circRNA

expression seems to be highly regulated during EMT, future questions should address their mechanistic function in EMT/MET cell plasticity and tumor progression.

2 Aim of the study

The gradual and highly plastic process of an EMT is essential during embryogenesis to allow the formation of various tissues and organs. During malignant tumor progression, EMT induces epithelial tumor cell dedifferentiation and thereby promotes tumor cell invasion and dissemination into the surrounding tissue – the initial steps for metastatic spread. In the past years, intensive research has demonstrated that interconnected transcriptional, post-transcriptional and post-translational networks can orchestrate the dynamic nature of EMT/MET processes in pathological contexts. These findings highlight the importance of a tight regulation of these networks and their aberrant activation can induce malignant tumor progression.

This work has aimed at identifying and characterizing EMT-regulatory mechanisms, which are implicated and conserved during normal and cancer-associated cell dedifferentiation and contribute to cancer metastasis *in vivo*. We have used normal and tumorigenic mammary epithelial cell lines, which undergo a stepwise EMT upon TGFβ-treatment, and have followed the morphological and transcriptional changes over time in order to:

- Identify critical factors, particularly TFs and miRNAs, regulating EMT plasticity
- Understand their molecular regulation during EMT, downstream targets and EMT-related functional outputs *in vitro*
- Test their impact on tumor progression and metastasis *in vivo*

Tumor metastasis is the main cause of cancer-related death in humans. Understanding the underlying cellular mechanisms of tumor cell dissemination and outgrowth at distant sites enables us to pharmacologically interfere and restrict these processes in the future.

3 Results

3.1 Tead2 expression levels control the subcellular distribution of Yap and Taz, zyxin expression and epithelial-mesenchymal transition

Maren Diepenbruck[1]*, Lorenz Waldmeier[1]*, Robert Ivanek[1], Philipp Berninger[2], Phil Arnold[2], Erik van Nimwegen[2], and Gerhard Christofori[1]

[1] Department of Biomedicine, University of Basel, 4058 Basel, Switzerland

[2] Biozentrum, University of Basel, and Swiss Institute of Bioinformatics, 4056 Basel, Switzerland

* Equally contributing authors.

- published -

April 2014 issue of Journal of Cell Science

3.1.1 Abstract

The cellular changes during an epithelial–mesenchymal transition (EMT) largely rely on global changes in gene expression orchestrated by transcription factors. Tead transcription factors and their transcriptional co-activators Yap and Taz have been previously implicated in promoting an EMT; however, their direct transcriptional target genes and their functional role during EMT have remained elusive. We have uncovered a previously unanticipated role of the transcription factor Tead2 during EMT. During EMT in mammary gland epithelial cells and breast cancer cells, levels of Tead2 increase in the nucleus of cells, thereby directing a predominant nuclear localization of its co-factors Yap and Taz via the formation of Tead2–Yap–Taz complexes. Genome-wide chromatin immunoprecipitation and next generation sequencing in combination with gene expression profiling revealed the transcriptional targets of Tead2 during EMT. Among these, Zyxin contributes to the migratory and invasive phenotype evoked by Tead2. The results demonstrate that Tead transcription factors are crucial regulators of the cellular distribution of Yap and Taz, and together they control the expression of genes critical for EMT and metastasis.

3.1.2 Introduction

Epithelial–mesenchymal transition (EMT) is a cell-biological program that is required at various stages of embryonic development. Activation of EMT in epithelial cells induces a loss of cell–cell adhesions and apical-basal polarity, and promotes trans-differentiation into a mesenchymal state, which is characterized by a migratory and invasive phenotype [19, 349]. During solid tumor progression, a reactivation of some of these features in epithelial tumor cells (oncogenic EMT) is regarded as one of the mechanisms that can facilitate metastatic spread [2, 26]. Oncogenic EMT not only provides tumor cells with invasive properties that permit dissemination from the primary tumor, but also results in the acquisition of stem-cell-like traits, which has implications for cancer therapy and might also be important for colonization at distant organs [26, 350-352]. Among the many genes and signaling pathways active during EMT, transcription factors are the master coordinators of the EMT program [21, 349, 353].

The Hippo tumor suppressor signaling pathway plays a critical role in restricting organ size by antagonizing the oncogenic transcriptional co-activators Yap and Taz [224, 354]. A complex network of cell adhesion and signaling molecules, including the tumor suppressor neurofibromin-2/Merlin, regulates the Hippo kinase cascade, leading from the protein kinases Mst1 and Mst2 via the protein kinases Lats1 and Lats2 to the transcriptional co-factors Yap and Taz. When the Hippo pathway is active, Yap and Taz are phosphorylated by Lats1 and Lats2, and phosphorylated Yap and Taz are retained in the cytoplasm. In the absence of activated Hippo signaling, unphosphorylated Yap and Taz are imported into the nucleus where they, together with Tead DNA-binding transcription factors, activate the expression of proliferative and anti-apoptotic genes.

In mammals, Tead transcription factors comprise a family of four members (Tead1–Tead4). They are ubiquitously expressed [155, 156] and exert partially redundant roles in regulating the development of various embryonic tissues, including neural crest, whose formation depends on EMT [2], notochord and trophectoderm [159-161, 355]. Transcriptional activity of the DNA-binding Teads requires their physical association with the transcriptional co-activators Yap or Taz [169, 356]. Upon Yap- or Taz-mediated activation, Teads can exert multiple functions. For example, they control proliferation in epithelial cells and fibroblasts [188, 209, 210]. Moreover, Yap and Taz are sufficient to induce EMT of MCF10A human breast epithelial cells in a Tead-dependent manner [208-210], and the nuclear accumulation of Yap and Taz in EpH4 murine mammary epithelial cells is required for these cells to undergo TGFβ-induced EMT [218]. Finally, elevated levels of Yap trigger increased tumor growth and a pro-metastatic phenotype through binding to Tead in breast cancer and melanoma cells [212]. Even though these studies clearly demonstrate a crucial role for Teads, Yap and Taz in mediating EMT induction and cancer progression, the mechanisms involved in the regulation of Tead transcriptional activity and the direct target genes during an EMT remain to be identified.

To delineate the mechanisms underlying the transcriptional activities of Teads and to identify their transcriptional target genes during EMT, we have utilized cellular model systems of EMT in non-transformed murine mammary gland epithelial cells and in murine breast cancer cells. We report that the expression of Tead family members is upregulated during EMT, concomitant with an overall increase in Tead transcriptional activity. We demonstrate that elevated levels of Tead2 lead to

increased nuclear localization of Yap and Taz, where they form a complex with Tead2. As a result, increased Tead2 transcriptional activity provokes the induction of EMT and a malignant tumor phenotype. Conversely, knockdown of Teads in cells undergoing an EMT prevents efficient subcellular redistribution of Yap and Taz and blocks EMT. Genome-wide chromatin immunoprecipitation and next generation sequencing (ChIP-Seq) in combination with gene expression profiling identified several EMT-relevant genes as direct transcriptional targets of Tead2 during EMT. Among these, the gene encoding Zyxin, a component of focal adhesions and an actin cytoskeleton remodeling protein, is required for EMT-related migration and invasion.

3.1.3 Results

3.1.3.1 Formation of a nuclear Tead2–Yap–Taz complex and its transcriptional activity during EMT

To identify crucial genes underlying the multiple stages of an EMT, we induced EMT in the untransformed normal murine mammary gland cell line NMuMG [357] by treatment with TGFβ for 0, 1, 4, 7, 10 and 20 days (data not shown). During this time course, the cells underwent progressive EMT and acquired a complete mesenchymal morphology [358]. Motif Activity Response Analysis (MARA) [359] of gene expression data derived from the EMT timecourse predicted that several transcription-factor-binding motifs were important regulators of the EMT expression dynamics, including a motif bound by Tead transcription factors (supplementary Figure S1A). This analysis suggested that genes that contain a species-conserved Tead-binding MCAT motif are upregulated during an EMT of NMuMG cells (supplementary Figure S1A).

To delineate the regulatory role of Tead transcription factors in the EMT process, we utilized NMuMG cells and Py2T murine breast cancer cells derived from a tumor of MMTV-PyMT transgenic mice [360], both of which underwent EMT upon treatment with TGFβ (Figure 1A). MTflECad cells have been established from a mammary tumor of a MMTV-Neu transgenic mouse carrying conditional (floxed) alleles of the E-cadherin gene (*Cdh1*). These cells undergo EMT upon Cre-mediated genetic ablation of the E-cadherin gene (MTΔECad) [358] (Figure 1A).

Gene expression profiling and quantitative RT-PCR revealed that the transcripts of all four Tead family members could be detected in the three model systems

before, during and after EMT, and that only Tead2 mRNA levels were upregulated across all model systems analyzed (supplementary Figure S1B). Based on the robust and reproducible expression of Tead2, we chose to focus on Tead2 and generated polyclonal antibodies specifically detecting Tead2 (supplementary Figure S1C-E). Analysis of endogenous Tead2 expression levels in all three EMT model systems by immunoblotting revealed that Tead2 protein levels increase upon EMT induction (Figure 1B). EMT-associated increased Tead2 expression was solely observed in Py2T cells stimulated with TGFβ, whereas other growth factors like EGF, FGF, HGF, IGF, PDGF or IL-6 were not able to induce Tead2 expression (supplementary Figure S1F). Furthermore, small interfering RNA (siRNA)-mediated ablation of Smad4 expression during TGFβ-induced EMT in NMuMG cells did not affect Tead2 expression levels, suggesting that TGFβ-induced Tead2 expression is independent of canonical TGFβ signaling (supplementary Figure S1G). Notably, siRNA-mediated ablation of Sox4 expression, a transcription factor known to be a crucial regulator of EMT [98], revealed that Tead2 expression is strictly dependent on Sox4 activity during TGFβ-induced EMT in NMuMG and Py2T cells (supplementary Figure S1H).

Immunofluorescence analysis showed that low levels of nuclear Tead2 were present in epithelial cells, yet a much stronger nuclear staining could be observed in cells that were in the process of undergoing EMT (Figure 1C; NMuMG and Py2T) or had undergone EMT (Figure 1C; MTΔEcad). Immunofluorescence staining of the Tead transcriptional co-activators Yap and Taz revealed that they were distributed throughout the cytoplasm and in the nuclei of epithelial cells, as expected in sparsely growing, proliferating cells [175] (Figure 1C). Conversely, NMuMG and Py2T cells undergoing TGFβ-induced EMT or stably mesenchymal MTΔEcad cells displayed predominantly a nuclear localization of Yap and Taz, highly similar to that of Tead2 (Figure 1C).

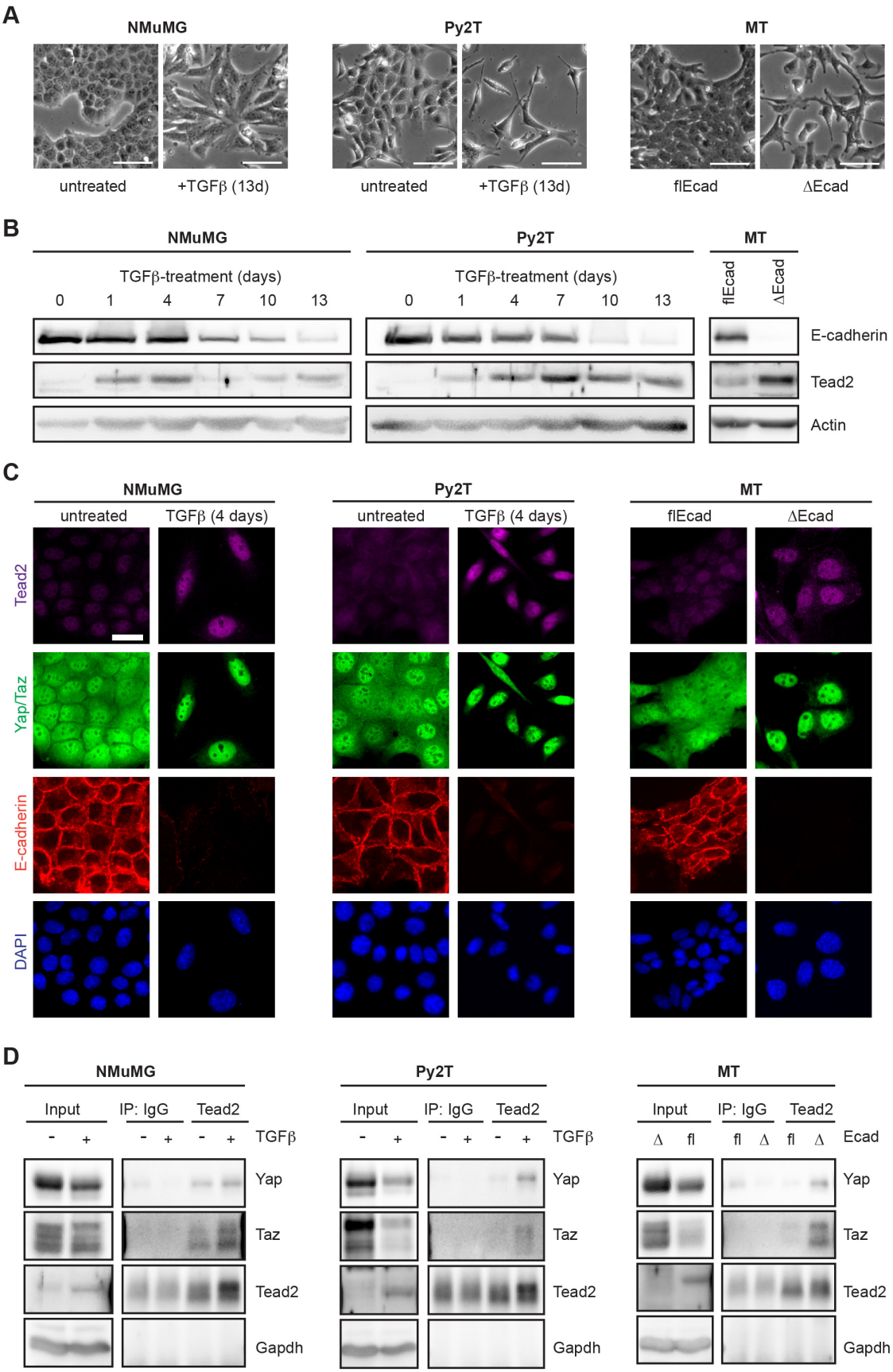


Figure 1: Tead2 upregulation and Yap and Taz subcellular redistribution during EMT. (A) Morphological differences between epithelial and mesenchymal counterparts of the three different cellular

EMT model systems used. Epithelial NMuMG and Py2T cells were treated with TGFβ for 13 days to induce EMT. Stably mesenchymal MTΔEcad cells were derived from epithelial MTflEcad cells by Cre-recombinase-mediated knockout of the E-cadherin gene. Scale bars: 50 µm. **(B)** Immunoblotting analysis of Tead2 and E-cadherin expression levels before, during and after TGFβ-induced EMT in NMuMG and Py2T cells, and by genetic deletion in MTflEcad cells. Actin served as a loading control. (**C)** Immunofluorescence staining of Tead2 and its co-factors Yap and Taz before and after induction of EMT. Yap and Taz were stained with an antibody that detects both proteins. E-cadherin staining served as a control, DAPI was used to visualize nuclei. Scale bar: 25 µm. **(D)** Interaction of Tead2 with Yap and Taz. Cells were treated with TGFβ for 4 days (NMuMG) or 7 days (Py2T). Co-immunoprecipitation (IP) was performed with an antibody against Tead2 or irrelevant IgG as a negative control. Levels of Yap, Taz and Tead2 were determined by immunoblotting analysis. Gapdh served as loading control.

We next investigated whether Tead2 binding to Yap and/or Taz was subject to change during an EMT. Co-immunoprecipitation experiments in total cell lysates (Figure 1D; Py2T, NMuMG) or in cytoplasmic and nuclear cell extracts (supplementary Figure S1I) revealed that, indeed, the binding of Yap and Taz to Tead2 increased exclusively in the nucleus upon TGFβ stimulation. Similarly, both Yap and Taz only bound to Tead2 in mesenchymal MTΔEcad cells and not in epithelial MTflEcad cells (Figure 1D; MT). From these data we conclude that formation of a nuclear Tead2–Yap–Taz complex is increased during EMT, even though total levels of Yap and Taz decrease upon EMT induction (Figure 1D; supplementary Figure S1I).

Consistent with these observations, we found that pan-Tead transcriptional activity increased upon EMT induction in a Sox4-dependent fashion (supplementary Figure S2A-C). Tead transcriptional activity in cells undergoing EMT was assessed by Tead-responsive luciferase reporter constructs bearing either MCAT core motifs (CATTCCT; supplementary Figure S2A) [361] or GTIIC core motifs (ACATTCCAC; supplementary Figure S2B) [188]. In addition, the transcriptional activity of a previously described Cyr61 promoter reporter containing a Tead-responsive MCAT motif [362] was induced during EMT of NMuMG cells, whereas mutation of the Tead-binding site resulted in a complete loss of the reporter activity (supplementary Figure S2D).

Taken together, these data demonstrate that, upon EMT induction, Tead2 expression levels are increased in a Sox4-dependent manner concomitant with enhanced formation of the Tead2–Yap–Taz complex in the nucleus. In accordance with these observations, overall Tead transcriptional activity increases during EMT.

3.1.3.2 Nuclear localization of Yap and Taz is mediated by Tead2 and is required for EMT

From the data presented above, we hypothesized that binding of the increased levels of Tead2 to Yap and Taz in the nucleus is required to activate Tead2-mediated transcriptional activity during EMT. To test this hypothesis, we first investigated whether overexpression of Tead2 in epithelial cells can provoke a subcellular redistribution of Yap and Taz. Immunofluorescence microscopy analysis revealed that Yap and Taz were evenly distributed between the cytoplasm and nucleus in control cells (Figure 2A, Vector) or in cells transfected to express a Tead2 mutant (Figure 2A, Tead2 Y440H), which is incapable of binding to Yap or Taz (Figure 2A) [210, 363-366]. Conversely, the forced expression of wildtype Tead2 (Tead2-WT) in NMuMG cells resulted in a marked concentration of Yap and Taz in the nucleus and reduced levels of cytoplasmic Yap and Taz, suggesting that Tead2 controls Yap and Taz subcellular distribution by direct binding. Consistent with these observations, Tead transcriptional activity was increased by the expression of wildtype Tead2 but not by the expression of mutant Tead2 Y440H (supplementary Figure S3A).

We next assessed whether upregulation of Teads during EMT is required for Yap and Taz redistribution and for EMT. NMuMG cells were transfected with control siRNA (siCtr) or with siRNAs against Tead1, 2 and 3 (siTead1-3) whose expression were upregulated during EMT in these cells (supplementary Figure S1B). As expected, TGFβ-treated cells transfected with control siRNA displayed predominantly nuclear localized Yap and Taz in response to TGFβ-treatment, whereas knockdown of Teads prevented this subcellular redistribution (Figure 2B). Importantly, depletion of Teads and the resulting failure of Yap and Taz subcellular redistribution also averted the disassembly of tight junctions and adherens junctions normally observed during EMT (supplementary Figure S3B). Moreover, immunoblotting analysis of EMT marker expression revealed an attenuation of EMT upon Tead ablation: downregulation of E-cadherin was inhibited and, conversely, upregulation of the mesenchymal markers fibronectin and N-cadherin were delayed (supplementary Figure S3C). Similarly, blocking Tead transcriptional activity by the inducible expression of a dominant-negative version of Tead2 (Tead2-EnR) attenuated the EMT process in TGFβ-treated Py2T cells (supplementary Figure S3D,E). These results indicate that Teads are crucial for regulating the cytoplasmic-nuclear

redistribution of Yap and Taz during EMT and that this regulatory process is required for EMT.

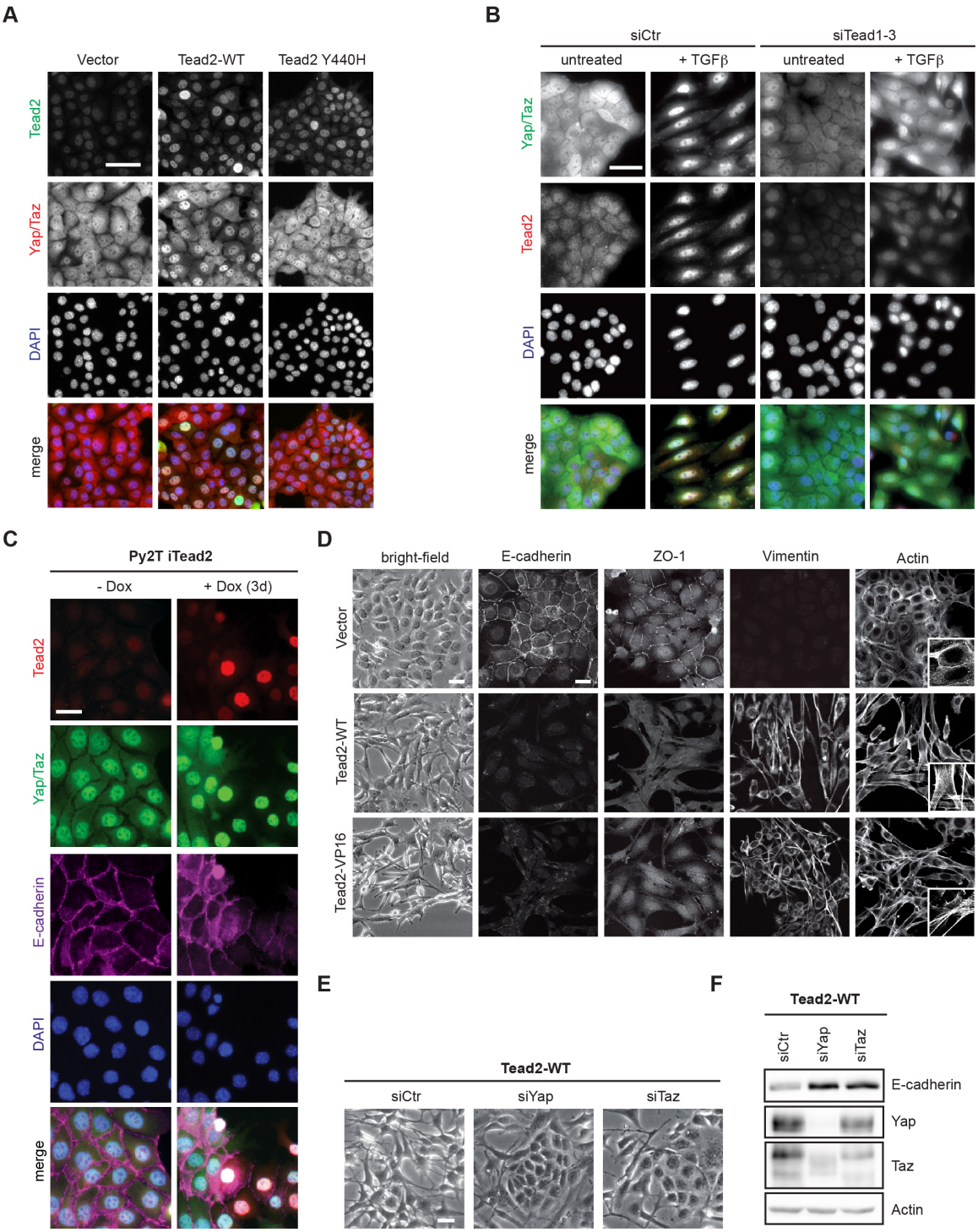


Figure 2: Elevated Tead2 levels induce a predominant nuclear localization of Yap and Taz and induce EMT. (A) Yap and Taz cellular localization is dependent on their direct binding to Tead2. A vector control (Vector), type Tead2 (Tead2-WT) or a Tead2 point mutant defective in Yap and Taz binding (Tead2 Y440H) were stably expressed in NMuMG cells, and Yap and Taz and Tead2 localization was assessed by immunofluorescence staining. DAPI was used to visualize nuclei. Scale bar: 20 µm. **(B)** Depletion of Tead expression prevents the reduction in cytoplasmic levels of Yap and Taz during EMT. NMuMG cells

transfected or not with siRNA pools targeting Tead1-3 were induced to undergo EMT by TGFβ-treatment for 4 days. Yap and Taz, and Tead2 were visualized as described in A. Scale bar: 25 µm. **(C)** Effect of acute Tead2 overexpression on Yap and Taz localization and epithelial differentiation. Tead2 was expressed in Py2T cells (Py2T-iTead2) in a doxycycline (Dox)-inducible fashion, and the localization of Tead2, Yap and Taz and E-cadherin were visualized by an immunofluorescence staining. Scale bar: 20 µm. **(D)** Effect of Tead2 gain-of-function on cell morphology and EMT. Py2T cells were stably transduced with constructs coding for Tead2 (Tead2-WT), a constitutively active version of Tead2 (Tead2-VP16) or an empty vector control. Overall morphological changes were visualized by phase-contrast microscopy and by immunofluorescence staining against E-cadherin, ZO-1, vimentin and the actin cytoskeleton (phalloidin staining). The insets show an enlarged view of F-actin staining. Scale bars: 15 µm. **(E,F)** Depletion of Yap or Taz expression prevents Tead2-induced EMT. Py2T cells stably overexpressing Tead2 were transfected with siRNA pools against Yap and Taz or with a control siRNA (siCtr). Overall cell morphology by phase-contrast microscopy **(E)** and immunoblotting analysis of E-cadherin, Yap and Taz expression **(F)** are shown. Scale bar: 15 µm.

We next asked whether experimentally increasing Tead2 levels and transcriptional activity in epithelial cells is sufficient to induce EMT. We generated Py2T cells that expressed wildtype Tead2 under the control of the doxycycline-inducible system. Treatment with doxycycline for 3 days led to a heterogeneous induction of Tead2 expression in these cells (Figure 2C). Non-induced cells expressing low endogenous levels of Tead2 showed Yap and Taz staining in the cytoplasm as well as in the nucleus. In doxycycline-treated cells, the elevated levels of nuclear Tead2 resulted in concentration of Yap and Taz in the nucleus (Figure 2C). By 3 days after doxycycline induction, Tead2-expressing cells had already started to lose E-cadherin expression; this phenotype was enforced with a prolonged doxycycline treatment (supplementary Figure S3F), indicating that the elevation of Tead2 levels was sufficient to induce EMT.

Yap and Taz nuclear localization, increased formation of Yap–Taz–Tead2 complexes and heightened Tead transcriptional activity was also observed in Py2T cells stably expressing wildtype Tead2 (Tead2-WT) (supplementary Figure S3G,H) [188]. Notably, although overexpression of a constitutively active version of Tead2 (Tead2-VP16) lacking a Yap- and Taz-binding site showed a transcriptional activity that was highly dependent on Tead (mediated by the *Herpes simplex* virus VP16 transactivation domain), but no increased nuclear localization of Yap and Taz (supplementary Figure S3G,I). Furthermore, forced expression of both wildtype Tead2 and constitutively active Tead2-VP16 also resulted in the morphological changes consistent with EMT [i.e. the loss of the epithelial markers E-cadherin and ZO-1, increased expression of the mesenchymal markers vimentin, Zeb1/2 and Slug, and a shift of the cytoskeleton from displaying cortical actin to actin stress fibers (Figure 2D; supplementary Figure S3J). Transient siRNA-mediated ablation of Yap or

Taz in Tead2-WT (siYap and siTaz, respectively) cells led to a mesenchymal-to-epithelial transition (MET), as indicated by an epithelial morphology and the increased expression of E-cadherin in Yap- or Taz-deficient cells (Figure 2E,F). Collectively, these results demonstrate that Tead2 is able to induce an EMT via the formation of transcriptionally active complexes with Yap and Taz.

3.1.3.3 Tead2 promotes tumor cell migration, invasion and metastasis

A hallmark of cells undergoing EMT is the acquisition of migratory and invasive properties [4, 349]. Consistent with their mesenchymal phenotype, cells expressing Tead2-WT exhibited an increased capability to migrate and to invade (Figure 3A). In contrast to control cells, which formed smooth spheres, Tead2-WT and Tead2-VP16 cells invaded there-dimensional Matrigel and projected filopodia into the extracellular matrix (Figure 3B). These results indicate that, consistent with its EMT-inducing activities, Tead2 also promotes cancer cell migration and invasion.

We next assessed whether Tead2-induced EMT, cell migration and cell invasion translated into a higher metastatic capability in Py2T murine breast cancer cells. Py2T cells stably expressing Tead2-WT, Tead2-VP16 or a control vector were injected into the tail veins of Balb/c nu/nu immune-deficient mice, and the formation of lung metastasis was scored 33 days after injection. Serial sectioning of paraffin-embedded lungs and staining by hematoxylin and eosin revealed that only one out of six mice injected with epithelial control cells displayed macroscopically visible tumor cell clusters. Notably, these nodules were found encapsulated within blood vessels (Figure 3C,D). In contrast, half of the mice injected with Tead2-WT cells and five out of six mice injected with Tead2-VP16 cells developed macroscopic metastases (Figure 3C,D), with a higher incidence per mouse as compared to control cells (Figure 3E).

Taken together, these results indicate that Tead2 promotes increased cancer cell migration/invasion and metastatic outgrowth of Py2T cells in the lung.

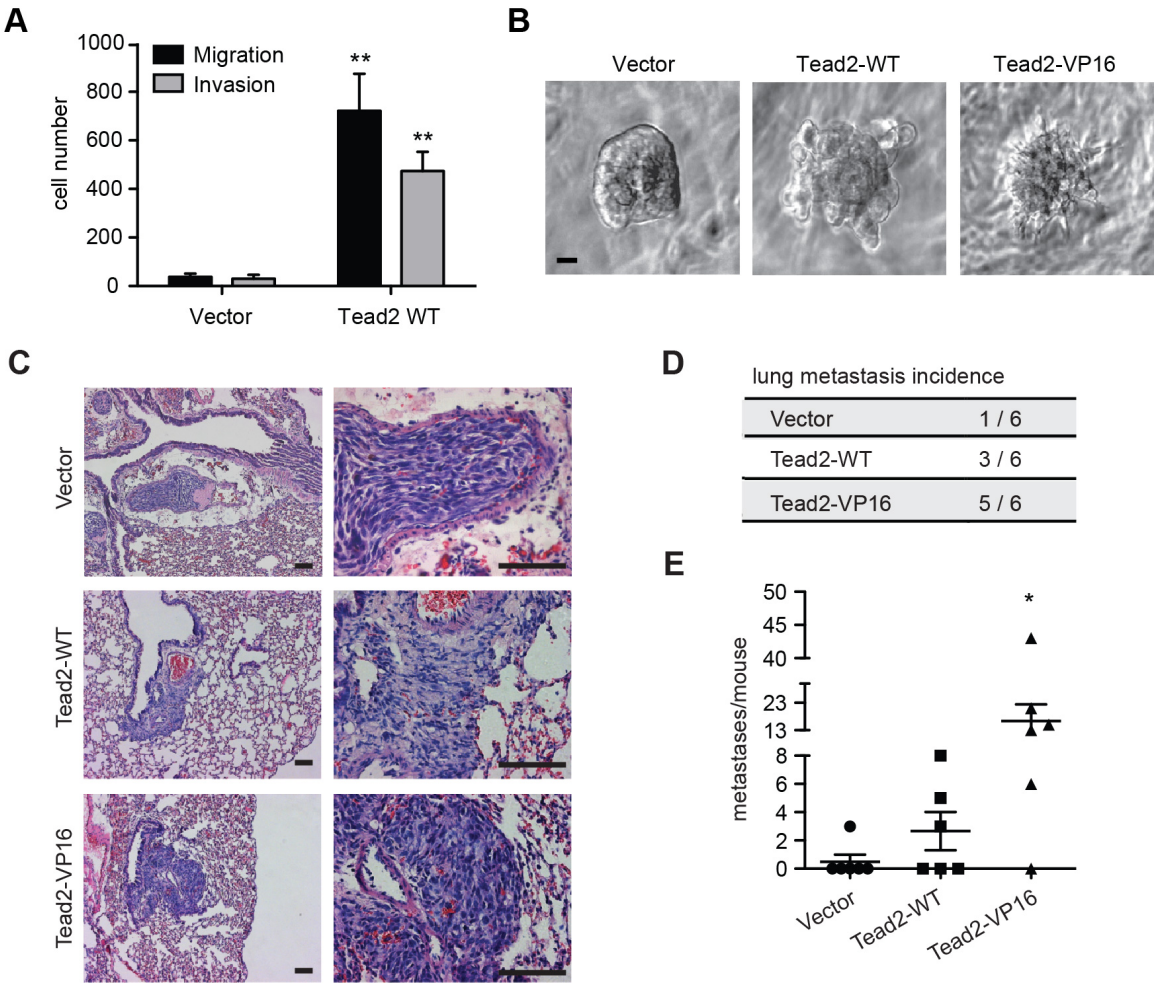


Figure 3: Tead2 promotes cell migration, invasion and metastasis. (A) Chemotactic migration and invasion of Py2T cells stably expressing Tead2-WT or an empty vector control. Transwell assays were performed utilizing cell culture inserts that were not coated (migration) or coated with Matrigel (invasion). Migrated and invaded cells were quantified. Data are shown as mean ± S.E.M. (n=3; **P<0.01). **(B)** Cell invasion in a 3D extracellular matrix. Py2T cells stably expressing Tead2-WT or Tead2-VP16 were embedded in Matrigel and allowed to grow for 5 days. Empty-vector-transduced Py2T cells served as a control. Scale bars: 50 µm. **(C)** Experimental metastasis. Py2T cells as described in B were injected into the tail veins of female Balb/c nu/nu mice. Mice were killed 33 days post-injection and lungs were sectioned and stained by hematoxilin and eosin (H&E). Higher magnifications are also shown (right). Scale bars: 100 µm. **(D,E)** Quantification of lung metastasis incidence and number of lung metastasis per mouse as determined by serial sectioning and microscopic analysis of lungs as described in C. The metastatic incidence was calculated as mice harboring metastases/total number of mice per group. Data are shown as mean ± S.E.M. (n=6 mice per group; *P<0.05).

3.1.3.4 The transcriptional Tead2 target genes during EMT

The results presented above suggest that Tead2 function is regulated differently during EMT, leading to a malignant cancer cell phenotype. To elucidate the corresponding downstream mechanisms we sought to identify genes that are transcriptionally regulated by Tead2 during the process of EMT. Using our Tead2

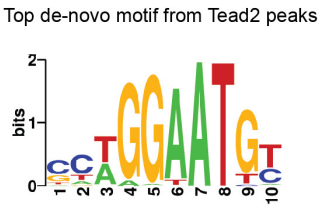
antibody we performed chromatin immunoprecipitation followed by next generation sequencing (ChIP-Seq) at 5 days of TGFβ-treatment, a time point at which Tead2 expression and activity was robustly increased in Py2T cells (Figure 1B; supplementary Figure S2A). Detection of known overrepresented transcription factor binding motifs by Hypergeometric Optimization Motif EnRichment (HOMER) revealed that the MCAT Tead-binding motif (supplementary Figure S1A) is the most significant motif found in Tead2 bound regions, followed by a motif bound by Jun-AP1 (Figure 4A). PhyloGibbs, an algorithm that *de novo* infers overrepresented and evolutionarily conserved sequence motifs from ChIP-Seq data, also revealed a motif that corresponds to the MCAT and to the highly similar GTIIC core Tead binding motifs (Figure 4B). These findings suggest that the known MCAT and GTIIC motifs are among the main sequence motifs targeted by Tead2 during TGFβ-induced EMT.

To identify direct target genes of Tead2 that showed a change in their expression during EMT, we performed genome-wide gene expression profiling in epithelial Py2T cells and in Py2T cells treated with TGFβ for 5 days. Genes that had a different expression level, passing a cutoff of at least 1.5-fold (adjusted $P<0.05$), were selected for an overlay with Tead2-bound genes as determined by ChIP-Seq analysis (Figure 4C; supplementary Table S1). Tead2-bound genes were defined as showing a ChIP-Seq peak at a distance less than 10 kb to their transcriptional start sites (TSS). The overlay analysis identified 132 genes bound by Tead2 in their promoter regions and showing increased or decreased expression during TGFβ-induced EMT (supplementary Table S2). Gene ontology analysis for cellular components by GOstats [367] revealed that Tead2 predominantly controls genes which encode for components of cell junctions and regulators of the actin cytoskeleton (Figure 4D).

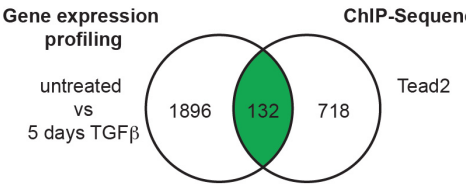
A

Rank	Motif	Name	p-Value
1		Tead	1E-126
2		Jun-AP1	1E-85
3		Myf	1E-30
4		Hlh-2; Hlh-8	1E-27
5		Fos	1E-23

B



C

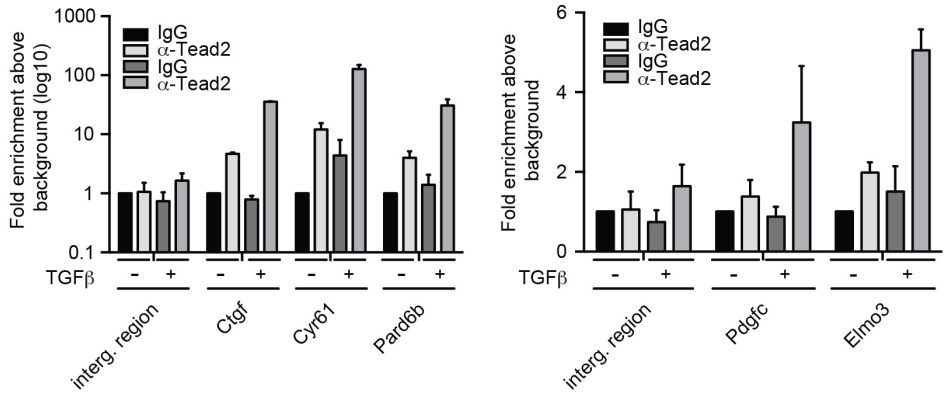


D

GOstats: Cellular Components

Term	pValue	Genes
perinuclear region of cytoplasm	3.24E-4	Clu, Ctgf, Jag2, Lamb1, Limk1, Myo6, Nppb, Tnfrsf1b, Sept9, Trim13
tight junction	5.21E-4	Ect2, Amotl2, Cldn9, Pard6b, Synpo
lysosomal membrane	2.84E-3	Myo7a, Rnf13, Atp6v1g1
actin cytoskeleton	2.99E-3	Csrp1, Myo6, Myo7a, Tnni2, Zyx, Sept9, Synpo
cell-cell junction	3.17E-3	Ect2, Zyx, Amotl2, Cldn9, Pard6b, Ahnak, Synpo

E



F

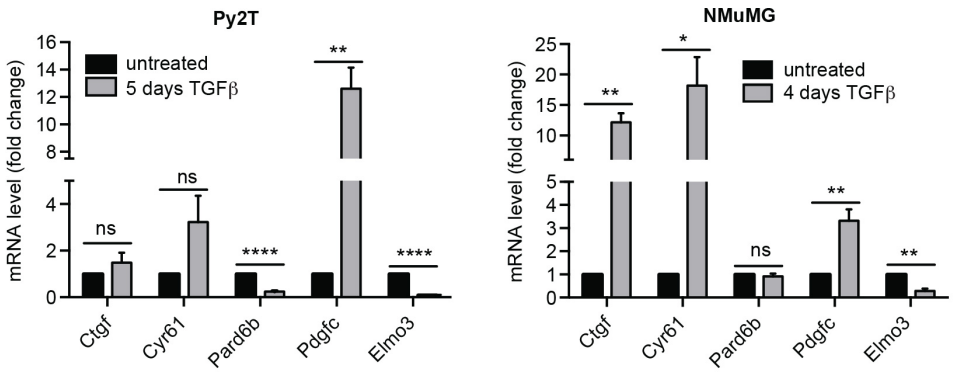


Figure 4: Identification of direct Tead2 transcriptional target genes during EMT. (A) DNA-binding motifs that are overrepresented in Tead2-binding regions. Py2T cells treated with TGFβ for 5 days were subjected to ChIP using an antibody for Tead2 followed by next generation sequencing (ChIP-Seq; *n*=2). The

sequencing data were subjected to Hypergeometric Optimization of Motif EnRichment analysis (HOMER). Shown are the motifs that are significantly enriched. **(B)** *De novo* generation of sequence motifs overrepresented in Tead2-binding regions using PhyloGibbs. Shown is the most significant Tead2-binding motif. **(C)** Determination of potential direct target genes of Tead2 during EMT. The Venn diagram depicts the number of genes from the Tead2 ChIP-Seq analysis, the number of genes that were regulated differently in Py2T cells before and after 5 days of TGFβ treatment and the number of overlapping genes. **(D)** Gene ontology analysis was performed on overlapping genes described in C using GOstats. The table shows the functional annotation clustering analysis for the top five cellular compartments, the associated genes per group and their *P*-values within the groups. **(E)** Validation of genes directly bound by Tead2 by quantitative PCR. Chromatin from the cells treated as described in A was subjected to qPCR using primer pairs spanning the Tead2-binding regions determined by ChIP-Seq. The data are presented as fold enrichment above background (IP over input) and were normalized to control IgG. An intergenic region was used as negative control. Data are represented as means ± S.E.M. ($n=2$). **(F)** Expression of Tead2-bound genes during EMT as determined by RT-qPCR. Py2T and NMuMG cells were treated with TGFβ for 5 or 4 days, respectively. Fold changes in mRNA expression are presented as means ± S.E.M. ($n=3$). *$P<0.05$, **$P<0.01$, ****$P<0.0001$.

Among these 132 genes, we validated known Tead target genes by ChIP-PCR, including CTGF and CYR61 [215], whose expression also increased upon TGFβ-stimulation in the three EMT model systems investigated (Figure 4E,F; supplementary Figure S4A,B). Furthermore, we identified other candidate Tead2 target genes that had been previously implicated in EMT and/or malignant tumor progression [74, 98, 350], such as AMOTL2, ESRP2, MAL, PDGFC, SERPINE1 and others (supplementary Table S2). We further ascertained that PARD6B, a gene encoding for a polarity complex protein, and ELMO3, a gene implicated in phagocytosis and cell migration, were Tead2 target gene candidates. The expression of these genes was decreased during TGFβ-induced EMT in Py2T and mesenchymal MTΔEcad cells, indicating that Tead2 is not exclusively a transcriptional activator but also a repressor of gene expression (Figure 4E,F; supplementary Figure S4B).

Given that, during EMT, Tead2 formed nuclear complexes with Yap and Taz (Figure 1D; supplementary Figure S1I), we assessed whether some of the above-mentioned Tead2-regulated genes were bound by a Tead2–Yap–Taz complex. ChIP experiments for Tead2, Yap and Taz in Py2T cells induced to undergo EMT and subsequent RT-PCR analysis revealed similar or increased binding of Yap and/or Taz to the same promoter regions targeted by Tead2, including the promoters of CTGF, CYR61, PARD6B, PDGFC (supplementary Figure S4C).

These results indicate that the Tead2–Yap–Taz complex regulates EMT-relevant genes by acting as a transcriptional activator or repressor, mainly by binding to promoter sequences containing Tead-specific MCAT or GTIIC motifs.

3.1.3.5 Zyxin is a Tead2–Taz target gene critical for EMT

ChIP-Seq also revealed a direct binding of Tead2 to an intronic region of the gene encoding for Zyxin (ZYX; supplementary Figure S4D). Close inspection of this region revealed the presence of a species-conserved MCAT motif (Figure 5A). Zyxin is an actin regulatory protein that localizes to sites of focal adhesions and stress fibers in response to mechanical cues to facilitate actin polymerization and generation of traction force [368-370]. In line with this functional role, Zyxin upregulation during EMT in NMuMG cells has been reported to be essential for cell migration [371], and given that forced expression of Tead2 promotes cancer cell migration or invasion, and metastatic outgrowth of Py2T cells in the lung (Figure 3), we decided to focus on Zyxin as a Tead2 target gene.

ChIP-PCR on Py2T cells treated with TGFβ for 5 days confirmed the binding of endogenous Tead2 to an intronic region of the ZYX gene (Figure 5B). Concomitantly, mRNA and protein levels of Zyxin were substantially increased during EMT in Py2T cells (Figure 5C). Higher Zyxin expression was also found in mesenchymal MTΔEcad cells as compared to MTflEcad cells (supplementary Figure S4E). Immunofluorescence microscopy analysis confirmed increased levels of Zyxin localized along stress fibers that are formed during TGFβ-induced EMT in Py2T cells and at focal adhesion sites (Figure 5D), as previously reported [371].

To evaluate whether increased expression of Zyxin during EMT is dependent on Tead activity, we induced EMT in a pool of Py2T cells expressing a dominant-negative version of Tead2 (Tead2-EnR) under the control of the doxycycline-inducible system. Immunoblotting and RT-PCR analysis demonstrated that expression of Tead2-EnR significantly attenuated Zyxin upregulation during EMT (Figure 5E,F). Similarly, siRNA-mediated knockdown of Tead1, 2 and 3 in NMuMG cells and knockdown of Taz, not Yap, in Py2T cells prevented increased expression of Zyxin during EMT (Figure 5G,H).

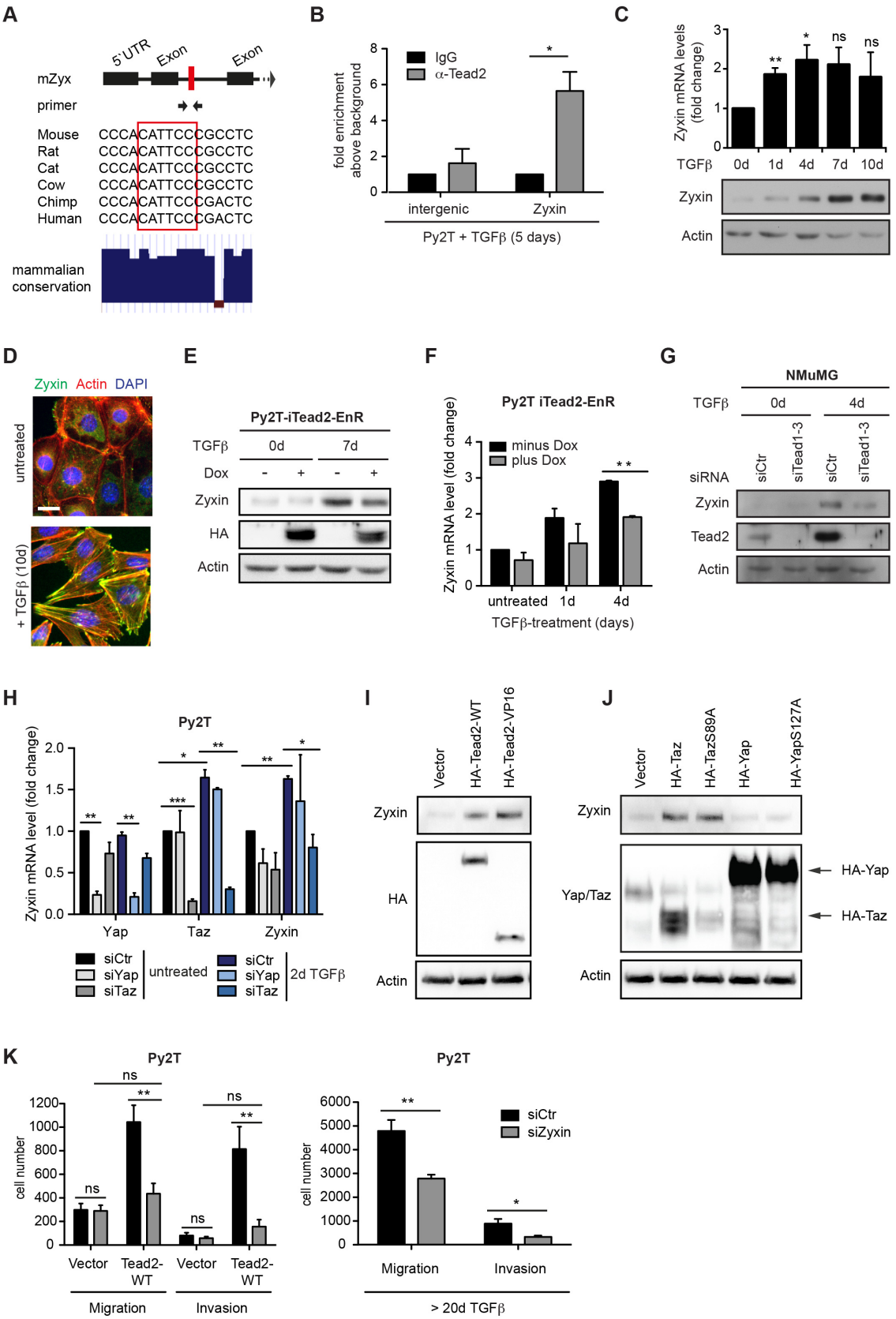


Figure 5: Zyxin is a direct Tead2 target gene critical for cell migration. **(A)** Top: scheme of the zyxin gene (ZYX) showing the location of the Tead2-binding site (red). Arrows denote primers used in B. Middle: species conservation of the Tead2-binding region. The red box denotes the core MCAT motif. Bottom: mammalian conservation plot encompassing 32 species (derived from UCSC genome browser). **(B)**

Validation of Tead2-binding to the *Zyx* gene by ChIP-qPCR in Py2T cells treated with TGFβ for 5 days. The qPCR data ($n=3$; mean ± S.E.M.; *$P<0.05$) indicate fold enrichment above background and were normalized to irrelevant IgG as negative control. **(C)** Zyxin expression during EMT. Py2T cells were treated with TGFβ for the indicated durations. Regulation of mRNA (top) and protein level (bottom) was determined by RT-qPCR and immunoblotting analysis. Results are presented as means ± S.E.M. ($n=3$). *$P<0.05$; **$P<0.01$. **(D)** Immunofluorescence staining of zyxin and F-actin in epithelial and mesenchymal Py2T cells. Cells were treated with TGFβ as indicated. F-actin was stained with phalloidin coupled to a fluorophore. Scale bar: 15 µm. **(E,F)** Zyxin protein and mRNA levels in Py2T cells undergoing EMT without (−) or with (+) doxycycline (Dox)-induced expression of a HA-tagged dominant-negative version of Tead2 (Tead2-EnR) before and during a TGFβ-induced EMT. Results are presented as means ± S.E.M. ($n=2$). **$P<0.01$. **(G)** Immunoblotting analysis of zyxin during TGFβ-induced EMT of NMuMG cells transfected with siRNA pools targeting Tead1, 2 and 3 or control siRNA. Membranes shown in supplementary Figure S3C were reprobed with an antibody against zyxin. **(H)** Yap, Taz and zyxin mRNA expression levels in Py2T cells. Cells were transfected with siRNA pools against Yap, Taz or a control siRNA and cultured in the absence and presence of TGFβ. Results are presented as means ± S.E.M. ($n=2$). *$P<0.05$, **$P<0.01$, ***$P<0.001$. **(I,J)** Immunoblotting analysis of zyxin expression in Py2T cells overexpressing HA-tagged wildtype Tead2 (HA–Tead2-WT), constitutively active Tead2 (HA–Tead2-VP16), and wildtype or Yap and Taz mutants (HA–TazS89A, HA–YapS127A). **(K)** Contribution of zyxin to mesenchymal migration and invasion induced by Tead2 overexpression (Tead2-WT; left panel) or by 20 day TGFβ-treatment (right panel). Py2T cells were transfected with siRNA pools against zyxin or control siRNA pools and subjected to chemotactic transwell migration and invasion assays as described in Figure 3A. Data are represented as means ± S.E.M. ($n=3$). *$P<0.05$, **$P<0.01$.

We next assessed whether Zyxin expression was indeed controlled by Tead2. We thus evaluated Zyxin expression in response to modulating Tead2 function by the expression of wildtype Tead2 and Tead2-VP16, by the expression of wildtype Yap and Taz and by the expression of Hippo-signaling-insensitive mutant versions of Yap and Taz [372]. Immunoblotting analysis revealed that Zyxin expression was induced by forced expression of Tead2-WT, Tead2-VP16 and both versions of Taz, but not by Yap (Figure 5I,J). These results demonstrate that Tead2 mediates *Zyx* gene expression via Taz co-activation.

To determine the functional contribution of increased Zyxin expression levels to cell migration and invasion, we tested whether the increased migratory and invasive phenotype of Py2T cells treated with TGFβ (>20 days) or of Py2T cells forced to express Tead2-WT (Figure 3A) could be reversed by siRNA-mediated ablation of Zyxin expression (supplementary Figure S4G). Indeed, knockdown of Zyxin in cells expressing wildtype Tead2 or cells treated with TGFβ led to a reduction in cell migration and invasion (Figure 5K). Interestingly, knockdown of Zyxin in Py2T cells overexpressing wildtype Tead2 had no effect on the mesenchymal morphology of these cells (supplementary Figure S4H), indicating that Zyxin is only required for cell migration and invasion and not for the morphogenic process of EMT. The results demonstrate that Tead2 directly regulates Zyxin expression during TGFβ-induced

EMT via its co-factor Taz and thereby induces a Zyxin-driven migratory and invasive cell phenotype.

3.1.4 Discussion

The transdifferentiation of epithelial cells into a mesenchymal state involves a thorough remodeling of cell architecture, such as the loss of apical-basal cell polarity accompanied by a breakdown of tight and adherens junctions. Concomitantly, the establishment of mesenchymal traits involves the generation of front–rear polarity via remodeling of the actin cytoskeleton and the activation of cell migration. These changes largely rely on global changes in gene expression orchestrated by transcription factors [373]. Here, we have uncovered a crucial mechanistic role of Tead transcription factors, in particular Tead2, during the process of EMT.

We demonstrate that Tead2 expression is increased during EMT and that elevated nuclear levels of Tead2 lead to nuclear complex formation with Yap and Taz and to increased nuclear localization of these cofactors. We further demonstrate that increased Tead2 expression during EMT is accompanied by enhanced Tead2-mediated transcriptional activity. Notably, experimental elevation of Tead2 transcriptional activity is sufficient to induce an EMT and the establishment of lung metastasis. The EMT-associated upregulated expression of Tead2 is independent of canonical TGFβ signaling, because the ablation of Smad4 during TGFβ-induced EMT had no substantial impact on Tead2 expression. Moreover, our studies indicate that the TEAD2 gene is a downstream target of the transcription factor Sox4, a crucial epigenetic regulator of EMT [98]. Whether TEAD2 gene expression is directly activated by Sox4 during EMT remains to be addressed. Recent studies on SoxC family members, which include Sox4, Sox11 and Sox12, demonstrate a direct regulation of the TEAD2 gene promoter by Sox4 and Sox11 in the fibroblast-like cell line Cos-1 [374]. Given that Sox4 controls the levels of Tead2, which are crucial for Tead2 binding to the Hippo downstream targets Yap and Taz in nucleus, it will be interesting to assess whether Sox4 is an integrator of non-canonical TGFβ signaling and Hippo signaling.

Several recent studies have suggested that elevated Tead activity mediated by a gain-of-function of Yap or Taz can evoke a malignant phenotype. Indeed, Taz is

highly expressed in ~20% of breast cancers, most of which represent invasive ductal carcinomas (IDCs), and expression of this Tead co-activator is responsible for migration and invasiveness of cultured breast cancer cell lines [203]. Recently, another study has demonstrated that overexpression of a Hippo-signaling-insensitive version of Yap (YapS127A) renders breast cancer and melanoma cells pro-metastatic [212]. In accordance with our results, this effect depends on an intact interaction domain for Yap in Tead, and Tead transcriptional activity correlates with the metastatic potential of various cancer cell lines. Our results are also consistent with previous reports demonstrating that ectopic expression of Yap and Taz in MCF10A normal breast epithelial cells induces EMT via Tead transcriptional activity [208, 210].

As expected from cells that have undergone an EMT, we report that Tead2-overexpressing cells display increased migration, invasion and metastasis in comparison to control cells. Conversely, an attenuation of TGFβ-induced EMT is observed when a dominant-negative version of Tead2 (*Tead2-EnR*) is expressed. This effect is also observed upon siRNA-mediated depletion of Tead family members. These results demonstrate that the presence and transcriptional activity of Teads is required for the proper execution of a TGFβ-induced EMT program.

Tead transcriptional activity is mainly controlled by the direct binding of the co-factors Yap and Taz [169, 209, 356, 375]. Here, we have investigated whether this is also true during EMT in normal mammary epithelial cells and in breast cancer cells. In epithelial cells, Yap and Taz are almost equally distributed between the cytoplasm and the nucleus. However, during EMT Yap and Taz are predominantly localized to the nucleus and this subcellular redistribution is dependent on their binding to Tead2. We also observe a membranous staining pattern of Yap and Taz in epithelial cells reminiscent of tight junctions (data not shown). This observation is consistent with a series of studies reporting that Yap and Taz clusters with apical junction proteins, such as Crumbs, PATJ, PALS, angiomotins and α-catenin, thereby counteracting their nuclear entry [218, 220, 222, 224-226]. On the basis of published observations and our own results, it is tempting to speculate that, in epithelial cells, membrane-bound Yap and Taz are released by EMT-induced disassembly of junctional complexes, and increased nuclear levels of a Tead2–Yap–Taz complex in the nucleus to finally promote Tead transcriptional activity.

To the best of our knowledge, this is the first report on genes that are transcriptionally regulated by activated Tead2 during EMT. ChIP-Seq analysis has identified genes that are directly bound by Tead2 during TGFβ-induced EMT (supplementary Table S1). Our results indicate that in cells undergoing EMT, Tead2 predominantly binds to regions harboring MCAT Tead-binding motifs [376]. A closer inspection of Tead2-bound genes and their changes in gene expression during EMT indicates that Tead2 acts as a transcriptional activator as well as a repressor. Because Tead transcription factors lack a transactivation domain, their function depends on transcriptional co-factors. Thus far, Yap and Taz have been reported to exclusively act as transcriptional co-activators of Teads, for example in regulating the expression of the CTGF and CYR61 genes [210, 215]. Our analysis identified new Tead2 target genes like PARD6B and ELMO3, whose expression are significantly reduced during EMT. The same promoter regions were also found to be bound by Taz upon TGFβ-treatment, which would implicate that the Tead2–Taz complex could also function as a transcriptional repressor.

We further report that Taz but not Yap acts as a co-activator of the Tead2 target gene ZYX. Zyxin protein expression is upregulated during EMT in a Tead-dependent manner. Zyxin associates with the actin cytoskeleton and therefore can be localized at the apical membrane, in focal adhesions and on stress fibers [368]. Notably, Zyxin is required for the rearrangement of the actin cytoskeleton during TGFβ-induced EMT of NMuMG cells, thereby enabling cell migration [371]. Zyxin upregulation could therefore be at least one of the mechanisms by which Tead transcriptional activity controls the induction of a migratory and invasive cellular phenotype during EMT.

In summary, our results establish a crucial regulatory role for Tead transcription factors during the process of EMT and suggest that Teads are not only mere executors of Yap and Taz functions in promoting EMT and metastasis, but also play a crucial regulatory role in controlling Yap and Taz subcellular localization and activity during these processes.

3.1.5 Materials and methods

Antibodies and reagents

Antibodies used were as against the following proteins: actin (Santa Cruz Biotechnology, Dallas, TX), E-cadherin (BD, San Jose, CA, USA), N-cadherin (Takara Bio, Otsu, Shiga, Japan), fibronectin (Sigma-Aldrich, St Louis, MO, USA), GAPDH (Sigma-Aldrich), vimentin (Sigma-Aldrich), ZO-1 (Zymed, San Francisco, CA), Yap and Taz (Santa Cruz Biotechnology), Taz (IMGENEX, San Diego, CA, USA), zyxin (Santa Cruz Biotechnology), HA (Covance, Princeton, NJ, USA). Affinity purified, polyclonal rabbit anti-Tead2 antibody was generated by immunizing rabbits with a peptide corresponding to the N-terminus of Tead2 (amino acids 16–32); phalloidin–Alexa-Fluor-568 was from Invitrogen (Carlsbad, CA). Reagents used were: recombinant human TGFβ1 (R&D Systems, Minneapolis, MN, USA); Matrigel, growth factor reduced (BD); Doxycycline (Clontech/Takara).

Cell culture and cell lines

The subclone of NMuMG cells (NMuMG/E9) was as previously described [357] and was originally obtained from the American Type Culture Collection (ATCC, Manassas, VA). MTflEcad and MTΔEcad cells and Py2T cells are as described previously [358, 360]. All cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with glutamine, penicillin, streptomycin and 10% FBS (Sigma-Aldrich). Cells were treated with 2 ng/ml TGFβ.

Plasmids

The GTIIC reporter was generated by subcloning the GTIIC Tead response elements including the basal promoter from pδ51-LucII (kindly provided by Hiroshi Sasaki, RIKEN Center for Developmental Biology, Kobe, Japan) into pGL4 (Promega, Madison, WI, USA) [167, 188]. The MCAT reporter was derived from this construct by replacement of 8×GTIIC with eight copies of the sequence CCTGACACACATTCCTCAGCT (8×MCAT), where the MCAT core motif is underlined, and flanking sequences were according to Larkin et al. [361]. The Cyr61prom WT and Cyr61prom TeadMut reporters were previously described [362]

and kindly provided by Xiaolong Yang (Pathology and Molecular Medicine, Queen's University, Ontario, Canada). The *Renilla* luciferase expressing vector (pRL-CMV) was from Promega. Murine Tead1-4 in pcDNA3.1 were kindly provided by Jaime Carvajal (The Institute of Cancer Research, London, UK). Retroviral Tead2-WT and Tead2-VP16 constructs were as described previously [188]. Retroviral HA-tagged Tead2, Tead2-VP16, Tead2-EnR, Taz and Yap were created by inserting the respective cDNAs into the pBabe-derived retroviral vector pRFTO containing an N-terminal HA tag (kindly provided by Reto Kohler, FMI, Basel, Switzerland). Original cDNAs were kind gifts from H. Sasaki (Tead2, Tead2-VP16), R. Kohler (Yap, YapS127A) and Kun-Liang Guan (Department of Pharmacology and Moores Cancer Center, UCSD, La Jolla, USA) (Taz, TazS89A). The lentiviral, doxycycline-inducible HA–Tead2 and HA–Tead2-EnR constructs were generated by subcloning from pRFTO into pLVX-tight-puro (Clontech, Madison, WI, USA). Retroviral HA-tagged Tead2 Y440H was generated by PCR-mediated site-specific mutagenesis.

RNA interference

10 nM siGENOME smart pool siRNAs (Dharmacon, Lafayette, CO, USA) against murine Tead1, Tead2, Tead3, Tead4, Yap, Taz and Zyxin were used for transient knockdown experiments. Transfection was performed with Lipofectamine RNAiMax (Invitrogen) according to the manufacturer's instructions.

RNA isolation and RT-PCR

Total RNA was prepared using TriReagent (Sigma-Aldrich), reverse transcribed with M-MLV reverse transcriptase (Promega), and transcripts were quantified by PCR using SYBR-green PCR MasterMix (Invitrogen). Riboprotein L19 primers were used for normalization. PCR assays were performed in triplicate, and fold induction was calculated using the comparative Ct method (ΔΔ Ct). Primers used for quantitative RT-PCR are listed in supplementary Table S3.

Luciferase reporter assay

Cells were plated in triplicate in 24-well plates. At 1 day after plating, cells were transfected with 800 ng reporter and 10 ng of plasmids encoding *Renilla* luciferase

using Lipofectamine 2000 (Invitrogen). Cells were lysed using 1× passive lysis buffer (Promega) and lysates were analyzed using the Dual-Luciferase reporter assay system (Promega) and a Berthold Luminometer LB960. Firefly luciferase values were normalized to internal *Renilla* luciferase control values.

Cell fractionation and co-immunoprecipitation

Nuclear and cytoplasmic extracts were prepared with NE-PER nuclear and cytoplasmic extraction reagents (Thermo Scientific, Waltham, MA, USA) according to the manufacturer's instructions. Protein-G-coupled Dynabeads (Invitrogen) were used for co-immunoprecipitation experiments on total, nuclear and cytoplasmic cell extracts according to the manufacturer's instructions. Samples were analyzed by SDS-PAGE.

Immunoblotting, immunofluorescence staining, retroviral infection and Affymetrix gene expression profiling

These were performed as described previously [360].

Lentiviral infection

Stable pools of cells expressing Tead2-WT or Tead2-EnR in a doxycycline-inducible fashion were generated using the Lenti-X Tet-On Advanced system (Clontech). Lentiviral particles were produced according to the manufacturer's instructions using the helper vectors pHDM-HGPM2, pHDM-Tat1b, pRC-CMV-RaII and the envelope-encoding vector pVSV. For infection, viral supernatants were added to target cells in the presence of polybrene (8 µg/ml). Cells were spun for 1 hour at 30°C at 1000 *g* and were subsequently incubated at 37°C under 5% CO_2 in a tissue culture incubator for 2 hours. Viral supernatant was replaced by normal growth medium and after 1 day, antibiotic selection was performed.

Transwell migration and invasion assay

Trypsinized and washed cells were resuspended in growth medium containing 0.2% FBS and 2 ng/ml TGFβ, where appropriate. 2.5×10^4 cells in 500 µl medium

were seeded into cell culture insert chambers (BD). Bottoms of the chambers contained 700 µl of growth medium supplemented with 20% FBS. After 24 hours in a tissue culture incubator at 37°C under 5% CO_2, cells were fixed with 4% PFA in PBS for 10 minutes. Cells that had not crossed the membrane were removed with a cotton swab, and cells on the bottom of the membrane were stained with DAPI and quantified using a Leica DMI 4000 microscope.

3D Matrigel culture

Growth factor-reduced Matrigel (BD) was diluted to 4 mg/ml protein with serum-free growth medium. 2500 cells in 10 µl Matrigel were transferred to one well of a µ-slide angiogenesis microscopy slide (ibidi, Martinsried, Germany). After 20 minutes of gel solidification in a tissue culture incubator, 50 µl of normal growth medium was added. Growth medium was replenished after 3 days. After 5 days of growth, structures were photographed using a Leica DMIL microscope.

Tail vein injection, tissue processing and H&E staining

0.5×10^6 Py2T cells were injected orthotopically into the tail vein of three months old female Balb/c nude mice. Mice were killed at 33 days post injection and lungs were isolated. Tissue processing and H&E staining were performed as previously described [360]. All studies involving mice were approved by the Swiss Federal Veterinary Office (SFVO) and the regulations of the Cantonal Veterinary Office of Basel Stadt (licences 1878, 1907, 1908).

Chromatin immunoprecipitation

ChIP experiments were performed as previously described [377]. In brief, crosslinked chromatin was sonicated to achieve an average fragment size of 500 bp. Starting with 150 µg of chromatin and 5 µg of antibody, 1/40 of the ChIP sample and a 1:100 dilution of input DNA were used for quantitative PCR per reaction. Fold enrichment of specific target genes was calculated by IP over Input samples and was normalized to IgG negative control. Primers used for ChIP-qPCR are listed in supplementary figure S4.

ChIP-Sequencing

ChIP libraries were prepared using the ChIP-Seq sample Prep Kit from Illumina (San Diego, CA, USA) and were sequenced using Illumina HiSeq 2000 according to the manufacturer's protocol. ChIP-Seq data were processed as described [378, 379]. Briefly, reads were mapped against the *Mus musculus* genome (UCSC, mm9) using bowtie software (version 0.9.9.1) with parameters -v 2 -a -m 100, tracking up to hundred best alignment positions per read and allowing at most two mismatches. Each alignment was weighted by the inverse of the number of hits. All quantifications were based on weighted alignments. Chipcor software (http://sourceforge.net/projects/chip-seq) was used to estimate the fragment length. Alignments from ChIP-Seq experiments were shifted by half of the estimated fragment length towards their 3′ end. Clusters of ChIP-Seq read alignments were identified employing MACS software (v.1.3.7.1) [380] using IP and input samples with following parameters: nomodel, gsize=2700000000, tsize=50, pvalue=1e-5 and shiftsize=(chipcor estimate)/2. Resulting peak candidates were further filtered based on the enrichment over the input chromatin and only peaks with at least 4-fold enrichment were used.

Motif analysis

For HOMER analysis, motif detection in the vicinity of Tead2-binding sites was carried out using the Hypergeometric Optimization of Motif EnRichment tools (HOMER; v4.1) [381]. For PhyloGibbs, *de novo* motif generation was used as described previously [382].

Motif activity response analysis

Microarray raw data (.CEL files) were uploaded to http://ismara.unibas.ch/fcgi/mara for analysis.

Statistical analysis

Statistical analysis and graphs were generated using the GraphPad Prism software (GraphPad Software Inc.). Statistical analyses were performed as indicated in figure legends.

Accesssion numbers

Gene expression data of Py2T untreated versus five days TGFβ-treated cells and ChIP-Seq data of Tead2 ChIP during TGFβ-induced EMT in Py2T cells are deposited at Gene Expression Omnibus (GEO accession number: GSE55711).

3.1.6 Supplementary data

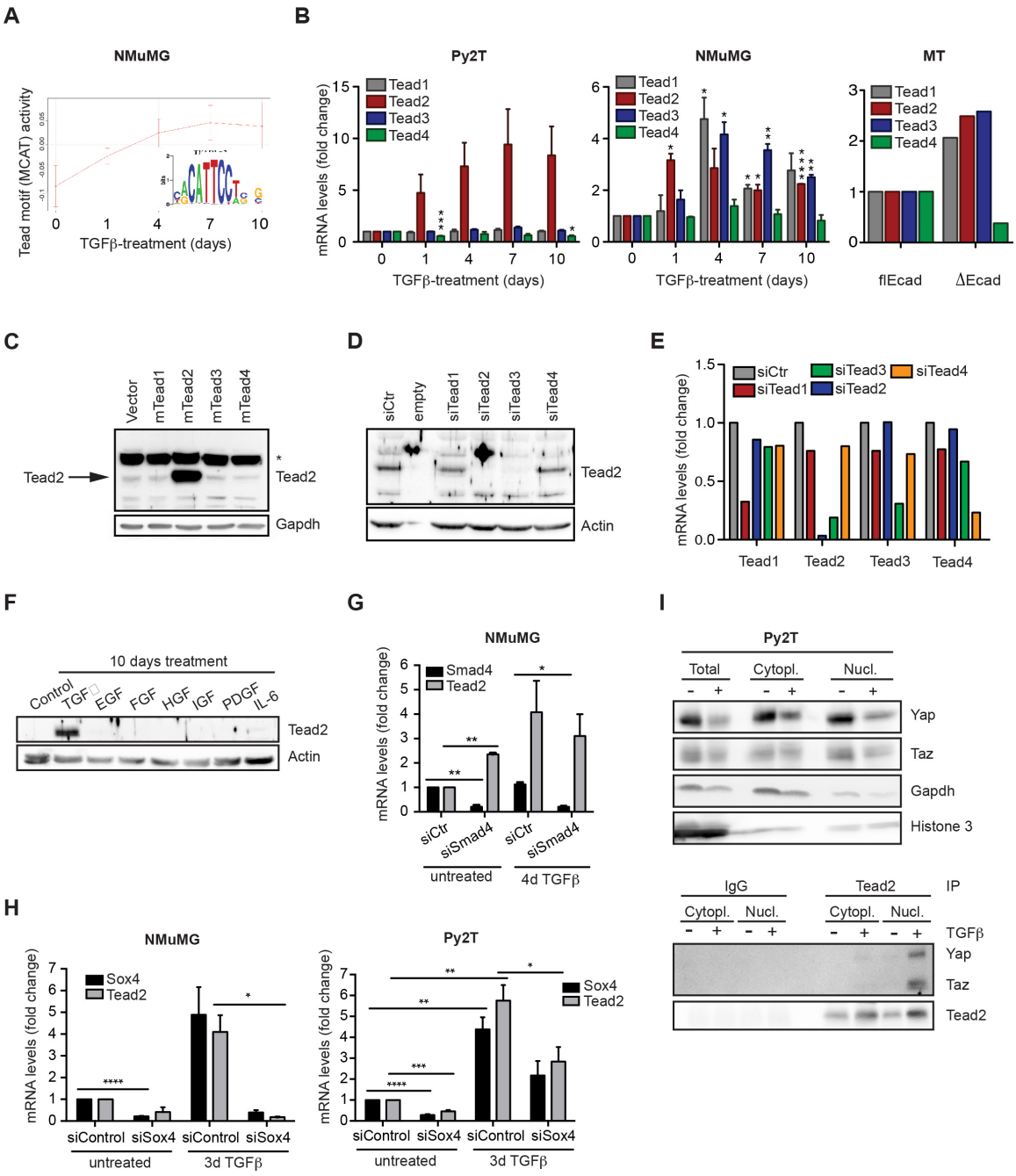


Figure S1. Tead activity and Tead2 levels are consistently upregulated in all 3 EMT models. (A) Motif Activity Response Analysis (MARA) predicts Tead activity during EMT in NMuMG cells. Shown is the Tead consensus binding motif (MCAT) used for the identification of Tead regulated genes and the increase in the expression of Teadregulated genes (activity) during TGFβ-induced EMT in NMuMG cells. **(B)** Changes in mRNA expression of Tead1-4 during EMT. Time course experiments of TGFβ-induced EMT were performed as described in Fig. 1B and mRNA levels were determined by quantitative RT-PCR. Results are presented as mean fold difference ± S.E.M. (n = 3; * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; Py2T and NMuMG). Expression changes in MT cells were determined once in triplicates. (C, D) α-Tead2 antibodies do not detect other family members than Tead2. Murine Tead 1-4 were transiently overexpressed in HEK293 cells **(C)** and siRNA pools against the different family members were transfected in NMuMG cells **(D)**. Cell lysates were probed with α-Tead2 by immunoblotting. The asterisk marks a band that is not specific for murine Tead2. **(E)**

Quantification of siRNA efficiency in the experiment described in (D) by quantitative RT-PCR. The downregulation of Teads mRNA levels was specific for the targeted family members, with the exception of siTead3, which depleted the expression of Tead2 and Tead3. **(F)** Localization of Yap and Taz in epithelial NMuMG cells was visualized by immunofluorescence staining. A higher magnification of the boxed area in the left panel is shown in the right panel. Arrows indicate junctional localization of Yap and Taz. Scale bar, 20 µm. **(G, H)** Measurement of Tead activity in the indicated EMT systems by luciferase reporter assays. NMuMG and Py2T cells were treated with TGFβ for the indicated durations prior to cell lysis and luminescence analysis. Reporter constructs were transfected along with a Renilla luciferase construct for normalization. Results are shown as mean ± S.E.M. (NMuMG: n = 3; * = p < 0.05; Py2T: n = 5; MT: n = 2; ** = p < 0.01). Top: Scheme of the luciferase reporter constructs. Eight repeats of the core MCAT (G) or GTIIC (H) Tead DNA binding motif was cloned in front of a basal promoter followed by the Firefly luciferase gene. A construct lacking a Tead-binding motif served as control. **(I)** Tead activity was determined using a luciferase reporter construct harboring the promoter of the Cyr61 gene (Cyr61prom WT). The same construct with a mutated Tead binding site (Cyr61prom TeadMut) was used as a control. Results are shown as mean ± S.E.M. (n = 3; * = p < 0.05).

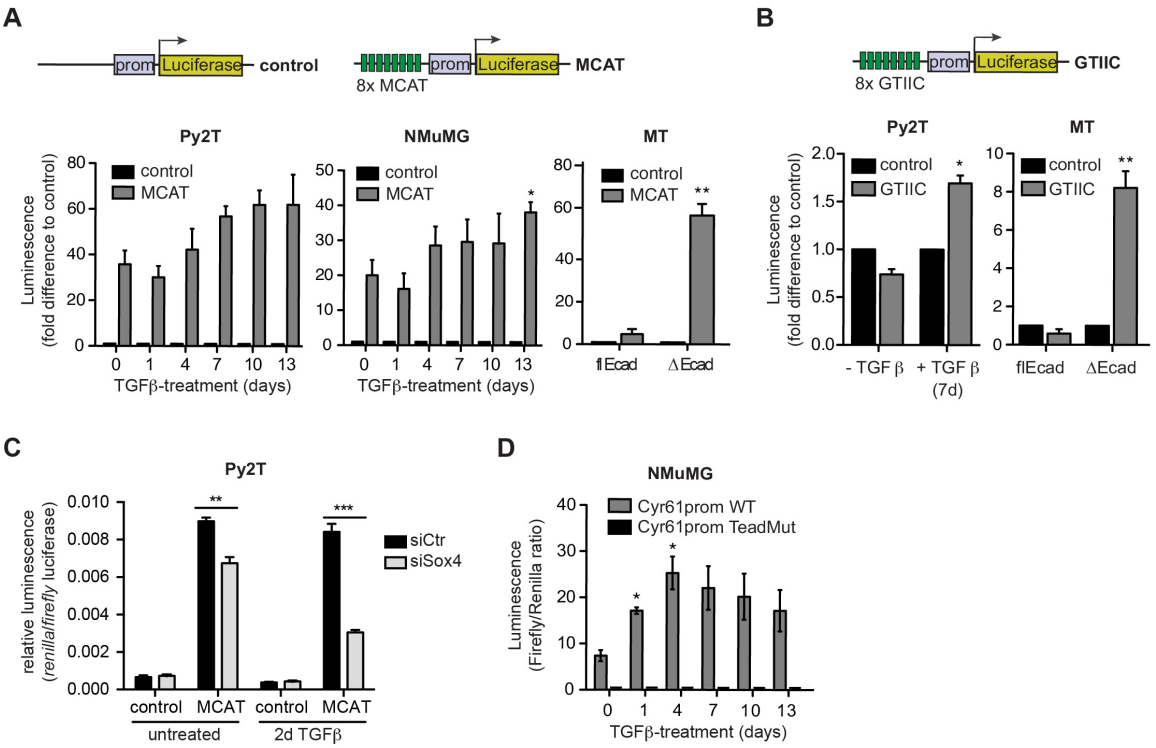


Figure S2. Tead2 transcriptional activity increase during EMT. (A,B) Measurement of Tead activity in the three cellular EMT systems by luciferase reporter assays. Top panels: Scheme of the luciferase reporter constructs. Eight repeats of the core MCAT (A) or GTIIC (B) Tead DNA binding motifs were cloned in front of a basal promoter followed by the Firefly luciferase gene. A construct lacking a Tead-binding motif served as control. Bottom panels: NMuMG and Py2T cells were treated with TGFβ for the indicated durations prior to luminescence analysis. Reporter constructs were transfected along with a Renilla luciferase construct for normalization. Results are shown as mean ± S.E.M. (NMuMG: n = 3; Py2T: n = 5; MT: n = 2; *p < 0.05; **p < 0.01). **(C)** Analysis of Tead transcriptional activity after Sox4 depletion. Py2T cells were transfected with siRNA pools against Sox4 or a control and treated with TGFβ as indicated. Relative Tead activity was determined by using the Tead-responsive MCAT luciferase reporter construct as described in supplementary material Fig. S2A. Results are shown as mean ± S.E.M. (n = 3; **p < 0.01, ***p < 0.001). **(D)** Tead activity was determined using a luciferase reporter construct harboring the promoter of the Cyr61 gene (Cyr61prom WT). The same construct with a mutated Tead binding site (Cyr61prom TeadMut) was used as control. Results are shown as mean ± S.E.M. (n = 3; *p < 0.05).

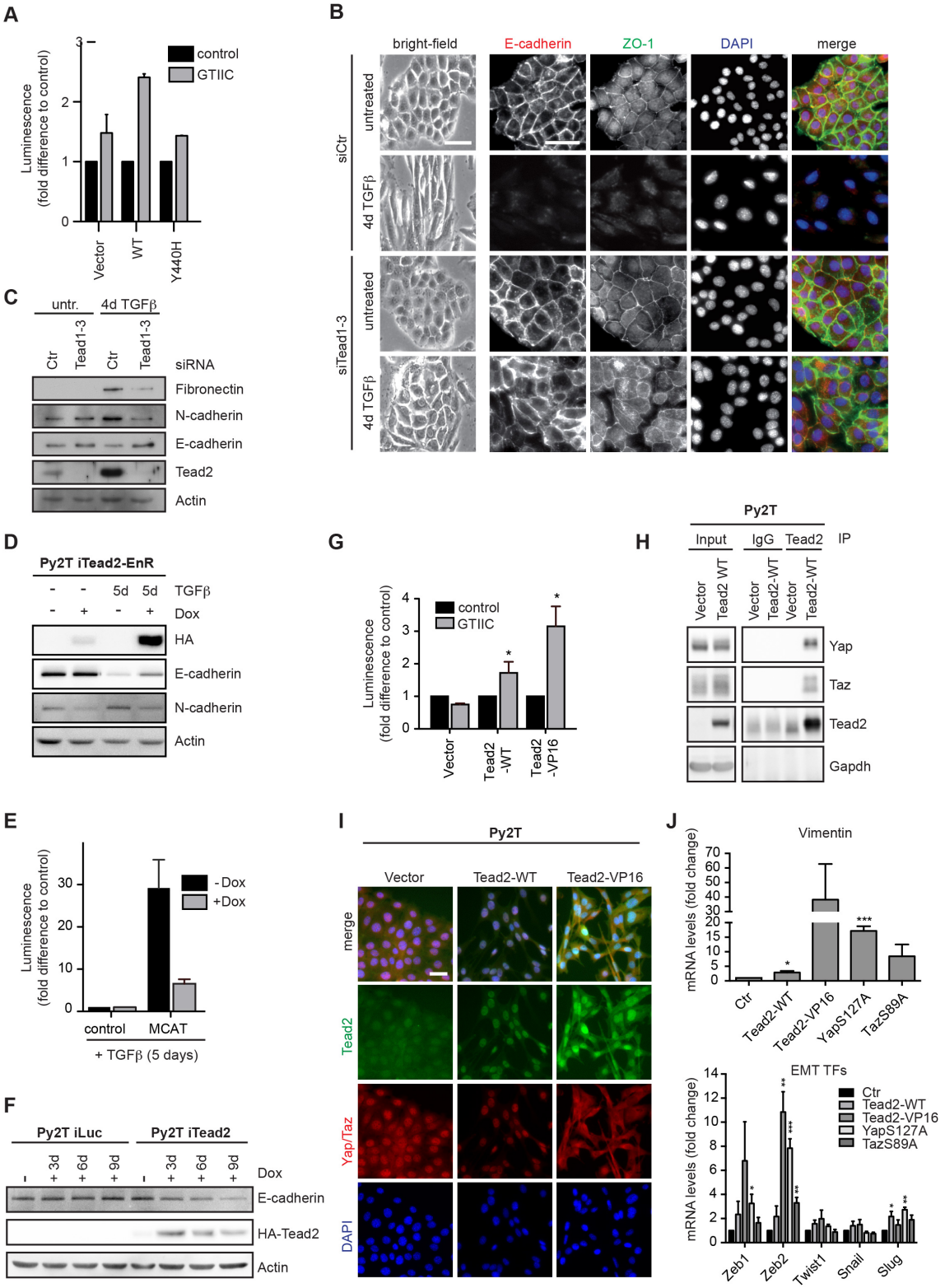


Figure S3. Tead2 upregulation is required for EMT. (A) Effect of Tead2 WT or Tead2 Y440H overexpression on Tead activity. Measurement of Tead activity in NMuMG cells stably expressing wildtype Tead2 (WT) or a point mutant unable to interact with Yap/Taz (Y440H) using the Teadresponsive GTIIC luciferase reporter construct as described in supplementary Figure S2B. Data represent mean ± S.E.M. (n = 2). **(B)** Effect of Tead knockdown on cell morphology and the disassembly of adherens and tight junctions. NMuMG cells were transfected with siRNA pools targeting each Tead1, 2, and 3. One day after transfection, cells were treated or not with TGFβ as indicated. Overall cell morphology (phase contrast) and

immunofluorescence staining of adherens junctions (E-cadherin) and tight-junctions (ZO-1) are shown. DAPI staining visualizes cell nuclei. Scale bars, 50 µm. **(C)** Immunoblotting analysis for the expression of the mesenchymal markers Fibronectin and N-cadherin and the epithelial marker E-cadherin in the experiment described in (B). Immunoblotting for Tead2 was used to assess knockdown efficiency, and immunoblotting for actin was used as a loading control. **(D)** Interference with Tead transcriptional activity in Py2T cells undergoing EMT. A HA-tagged, dominant-negative version of Tead2 (Tead2-EnR) was expressed in Py2T cells in a doxycycline-inducible fashion (Py2T-iTead2-EnR). The Tead2-EnR construct encompasses the Tead2 DNA-binding domain fused to the transcriptional repression domain of Drosophila Engrailed. Py2T-iTead2-EnR cells were induced to undergo EMT by TGFβ treatment for five days in the presence or absence of doxycycline (Dox), and the expression of Tead2-EnR (HA) and of the EMT markers E-cadherin and N-cadherin were analyzed by immunoblotting. Actin was used as loading control. **(E)** In the experiment described in (D), Tead transcriptional activity was determined by using the MCAT firefly luciferase reporter as described in supplemental Figure S1F. Results are presented as mean ± S.E.M. (n = 2). **(F)** Impact of prolonged Tead2 expression on E-cadherin protein levels. Py2T cells were induced to express HA-tagged Tead2-WT (iTead2) or firefly luciferase (iLuc) as a control upon doxycycline treatment for indicated periods. Tead2 (HA) and Ecadherin expression are shown by immunoblotting analysis. Actin served as loading control. **(G)** Analysis of Tead transcriptional activity in Py2T cells expressing wildtype Tead2 (Tead2-WT) or a constitutively active version of Tead2 (Tead2-VP16) using the GTIIC Tead-responsive Firefly luciferase reporter construct (supplemental Figure S1G). The results shown represent mean ± S.E.M. (n = 2; *$p < 0.05$). **(H)** Determination of the amount of Yap/Taz/Tead2 complexes in control (vector) and Tead2-expressing Py2T cells (Tead2-WT). Lysates were subjected to immunoprecipitation with a Tead2 antibody or with irrelevant IgG as negative control (IP). The presence of Tead2 and of co-immunoprecipitated Yap and Taz was analyzed by immunoblotting. The levels of these proteins in total cell lysates were also assessed (Input). Gapdh was used as loading control. **(I)** Immunofluorescence staining of Yap/Taz and Tead2 in Py2T cells expressing wildtype Tead2 (Tead2-WT), constitutively active Tead2 (Tead2-VP16) or transduced with an empty vector control in the absence of TGFβ. Scale bar, 25 µm. **(J)** Analysis of mRNA expression levels of vimentin and EMT transcription factors in Py2T cells stably expressing wildtype Tead2 (Tead2-WT), constitutively active Tead2 (Tead2-VP16) or the Hippo-insensitive mutants of Yap (Yap S127A) and Taz (Taz S89A). The results shown represent mean ± S.E.M. (n = 3; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

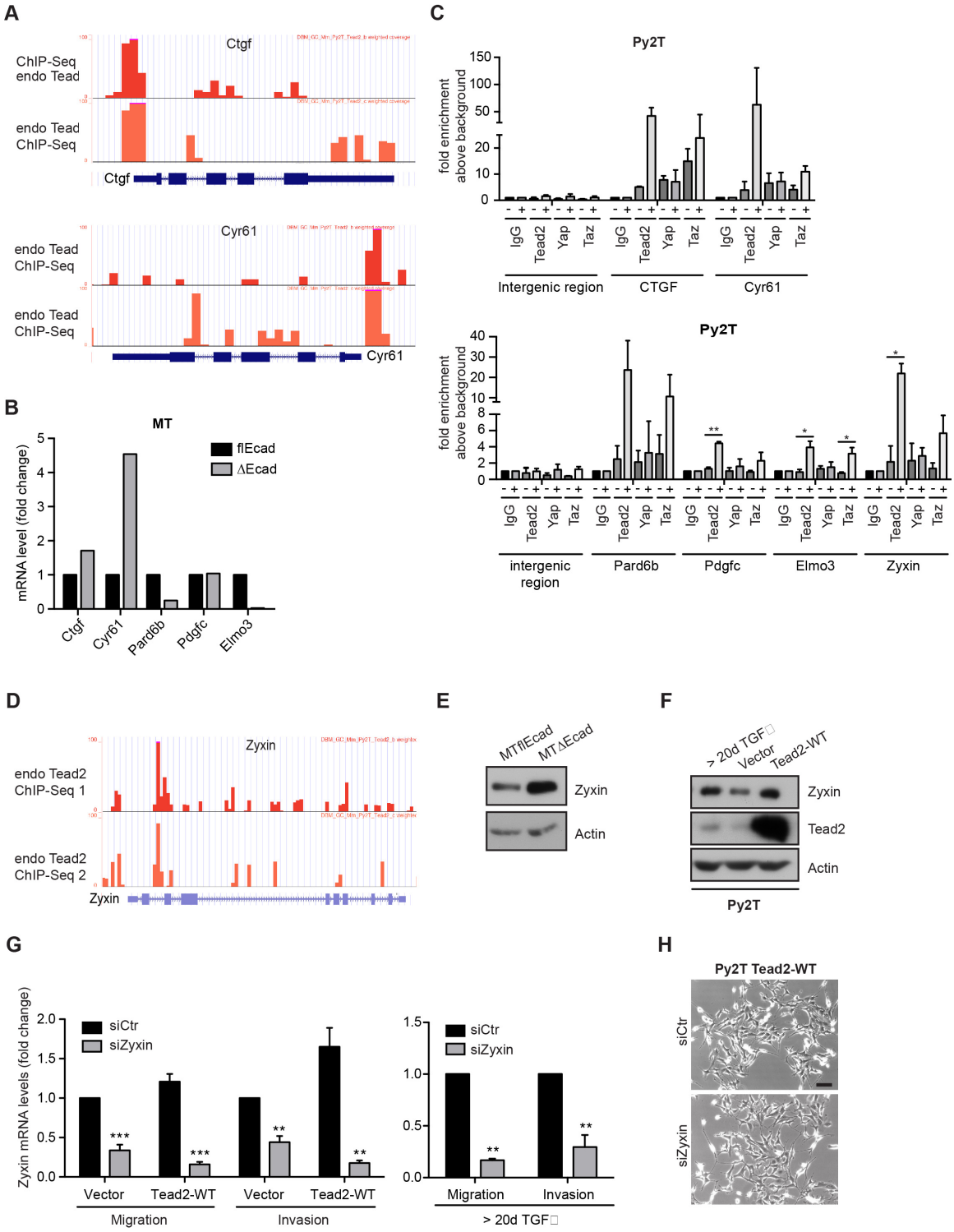


Figure S4. Identification and validation of direct Tead2 target genes. (A) Genome browser view of Tead2 ChIP-Seq data showing the binding of Tead2 at promoter regions of the known Tead target genes Ctgf and Cyr61. Py2T cells were treated with TGFβ for five days, chromatin was prepared and subjected to chromatin immunoprecipitation (ChIP) using a Tead2 antibody. Precipitated chromatin was then subjected to next generation sequencing. Data of two independent biological replicates are shown on separate wiggle-tracks. **(B)** Expression of Tead2 target genes in the MTflEcad/MTΔEcad model of EMT as determined by quantitative RT-PCR. **(C)** Binding of Yap and/or Taz to direct Tead2 target genes. Py2T cells were treated with TGFβ for five days, chromatin was isolated and subjected to ChIP using a Tead2, Yap, Taz or IgG antibody as negative control. ChIP-qPCR analysis was performed as shown in Figure 4E. Data are presented as mean ± S.E.M. (n = 2, *p < 0.05; **p < 0.01). **(D)** Genome browser view of the Tead2 ChIP-Seq

data showing the binding of Tead2 at an intronic region in the zyxin gene (Zyx). **(E)** Immunoblotting analysis of zyxin expression levels in the MTflEcad/MT Ecad model. Actin served as a loading control. **(F)** Immunoblotting analysis of zyxin and Tead2 expression levels in Py2T cells transduced with an empty vector (Vector) or overexpressing wildtype Tead2 (Tead2-WT), and Py2T cells treated with TGFβ for at least 20 days (> 20d TGFβ). **(G)** Efficiency of siRNA-mediated ablation of zyxin expression. Quantification of zyxin mRNA levels by quantitative RT-PCR in control Py2T cells (Vector) or cells expressing wildtype Tead2 (Tead2-WT; left) and in Py2T cells treated with TGFβ for at least 20 days (> 20d TGFβ; right) and treated with siRNA against zyxin (siZyxin) or control siRNA (siCtr). Measurements were performed in parallel to the transwell migration/invasion assays shown in Figure 5K. (n = 3 mean ± S.E.M.; **p < 0.01, ***p < 0.001) **(H)** Cell morphology of Py2T cells stably overexpressing Tead2-WT after siRNAmediated depletion of zyxin. Phase-contrast images of Py2T cells overexpressing Tead2-WT two days after transfection of siRNA pools against zyxin or a control siRNA. Scale bar, 100 µm.

Table S1 is available upon request

Table S2. Genes that are directly regulated by Tead2 during EMT

Listed are genes, which are directly bound by Tead2 and change their expression after five days of TGFβ-induced EMT in Py2T cells. Genes are represented by their gene names, their log2 fold changes (log2FC) and by statistical significance (p-value).

Gene	log2 FC	pValue
Il11	5.56	9.21E-04
Serpine1	3.35	1.51E-03
Olfm2	3.04	5.49E-03
Pdgfc	2.98	1.43E-03
Pgf	2.84	1.58E-03
Itga5	2.80	1.06E-03
Marcksl1	2.54	5.21E-03
Clu	2.53	2.12E-03
Adamts4	2.07	1.47E-03
9930111J21Rik2	2.01	6.11E-03
Hbegf	1.89	2.94E-03
Kctd11	1.87	1.13E-02
Cyr61	1.82	1.71E-03
Rhob	1.76	1.34E-02
Sfn	1.73	3.53E-03
Sfn	1.72	6.36E-03
Ccrn4l	1.61	2.52E-03
Mir181b-2	1.50	2.99E-02
Nppb	1.47	1.90E-02
Ctgf	1.36	2.49E-03
Itpripl2	1.33	2.52E-03
Tnfaip6	1.33	3.53E-03
Rnf122	1.33	3.09E-02
Cd97	1.32	1.21E-02
Ctla2b	1.31	1.89E-02
Tmprss11f	1.30	1.07E-02
Myo7a	1.22	8.21E-03
Smad7	1.21	4.44E-03

Limk1	1.21	7.45E-03
Itpripl2	1.20	6.44E-03
Atg4c	1.19	7.18E-03
Creb3l2	1.12	2.01E-02
Cttnbp2nl	1.08	4.88E-03
Tmcc3	1.08	3.54E-02
Zyx	1.04	2.84E-02
Plaur	1.04	6.24E-03
Ier3	1.03	6.21E-03
Mfhas1	1.03	4.02E-03
Cnn3	1.01	6.17E-03
Podnl1	1.01	9.48E-03
Klk9	0.99	2.75E-02
Atp11a	0.99	5.39E-03
Amotl2	0.96	1.21E-02
Schip1	0.94	6.33E-03
Fbln2	0.94	9.52E-03
Sdcbp2	0.92	1.94E-02
Klf6	0.91	8.44E-03
Zfp451	0.90	2.90E-02
Soga1	0.89	1.35E-02
Egr3	0.89	1.41E-02
Gne	0.87	3.17E-02
Acvr1	0.86	8.82E-03
Adcy7	0.86	2.19E-02
Cnn3	0.83	5.74E-03
Lrrc8c	0.82	1.19E-02
Trim13	0.81	2.25E-02
Mir22hg	0.80	3.80E-02
B230120H23Rik	0.80	1.73E-02
Zfp658	0.79	9.73E-03
Ugdh	0.78	1.43E-02
Loxl3	0.76	1.36E-02
E030011O05Rik	0.75	1.69E-02
Synpo	0.75	1.89E-02
Ercc1	0.75	6.86E-03
Ncor2	0.74	4.56E-02
Csrp1	0.73	2.74E-02
Atp6v1g1	0.72	4.26E-02
H2-Ab1	0.72	1.30E-02
Tmem150a	0.71	3.27E-02
Ankrd1	0.71	9.96E-03
Sept9	0.69	1.50E-02
Zfp568	0.68	4.68E-02
Mapk6	0.66	1.37E-02
Atp6v1g1	0.66	3.74E-02

Fxr2	0.66	1.48E-02
Lamb1	0.63	2.60E-02
Mat2a	0.62	3.97E-02
Cdc42ep1	0.61	2.42E-02
Fpgt	0.60	1.45E-02
Zkscan5	0.59	3.11E-02
Dok1	0.58	2.64E-02
Pcif1	-0.59	1.81E-02
Ect2	-0.62	1.69E-02
Tnfrsf1b	-0.63	1.19E-02
Zfp672	-0.67	2.64E-02
Tnni2	-0.68	3.15E-02
Ift80	-0.68	1.07E-02
Cep68	-0.69	1.05E-02
Tmem106b	-0.70	1.90E-02
Zcwpw1	-0.70	1.77E-02
Ahnak	-0.70	9.37E-03
Unc13d	-0.77	3.39E-02
A430105I19Rik	-0.78	7.18E-03
Tspan4	-0.78	9.58E-03
Ly6f	-0.80	1.91E-02
Rere	-0.83	9.26E-03
Arrdc1	-0.83	9.26E-03
Sh3rf2	-0.88	1.91E-02
Myo6	-0.91	1.83E-02
Nbeal2	-0.93	4.17E-03
Hist2h3c2	-0.99	4.65E-02
Pkhd1l1	-1.00	1.56E-02
Tnfrsf26	-1.02	1.82E-02
Rnf13	-1.04	1.33E-02
Malat1	-1.05	2.49E-02
Yipf2	-1.11	4.02E-02
March9	-1.14	4.55E-02
Tmtc2	-1.15	4.09E-03
Entpd3	-1.16	4.66E-02
Jag2	-1.16	4.07E-03
Tinagl1	-1.18	4.29E-03
Airn	-1.20	6.26E-03
Prss22	-1.20	7.39E-03
Pard6b	-1.21	1.45E-02
Syt8	-1.27	4.40E-03
Capn5	-1.27	1.59E-02
H1f0	-1.31	4.77E-03
Emp2	-1.35	4.61E-02
Clip4	-1.36	4.28E-03
Tlr5	-1.40	3.10E-03

Dcaf4	-1.45	5.04E-03
Atp6v0a1	-1.45	1.33E-02
Esrp2	-1.54	6.11E-03
Ckb	-1.68	5.39E-03
Uap1l1	-1.73	3.75E-02
Atp1b1	-1.82	1.43E-03
Blnk	-1.84	2.04E-03
Rbm47	-2.00	2.87E-02
Mal	-2.07	3.83E-03
Gpr97	-2.08	1.34E-02
Elmo3	-2.12	4.99E-03
Tenc1	-2.18	1.43E-03
Afap1l2	-2.39	2.30E-03
Klk10	-3.24	3.83E-03
Cldn9	-3.89	9.11E-04
Cp	-4.01	3.45E-02

Table S3. Primer sequences used for RT-pPCR

Name	**Sequence (5'-3')**
Tead1	ccaggatcctcacaagacg
	gaatgggggctgtgactg
Tead2	ctgaggacagggaagacgag
	cttcgagccaaaacctgaat
Tead3	gagctgattgcccgctac
	tgtatgtggctggacacctg
Tead4	tcaaaacacctaccctgtcca
	gccctgcaggagactcaa
Zyxin	aagagaagcagcacccaca
	ctctacctccttcagggtcaag
E-cadherin	cgaccctgcctctgaatcc
	tacacgctgggaaacatgagc
Cyr61	ctggcatctccacacgagttac
	tgccctttttaggctgctg
Rpl19	ctcgttgccggaaaaaca
	tcatccaggtcaccttctca
Vimentin	ccaacctttttcttccctgaa
	ttgagtgggtgtcaaccaga
Zeb1	gccagcagtcatgatgaaaa
	tatcacaatacgggcaggtg
Zeb2	ggaggaaaaacgtggtgaactat
	gcaatgtgaagcttgtcctctt
Twist	gccggagacctagatgtcattg
	cacgccctgattcttgtgaa

Snail	ctctgaagatgcacatccgaa
	ggcttctcaccagtgtgggt
Slug	tgtgtctgcaagatctgtggc
	tccccagtgtgagttctaatgtg
Ctgf	ctctcgtcgcctctgcac
	cagtcctggcccatagca
Pard6b	tgggctacgcagatatcca
	ggcactgtagtcagcttcttcc
Elmo3	ctcagcactgccccagat
	gcacccttcacgacttgtatt
Pdgfc	tgtgtcccacgtaaagttacaaa
	tcagtgagtgacttatgcaatcc

Table S4. Primer sequences used for ChIP-qPCR

Name	**Sequence (5'- 3')**
Intergenic region	gctccgggtcctattcttgt
	tcttggtttccaggagatgc
Ctgf	caatccggtgtgagttgatg
	ggcgctggcttttatacg
Cyr61	ctctgatggatctgagaagagg
	gccctttataatgcctgccta
Pard6b	ccacctactttccagcagaga
	gctattttgaggcattactcagc
Elmo3	tgcctctgagaatagctgga
	ggagatccagaactgaaagagaa
Pdgfc	tgctctgtggattgcttcac
	ctccaagaagaagcctgtcc
Zyxin	ccctgtcctgagcagatgtt
	agaacgagccaggttgaaga

3.2 Identification and characterization of miR-1199-5p in EMT and mesenchymal migration

Maren Diepenbruck[1], Nathalie Meyer-Schaller[1], Ryan Goosen[1], Meera Saxena[1] and Gerhard Christofori[1]

[1] Department of Biomedicine, University of Basel, 4058 Basel, Switzerland

- in preparation -

3.2.1 Abstract

Epithelial to mesenchymal transition (EMT) is an essential developmental process, which in part is reactivated by epithelial tumor cells and enables them to convert into metastatic cancer cells. Upon diverse extracellular stimuli, EMT is initiated and driven by transcriptional and post-transcriptional alterations.

Here we focused on post-transcriptional regulation of EMT by microRNAs (miRNAs) and their interaction with transcription factors mediating this dynamic process. We used normal murine mammary gland (NMuMG) cells to induce gradual EMT by exposing the cells to the cytokine transforming growth factor beta (TGFβ). Analyzing different cell dedifferentiation stages by deep sequencing allowed us to obtain detailed insights into global transcriptional changes during this transition in a time-resolved manner. We focused on highly differentially expressed miRNAs and, subsequently, tested their functionality during EMT and mesenchymal tumor cell migration. We identified miR-1199-5p as novel EMT-regulatory miRNA whose transcript levels and promoter activity were continuously downregulated during TGFβ-induced EMT. Ectopic expression of miR-1199-5p prevented TGFβ-induced dedifferentiation of mouse and human untransformed mammary gland cells and decreased the migratory and invasive behavior of mesenchymal breast cancer cells. We further report a negative reciprocal regulation of miR-1199-5p and the key EMT transcription factor Zeb1 during EMT.

These results suggest that miR-1199-5p-Zeb1 interaction play a critical role during EMT and might control cell migration and invasion during breast cancer progression.

3.2.2 Introduction

Epithelial to mesenchymal transition (EMT) as well as its reversal process mesenchymal to epithelial transition (MET) reflect two gradual, well-controlled processes during embryogenesis and wound healing in adults to promote tissue and organ formation [2, 21]. The process is defined by a collapse of cell-cell adhesion and tight junctions, loss of apical-basal cell polarity, a rearrangement of the actin cytoskeleton and a gain of mesenchymal characteristics. These structural changes

further lead to increased cell motility, survival and decreased proliferation [13, 19, 138]. Therefore, this process allows cells to switch forth and back between two different cell types and behaviors to endow the necessity of tissue plasticity.

In the context of malignant tumor progression, epithelial tumor cells can revive an EMT upon diverse extracellular stimuli and consequently gaining metastatic properties [383]. This allows tumor cells to disseminate from the primary tumor, intravasate and survive in the blood circulation. At a distant organ, cells extravasate into the parenchyma and eventually grow out as lethal metastasis possibly promoted by a MET [26, 39, 384-386]. The cytokine TGFβ is an inducer and driver of EMT in transformed epithelial tumor cells. At the primary tumor site it stimulates tumor growth, cell invasion and metastasis [70]. High levels of TGFβ1 are often associated with malignant breast tumor progression [73, 387]. Epithelial tumor cells, once exposed to TGFβ, activate the dedifferentiation process, which involves global changes in the transcriptional and post-transcriptional network [98] [150, 153, 268].

The control and regulation of this highly dynamic process involves epigenetic and mRNA splicing alterations, transcription factor (TF) (in)activation as well as changes in the expression of non-coding RNAs, in particular microRNAs (miRNAs) [74]. MiRNAs represent a class of ~ 22 nucleotide-long RNAs which can regulate EMT-MET cell plasticity on the post-transcriptional level by either promoting mRNA target degradation or preventing mRNA translation [246]. Up to date, several studies have identified differentially regulated miRNAs during cell dedifferentiation. The miR-200 family, which consists of five members (miR-200a/b/c, miR-141 and miR-429), was one of the first miRNA families discovered to play a critical role during EMT. Several research groups have demonstrated that miR-200 family members are required to maintain an epithelial cell morphology for instance by controlling the transcripts of the key EMT TFs Zeb1 and Zeb2. MiR-200 family members bind to specific seed sequences in the Zebs mRNA 3` untranslated region (3`UTR) via classical Watson-Crick base-pairing and thus destabilize Zeb1 and Zeb2 mRNA levels, consequently, strengthening epithelial cell-cell E-cadherin adhesion junctions. Upon activation of an EMT program, miR-200 family members are downregulated, which drives the expression of Zeb1 and Zeb2. In turn, both TFs are reported to directly suppress transcription of the miR-200 family by binding to E-box elements in their promoter region. In sum, these results describe a negative reciprocal regulation of a TF and a miRNA family, which tightly controls tumor-related EMT and breast

cancer progression [254, 265, 268, 269, 388]. These and other studies demonstrating the regulation of EMT/MET plasticity and tumor metastasis by miRNAs highlight the potential of this class of short, non-coding RNAs and illustrate once more their prognostic and therapeutic value for cancer treatment [389].

The regulation of EMT and MET seems to be highly context-dependent and various transcriptional and post-transcriptional changes may vary in the initiation and regulation of these processes. Therefore, analyzing the main checkpoints controlling EMT in physiological versus pathological context may deepen our understanding of cancer-associated pathways. Here we focused on the post-transcriptional regulation of EMT by miRNAs and their interaction with TFs mediating the initial steps of the metastatic cascade. As a model, we utilized normal murine mammary gland (NMuMG) cells and induced a gradual transition by treating the cells with TGFβ for different time points. Deep sequencing analysis of various dedifferentiation stages provided a detailed insight into the transcriptional changes during the initiation of EMT in a time-resolved manner. We focused on highly differentially expressed miRNAs during this dynamic process and, subsequently, tested their functionality in normal and breast cancer cells during EMT and mesenchymal tumor cell migration. We identified miR-1199-5p, which exhibited decreased expression and promoter activity during TGFβ-induced EMT. Forced expression of miR-1199-5p in mouse and human untransformed mammary gland cells blocked EMT and migration and invasion of mesenchymal breast tumor cells. We further report that miR-1199-5p controlled post-transcriptionally the mRNA levels of the key EMT TF Zeb1. Zeb1, in turn, regulated the promoter activity and transcript levels of miR-1199-5p during EMT. Our work demonstrates a new critical double-negative feedback loop between Zeb1 and a miRNA during EMT, which might control EMT-derived metastasis in breast cancer.

3.2.3 Results

3.2.3.1 Identification of new regulatory miRNAs in EMT and mesenchymal tumor cell migration

In the context of malignant tumor progression epithelial tumor cells can gain metastatic properties via an EMT to allow dissemination from the primary tumor and extravasation at a distant organ followed by the outgrowth as a metastasis [26]. To identify new regulatory miRNAs involved in the gradual process of EMT, we performed RNA-sequencing on normal murine mammary gland cells (NMuMG subclone E9; NMuMG/E9) induced to undergo a gradual EMT upon treatment with the cytokine TGFβ. We cultured NMuMG/E9 cells in the presence of TGFβ and isolated total RNA at the following time points: 2, 6, 12, 24, 36, 48, 60, 72 and 96 hours. We then analyzed the differential expression of miRNAs in TGFβ-treated versus epithelial, untreated cells (Figure 1A). This cellular model enabled us to obtain a detailed insight into the kinetic of transcriptional changes during normal epithelial cell transition in a time-resolved manner. We found 32 differentially regulated miRNAs, which displayed an at least 4-fold up- or downregulation (in at least one of the time points with $p < 0.05$) in their expression compared to untreated cells within four days of TGFβ treatment. Unsupervised hierarchical clustering illustrated that roughly half of these differentially expressed miRNAs showed a continuous increase in their expression, whereas the other half a continuous decrease in their expression (Figure 1A, D).

To extract those differentially expressed miRNAs, which have a functional impact on TGFβ-induced EMT we came up with a combination of functional screens. First, we included these 32 differentially regulated miRNAs in an EMT phenotypic microscopy-based screen (Figure 1B). Here, mesenchymal cell characteristics like secretion of the extracellular matrix protein Fibronectin, formation of focal adhesions and actin stress fibers were monitored by fluorescence microscopy and quantified. Using miRNA mimics to overexpress individual miRNAs in NMuMG/E9 cells cultured in the absence or presence of TGFβ (four days) allowed us to test their contribution to EMT. With this screen we identified ten miRNAs which were able to block/delay EMT when overexpressed in cells treated with TGFβ: miR-125b-5p, miR-6944-3p, miR-181b-2-3p, miR-1247-3p, miR-200a-3p, miR-200b-3p, miR-429-3p, miR-1199-5p, miR-145a-3p and miR-504-3p. One miRNA induced EMT in epithelial NMuMG/E9

cells: miR-145a-5p (Figure 1D). Additional analysis of cell morphology and localization of the epithelial adhesion junction protein E-cadherin further confirmed the impact of different miRNAs on EMT (supplementary Figure 1A, B).

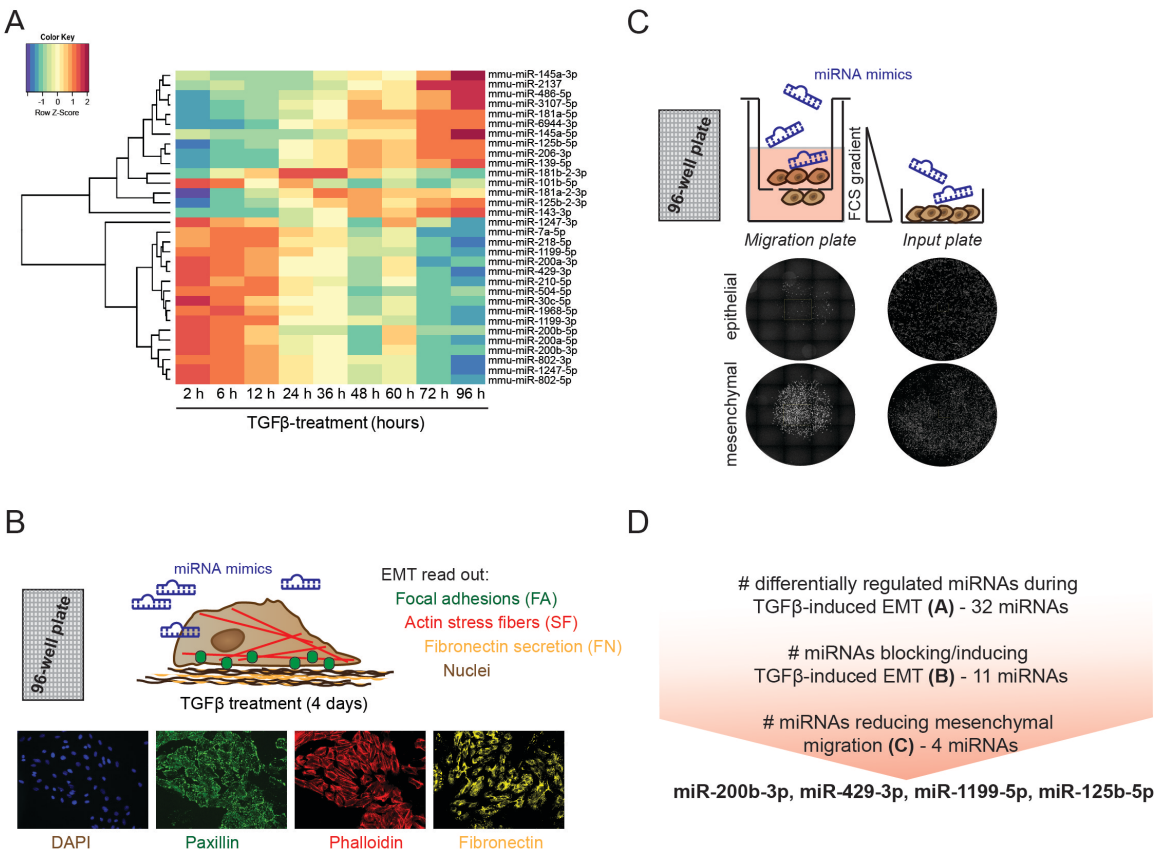


Figure 1: Identification of regulatory miRNAs during TGTβ-induced EMT and mesenchymal migration. **(A)** Expression profiling of miRNAs during TGFβ-induced EMT in NMuMG/E9 cells. NMuMG/E9 cells were treated with TGFβ for the indicated time points and total RNA was isolated for RNA-sequencing. The heat map summarizes the hierarchical clustering of differentially expressed miRNAs (fold change log2 > 2, $p < 0.05$ compared to epithelial, untreated cells) during EMT according to the indicated color scale. Red represents increased expression; blue represents reduced expression. **(B)** Identification of miRNAs controlling EMT. NMuMG/E9 cells were transfected with miRNA mimics for individual miRNAs and analyzed for mesenchymal characteristics in the absence and presence of TGFβ (4 days). Formation of focal adhesions (green), actin cytoskeleton reorganization to actin stress fibers (red), Fibronectin secretion (yellow) and nuclei (brown) were visualized by high-content fluorescence screening microscopy and the different EMT read outs were quantified. **(C)** Identification miRNAs regulating mesenchymal migration. Mesenchymal, migratory Py2T cells (>20 days TGFβ) were transfected with different miRNA mimics and, 2 days later, replated in a Boyden chamber migration insert within a FCS gradient and in parallel on an input plate (96-well plate formats). After 18 hours, cell nuclei were visualized with DAPI and the bottom of the migration inserts and the input plate were imaged using a fluorescence screening microscope. Nuclei were quantified and the number of migrated cells was normalized to the total cell number. **(D)** Overview of the functional miRNA screening process. Depicted are the different screening methods (A-C) along with the number of miRNAs showing differential expression during EMT and the number of miRNAs controlling EMT and mesenchymal migration.

Mesenchymal cell migration is one functional output of EMT, which allows tumor cell dissemination and invasion [20]. In a second screen, we focused on those miRNAs, which had an impact on EMT and tested whether their overexpression in mesenchymal migratory tumor cells, such as Py2T cells treated for > 20 days with TGFβ (long-term TGFβ treated; LT), would influence their trans-well migration behavior (Figure 1C). We transfected Py2T LT cells with different miRNA mimics, two days later, cells were replated within a FCS gradient in a 96-well plate Boyden chamber migration insert (migration plate) and in parallel in a 96-well input plate as a reference for total cell number. 18 hours later, migrated as well as total cell numbers were quantified and normalized to each other. Trans-well migration behavior of Py2T LT cells was significantly reduced (< 1.5-fold) by the expression of four out of eleven miRNAs: miR-200b-3p, miR-429-3p, miR-1199-5p and miR-125b-5p (< 1.35-fold) (Figure 1D; supplementary Figure 2 A, B).

Our screen identified miR-1199-5p as entirely novel regulatory miRNA, which seemed to be involved in TGFβ-induced EMT and mesenchymal tumor cell migration (Figure 1A-D). Several groups have extensively studied the function of the two miR-200 family members miR-200b-3p and miR-429-3p in EMT [254, 268]. The role of miR-125b-5p has also been investigated during tumor cell migration and invasion [390, 391]. Therefore, we decided to focus on the mechanistic function of miR-1199-5p during EMT.

3.2.3.2 Ectopic expression of miRNA-1199-5p inhibits TGFβ-induced EMT

In the course of a TGFβ-induced EMT, miR-1199-5p transcripts were found to be continuously downregulated in NMuMG/E9 cells. This observation was additionally confirmed by a luciferase promoter reporter assay, showing a significant decrease in miR-1199 promoter activity during EMT (Figure 2A, Supplementary Figure 3A). Overexpression of miR-1199-5p in NMuMG/E9 cells verified a more epithelial morphology of cells cultured for four days in the presence of TGFβ compared with cells transfected with a miR-Ctr mimic (Figure 2B). These cells maintained their cell-cell adhesion junction protein E-cadherin and tight junction protein ZO-1 at the plasma membrane while miR-Ctr transfected cells lose these epithelial characteristics (Figure 2C *top*, Figure 2D, E). The reorganization of filamentous actin from cortical to stress fibers as well as formation of focal adhesions are features of mesenchymal cells. In miR-1199-5p-overexpressing cells, these rearrangements

were mostly prevented, here analyzed with Phalloidin to stain filamentous actin and Paxillin, a focal adhesion complex protein (Figure 2C *bottom*). Additional structural mesenchymal markers like extracellular matrix protein Fibronectin and intermediate filament Vimentin are repressed on the mRNA (Figure 2E) and protein (Figure 2D) levels in miR-1199-5p mimic-transfected cells. We further examined the effect of miR-1199-5p on some key EMT TFs, drivers of the transcriptional changes occurring during EMT. We found that miR-1199-5p induced a decrease in Zeb1 mRNA levels, but did not affect the levels of the zinc-finger TF family member Zeb2 [392] or Sox4 [98] (Figure 2E).

A similar phenotype was observed in another cellular TGFβ-inducible EMT system, human untransformed MCF10A mammary epithelial cells. MCF10A cells also maintained their epithelial cell morphology and showed a strong sustained E-cadherin expression at the plasma membrane in cells forced to express miR-1199-5p in the presence of TGFβ. The mesenchymal markers Fibronectin, N-cadherin and Zeb1 were repressed (supplementary Figure 3B-E). In order to test whether miR-1199-5p affects EMT in transformed tumor cells, we ectopically expressed miR-1199-5p in Py2T cells and cultured them for three days in the presence of TGFβ. As shown in supplementary Figure 3F-I, overexpression of miR-1199-5p stabilized E-cadherin expression on the mRNA and protein level. However, it had only minor effects on overall cell morphology and mesenchymal marker expression, such as Fibronectin and Vimentin.

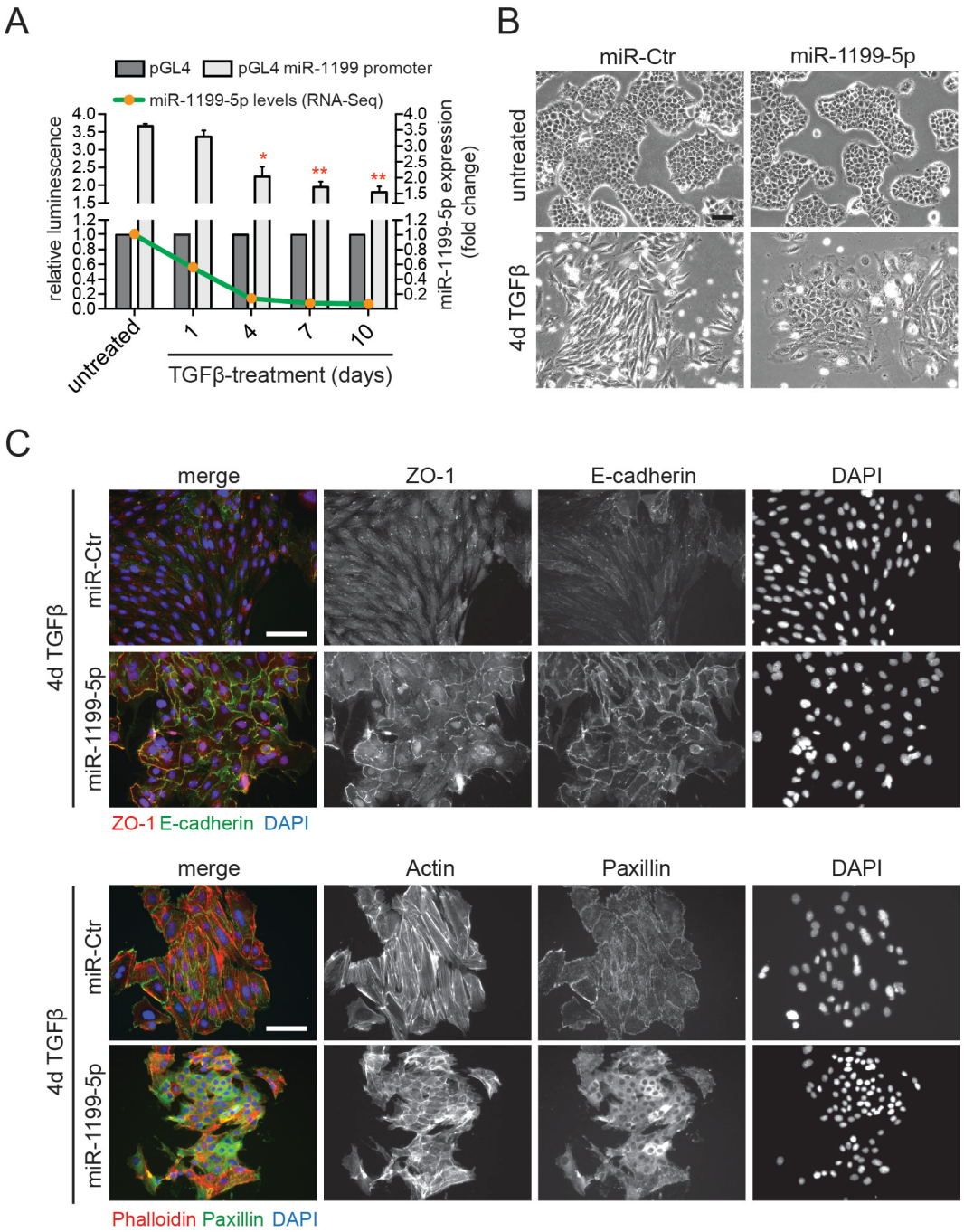

A
pGL4
pGL4 miR-1199 promoter
miR-1199-5p levels (RNA-Seq)
relative luminescence
miR-1199-5p expression (fold change)
untreated
1
4
7
10
TGFβ-treatment (days)
B
miR-Ctr
miR-1199-5p
untreated
4d TGFβ
C
merge
ZO-1
E-cadherin
DAPI
4d TGFβ
miR-Ctr
miR-1199-5p
ZO-1 E-cadherin DAPI
merge
Actin
Paxillin
DAPI
Phalloidin Paxillin DAPI
D
untreated
4d TGFβ
miR-Ctr
miR-1199-5p
Fibronectin
E-cadherin
Vimentin
GAPDH
E
mRNA expression (fold change)
miR-Ctr
miR-1199-5p
E-cadherin
Fibronectin
Zeb1
Zeb2
Sox4

Figure 2: Overexpression of miRNA-1199-5p inhibits TGFβ-induced EMT. (A) miR-1199-5p transcripts as well as miR-1199 promoter activity are reduced during TGFβ-induced EMT. *Green*: Expression profile of miR-1199-5p during EMT in NMuMG/E9 cells measured by RNA-sequencing (Figure 1A). *Grey*: miR-1199 promoter activity during EMT. NMuMG/E9 cells were transfected with a miR-1199 *Firefly* luciferase promoter reporter (pGL4 miR-1199 promoter) or a control reporter (pGL4), which lacks a promoter cassette (supplementary Figure 3A). Additionally, cells were co-transfected with a *Renilla* luciferase reporter and treated with TGFβ for the indicated time points. Relative luminescence (*Firefly*/*Renilla*) was calculated and normalized to the vector control (pGL4). Results are presented as mean fold change +/- S.E.M. of two independent experiments. **(B)** Bright field images of NMuMG/E9 cells transfected with either a control miRNA mimic (miR-Ctr) or a miRNA mimic for miR-1199-5p. The cells were cultured in the absence (untreated) or presence of TGFβ for 4 days. Scale bar: 100 µm. **(C)** Immunofluorescence images of NMuMG/E9 cells transfected with miR-Ctr or miR-1199-5p mimic after 4 days of TGFβ treatment. Immunofluorescence analysis was performed to visualize differences in E-cadherin (green), ZO-1 (red) localization (*top*) and actin cytoskeleton reorganization (Phalloidin in red) and focal adhesion formation (Paxillin in green; *bottom*). DAPI was used to label nuclei (blue). Scale bar: 100 µm. **(D)** Immunoblot analysis for the expression of indicated EMT markers in NMuMG/E9 cells transfected with miR-Ctr or miR-1199-5p mimic and cultured in the absence (untreated) and presence (4 days) of TGFβ. GAPDH served as loading control. **(E)** mRNA expression levels of indicated EMT markers. NMuMG/E9 cells were transfected with miR-Ctr or miR-1199-5p mimic and cultured for 4 days in the presence of TGFβ. qRT-PCR was used to measure the expression of different mRNAs. Results are presented as mean fold change +/- S.E.M of six independent experiments.

In summary, miR-1199-5p transcript levels and its promoter activity are reduced during TGFβ-induced EMT in NMuMG/E9 cells. Forced expression of miR-1199-5p preserved an epithelial cell morphology of NMuMG/E9 and MCF10A cells exposed to TGFβ. In tumor-derived Py2T cells ectopic expression of miR-1199-5p rather induced a delay than a complete block in EMT upon dedifferentiation.

3.2.3.3 Expression of miR-1199-5p reduces mesenchymal tumor cell migration and invasion

A gain in cell migration and invasion is one of the consequences of TGFβ-induced EMT and allows tumor cells to leave the primary tumor and intravasate the blood circulation or extravasate at a distant organ [138].

We tested the influence of miR-1199-5p on mesenchymal tumor cell migration by using a trans-well Boyden chamber assay (Figure 1C). Only a few epithelial, TGFβ-untreated Py2T cells transfected with miR-Ctr mimic were able to trans-migrate towards a higher FCS concentration. Their mesenchymal, TGFβ-treated counterparts on the other hand showed an around 4-fold increase in tumor cell migration. These migratory properties were in turn significantly reduced by an overexpression of miR-1199-5p (Figure 3A, B). This inhibitory effect of miR-1199-5p on cell migration was also observed in 4T1 cells treated with TGFβ (LT) (Figure 3C), another migratory and

invasive breast cancer cell line with a high metastatic potential *in vivo* [393]. Moreover, forced expression of miR-1199-5p in Py2T LT and 4T1 LT cells also significantly decreased trans-well invasion of these cells (Figure 3D, E).

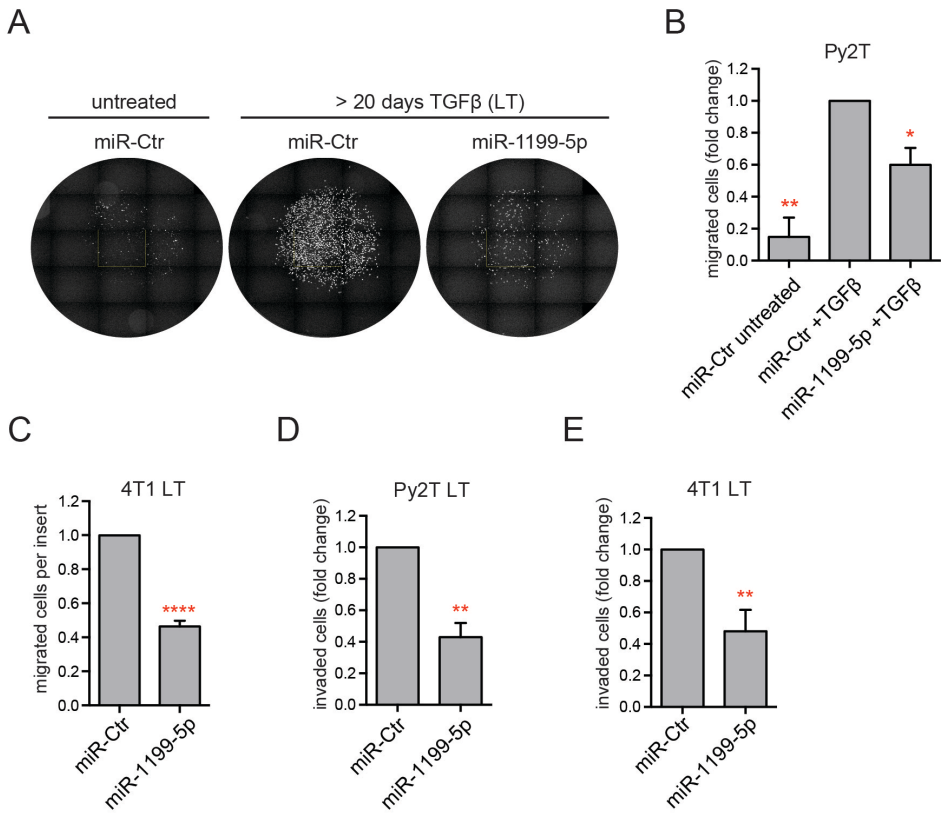


Figure 3: Expression of miR-1199-5p reduces mesenchymal tumor cell migration and invasion. (A) Representative immunofluorescence images of migrated epithelial (untreated) and mesenchymal (long term (LT) TGFβ-treated (> 20 days)) Py2T cells. Cells were transfected with miR-Ctr or miR-1199-5p mimic and, after two days, replated within a FCS gradient in a 96-well Boyden chamber trans-well migration insert and a 96-well reference plate. After 18 hours, cell nuclei were stained with DAPI and cells on the bottom of the inserts as well as the cells in the reference well were imaged and quantified. **(B)** Quantification of migrated Py2T cells from the experiment described in A. The graph represents the number of migrated cells per insert (fold change) normalized to the total cell number. Data are shown as mean +/- S.E.M. of three independent experiments. **(C)** Mesenchymal 4T1 cells were transfected with miR-Ctr or miR-1199-5p mimic and replated within a FCS gradient in a Boyden chamber trans-well migration insert (24-well plate format). After 18 hours, migrated cells were quantified. Graph represents the number of migrated cells per insert (fold change) and data are shown as mean +/- S.E.M. of three independent experiments. **(D, E)** Boyden chamber trans-well invasion assays with mesenchymal Py2T (D) and 4T1 (E) cells. Experiments were performed as described in C. Data represent the mean +/- S.E.M. of three (D) and four (E) independent experiments.

3.2.3.4 miR-1199-5p post-transcriptionally controls the expression of the key EMT transcription factor Zeb1

To understand how miR-1199-5p is able to maintain an epithelial cell morphology, we utilized the online database miRWalk [394] to identify miR-1199-5p mRNA targets that might play a role during EMT. MiRWalk, which itself also predicts mRNA targets with its own algorithm, combines several established miRNA:mRNA databases in order to reduce unwanted false positive mRNA target predictions. Zeb1 turned out to be the top predicted target of mmu-miR-1199-5p in five of eight different databases (TargetScan, miRanda, miRDB, PITA and miRWalk). One potential miR-1199-5p binding site within the mouse Zeb1 3`UTR at position 239-246 was predicted by this *in situ* analysis. Additional sequence alignments of Zeb1's 3`UTR region displayed a well-conserved 8-mer seed sequence in different species (Figure 4A).

In our EMT systems, Zeb1 mRNA expression was increased in NMuMG/E9, Py2T and MCF10A cells within a TGFβ time course experiment (Supplementary Figure 4A). Furthermore, siRNA-mediated knockdown of Zeb1 conserved an epithelial cell morphology of MCF10A cells treated with TGFβ for five days (supplementary Figure 4B) and significantly reduced the migratory behavior of Py2T LT and 4T1 LT cells (supplementary Figure 4C).

In order to examine whether miR-1199-5p indeed regulates Zeb1 levels during EMT, we overexpressed miR-1199-5p in NMuMG/E9 and Py2T cells and analyzed Zeb1 abundance in the absence and presence of TGFβ by immunofluorescence analysis. Ectopic expression of miR-1199-5p in NMuMG/E9 and Py2T cells treated with TGFβ for four or three days, respectively, resulted in the stabilization of cell junction protein E-cadherin, while nuclear levels of Zeb1 were reduced (Figure 4B, supplementary Figure 4D). This result was further confirmed by immunoblotting analysis for Zeb1 in NMuMG/E9, Py2T and MCF10A cells transfected with either miR-1199-5p mimic. Here, Zeb1 proteins levels were diminished upon forced expression of miR-1199-5p in all three cellular EMT model systems treated with TGFβ (Figure 4C, supplementary Figure 4E, F).

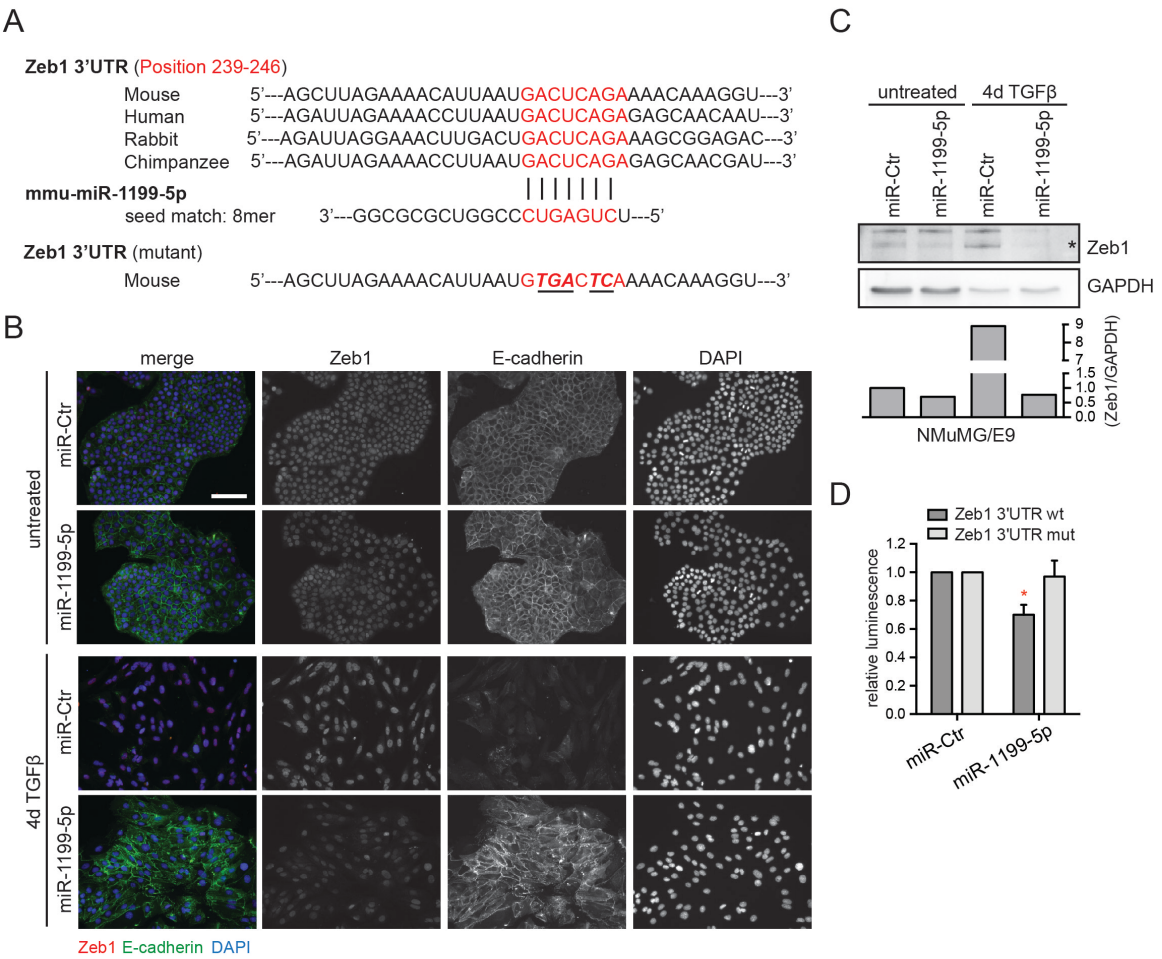


Figure 4: miR-1199-5p post-transcriptionally controls the cellular levels of the transcription factor Zeb1. (A) *Top:* Schematic presentation of the predicted mouse miR-1199-5p target site (red: position 239-249 bps) and its conservation through different species within the 3`UTR of Zeb1 mRNA. *Middle:* Sequence alignment of mouse miR-1199-5p-Zeb1 seed sequence. *Bottom:* Mutated seed sequence (red) within mouse Zeb1 3`UTR. Exchanged nucleotides are underlined. **(B)** miR-1199-5p reduces Zeb1 protein levels. Immunofluorescence images of NMuMG/E9 cells transfected with miR-Ctr or miR-1199-5p mimic. Cells were cultured in the absence (untreated) or presence of TGFβ (4 days) and the localization of epithelial marker E-cadherin (green) and levels of Zeb1 (red) were analyzed. Nuclei were stained with DAPI (blue). Scale bar: 100 µm. **(C)** Immunoblot analysis of Zeb1 protein levels in NMuMG/E9 cells. Experiment was performed as described in B. GAPDH was used as loading control. Levels of Zeb1 in indicated samples were quantified and normalized to GAPDH. **(D)** Post-transcriptional regulation of Zeb1 by miR-1199-5p. NMuMG/E9 cells were transfected with the indicated miRNA mimics, with either a wild type (wt) or a mutated (mut) Zeb1 3`UTR *Firefly* luciferase reporter construct (A) and further a R*enilla* luciferase reporter. After three days, relative luminescence intensities (*Firefly*/*Renilla)* were calculated and normalized to miR-Ctr transfected cells. Data are represented as mean +/- S.E.M. of three independent experiments.

To proof a direct regulation of Zeb1 by miR-1199-5p on the post-transcriptional level, we generated two Zeb1 3`UTR luciferase reporters. One exhibited the wild type miR-1199-5p seed sequence (Zeb1 3`UTR wt) and the other one possessed a mutated seed sequence where five nucleotides have been exchanged (Zeb1 3`UTR mut; Figure 4A *bottom*). NMuMG/E9 cells overexpressing either miR-1199-5p or a

miRNA negative-control were transfected with the Zeb1 3`UTR luciferase reporters mentioned before. Expression of miR-1199-5p mimic in epithelial NMuMG/E9 cells transfected with Zeb1 3`UTR wt reporter led to a significant reduction in relative luminescence compared to miR-Ctr transfected cells. This reduction in luminescence was not observed in cells transduced with 3`UTR Zeb1 mut reporter (Figure 4D).

In summary, we identified Zeb1 as a direct target of miR-1199-5p, which binds a conserved seed sequence in the Zeb1 3`UTR and, therefore, negatively regulates its expression on the post-transcriptional level during EMT.

3.2.3.5 Zeb1 controls the expression of miR-1199-5p on the transcriptional level

Zeb1 has been shown to be directly regulated by different miRNAs [254, 286, 395, 396], and the interplay between Zeb1 and miR-200 family members is probably the best characterized during EMT. Several groups have shown that Zeb1 and miR-200 family members control their transcripts in a double negative feedback loop [272].

To check whether Zeb1 also regulates expression of miR-1199-5p during EMT, we examined the mouse miR-1199 promoter region on chromosome eight. We found several E-box binding motifs (CANNTG; N = G or C) as potential binding sites for Zeb1 (Figure 5A). This encouraged us to test our hypothesis and utilized a miR-1199 luciferase promoter reporter (supplementary Figure 3A). We performed a siRNA-mediated knock down of Zeb1 in NMuMG/E9 cells cultured in the presence of TGFβ for four days and transfected the cells with either a miR-1199 promoter reporter (pGL4 miR-1199 promoter) or a control promoter reporter (pGL4). Luciferase activity analysis revealed a significant increase in miR-1199 promoter activity upon loss of Zeb1 (Figure 5B). Conversely, overexpression of 6xMyc-tagged Zeb1 in epithelial NMuMG/E9 cells significantly decreased miR-1199 promoter activity compared to miR-1199 promoter activity of cells transfected with 6xMyc-tag as negative control (Figure 5C). Furthermore, the regulation of miR-1199 promoter activity by loss or gain of function experiments of Zeb1 correlated with an increase or decrease in miR-1199-5p transcripts, respectively (Figure 5D, E).

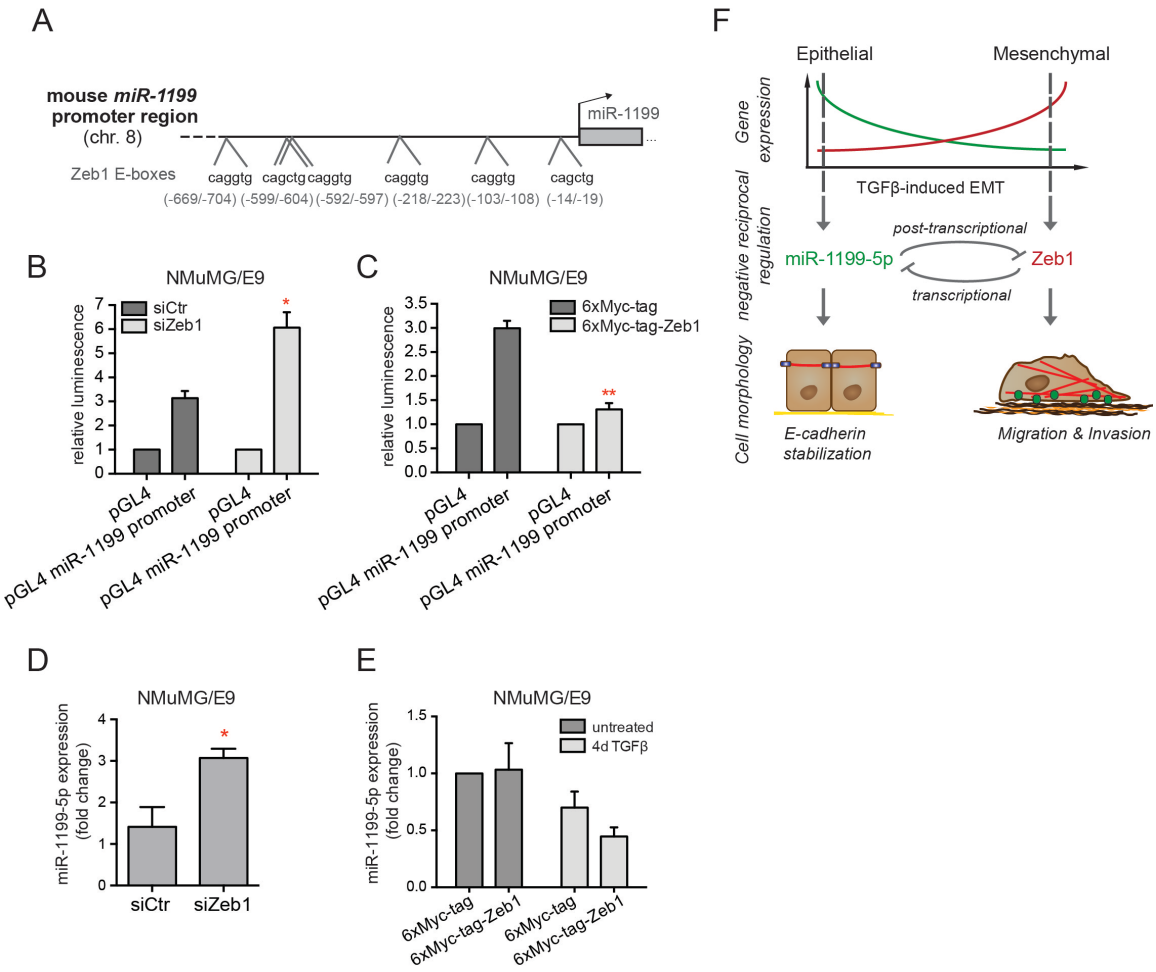


Figure 5: Zeb1 controls the expression of miR-1199-5p on the transcriptional level. (A) Schematic presentation of the mouse miR-1199 promoter region on chromosome eight. *Grey:* Localization of E-boxes (CANNTG, N = G or C) as potential binding sites for Zeb1. *Red:* Localization of primer pairs within the miR-1199 promoter region used for ChIP-PCR analysis in F. **(B, C)** Zeb1 controls the promoter activity of miR-1199. NmuMG/E9 cells were transfected with the *Firefly* luciferase reporter constructs described in supplementary Figure 3A, a *Renilla* reporter and further with either siRNAs against Zeb1 or a control (B) or 6xMyc-tag-Zeb1 or 6xMyc-tag as negative control (C). Cells transfected with siRNAs were cultured for four days with TGFβ and cells transfected with 6xMyc-tag or 6xMyc-tag-Zeb1 were cultured in the absence of the growth factor. Relative luminescence intensities (*Firefly*/*Renilla*) were measured and normalized to the control reporter (pGL4). **(D)** miR-1199-5p expression increases upon loss of Zeb1. NMuMG/E9 cells were transfected with siRNAs against Zeb1 or a negative control and cultured in the presence of TGFβ for four days. miR-1199-5p levels were analyzed by qRT-PCR. Data are represented as mean +/- S.E.M. of three independent experiments. **(E)** miR-1199-5p expression decreases upon overexpression of Zeb1. NMuMG/E9 cells were transiently transfected with 6xMyc-tag-Zeb1 or 6xMyc-tag as negative control and cultured in the absence (untreated) and presence of TGFβ (4 days). miR-1199-5p levels were analyzed by qRT-PCR. Data are represented as mean +/- S.E.M. of three independent experiments. **(F)** Schematic model depicting the mechanistic function of miR-1199-5p during TGFβ-induced EMT in mammary (tumor) cells. Upon EMT induction the transcript levels of miR-1199-5p (green) decrease, whereas Zeb1 levels (red) increase. MiR-1199-5p and Zeb1 control each other in a double negative feedback loop. Expression of miR-1199-5p prevents TGFβ-induced EMT, particularly, stabilizes E-cadherin at the plasma membrane and reduces tumor cell migration and invasion.

These results suggested that Zeb1 functions as a transcriptional repressor of miR-1199, and that miR-1199-5p and Zeb1 indeed regulate each other in a reciprocal negative feedback loop during TGFβ-induced EMT.

3.2.4 Discussion

Cancer-associated EMT can drive early steps of the metastatic cascade by converting epithelial tumor cells into metastatic cancer cells. This developmental dedifferentiation process allows tumor cell dissemination and invasion at the primary site. Mesenchymal tumor cells can enter the blood circulation and extravasate at a distant organ to, eventually, grow out as lethal macro metastasis promoted by a MET [2, 13, 26]. In order to discover conserved EMT-regulatory mechanisms during normal and cancer-associated EMT we utilized the untransformed NMuMG cells, a frequently used model to study TGFβ-induced EMT [357]. A detailed TGFβ time course experiment and subsequent deep sequencing analysis displayed the kinetic transcriptional changes during EMT. Hereby, we established a basis to understand the molecular mechanisms during EMT to potentially interfere and prevent EMT-driven cancer cell dissemination from tumors of an epithelial origin.

In this study, we concentrated on post-transcriptional regulation of EMT by miRNAs and their interaction with transcription factors mediating this dynamic process. Using deep sequencing analysis, we extracted 32 highly differentially regulated miRNAs (> 4-fold; $p < 0.05$) during TGFβ-induced EMT in NMuMG/E9 cells (Figure 1A). We found transcripts of miR-200 family members (miR-200a-3p/-5p, miR-200b-3p/-5p and miR-429-3p) significantly downregulated which is consistent with other studies in NMuMG and other EMT model systems [254, 264, 268]. Kong and colleges have published a similar approach in 2008. They have profiled an EMT miRNA signature induced by canonical TGFβ signaling (24 hours) in NMuMG cells by microarray analysis. They have found miR-155 to be upregulated by the TGFβ/Smad pathway. The authors have convincingly demonstrated that elevated levels of miR-155 promotes EMT through targeting RhoA GTPase, a regulator of cell polarity and tight junction stability [262]. In our deep sequencing analysis we used a very stringent cut-off to select for highly regulated miRNAs during TGFβ-induced EMT. Although miR-155 was not in our final miRNA list, we still observed an upregulation of miR-155

by 2-fold after 24 hours and almost 3-fold after 96 hours of TGFβ treatment (data not shown). However, comparing our miRNA EMT signature with other published signatures it is striking that we did not observe a complete overlap in miRNA expression within the different studies [254, 262-264, 266]. These discrepancies in miRNA expression during EMT might have diverse reasons. They can be due to, for example, different cell model systems, duration of TGFβ treatment, other EMT-inducing stimuli or miRNA expression cut off stringencies. Additionally, it should be noted that the other published miRNA signatures were based on microarray technology. We utilized deep sequencing analysis to study miRNA kinetics during cell dedifferentiation. This method has the advantage to provide better specificity and sensitivity in miRNA detection and allows the identification of so far unknown miRNAs. To our knowledge, this is the first time-resolved miRNA profiling of TGFβ-induced EMT based on deep sequencing analysis. However, these two different miRNA detection technologies can cause divergent miRNA signatures between the different studies.

We further tested the functionality of the 32 highly differentially expressed miRNAs during EMT and mesenchymal tumor cell migration by two different approaches (Figure 1B, C). From this analysis, we extracted four miRNAs, which prevented cells to undergo EMT upon TGFβ stimulation and reduced mesenchymal tumor cell migration. We found the two miR-200 family members miR-200b-3p and miR-429-3p and confirmed their importance for an epithelial differentiated cell phenotype, which has been observed by several other groups before [254, 265, 268, 269]. We also identified miR-125b-5p in our screen, which has been reported by others to facilitate tumor cell migration and invasion [390, 391]. Finally, we discovered miR-1199-5p as a novel EMT-regulatory miRNA with a functional impact on mesenchymal cancer cell migration (Figure 1D).

The transcript levels of miR-1199-5p as well as its promoter activity were continuously downregulated during TGFβ-induced EMT in NMuMG/E9 cells (Figure 1A, Figure 2A). Many studies have shown that epigenetic alterations, especially CpG island hypermethylation-associated miRNA siliencing, could be responsible for the observed deregulation of miRNAs in human cancers [260, 397]. In 2012, Davalos and colleges investigated the reversibility of CpG island hypermethylation of the two miR-200 family gene cluster promoters during EMT. Induction of EMT by TGFβ in MDCK cells correlated with increased CpG island hypermethylation and decreased

miR-200 family expression. Subsequently, TGFβ withdrawal led to MET and a removal of the methylation marks in the miR-200 promoter regions. Furthermore, these methylation patterns in the miR-200 family gene clusters correlated with epithelial or mesenchmyal cell morphology of different human cancer cell lines [102]. MiR-1199-5p and miR-200 family members seem to display a similar regulation and function during EMT. MiR-1199 is located on chromosome 8 in the mouse and on chromosome 19 in the human genome within CpG islands. Whether the expression of miR-1199 is indeed controlled by dynamic hypermethylation during TGFβ-induced EMT and/or metastatic breast cancer cells, will be investigated in the future.

Forced expression of miR-1199-5p in untransformed mouse and human mammary cells prevented cells to undergo EMT upon TGFβ treatment (Figure 2B-E, supplementary Figure 3B-E). The epithelial adhesion junction protein E-cadherin was strongly stabilized in both untransformed cell lines. However, sustained E-cadherin expression upon miR-1199-5p overexpression in the breast cancer cell line Py2T did not prevent a mesenchymal cell morphology, it rather delayed this dedifferentiation process (supplementary Figure 3F-I). Nonetheless, forced expression of miR-1199-5p in mesenchymal Py2T and 4T1 breast cancer cells significantly reduced their migratory and invasive cell behavior (Figure 3A-E). Increased mesenchymal tumor cell motility is an output of cancer-associated EMT, which promotes metastatic spread *in vivo* [138]. Whether gain or loss of function of miR-1199-5p would reduce or increase, respectively, the metastatic potential of orthotopically transplanted 4T1 cells i*n vivo* has to be investigated in the future. To address these questions, we plan to establish an inducible 4T1 cell line, which expresses miR-1199 upon doxycycline treatment. In parallel, genomic excision of the miR-1199 gene in 4T1 cells by using the CRISPR/Cas9 system (Clustered Regularly Interspaced Short Palindromic Repeats) will further help us to understand the role of miR-1199 in the context of malignant tumor progression.

Transcriptional and posttranscriptional networks built by TFs and miRNAs represent two different yet interconnected mechanisms regulating EMT-MET plasticity [143, 269, 290, 291]. The well-described TF families of Snail, Zeb and Twist are considered as key EMT drivers guiding the initiation and maintenance of the cellular dedifferentiation process [138]. Here, we demonstrated a new negative reciprocal regulation of miR-1199-5p and Zeb1 during TGFβ-induced EMT in different breast model systems (Figure 4, Figure 5, supplementary Figure 4). Other

than miR-1199-5p, Zeb1 mRNA levels were upregulated during EMT (supplementary Figure 4A). RNAi-mediated knockdown of Zeb1 prevented TGFβ-induced EMT in MCF10A cells and reduced mesenchymal tumor cell migration of Py2T and 4T1 cells (supplementary Figure 4B, C). Zeb1, together with Zeb2, belongs to the zinc finger E-box-binding homeobox TF family. They bind to E-box promoter elements and directly repress transcription of miR-200 family members or E-cadherin, whose loss is a hallmark of EMT [269, 398, 399]. Whether Zeb1 binds to existing E-box elements in the miR-1199 promoter region (Figure 5A) to directly repress its transcription has to be investigated. Like Zeb1, Zeb2 also inhibited miR-1199 promoter activity and reduced miR-1199-5p transcript levels (data not shown). However, Zeb2 was not a predicted downstream target of miR-1199-5p and its mRNA levels were unchanged upon miR-1199-5p overexpression in NMuMG/E9 cells (Figure 2E). Therefore, both Zeb1 and Zeb2 transcriptionally control miR-1199 expression, but only Zeb1 seems to be post-transcriptionally regulated by miR-1199-5p.

In summary, we identified miR-1199-5p as a regulator of EMT in normal and mammary tumor cells. MiR-1199-5p is expressed in epithelial cells and is downregulated during cell dedifferentiation. Forced expression of miR-1199-5p prevented or delayed TGFβ-induced EMT and significantly reduced mesenchymal tumor cell migration and invasion. We further report a new negative reciprocal regulation of miR-1199 and Zeb1 during EMT (Figure 5F). Several research groups have described double negative feedback loops between different miRNAs and EMT master TFs like Zeb1/2, Snail and Slug, which together tightly control EMT and malignant tumor progression [139, 143, 269, 288, 400].

These feedback loops involving major TFs and miRNAs suggest a crucial role of post-transcriptional regulation during the highly dynamic process of EMT and perhaps cell fate. It is fascinating to think of this kind of regulation as a remnant of tissue development and embryogenesis, when this kind of cell plasticity is required. If so, are these superordinate functional units that facilitate the dynamic switch between an epithelial and mesenchymal cell type? Do they endow cells with a high degree of cellular plasticity, which is necessary during embryogenesis and revived or disturbed during malignant tumor progression? To fully understand the function and regulation of TF-miRNA interactions in EMT-associated tumor progression, additional investigation is required. However, as this study demonstrates cell fate is highly dynamic and is closely regulated on the post-transcriptional level.

3.2.5 Materials and methods

Reagents and Antibodies

Antibodies

The following antibodies were used: E-cadherin (BD Transduction Labs, 610182), N-cadherin (Takara, M142), ZO-1 (Zymed, 617300), Paxillin (BD, 610052), Fibronectin (Sigma-Aldrich, F3648), Vimentin (Sigma-Aldrich, V2258), GAPDH (Sigma-Aldrich, G8795), Zeb1 (Cell Signaling, 3396 or Santa Cruz Biotechnology, sc-25388), Alexa-Fluor 488 and 568 (Molecular Probes), secondary HRP-conjugated antibodies against mouse and rabbit (Jackson ImmunoResearch)

Reagents

Recombinant human TGFβ1 (R&D Systems, 240-B), 4',6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich), Phalloidin Alexa-Fluor 568 (Molecular Probes, A12380)

Cell lines and cell culture

A subclone of normal murine mammary gland cells (NMuMG/E9) has been described previously [357]. The human epithelial mammary gland cells MCF10A have been described earlier [401]. Py2T cells have been established from a MMTV-PyMT mammary tumor and have been described previously [360]. The murine metastatic cell line 4T1 has been described earlier [393].

NMuMG/E9, Py2T and 4T1 cells were cultured in Dulbecco`s modified eagle medium (DMEM; Sigma-Aldrich) supplemented with fetal calf serum (FCS, 10 %; Sigma-Aldrich), glutamine (2 mM; Sigma-Aldrich), penicillin (100 U; Sigma-Aldrich) and streptomycin (0.2 mg/l; Sigma-Aldrich). MCF10A cells were cultured in DMEM/F12 medium supplemented with horse serum (5 %; Bioconcept Amimed), insulin (10 µg/ml; Sigma-Aldrich), hydrocortisone (0.5 µg/ml; Sigma-Aldrich), human EGF (0.02 µg/ml; Invitrogen) and cholera toxin (0.01 µg/ml; Sigma-Aldrich). All cell lines were grown at 37 °C, 5 % CO_2, 95 % humidity.

Transient transfections

miRNA mimic transfection

miRNA mimics are chemically modified, double-stranded RNA molecules and were used for ectopic expression of mature miRNAs. NMuMG/E9, MCF10A, Py2T and 4T1 cells were transfected with 20 nM or 50 nM of mouse or human pre-miR miRNA precursors (Ambion; mmu-miR-1199-5p: PM13577; hsa-miR-1199-5p: PM26554) by using Lipofectamine RNAiMax (Invitrogen) according to the manufacture`s instructions. A random sequence has been used as negative control (Ambion, AM17110 control #1). To maintain overexpression of miR-1199-5p, the transfection was repeated every third day.

siRNA transfection

20 nM of siRNA against murine/human Zeb1 (AM16708) or a scrambled control siRNA (Ambion) was used for transient knock down experiments in NMuMG/E9 and MCF10A cells. Lipofectamine RNAiMax (Invitrogen) was used for the transfection according to the manufacture`s instructions and the transfection was repeated every third day to maintain a knock down of gene transcripts during the experiments.

Plasmids

6xMyc-tagged Zeb1

pCS3-6xMyc-tag-Zeb1 was a kind gift from T. Brabletz (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany). The proper control plasmid pCS3-6xMyc-tag was generated by digesting pCS3-6xMyc-tag-Zeb2 with EcoRI/XbaI, thereby removing the Zeb2 gene but retaining the 6xMyc tag. pCS3-6xMyc-tag-Zeb2 was also a gift from T. Brabletz (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany).

miR-1199 promoter reporter

The genomic promoter region of miR-1199 (800 nucleotides upstream of TSS) of NMuMG/E9 cells was PCR-amplified with the following primers: 5`-tcgactcgaggaccggggaaacactctgta-3` and 5`-tcgaaagcttgcgtctccatctgcaattccgc-3`, which exhibit restriction enzyme sites for XhoI and HindIII, respectively (underlined).

The PCR amplicon was digested with XhoI and HindIII and subcloned into the pGL4.10[luc2] *firefly* luciferase reporter vector purchased from Promega.

Zeb1 3`UTR reporter

A mouse Zeb1 3`UTR reporter was generated by subcloning the Zeb1 3`UTR sequence (166-302 bps) containing a wild type (wt: GACTCAGA) or mutated miR-1199-5p binding site (mut: GTGACTCA) into the pMIR-REPORT *Firefly* luciferase vector (Applied Biosystems) via SpeI/HindIII sites. Integrated DNA Technologies synthesized the different DNA fragments.

Luciferase reporter Assay

A Dual-Luciferase reporter assay (Promega) was used to measure the miR-1199 promoter activity and posttranscriptional repression of Zeb1 3`UTR by miR-1199-5p.

miR-1199 promoter reporter:

In a TGFβ time course experiment, NMuMG/E9 cells were plated in triplicates in a 24-well plate and treated with TGFβ for zero, one, four, seven and ten days. Three days before measuring the luciferase activity, cells were transfected with 0.5 ug of pGL4-miR-1199 *Firefly* luciferase promoter reporter or pGL4 reporter (lacking a promoter cassette) by using Lipofectamine 3000 (Invitrogen). Additionally, cells were transfected with 10 ng of pRL-CMV (Promega), which encodes a *Renilla* luciferase and was used for subsequent cell number normalization.

To test miR-1199 promoter activity upon Zeb1 gain and loss of function studies, NMuMG/E9 cells were plated as described before and reverse transfected with 20 nM of siRNAs against Zeb1 or a negative control (loss of function). On the next day, cells were transfected with the different reporter constructs mentioned above and treated with TGFβ for three days. For Zeb1 gain of function studies, NMuMG/E9 cells were forward transfected with 50 ng of pCS3-6xMyc-tag or pCS3-6xMyc-tag-Zeb1 along with the different luciferase reporters.

Zeb1 3`UTR reporter:

NMuMG/E9 cells were plated as described above and reverse transfected with 20 nM of miR-1199-5p or a negative control mimic. On the next day, cells were

transfected with 0.2 µg of pMIR-Report-Zeb1 3`UTR wt or pMIR-Report-Zeb1 3`UTR mut along with 10 ng of pRL-CMV by using Lipofectamine 3000 (Invitrogen).

Three days after reporter plasmid transfection, cells were washed twice with PBS, 100 µl of 1x passive lysis buffer (Promega) was added per well and incubated for 30 minutes at room temperature. 6 µl of supernatant was transferred to a 96-well plate and *Firefly*/*Renilla* luciferase activity was measured with a luminometer (Berthold Technologies; Centro LB 960).

Quantitative RT-PCR

In order to quantify gene transcripts, total RNA was isolated using Tri Reagent (Sigma-Aldrich) following the manufacture's instruction. cDNA was generated by reverse transcription of RNA using M-MLV reverse transcriptase (Promega) and was quantified by real-time PCR using Mesa Green qPCR MasterMix plus (Eurogentec). Riboprotein L19 was used as internal normalization control. Fold changes were calculated using the comparative Ct method ($\Delta\Delta$Ct).

Primers used for quantitative RT-PCR:

Table 1: List of oligonucleotides used for quantitative RT-PCR

Name	**fwd. primer (5`- 3`)**	**rev. primer (5`- 3`)**
Murine genes		
mRpl19	ctcgttgccggaaaaaca	tcatccaggtcaccttctca
mCdh1	cgaccctgcctctgaatcc	tacacgctgggaaacatgagc
mFibronectin	cccagacttatggtggcaatt	aatttccgcctcgagtctga
mVimentin	ccaacctttttcttccctgaa	ttgagtgggtgtcaaccaga
mCdh2	caatgacgtccaccctgttct	ctgccatgactttctacggaga
mZeb1	gccagcagtcatgatgaaaa	tatcacaatacgggcaggtg
mSox4	cctcgctctcctcgtcct	tcgtcttcgaactcgtcgt
Human genes		
hRpl19	gatgccggaaaaacaccttg	tggctgtacccttccgctt
hCdh1	agaacgcattgccacatacact	tctgatcggttaccgtgatcaa
hFibronectin	gaactatgatgccgaccagaa	ggttgtgcagatttcctcgt
hCdh2	tagtcaccgtggtcaaaccaat	gtgctgaattcccttggctaat
hZeb1	gccaacagaccagacagtgtt	tcttgcccttcctttcctg

For quantitative miRNA analysis, total RNA was isolated using miRNeasy kit (Qiagen). Mature miRNAs were reverse transcript and measured by using the miRCURY LNA Universal RT microRNA PCR kit from Exiqon. U6 snRNA was utilized as internal normalization control during quantitative RT-PCR. Fold changes were calculated using the comparative Ct method ($\Delta\Delta$Ct).

Primers used for quantitative RT-PCR of miRNAs:

- Mouse U6 miRCURY LNA UniRT PCR reference primer mix (Exiqon; 203907)
- mmu-miR-1199-5p miRCURY LNA UniRT PCR primer mix (Exiqon; 206004)

Immunoblotting

Cells were lysed in RIPA buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 10% Glycerol, 1 % NP40, 0.5 % NaDOC, 0.1 % SDS, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 1 mM DTT, 1 mM NaF, 2 mM Na_3VO_4 and 1x protease inhibitor cocktail (Sigmal-Aldrich)) on ice for 20 minutes, centrifuged and protein concentrations were determined via Biorad Bradford solution according to the manufacturer's instructions. Proteins were mixed with SDS-PAGE loading buffer (10% glycerol, 2% SDS, 65 mM Tris, 1 mg/100ml bromophenol blue, 1 % β-mercaptoethanol) and equally loaded on a SDS polyacrylamide gel. After size fractionation, proteins were transferred on an Immobilon-P PVDF membrane (Millipore) using an electrophoretic transfer cell (Biorad) and blocked with 5 % skim milk powder dissolved in TBS/0.05 % Tween for one hour. The membrane was incubated with the primary antibody (listed above) over night at 4 °C or one hour at room temperature. HRP-conjugated secondary antibodies were used to visualize specific proteins by the Fusion Fx7 chemoluminescence reader (Vilber Lourmat).

Immunofluorescence

Cells were grown on uncoated glas cover slips (Menzel-Glaser), fixed with 4 % paraformaldehyde/PBS for 20 minutes, permeabilized with 0.5 % NP40 for 5 minutes and blocked with 3 % BSA/0.01 % Triton-X-100/PBS for 30 minutes. The primary antibodies (listed above) were diluted in 3 % BSA/0.01 % Triton-X-100/PBS and incubated with the cells at room temperature for 1.5 hours, followed by an incubation with a fluorophore-coupled secondary antibody (Alexa Fluor, Invitrogen). Cell nuclei were visualized with DAPI (Sigma-Aldrich). After staining, the cells were mounted

with fluorescence mounting medium (Dako) on microscope slides and imaged using fluorescence microscopy (Leica DMI 4000).

EMT phenotypic microscopy-based screen

NMuMG/E9 cells were plated as duplicates in a 96-well plate and reverse transfected with 20 nM miRNA mimics using Lipofectamine RNAiMax (Invitrogen) prior to four days TGFβ treatment or not. For immunofluorescence analysis, cells were processed as described above and mesenchymal characteristics like formation of focal adhesions (Paxillin), actin stress fibers (Phalloidin) and secretion of Fibronectin were stained with antibodies described before. Images of cells in a 96-well plate were taken with the fluorescence Operetta HCS microscope (Perkin Elmers). EMT features were quantified (Columbus software, version 2.5.0) and the median of two replicates was calculated, normalized to the number of nuclei (DAPI) and compared to the median value of miR-Ctr transfected cells. Standard deviations (SD) were estimated. MiRNAs showing an at least 3x SD for all three EMT read outs or 4x SD in two EMT read outs aside from the median value of miR-Ctr transfected cells were considered as hits to block or induce EMT. Some miRNAs, which turned out to be no hits in the EMT phenotypic microscopy-based screen, but maintained an epithelial cell morphology of NMuMG/E9 cells cultured in the presence of TGFβ (judged by eye), were also considered as hits.

Transwell Assays

Boyden Chamber migration and invasion Assay (24-well plate format)

Mesenchymal Py2T and 4T1 cells were reverse transfected with 50 nM of miRNA mimics in a 6-well plate. Two days later, the cells were trypsinized, washed once with PBS and counted. 25`000 cells were plated as duplicates in a 24-trans-well migration (8 µm pore membrane) or invasion (8 µm pore membrane, covered with a layer of growth factor reduced matrigel) insert (BD Biosciences). After 18 hours, the cells were fixed with 4 % paraformaldehyde/PBS, nuclei were stained with DAPI (Sigma-Aldrich) and non-migrated/invaded cells on the upper surface of the membrane were removed with a cotton swab. Migrated/invaded cells on the bottom of the membrane were imaged with a fluorescence microscope (Leica DMI 4000) and quantified. As chemoattractant a gradient of 0.2 % - 20 % FCS was used.

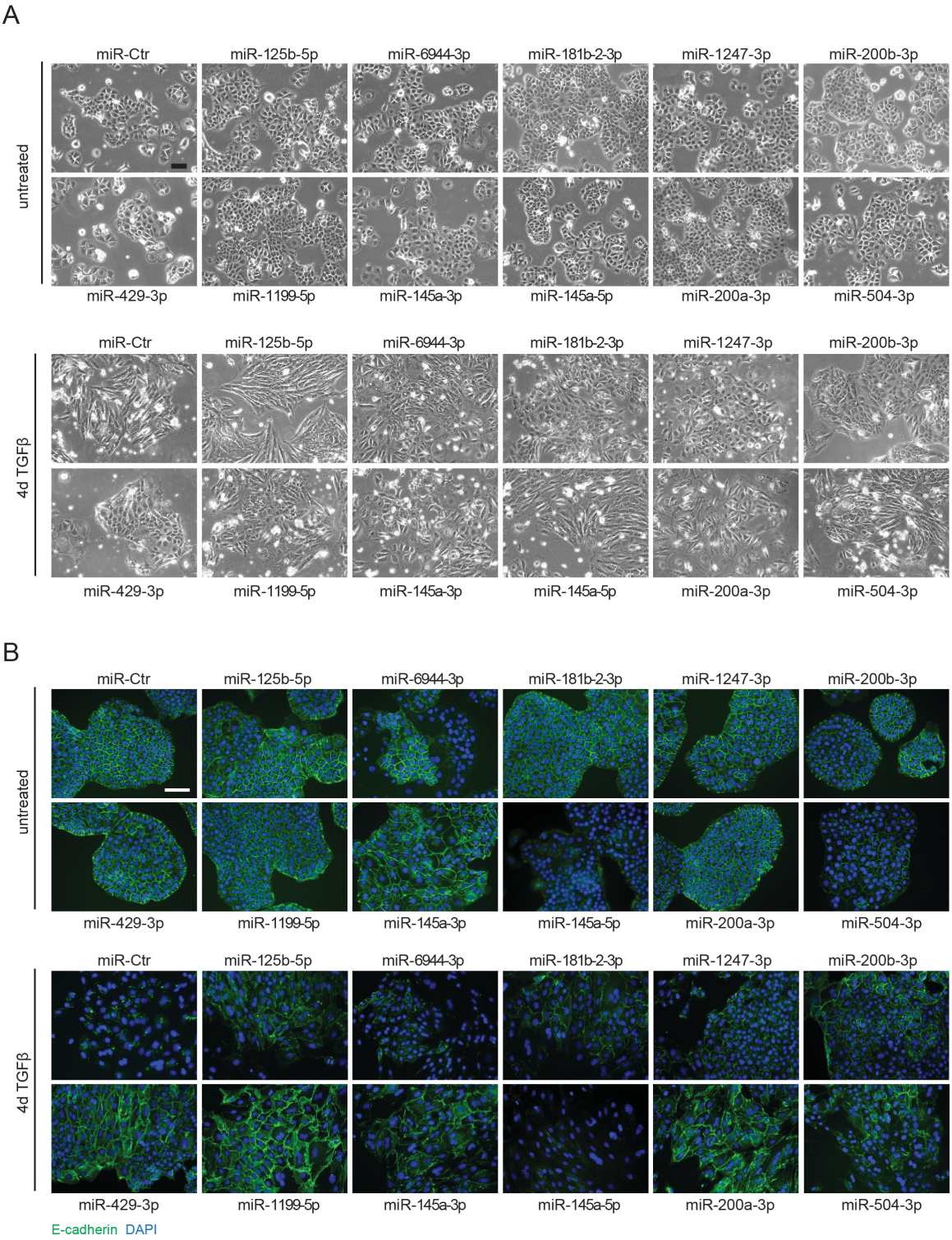
Boyden Chamber migration Assay (96-well plate format)

Mesenchymal Py2T cells were reverse-transfected with 50 nM of miRNA mimics in a 96-well plate. After two days, cells were trypsinized and washed with PBS in a round-bottom 96-well plate (Corning). MiR-Ctr-transfected cells were counted and 6`000 cells were plated as duplicates in the inserts of a 96-well FluoroBlok fluorescent-blocking High Density PET insert (BD Biosciences) and, in parallel, in a 96-well reference plate (Cell Carrier 96, Perkin Elmers). The same cell suspension volume was used to plate cells transfected with other miRNA mimics. As chemoattractant a gradient of 0.2 % - 20 % FCS was used. Migrated cells (bottom of membrane of migration insert) as well as cells in the reference plate were imaged with the Operetta HCS microscope (Perkin Elmers), quantified (Columbus software, version 2.5.0) and normalized to each other.

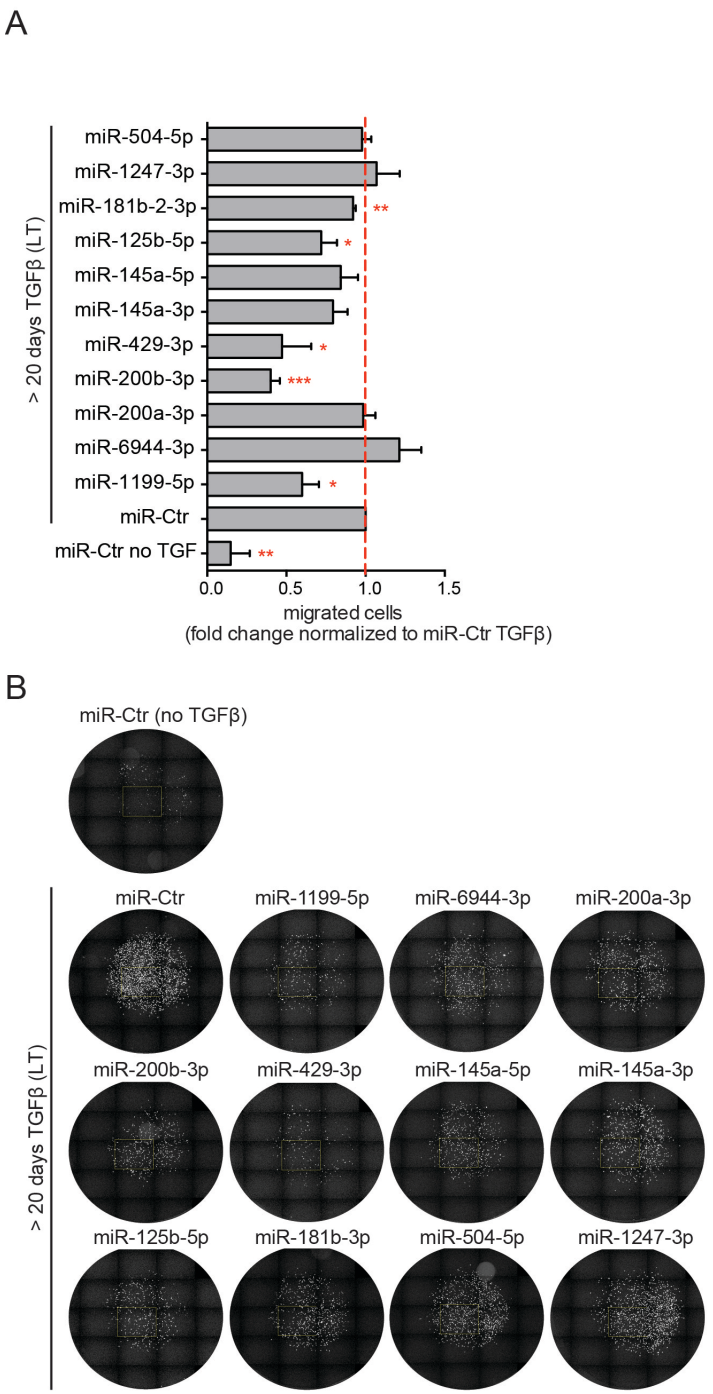
Statistical analysis

Statistical analyses and graphs were generated using GraphPad Prism software (version 6). Statistical analyses were performed by an unpaired, two-sided student t-test with * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

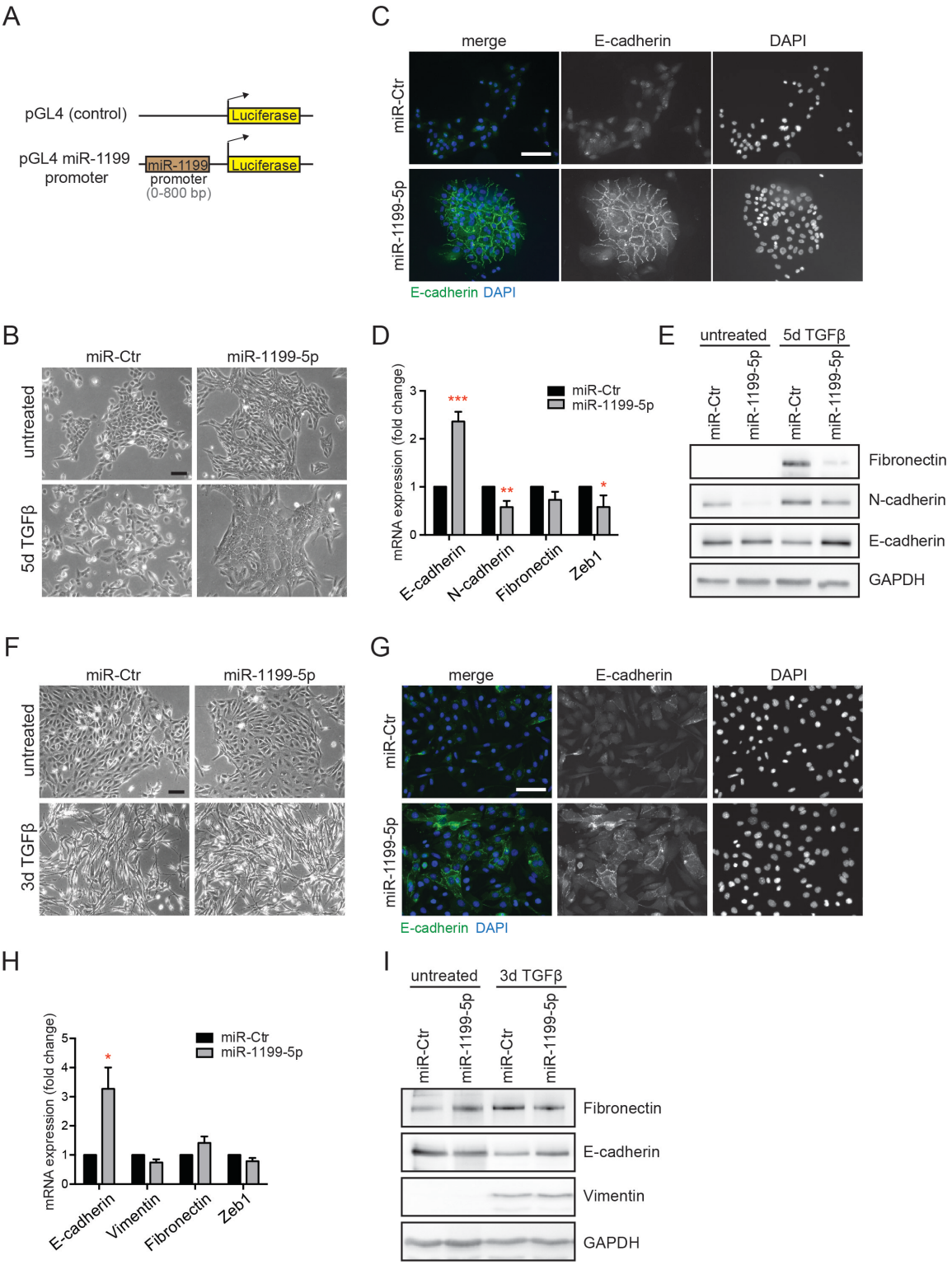
3.2.6 Supplementary data



Supplementary Figure 1: Identified miRNAs affecting TGFβ-induced EMT. (A) Bright field images of NMuMG/E9 cells transfected with the indicated miRNA mimics. *Top*: NMuMG/E9 cells cultured in the absence of TGFβ. *Bottom*: NMuMG/E9 cells cultured in the presence of TGFβ for 4 days. Scale bar: 100 µm. **(B)** Immunofluorescence images of NMuMG/E9 cells cultured and treated as described in A. Localization of the epithelial marker E-cadherin (green) was analyzed. DAPI was used to stain nuclei. Scale bar: 100 µm.

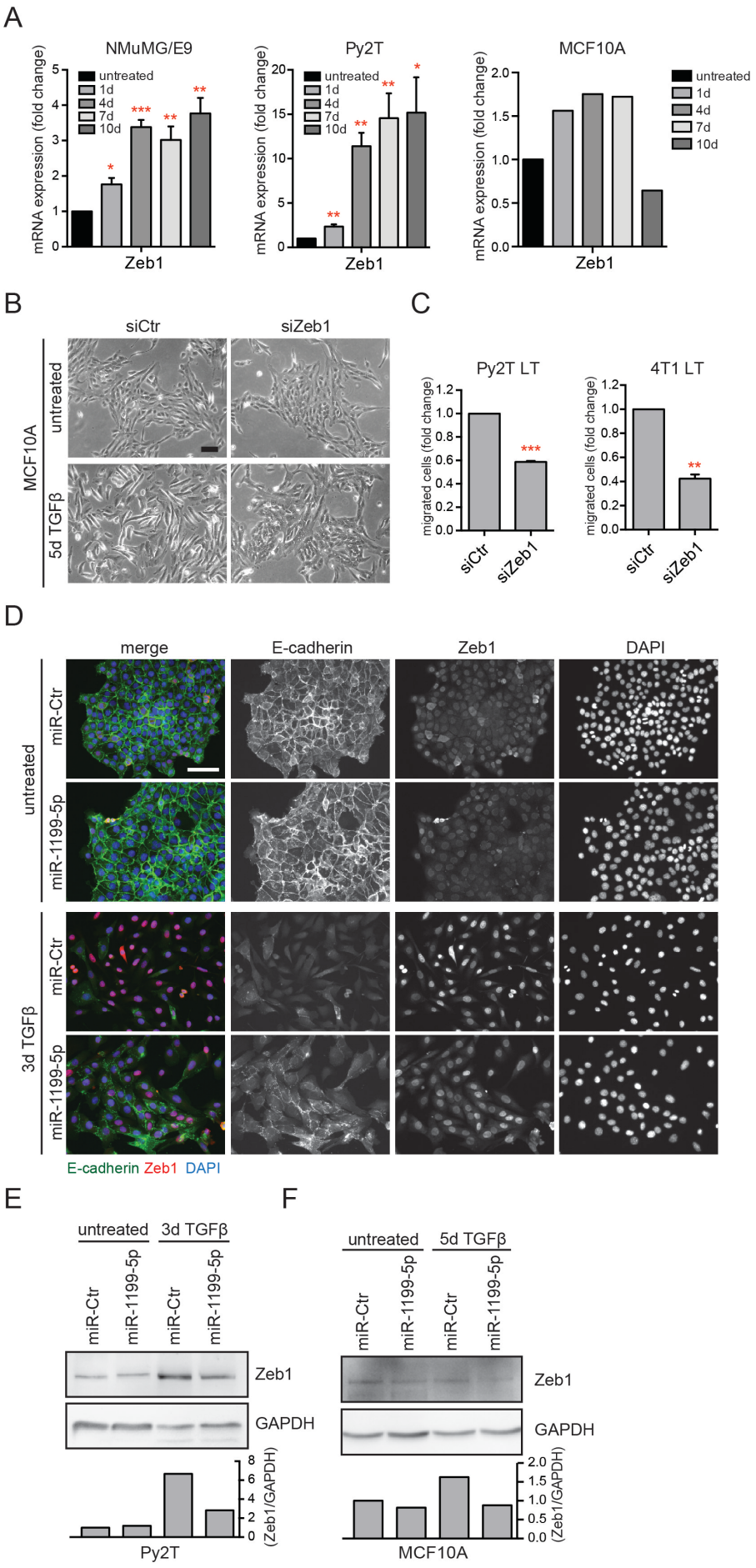


Supplementary Figure 2: Identification of miRNAs regulating mesenchymal tumor cell migration. **(A)** Quantification of migrated epithelial (no TGFβ) and mesenchymal (> 20 days TGFβ; LT) Py2T cells transfected with the indicated miRNA mimics in a 96-well plate format. Graph summarizes the number of migrated cells per insert normalized to the total cell number (input plate). Fold changes of migrated cells were calculated by normalization to the cell number of migrated Py2T LT cells transfected with miR-Ctr mimic (red dashed line). Data are shown as mean +/- S.E.M. of three independent experiments. **(B)** Representative immune-fluorescence images of migrated epithelial (no TGFβ) and mesenchymal (> 20 days TGFβ; LT) Py2T cells from the experiment described in A. Nuclei were visualized with DAPI and the bottom of the trans-well migration inserts were imaged by fluorescence microscopy.



Supplementary Figure 3: Ectopic expression of miR-1199-5p affects EMT in Py2T and MCF10A cells **(A)** Scheme of miR-1199 promoter luciferase reporters. miR-1199 promoter region (-1 → -800 bps) was cloned in front of a *Firefly* luciferase gene (pGL4 miR-1199 promoter). A construct without a promoter region was used as control (pGL4). **(B)** Bright field images of MCF10A cells transfected with miR-Ctr or miR-1199-5p mimic. Cells were cultured in the absence (untreated) and presence (5 days) of TGFβ. Scale bar: 100 µm. **(C)** Immunofluorescence images of MCF10A cells transfected with the indicated miRNA mimics and treated with TGFβ for 5 days. Localization of E-cadherin (green) was analyzed. DAPI was used to stain nuclei. Scale bar: 100 µm. **(D)** mRNA levels of indicated EMT markers in MCF10A cells transfected with indicated miRNA

mimics and treated with TGFβ for 5 days. qRT-PCR analysis was used to determine mRNA expression and data are presented as mean fold change +/- S.E.M. of three independent experiments. **(E)** Immunoblot analysis for indicated EMT markers in MCF10A cells cultured as described in D. GAPDH was used as loading control. **(F)** Bright field images of Py2T cells transfected with miR-Ctr or miR-1199-5p mimic. Cells were cultured in the absence (untreated) and presence (3 days) of TGFβ. Scale bar: 100 μm. **(G)** Immunofluorescence images of Py2T cells transfected with the indicated miRNA mimics and treated with TGFβ for 3 days. Localization of E-cadherin (green) was analyzed. DAPI was used to stain cell nuclei. Scale bar: 100 μm. **(H)** mRNA levels of indicated EMT markers in Py2T cells transfected with the indicated miRNA mimics and treated with TGFβ for 3 days. qRT-PCR analysis was used for the measurement and data are presented as mean fold change +/- S.E.M. of three independent experiments. **(I)** Immunoblot analysis for indicated EMT markers in Py2T cells cultured as described in H. GAPDH was used as loading control.


A
NMuMG/E9
Py2T
MCF10A
mRNA expression (fold change)
untreated
1d
4d
7d
10d
Zeb1
B
siCtr
siZeb1
MCF10A
untreated
5d TGFβ
C
Py2T LT
4T1 LT
migrated cells (fold change)
siCtr
siZeb1
D
merge
E-cadherin
Zeb1
DAPI
untreated
3d TGFβ
miR-Ctr
miR-1199-5p
E-cadherin Zeb1 DAPI
E
F
untreated
3d TGFβ
5d TGFβ
miR-Ctr
miR-1199-5p
Zeb1
GAPDH
(Zeb1/GAPDH)
Py2T
MCF10A

Supplementary Figure 4: The transcription factor Zeb1 controls TGFβ-induced EMT and mesenchymal tumor cell migration and its expression is further controlled by miR-1199-5p in Py2T and MCF10A cells. (A) Zeb1 mRNA levels were determined during TGFβ-induced EMT in NMuMG/E9 (*left*), Py2T (*middle*) and MCF10A cells (*right*) by qRT-PCR analysis. Data are presented as mean fold changes +/- S.E.M. of one (MCF10A) or three (NMuMG/E9, Py2T) independent experiments. **(B)** Bright field images of MCF10A cells transfected with siRNAs against Zeb1 (siZeb1) or a negative control (siCtr) and cultured in the absence (untreated) and presence (5 days) of TGFβ. Scale bar: 100 µm. **(C)** Trans-well Boyden chamber migration assay. Mesenchymal Py2T (*left*) and 4T1 (*right*) cells were transfected with siRNAs against Zeb1 or a negative control. Two days later, cells were replated in a trans-well migration insert (24-well plate format) and migrated cells were quantified after 18 hours. Data are represented as mean fold changes of migrated cells per insert +/- S.E.M. of three independent experiments. **(D)** miR-1199-5p regulates the protein levels of Zeb1 in Py2T cells. Immunofluorescence images of Py2T cells transfected with miR-Ctr or miR-1199-5p mimic. Cells were cultured in the absence (untreated; *top*) or presence of TGFβ (3 days TGFβ; *bottom*). Localization of E-cadherin (green) and levels of Zeb1 (red) were analyzed. DAPI was used to stain cell nuclei (blue). Scale bar: 100 µm. **(E)** Immunoblot analysis of Zeb1 protein levels in Py2T cells. Cells were cultured and treated as described in D. Zeb1 levels were quantified and normalized to GAPDH levels, which was used as loading control. **(F)** Immunoblot analysis of Zeb1 protein levels in MCF10A cells. Cells were transfected with miR-Ctr or miR-1199-5p mimics and cultured in the absence (untreated) or presence of TGFβ for 5 days. GAPDH was used as loading control. Zeb1 protein levels were quantified and normalized to GAPDH protein levels.

4 General conclusions and future plans

The dynamic nature of cellular EMT and MET processes are essential during embryogenesis and wound healing since they promote organ formation and tissue regeneration. Upon diverse extracellular signals, interconnected molecular networks consisting of transcription factors, epigenetic regulators, splicing factors and non-coding RNAs can orchestrate EMT/MET processes to endow the necessity of tissue plasticity in these contexts. However, during tumorigenesis aberrant (in)activation of these networks can induce oncogenic EMT, which promotes tumor cell invasion and malignant tumor progression.

The work presented here has aimed to provide new insights into the molecular mechanisms orchestrating an EMT and its functional consequences in normal and tumorigenic mammary epithelial cells. We have reported the identification and characterization of the transcription factor Tead2. As a transcriptional effector of the Hippo pathway, Tead2 is required during TGFβ-induced EMT and promotes mesenchymal tumor cell migration, invasion and lung colonization *in vivo*. Since Tead2 induces an oncogenic EMT via direct binding to its cofactors Yap and Taz, pharmacological inhibition of the Tead2/Yap/Taz interface might be beneficial for breast cancer treatment. In this case, disruption of Tead2 binding to Yap and/or Taz would prevent oncogenic EMT and might reduce primary tumor cell invasion and their systemic dissemination. Considering the pro-proliferative and anti-apoptotic cellular effects induced by Tead/Yap/Taz transcriptional activity, their pharmacological inhibition might further restrict primary tumor as well as metastatic growth. Some molecules have been reported to disrupt the Tead/Yap/Taz transcriptional complex and future investigations will prove whether this strategy can be translated into the clinics and provide the expected beneficial outcome for cancer patients.

Aside from transcriptional control of EMT by Tead2, we have further reported the identification and characterization of miR-1199-5p - a novel and critical EMT-regulatory miRNA that mechanistically functions in a double-negative feedback regulation with the EMT transcription factor Zeb1. However, whether Zeb1 directly targets existing E-box sequences in the miR-1199 promoter and thereby represses its activity, will be tested by Zeb1 chromatin immunoprecipitation (ChIP) followed by

quantitative RT-PCR. Furthermore, transcriptional downregulation of miR-1199-5p seems to be required for a TGFβ-induced EMT and mesenchymal mammary tumor cell migration and invasion *in vitro*. However, whether its downregulation in epithelial cells is also sufficient to induce EMT and increase cell migration and invasion has to be examined. Additionally, RNA deep sequencing analysis of cells forced to express miR-1199-5p after TGFβ-treatment should help to identify other EMT-relevant targets of miR-1199-5p aside from Zeb1 and further embed miR-1199-5p in an EMT molecular network. Regarding miR-1199-5p's impact on breast cancer progression, the following questions arise and shall be addressed in the future: (a) What are the consequences of miR-1199-5p overexpression or its genetic deletion in the context of breast cancer progression *in vivo*? Does a stable overexpression of miR-1199-5p in breast tumor cells lead to a reduction and its loss to an increase in breast cancer metastases of orthotopically transplanted mouse 4T1 cells? (b) Does miR-1199-5p expression correlate with specific breast cancer subtypes in patients? (c) How is miR-1199 expression regulated during EMT and during tumor progression? Strikingly, the MIR-1199 gene is located within CpG islands in the mouse and human genome, therefore its transcription might be further controlled by DNA hypermethylation during TGFβ-induced EMT and/or in metastatic breast cancer cells. The identification and initial characterization of miR-1199-5p during EMT raises many exciting questions, which should be investigated in the future.

5 References

1. Hay, E.D., *Organization and fine structure of epithelium and mesenchyme in the developing chick embryo*. Epithelial-mesenchymal interactions, ed. R.F.a.R.E. Billingham. 1968, Baltimore, Maryland, USA: Wiliams and Wilkins.
2. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.
3. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol, 2014. **15**(3): p. 178-96.
4. Yilmaz, M. and G. Christofori, *EMT, the cytoskeleton, and cancer cell invasion.* Cancer Metastasis Rev, 2009. **28**(1-2): p. 15-33.
5. Huang, R.Y., P. Guilford, and J.P. Thiery, *Early events in cell adhesion and polarity during epithelial-mesenchymal transition.* J Cell Sci, 2012. **125**(Pt 19): p. 4417-22.
6. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nat Rev Cancer, 2004. **4**(2): p. 118-32.
7. Vleminckx, K., et al., *Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role.* Cell, 1991. **66**(1): p. 107-19.
8. Wheelock, M.J., et al., *Cadherin switching.* J Cell Sci, 2008. **121**(Pt 6): p. 727-35.
9. Yilmaz, M. and G. Christofori, *Mechanisms of motility in metastasizing cells.* Mol Cancer Res, 2010. **8**(5): p. 629-42.
10. Savagner, P., K.M. Yamada, and J.P. Thiery, *The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition.* J Cell Biol, 1997. **137**(6): p. 1403-19.
11. Yang, X., et al., *Regulation of beta 4-integrin expression by epigenetic modifications in the mammary gland and during the epithelial-to-mesenchymal transition.* J Cell Sci, 2009. **122**(Pt 14): p. 2473-80.
12. Maschler, S., et al., *Tumor cell invasiveness correlates with changes in integrin expression and localization.* Oncogene, 2005. **24**(12): p. 2032-41.
13. Thiery, J.P. and J.P. Sleeman, *Complex networks orchestrate epithelial-mesenchymal transitions.* Nat Rev Mol Cell Biol, 2006. **7**(2): p. 131-42.
14. Nistico, P., M.J. Bissell, and D.C. Radisky, *Epithelial-mesenchymal transition: general principles and pathological relevance with special emphasis on the role of matrix metalloproteinases.* Cold Spring Harb Perspect Biol, 2012. **4**(2).
15. Sheppard, D., *Integrin-mediated activation of latent transforming growth factor beta.* Cancer Metastasis Rev, 2005. **24**(3): p. 395-402.
16. Xu, J., S. Lamouille, and R. Derynck, *TGF-beta-induced epithelial to mesenchymal transition.* Cell Res, 2009. **19**(2): p. 156-72.
17. Hall, A., *Rho GTPases and the control of cell behaviour.* Biochem Soc Trans, 2005. **33**(Pt 5): p. 891-5.
18. Mendez, M.G., S. Kojima, and R.D. Goldman, *Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition.* FASEB J, 2010. **24**(6): p. 1838-51.
19. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* J Clin Invest, 2009. **119**(6): p. 1420-8.
20. Nieto, M.A., *Epithelial plasticity: a common theme in embryonic and cancer cells.* Science, 2013. **342**(6159): p. 1234850.
21. Acloque, H., et al., *Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease.* J Clin Invest, 2009. **119**(6): p. 1438-49.
22. Gupta, G.P. and J. Massague, *Cancer metastasis: building a framework.* Cell, 2006. **127**(4): p. 679-95.
23. Klein, C.A., *Parallel progression of primary tumours and metastases.* Nat Rev Cancer, 2009. **9**(4): p. 302-12.
24. Chambers, A.F., A.C. Groom, and I.C. MacDonald, *Dissemination and growth of cancer cells in metastatic sites.* Nat Rev Cancer, 2002. **2**(8): p. 563-72.
25. Sosa, M.S., P. Bragado, and J.A. Aguirre-Ghiso, *Mechanisms of disseminated cancer cell dormancy: an awakening field.* Nat Rev Cancer, 2014. **14**(9): p. 611-22.
26. Chaffer, C.L. and R.A. Weinberg, *A perspective on cancer cell metastasis.* Science, 2011. **331**(6024): p. 1559-64.
27. Brabletz, T., et al., *Invasion and metastasis in colorectal cancer: epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin.* Cells Tissues Organs, 2005. **179**(1-2): p. 56-65.
28. Perl, A.K., et al., *A causal role for E-cadherin in the transition from adenoma to carcinoma.* Nature, 1998. **392**(6672): p. 190-3.
29. Yu, M., et al., *Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition.* Science, 2013. **339**(6119): p. 580-4.

30. Kallergi, G., et al., *Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients.* Breast Cancer Res, 2011. **13**(3): p. R59.
31. Rhim, A.D., et al., *EMT and dissemination precede pancreatic tumor formation.* Cell, 2012. **148**(1-2): p. 349-61.
32. Prat, A., et al., *Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.* Breast Cancer Res, 2010. **12**(5): p. R68.
33. Tsai, J.H., et al., *Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis.* Cancer Cell, 2012. **22**(6): p. 725-36.
34. Ocana, O.H., et al., *Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1.* Cancer Cell, 2012. **22**(6): p. 709-24.
35. Stankic, M., et al., *TGF-beta-Id1 signaling opposes Twist1 and promotes metastatic colonization via a mesenchymal-to-epithelial transition.* Cell Rep, 2013. **5**(5): p. 1228-42.
36. Tran, H.D., et al., *Transient SNAIL1 expression is necessary for metastatic competence in breast cancer.* Cancer Res, 2014. **74**(21): p. 6330-40.
37. Brabletz, T., et al., *Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10356-61.
38. Kowalski, P.J., M.A. Rubin, and C.G. Kleer, *E-cadherin expression in primary carcinomas of the breast and its distant metastases.* Breast Cancer Res, 2003. **5**(6): p. R217-22.
39. Brabletz, T., *To differentiate or not--routes towards metastasis.* Nat Rev Cancer, 2012. **12**(6): p. 425-36.
40. Tarin, D., E.W. Thompson, and D.F. Newgreen, *The fallacy of epithelial mesenchymal transition in neoplasia.* Cancer Res, 2005. **65**(14): p. 5996-6000; discussion 6000-1.
41. Entenberg, D., et al., *Imaging tumor cell movement in vivo.* Curr Protoc Cell Biol, 2013. **Chapter 19**: p. Unit19 7.
42. Raimondi, C., et al., *Epithelial-mesenchymal transition and stemness features in circulating tumor cells from breast cancer patients.* Breast Cancer Res Treat, 2011. **130**(2): p. 449-55.
43. Kang, Y. and K. Pantel, *Tumor cell dissemination: emerging biological insights from animal models and cancer patients.* Cancer Cell, 2013. **23**(5): p. 573-81.
44. Polyak, K., *Heterogeneity in breast cancer.* J Clin Invest, 2011. **121**(10): p. 3786-8.
45. Sinn, H.P. and H. Kreipe, *A Brief Overview of the WHO Classification of Breast Tumors, 4th Edition, Focusing on Issues and Updates from the 3rd Edition.* Breast Care (Basel), 2013. **8**(2): p. 149-54.
46. Bertos, N.R. and M. Park, *Breast cancer - one term, many entities?* J Clin Invest, 2011. **121**(10): p. 3789-96.
47. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
48. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
49. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
50. Herschkowitz, J.I., et al., *Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors.* Genome Biol, 2007. **8**(5): p. R76.
51. Sarrio, D., et al., *Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype.* Cancer Res, 2008. **68**(4): p. 989-97.
52. Weigelt, B., B. Kreike, and J.S. Reis-Filho, *Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis.* Breast Cancer Res Treat, 2009. **117**(2): p. 273-80.
53. Creighton, C.J., et al., *Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13820-5.
54. Feng, X.H. and R. Derynck, *Specificity and versatility in tgf-beta signaling through Smads.* Annu Rev Cell Dev Biol, 2005. **21**: p. 659-93.
55. Shi, Y. and J. Massague, *Mechanisms of TGF-beta signaling from cell membrane to the nucleus.* Cell, 2003. **113**(6): p. 685-700.
56. Hoot, K.E., et al., *Keratinocyte-specific Smad2 ablation results in increased epithelial-mesenchymal transition during skin cancer formation and progression.* J Clin Invest, 2008. **118**(8): p. 2722-32.
57. Shirakihara, T., M. Saitoh, and K. Miyazono, *Differential regulation of epithelial and mesenchymal markers by deltaEF1 proteins in epithelial mesenchymal transition induced by TGF-beta.* Mol Biol Cell, 2007. **18**(9): p. 3533-44.
58. Kang, Y., C.R. Chen, and J. Massague, *A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells.* Mol Cell, 2003. **11**(4): p. 915-26.
59. Vincent, T., et al., *A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF-beta mediated epithelial-mesenchymal transition.* Nat Cell Biol, 2009. **11**(8): p. 943-50.
60. Tian, X.J., H. Zhang, and J. Xing, *Coupled reversible and irreversible bistable switches underlying TGFbeta-induced epithelial to mesenchymal transition.* Biophys J, 2013. **105**(4): p. 1079-89.
61. Lu, M., et al., *MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination.* Proc Natl Acad Sci U S A, 2013. **110**(45): p. 18144-9.
62. Zhang, J., et al., *TGF-beta-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops.* Sci Signal, 2014. **7**(345): p. ra91.

63. Edlund, S., et al., *Transforming growth factor-beta-induced mobilization of actin cytoskeleton requires signaling by small GTPases Cdc42 and RhoA.* Mol Biol Cell, 2002. **13**(3): p. 902-14.
64. Bhowmick, N.A., et al., *Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism.* Mol Biol Cell, 2001. **12**(1): p. 27-36.
65. Ozdamar, B., et al., *Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity.* Science, 2005. **307**(5715): p. 1603-9.
66. Bakin, A.V., et al., *Phosphatidylinositol 3-kinase function is required for transforming growth factor beta-mediated epithelial to mesenchymal transition and cell migration.* J Biol Chem, 2000. **275**(47): p. 36803-10.
67. Yu, L., M.C. Hebert, and Y.E. Zhang, *TGF-beta receptor-activated p38 MAP kinase mediates Smad-independent TGF-beta responses.* EMBO J, 2002. **21**(14): p. 3749-59.
68. Hocevar, B.A., T.L. Brown, and P.H. Howe, *TGF-beta induces fibronectin synthesis through a c-Jun N-terminal kinase-dependent, Smad4-independent pathway.* EMBO J, 1999. **18**(5): p. 1345-56.
69. Zavadil, J., et al., *Genetic programs of epithelial cell plasticity directed by transforming growth factor-beta.* Proc Natl Acad Sci U S A, 2001. **98**(12): p. 6686-91.
70. Massague, J., *TGFbeta in Cancer.* Cell, 2008. **134**(2): p. 215-30.
71. Levy, L. and C.S. Hill, *Alterations in components of the TGF-beta superfamily signaling pathways in human cancer.* Cytokine Growth Factor Rev, 2006. **17**(1-2): p. 41-58.
72. Pardali, K. and A. Moustakas, *Actions of TGF-beta as tumor suppressor and pro-metastatic factor in human cancer.* Biochim Biophys Acta, 2007. **1775**(1): p. 21-62.
73. Dalal, B.I., P.A. Keown, and A.H. Greenberg, *Immunocytochemical localization of secreted transforming growth factor-beta 1 to the advancing edges of primary tumors and to lymph node metastases of human mammary carcinoma.* Am J Pathol, 1993. **143**(2): p. 381-9.
74. De Craene, B. and G. Berx, *Regulatory networks defining EMT during cancer initiation and progression.* Nat Rev Cancer, 2013. **13**(2): p. 97-110.
75. Pan, Q., et al., *Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing.* Nat Genet, 2008. **40**(12): p. 1413-5.
76. Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. **456**(7221): p. 470-6.
77. Shapiro, I.M., et al., *An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype.* PLoS Genet, 2011. **7**(8): p. e1002218.
78. Warzecha, C.C., et al., *ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing.* Mol Cell, 2009. **33**(5): p. 591-601.
79. Partanen, J., L. Schwartz, and J. Rossant, *Opposite phenotypes of hypomorphic and Y766 phosphorylation site mutations reveal a function for Fgfr1 in anteroposterior patterning of mouse embryos.* Genes Dev, 1998. **12**(15): p. 2332-44.
80. Eswarakumar, V.P., et al., *The IIIc alternative of Fgfr2 is a positive regulator of bone formation.* Development, 2002. **129**(16): p. 3783-93.
81. De Moerlooze, L., et al., *An important role for the IIIb isoform of fibroblast growth factor receptor 2 (FGFR2) in mesenchymal-epithelial signalling during mouse organogenesis.* Development, 2000. **127**(3): p. 483-92.
82. Horiguchi, K., et al., *TGF-beta drives epithelial-mesenchymal transition through deltaEF1-mediated downregulation of ESRP.* Oncogene, 2012. **31**(26): p. 3190-201.
83. Reinke, L.M., Y. Xu, and C. Cheng, *Snail represses the splicing regulator epithelial splicing regulatory protein 1 to promote epithelial-mesenchymal transition.* J Biol Chem, 2012. **287**(43): p. 36435-42.
84. Warzecha, C.C., et al., *An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition.* EMBO J, 2010. **29**(19): p. 3286-300.
85. Brown, R.L., et al., *CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression.* J Clin Invest, 2011. **121**(3): p. 1064-74.
86. Keirsebilck, A., et al., *Molecular cloning of the human p120ctn catenin gene (CTNND1): expression of multiple alternatively spliced isoforms.* Genomics, 1998. **50**(2): p. 129-46.
87. Pino, M.S., et al., *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, 2008. **14**(15): p. 4943-50.
88. Davis, M.A., R.C. Ireton, and A.B. Reynolds, *A core function for p120-catenin in cadherin turnover.* J Cell Biol, 2003. **163**(3): p. 525-34.
89. Yanagisawa, M., et al., *A p120 catenin isoform switch affects Rho activity, induces tumor cell invasion, and predicts metastatic disease.* J Biol Chem, 2008. **283**(26): p. 18344-54.
90. Braeutigam, C., et al., *The RNA-binding protein Rbfox2: an essential regulator of EMT-driven alternative splicing and a mediator of cellular invasion.* Oncogene, 2014. **33**(9): p. 1082-92.
91. Chen, T. and S.Y. Dent, *Chromatin modifiers and remodellers: regulators of cellular differentiation.* Nat Rev Genet, 2014. **15**(2): p. 93-106.
92. Berger, S.L., et al., *An operational definition of epigenetics.* Genes Dev, 2009. **23**(7): p. 781-3.
93. Baylin, S.B. and P.A. Jones, *A decade of exploring the cancer epigenome - biological and translational implications.* Nat Rev Cancer, 2011. **11**(10): p. 726-34.

94. Hansen, K.D., et al., *Increased methylation variation in epigenetic domains across cancer types.* Nat Genet, 2011. **43**(8): p. 768-75.
95. Wu, C.Y., et al., *Epigenetic reprogramming and post-transcriptional regulation during the epithelial-mesenchymal transition.* Trends Genet, 2012. **28**(9): p. 454-63.
96. Malouf, G.G., et al., *Architecture of epigenetic reprogramming following Twist1-mediated epithelial-mesenchymal transition.* Genome Biol, 2013. **14**(12): p. R144.
97. McDonald, O.G., et al., *Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition.* Nat Struct Mol Biol, 2011. **18**(8): p. 867-74.
98. Tiwari, N., et al., *Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming.* Cancer Cell, 2013. **23**(6): p. 768-83.
99. Goll, M.G. and T.H. Bestor, *Eukaryotic cytosine methyltransferases.* Annu Rev Biochem, 2005. **74**: p. 481-514.
100. Graff, J.R., et al., *E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas.* Cancer Res, 1995. **55**(22): p. 5195-9.
101. Nass, S.J., et al., *Aberrant methylation of the estrogen receptor and E-cadherin 5' CpG islands increases with malignant progression in human breast cancer.* Cancer Res, 2000. **60**(16): p. 4346-8.
102. Davalos, V., et al., *Dynamic epigenetic regulation of the microRNA-200 family mediates epithelial and mesenchymal transitions in human tumorigenesis.* Oncogene, 2012. **31**(16): p. 2062-74.
103. Lennartsson, A. and K. Ekwall, *Histone modification patterns and epigenetic codes.* Biochim Biophys Acta, 2009. **1790**(9): p. 863-8.
104. Sparmann, A. and M. van Lohuizen, *Polycomb silencers control cell fate, development and cancer.* Nat Rev Cancer, 2006. **6**(11): p. 846-56.
105. Bracken, A.P., et al., *Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions.* Genes Dev, 2006. **20**(9): p. 1123-36.
106. Di Croce, L. and K. Helin, *Transcriptional regulation by Polycomb group proteins.* Nat Struct Mol Biol, 2013. **20**(10): p. 1147-55.
107. Jacobs, J.J., et al., *The oncogene and Polycomb-group gene bmi-1 regulates cell proliferation and senescence through the ink4a locus.* Nature, 1999. **397**(6715): p. 164-8.
108. Jacobs, J.J., et al., *Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF.* Genes Dev, 1999. **13**(20): p. 2678-90.
109. Haupt, Y., et al., *Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in E mu-myc transgenic mice.* Cell, 1991. **65**(5): p. 753-63.
110. Yang, M.H., et al., *Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition.* Nat Cell Biol, 2010. **12**(10): p. 982-92.
111. Wellner, U., et al., *The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs.* Nat Cell Biol, 2009. **11**(12): p. 1487-95.
112. Herranz, N., et al., *Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor.* Mol Cell Biol, 2008. **28**(15): p. 4772-81.
113. Iliopoulos, D., et al., *Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells.* Mol Cell, 2010. **39**(5): p. 761-72.
114. Zheng, M., et al., *Snail and Slug collaborate on EMT and tumor metastasis through miR-101-mediated EZH2 axis in oral tongue squamous cell carcinoma.* Oncotarget, 2015. **6**(9): p. 6797-810.
115. Dong, C., et al., *G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer.* J Clin Invest, 2012. **122**(4): p. 1469-86.
116. Dong, C., et al., *Interaction with Suv39H1 is critical for Snail-mediated E-cadherin repression in breast cancer.* Oncogene, 2013. **32**(11): p. 1351-62.
117. Rinn, J.L., et al., *Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs.* Cell, 2007. **129**(7): p. 1311-23.
118. Khalil, A.M., et al., *Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression.* Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11667-72.
119. Zhao, J., et al., *Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome.* Science, 2008. **322**(5902): p. 750-6.
120. Zhao, J., et al., *Genome-wide identification of polycomb-associated RNAs by RIP-seq.* Mol Cell, 2010. **40**(6): p. 939-53.
121. Pandey, R.R., et al., *Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation.* Mol Cell, 2008. **32**(2): p. 232-46.
122. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.
123. Ku, M., et al., *Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains.* PLoS Genet, 2008. **4**(10): p. e1000242.
124. Chaffer, C.L., et al., *Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity.* Cell, 2013. **154**(1): p. 61-74.
125. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.
126. Peinado, H., et al., *Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex.* Mol Cell Biol, 2004. **24**(1): p. 306-19.

127. Fu, J., et al., *The TWIST/Mi2/NuRD protein complex and its essential role in cancer metastasis.* Cell Res, 2011. **21**(2): p. 275-89.
128. Wu, M.Z., et al., *Interplay between HDAC3 and WDR5 is essential for hypoxia-induced epithelial-mesenchymal transition.* Mol Cell, 2011. **43**(5): p. 811-22.
129. Byles, V., et al., *SIRT1 induces EMT by cooperating with EMT transcription factors and enhances prostate cancer cell migration and metastasis.* Oncogene, 2012. **31**(43): p. 4619-29.
130. Eades, G., et al., *miR-200a regulates SIRT1 expression and epithelial to mesenchymal transition (EMT)-like transformation in mammary epithelial cells.* J Biol Chem, 2011. **286**(29): p. 25992-6002.
131. Aghdassi, A., et al., *Recruitment of histone deacetylases HDAC1 and HDAC2 by the transcriptional repressor ZEB1 downregulates E-cadherin expression in pancreatic cancer.* Gut, 2012. **61**(3): p. 439-48.
132. Meidhof, S., et al., *ZEB1-associated drug resistance in cancer cells is reversed by the class I HDAC inhibitor mocetinostat.* EMBO Mol Med, 2015.
133. Shi, Y., et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.* Cell, 2004. **119**(7): p. 941-53.
134. Lin, T., et al., *Requirement of the histone demethylase LSD1 in Snai1-mediated transcriptional repression during epithelial-mesenchymal transition.* Oncogene, 2010. **29**(35): p. 4896-904.
135. Lim, S., et al., *Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology.* Carcinogenesis, 2010. **31**(3): p. 512-20.
136. Ferrari-Amorotti, G., et al., *Inhibiting interactions of lysine demethylase LSD1 with snail/slug blocks cancer cell invasion.* Cancer Res, 2013. **73**(1): p. 235-45.
137. Metzger, E., et al., *LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription.* Nature, 2005. **437**(7057): p. 436-9.
138. Nieto, M.A. and A. Cano, *The epithelial-mesenchymal transition under control: global programs to regulate epithelial plasticity.* Semin Cancer Biol, 2012. **22**(5-6): p. 361-8.
139. Peinado, H., D. Olmeda, and A. Cano, *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nat Rev Cancer, 2007. **7**(6): p. 415-28.
140. Carver, E.A., et al., *The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition.* Mol Cell Biol, 2001. **21**(23): p. 8184-8.
141. Van de Putte, T., et al., *Mice lacking ZFHX1B, the gene that codes for Smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome.* Am J Hum Genet, 2003. **72**(2): p. 465-70.
142. Chen, Z.F. and R.R. Behringer, *twist is required in head mesenchyme for cranial neural tube morphogenesis.* Genes Dev, 1995. **9**(6): p. 686-99.
143. Siemens, H., et al., *miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions.* Cell Cycle, 2011. **10**(24): p. 4256-71.
144. Mironchik, Y., et al., *Twist overexpression induces in vivo angiogenesis and correlates with chromosomal instability in breast cancer.* Cancer Res, 2005. **65**(23): p. 10801-9.
145. Martin, T.A., et al., *Expression of the transcription factors snail, slug, and twist and their clinical significance in human breast cancer.* Ann Surg Oncol, 2005. **12**(6): p. 488-96.
146. Cheng, C.W., et al., *Mechanisms of inactivation of E-cadherin in breast carcinoma: modification of the two-hit hypothesis of tumor suppressor gene.* Oncogene, 2001. **20**(29): p. 3814-23.
147. Blanco, M.J., et al., *Correlation of Snail expression with histological grade and lymph node status in breast carcinomas.* Oncogene, 2002. **21**(20): p. 3241-6.
148. Elloul, S., et al., *Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma.* Cancer, 2005. **103**(8): p. 1631-43.
149. Spaderna, S., et al., *The transcriptional repressor ZEB1 promotes metastasis and loss of cell polarity in cancer.* Cancer Res, 2008. **68**(2): p. 537-44.
150. Tiwari, N., et al., *Klf4 is a transcriptional regulator of genes critical for EMT, including Jnk1 (Mapk8).* PLoS One, 2013. **8**(2): p. e57329.
151. Mani, S.A., et al., *Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers.* Proc Natl Acad Sci U S A, 2007. **104**(24): p. 10069-74.
152. Chang, C.J., et al., *p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs.* Nat Cell Biol, 2011. **13**(3): p. 317-23.
153. Diepenbruck, M., et al., *Tead2 expression levels control the subcellular distribution of Yap and Taz, zyxin expression and epithelial-mesenchymal transition.* J Cell Sci, 2014. **127**(Pt 7): p. 1523-36.
154. Kaneko, K.J., et al., *Transcription factor mTEAD-2 is selectively expressed at the beginning of zygotic gene expression in the mouse.* Development, 1997. **124**(10): p. 1963-73.
155. Kaneko, K.J. and M.L. DePamphilis, *Regulation of gene expression at the beginning of mammalian development and the TEAD family of transcription factors.* Dev Genet, 1998. **22**(1): p. 43-55.
156. Jacquemin, P., et al., *A novel family of developmentally regulated mammalian transcription factors containing the TEA/ATTS DNA binding domain.* J Biol Chem, 1996. **271**(36): p. 21775-85.
157. Chen, Z., G.A. Friedrich, and P. Soriano, *Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice.* Genes Dev, 1994. **8**(19): p. 2293-301.

158. Yoshida, T., *MCAT elements and the TEF-1 family of transcription factors in muscle development and disease.* Arterioscler Thromb Vasc Biol, 2008. **28**(1): p. 8-17.
159. Milewski, R.C., et al., *Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3.* Development, 2004. **131**(4): p. 829-37.
160. Sawada, A., et al., *Redundant roles of Tead1 and Tead2 in notochord development and the regulation of cell proliferation and survival.* Mol Cell Biol, 2008. **28**(10): p. 3177-89.
161. Yagi, R., et al., *Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development.* Development, 2007. **134**(21): p. 3827-36.
162. Simmonds, A.J., et al., *Molecular interactions between Vestigial and Scalloped promote wing formation in Drosophila.* Genes Dev, 1998. **12**(24): p. 3815-20.
163. Halder, G., et al., *The Vestigial and Scalloped proteins act together to directly regulate wing-specific gene expression in Drosophila.* Genes Dev, 1998. **12**(24): p. 3900-9.
164. Desai, C., et al., *A genetic pathway for the development of the Caenorhabditis elegans HSN motor neurons.* Nature, 1988. **336**(6200): p. 638-46.
165. Wu, J., A. Duggan, and M. Chalfie, *Inhibition of touch cell fate by egl-44 and egl-46 in C. elegans.* Genes Dev, 2001. **15**(6): p. 789-802.
166. Campbell, S., et al., *The scalloped gene encodes a novel, evolutionarily conserved transcription factor required for sensory organ differentiation in Drosophila.* Genes Dev, 1992. **6**(3): p. 367-79.
167. Davidson, I., et al., *The HeLa cell protein TEF-1 binds specifically and cooperatively to two SV40 enhancer motifs of unrelated sequence.* Cell, 1988. **54**(7): p. 931-42.
168. Yockey, C.E., et al., *cDNA cloning and characterization of murine transcriptional enhancer factor-1-related protein 1, a transcription factor that binds to the M-CAT motif.* J Biol Chem, 1996. **271**(7): p. 3727-36.
169. Vassilev, A., et al., *TEAD/TEF transcription factors utilize the activation domain of YAP65, a Src/Yes-associated protein localized in the cytoplasm.* Genes Dev, 2001. **15**(10): p. 1229-41.
170. Chan, S.W., et al., *TEADs mediate nuclear retention of TAZ to promote oncogenic transformation.* J Biol Chem, 2009. **284**(21): p. 14347-58.
171. Vaudin, P., et al., *TONDU (TDU), a novel human protein related to the product of vestigial (vg) gene of Drosophila melanogaster interacts with vertebrate TEF factors and substitutes for Vg function in wing formation.* Development, 1999. **126**(21): p. 4807-16.
172. Maeda, T., D.L. Chapman, and A.F. Stewart, *Mammalian vestigial-like 2, a cofactor of TEF-1 and MEF2 transcription factors that promotes skeletal muscle differentiation.* J Biol Chem, 2002. **277**(50): p. 48889-98.
173. Chen, H.H., S.J. Mullett, and A.F. Stewart, *Vgl-4, a novel member of the vestigial-like family of transcription cofactors, regulates alpha1-adrenergic activation of gene expression in cardiac myocytes.* J Biol Chem, 2004. **279**(29): p. 30800-6.
174. Belandia, B. and M.G. Parker, *Functional interaction between the p160 coactivator proteins and the transcriptional enhancer factor family of transcription factors.* J Biol Chem, 2000. **275**(40): p. 30801-5.
175. Zhao, B., et al., *The Hippo-YAP pathway in organ size control and tumorigenesis: an updated version.* Genes Dev, 2010. **24**(9): p. 862-74.
176. Wu, S., et al., *The TEAD/TEF family protein Scalloped mediates transcriptional output of the Hippo growth-regulatory pathway.* Dev Cell, 2008. **14**(3): p. 388-98.
177. Justice, R.W., et al., *The Drosophila tumor suppressor gene warts encodes a homolog of human myotonic dystrophy kinase and is required for the control of cell shape and proliferation.* Genes Dev, 1995. **9**(5): p. 534-46.
178. Xu, T., et al., *Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase.* Development, 1995. **121**(4): p. 1053-63.
179. Kango-Singh, M., et al., *Shar-pei mediates cell proliferation arrest during imaginal disc growth in Drosophila.* Development, 2002. **129**(24): p. 5719-30.
180. Tapon, N., et al., *salvador Promotes both cell cycle exit and apoptosis in Drosophila and is mutated in human cancer cell lines.* Cell, 2002. **110**(4): p. 467-78.
181. Pantalacci, S., N. Tapon, and P. Leopold, *The Salvador partner Hippo promotes apoptosis and cell-cycle exit in Drosophila.* Nat Cell Biol, 2003. **5**(10): p. 921-7.
182. Harvey, K.F., C.M. Pfleger, and I.K. Hariharan, *The Drosophila Mst ortholog, hippo, restricts growth and cell proliferation and promotes apoptosis.* Cell, 2003. **114**(4): p. 457-67.
183. Udan, R.S., et al., *Hippo promotes proliferation arrest and apoptosis in the Salvador/Warts pathway.* Nat Cell Biol, 2003. **5**(10): p. 914-20.
184. Edgar, B.A., *From cell structure to transcription: Hippo forges a new path.* Cell, 2006. **124**(2): p. 267-73.
185. Zhao, B., et al., *A coordinated phosphorylation by Lats and CK1 regulates YAP stability through SCF(beta-TRCP).* Genes Dev, 2010. **24**(1): p. 72-85.
186. Liu, C.Y., et al., *The hippo tumor pathway promotes TAZ degradation by phosphorylating a phosphodegron and recruiting the SCF{beta}-TrCP E3 ligase.* J Biol Chem, 2010. **285**(48): p. 37159-69.
187. Zhao, B., et al., *Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control.* Genes Dev, 2007. **21**(21): p. 2747-61.

188. Ota, M. and H. Sasaki, *Mammalian Tead proteins regulate cell proliferation and contact inhibition as transcriptional mediators of Hippo signaling.* Development, 2008. **135**(24): p. 4059-69.
189. Lu, L., et al., *Hippo signaling is a potent in vivo growth and tumor suppressor pathway in the mammalian liver.* Proc Natl Acad Sci U S A, 2010. **107**(4): p. 1437-42.
190. St John, M.A., et al., *Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian tumours and pituitary dysfunction.* Nat Genet, 1999. **21**(2): p. 182-6.
191. Dong, J., et al., *Elucidation of a universal size-control mechanism in Drosophila and mammals.* Cell, 2007. **130**(6): p. 1120-33.
192. Camargo, F.D., et al., *YAP1 increases organ size and expands undifferentiated progenitor cells.* Curr Biol, 2007. **17**(23): p. 2054-60.
193. Goulev, Y., et al., *SCALLOPED interacts with YORKIE, the nuclear effector of the hippo tumor-suppressor pathway in Drosophila.* Curr Biol, 2008. **18**(6): p. 435-41.
194. Takahashi, Y., et al., *Down-regulation of LATS1 and LATS2 mRNA expression by promoter hypermethylation and its association with biologically aggressive phenotype in human breast cancers.* Clin Cancer Res, 2005. **11**(4): p. 1380-5.
195. Roman-Gomez, J., et al., *Promoter hypermethylation of cancer-related genes: a strong independent prognostic factor in acute lymphoblastic leukemia.* Blood, 2004. **104**(8): p. 2492-8.
196. Seidel, C., et al., *Frequent hypermethylation of MST1 and MST2 in soft tissue sarcoma.* Mol Carcinog, 2007. **46**(10): p. 865-71.
197. Minoo, P., et al., *Prognostic significance of mammalian sterile20-like kinase 1 in colorectal cancer.* Mod Pathol, 2007. **20**(3): p. 331-8.
198. Lai, Z.C., et al., *Control of cell proliferation and apoptosis by mob as tumor suppressor, mats.* Cell, 2005. **120**(5): p. 675-85.
199. Piccolo, S., S. Dupont, and M. Cordenonsi, *The biology of YAP/TAZ: hippo signaling and beyond.* Physiol Rev, 2014. **94**(4): p. 1287-312.
200. Cordenonsi, M., et al., *The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells.* Cell, 2011. **147**(4): p. 759-72.
201. Vlug, E.J., et al., *Nuclear localization of the transcriptional coactivator YAP is associated with invasive lobular breast cancer.* Cell Oncol (Dordr), 2013. **36**(5): p. 375-84.
202. Wang, X., L. Su, and Q. Ou, *Yes-associated protein promotes tumour development in luminal epithelial derived breast cancer.* Eur J Cancer, 2012. **48**(8): p. 1227-34.
203. Chan, S.W., et al., *A role for TAZ in migration, invasion, and tumorigenesis of breast cancer cells.* Cancer Res, 2008. **68**(8): p. 2592-8.
204. Liu-Chittenden, Y., et al., *Genetic and pharmacological disruption of the TEAD-YAP complex suppresses the oncogenic activity of YAP.* Genes Dev, 2012. **26**(12): p. 1300-5.
205. Bao, Y., et al., *A cell-based assay to screen stimulators of the Hippo pathway reveals the inhibitory effect of dobutamine on the YAP-dependent gene transcription.* J Biochem, 2011. **150**(2): p. 199-208.
206. Wang, Z., et al., *Interplay of mevalonate and Hippo pathways regulates RHAMM transcription via YAP to modulate breast cancer cell motility.* Proc Natl Acad Sci U S A, 2014. **111**(1): p. E89-98.
207. Overholtzer, M., et al., *Transforming properties of YAP, a candidate oncogene on the chromosome 11q22 amplicon.* Proc Natl Acad Sci U S A, 2006. **103**(33): p. 12405-10.
208. Lei, Q.Y., et al., *TAZ promotes cell proliferation and epithelial-mesenchymal transition and is inhibited by the hippo pathway.* Mol Cell Biol, 2008. **28**(7): p. 2426-36.
209. Zhang, H., et al., *TEAD transcription factors mediate the function of TAZ in cell growth and epithelial-mesenchymal transition.* J Biol Chem, 2009. **284**(20): p. 13355-62.
210. Zhao, B., et al., *TEAD mediates YAP-dependent gene induction and growth control.* Genes Dev, 2008. **22**(14): p. 1962-71.
211. Xu, M.Z., et al., *AXL receptor kinase is a mediator of YAP-dependent oncogenic functions in hepatocellular carcinoma.* Oncogene, 2011. **30**(10): p. 1229-40.
212. Lamar, J.M., et al., *The Hippo pathway target, YAP, promotes metastasis through its TEAD-interaction domain.* Proc Natl Acad Sci U S A, 2012. **109**(37): p. E2441-50.
213. Ramalho-Santos, M., et al., *"Stemness": transcriptional profiling of embryonic and adult stem cells.* Science, 2002. **298**(5593): p. 597-600.
214. Lian, I., et al., *The role of YAP transcription coactivator in regulating stem cell self-renewal and differentiation.* Genes Dev, 2010. **24**(11): p. 1106-18.
215. Zhang, H., H.A. Pasolli, and E. Fuchs, *Yes-associated protein (YAP) transcriptional coactivator functions in balancing growth and differentiation in skin.* Proc Natl Acad Sci U S A, 2011. **108**(6): p. 2270-5.
216. Cao, X., S.L. Pfaff, and F.H. Gage, *YAP regulates neural progenitor cell number via the TEA domain transcription factor.* Genes Dev, 2008. **22**(23): p. 3320-34.
217. Schlegelmilch, K., et al., *Yap1 acts downstream of alpha-catenin to control epidermal proliferation.* Cell, 2011. **144**(5): p. 782-95.
218. Varelas, X., et al., *The Crumbs complex couples cell density sensing to Hippo-dependent control of the TGF-beta-SMAD pathway.* Dev Cell, 2010. **19**(6): p. 831-44.
219. Varelas, X., et al., *TAZ controls Smad nucleocytoplasmic shuttling and regulates human embryonic stem-cell self-renewal.* Nat Cell Biol, 2008. **10**(7): p. 837-48.

220. Silvis, M.R., et al., *alpha-catenin is a tumor suppressor that controls cell accumulation by regulating the localization and activity of the transcriptional coactivator Yap1.* Sci Signal, 2011. **4**(174): p. ra33.
221. Kim, N.G., et al., *E-cadherin mediates contact inhibition of proliferation through Hippo signaling-pathway components.* Proc Natl Acad Sci U S A, 2011. **108**(29): p. 11930-5.
222. Varelas, X., et al., *The Hippo pathway regulates Wnt/beta-catenin signaling.* Dev Cell, 2010. **18**(4): p. 579-91.
223. Wells, C.D., et al., *A Rich1/Amot complex regulates the Cdc42 GTPase and apical-polarity proteins in epithelial cells.* Cell, 2006. **125**(3): p. 535-48.
224. Zhao, B., et al., *Angiomotin is a novel Hippo pathway component that inhibits YAP oncoprotein.* Genes Dev, 2011. **25**(1): p. 51-63.
225. Chan, S.W., et al., *Hippo pathway-independent restriction of TAZ and YAP by angiomotin.* J Biol Chem, 2011. **286**(9): p. 7018-26.
226. Wang, W., J. Huang, and J. Chen, *Angiomotin-like proteins associate with and negatively regulate YAP1.* J Biol Chem, 2011. **286**(6): p. 4364-70.
227. Vogel, V. and M. Sheetz, *Local force and geometry sensing regulate cell functions.* Nat Rev Mol Cell Biol, 2006. **7**(4): p. 265-75.
228. Dupont, S., et al., *Role of YAP/TAZ in mechanotransduction.* Nature, 2011. **474**(7350): p. 179-83.
229. Aragona, M., et al., *A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors.* Cell, 2013. **154**(5): p. 1047-59.
230. Wada, K., et al., *Hippo pathway regulation by cell morphology and stress fibers.* Development, 2011. **138**(18): p. 3907-14.
231. Zhao, B., et al., *Cell detachment activates the Hippo pathway via cytoskeleton reorganization to induce anoikis.* Genes Dev, 2012. **26**(1): p. 54-68.
232. Sansores-Garcia, L., et al., *Modulating F-actin organization induces organ growth by affecting the Hippo pathway.* EMBO J, 2011. **30**(12): p. 2325-35.
233. Fernandez, B.G., et al., *Actin-Capping Protein and the Hippo pathway regulate F-actin and tissue growth in Drosophila.* Development, 2011. **138**(11): p. 2337-46.
234. Mo, J.S., et al., *Regulation of the Hippo-YAP pathway by protease-activated receptors (PARs).* Genes Dev, 2012. **26**(19): p. 2138-43.
235. Yu, F.X. and K.L. Guan, *The Hippo pathway: regulators and regulations.* Genes Dev, 2013. **27**(4): p. 355-71.
236. Djebali, S., et al., *Landscape of transcription in human cells.* Nature, 2012. **489**(7414): p. 101-8.
237. Eddy, S.R., *Non-coding RNA genes and the modern RNA world.* Nat Rev Genet, 2001. **2**(12): p. 919-29.
238. Mattick, J.S. and I.V. Makunin, *Non-coding RNA.* Hum Mol Genet, 2006. **15 Spec No 1**: p. R17-29.
239. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs.* Genome Res, 2009. **19**(1): p. 92-105.
240. Bartel, D.P., *MicroRNAs: genomics, biogenesis, mechanism, and function.* Cell, 2004. **116**(2): p. 281-97.
241. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-54.
242. Kozomara, A. and S. Griffiths-Jones, *miRBase: integrating microRNA annotation and deep-sequencing data.* Nucleic Acids Res, 2011. **39**(Database issue): p. D152-7.
243. Vasudevan, S., Y. Tong, and J.A. Steitz, *Switching from repression to activation: microRNAs can up-regulate translation.* Science, 2007. **318**(5858): p. 1931-4.
244. Orom, U.A., F.C. Nielsen, and A.H. Lund, *MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation.* Mol Cell, 2008. **30**(4): p. 460-71.
245. Henke, J.I., et al., *microRNA-122 stimulates translation of hepatitis C virus RNA.* EMBO J, 2008. **27**(24): p. 3300-10.
246. Bartel, D.P., *MicroRNAs: target recognition and regulatory functions.* Cell, 2009. **136**(2): p. 215-33.
247. Kim, V.N. and J.W. Nam, *Genomics of microRNA.* Trends Genet, 2006. **22**(3): p. 165-73.
248. Rodriguez, A., et al., *Identification of mammalian microRNA host genes and transcription units.* Genome Res, 2004. **14**(10A): p. 1902-10.
249. Yang, J.S. and E.C. Lai, *Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants.* Mol Cell, 2011. **43**(6): p. 892-903.
250. Ha, M. and V.N. Kim, *Regulation of microRNA biogenesis.* Nat Rev Mol Cell Biol, 2014. **15**(8): p. 509-24.
251. Lytle, J.R., T.A. Yario, and J.A. Steitz, *Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR.* Proc Natl Acad Sci U S A, 2007. **104**(23): p. 9667-72.
252. Grimson, A., et al., *MicroRNA targeting specificity in mammals: determinants beyond seed pairing.* Mol Cell, 2007. **27**(1): p. 91-105.
253. Lim, L.P., et al., *Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs.* Nature, 2005. **433**(7027): p. 769-73.
254. Gregory, P.A., et al., *The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1.* Nat Cell Biol, 2008. **10**(5): p. 593-601.
255. Tsang, J., J. Zhu, and A. van Oudenaarden, *MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals.* Mol Cell, 2007. **26**(5): p. 753-67.

256. Aguda, B.D., *Modeling microRNA-transcription factor networks in cancer.* Adv Exp Med Biol, 2013. **774**: p. 149-67.
257. He, L. and G.J. Hannon, *MicroRNAs: small RNAs with a big role in gene regulation.* Nat Rev Genet, 2004. **5**(7): p. 522-31.
258. Mendell, J.T., *MicroRNAs: critical regulators of development, cellular physiology and malignancy.* Cell Cycle, 2005. **4**(9): p. 1179-84.
259. Ambros, V., *The functions of animal microRNAs.* Nature, 2004. **431**(7006): p. 350-5.
260. Croce, C.M., *Causes and consequences of microRNA dysregulation in cancer.* Nat Rev Genet, 2009. **10**(10): p. 704-14.
261. Krol, J., I. Loedige, and W. Filipowicz, *The widespread regulation of microRNA biogenesis, function and decay.* Nat Rev Genet, 2010. **11**(9): p. 597-610.
262. Kong, W., et al., *MicroRNA-155 is regulated by the transforming growth factor beta/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA.* Mol Cell Biol, 2008. **28**(22): p. 6773-84.
263. Zavadil, J., et al., *Transforming growth factor-beta and microRNA:mRNA regulatory networks in epithelial plasticity.* Cells Tissues Organs, 2007. **185**(1-3): p. 157-61.
264. Diaz-Martin, J., et al., *A core microRNA signature associated with inducers of the epithelial-to-mesenchymal transition.* J Pathol, 2014. **232**(3): p. 319-29.
265. Park, S.M., et al., *The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2.* Genes Dev, 2008. **22**(7): p. 894-907.
266. Gebeshuber, C.A., K. Zatloukal, and J. Martinez, *miR-29a suppresses tristetraprolin, which is a regulator of epithelial polarity and metastasis.* EMBO Rep, 2009. **10**(4): p. 400-5.
267. Castilla, M.A., et al., *Micro-RNA signature of the epithelial-mesenchymal transition in endometrial carcinosarcoma.* J Pathol, 2011. **223**(1): p. 72-80.
268. Korpal, M., et al., *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, 2008. **283**(22): p. 14910-4.
269. Burk, U., et al., *A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells.* EMBO Rep, 2008. **9**(6): p. 582-9.
270. Bracken, C.P., et al., *A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition.* Cancer Res, 2008. **68**(19): p. 7846-54.
271. Bracken, C.P., et al., *Genome-wide identification of miR-200 targets reveals a regulatory network controlling cell invasion.* EMBO J, 2014. **33**(18): p. 2040-56.
272. Brabletz, S. and T. Brabletz, *The ZEB/miR-200 feedback loop--a motor of cellular plasticity in development and cancer?* EMBO Rep, 2010. **11**(9): p. 670-7.
273. Onder, T.T., et al., *Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways.* Cancer Res, 2008. **68**(10): p. 3645-54.
274. Derksen, P.W., et al., *Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis.* Cancer Cell, 2006. **10**(5): p. 437-49.
275. Ma, L., et al., *miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis.* Nat Cell Biol, 2010. **12**(3): p. 247-56.
276. Zhang, L., et al., *MicroRNA-10b Triggers the Epithelial-Mesenchymal Transition (EMT) of Laryngeal Carcinoma Hep-2 Cells by Directly Targeting the E-cadherin.* Appl Biochem Biotechnol, 2015.
277. Cao, M., et al., *MiR-23a regulates TGF-beta-induced epithelial-mesenchymal transition by targeting E-cadherin in lung cancer cells.* Int J Oncol, 2012. **41**(3): p. 869-75.
278. Chen, Z.L., et al., *microRNA-92a promotes lymph node metastasis of human esophageal squamous cell carcinoma via E-cadherin.* J Biol Chem, 2011. **286**(12): p. 10725-34.
279. Ma, L., J. Teruya-Feldstein, and R.A. Weinberg, *Tumour invasion and metastasis initiated by microRNA-10b in breast cancer.* Nature, 2007. **449**(7163): p. 682-8.
280. Liu, Y., et al., *MicroRNA-10b targets E-cadherin and modulates breast cancer metastasis.* Med Sci Monit, 2012. **18**(8): p. BR299-308.
281. Meng, Z., et al., *miR-194 is a marker of hepatic epithelial cells and suppresses metastasis of liver cancer cells in mice.* Hepatology, 2010. **52**(6): p. 2148-57.
282. Arora, H., R. Qureshi, and W.Y. Park, *miR-506 regulates epithelial mesenchymal transition in breast cancer cell lines.* PLoS One, 2013. **8**(5): p. e64273.
283. Cheng, C.W., et al., *MicroRNA-30a inhibits cell migration and invasion by downregulating vimentin expression and is a potential prognostic marker in breast cancer.* Breast Cancer Res Treat, 2012. **134**(3): p. 1081-93.
284. Shan, S.W., et al., *Mature miR-17-5p and passenger miR-17-3p induce hepatocellular carcinoma by targeting PTEN, GalNT7 and vimentin in different signal pathways.* J Cell Sci, 2013. **126**(Pt 6): p. 1517-30.
285. Furuta, M., et al., *miR-124 and miR-203 are epigenetically silenced tumor-suppressive microRNAs in hepatocellular carcinoma.* Carcinogenesis, 2010. **31**(5): p. 766-76.
286. Dong, P., et al., *Mutant p53 gain-of-function induces epithelial-mesenchymal transition through modulation of the miR-130b-ZEB1 axis.* Oncogene, 2013. **32**(27): p. 3286-95.
287. Zhang, J., et al., *miR-30 inhibits TGF-beta1-induced epithelial-to-mesenchymal transition in hepatocyte by targeting Snail1.* Biochem Biophys Res Commun, 2012. **417**(3): p. 1100-5.

288. Liu, Y.N., et al., *MiR-1 and miR-200 inhibit EMT via Slug-dependent and tumorigenesis via Slug-independent mechanisms.* Oncogene, 2013. **32**(3): p. 296-306.
289. Nairismagi, M.L., et al., *Translational control of TWIST1 expression in MCF-10A cell lines recapitulating breast cancer progression.* Oncogene, 2012. **31**(47): p. 4960-6.
290. Shi, L., et al., *p53-induced miR-15a/16-1 and AP4 form a double-negative feedback loop to regulate epithelial-mesenchymal transition and metastasis in colorectal cancer.* Cancer Res, 2014. **74**(2): p. 532-42.
291. Kundu, S.T., et al., *The miR-200 family and the miR-183~96~182 cluster target Foxf2 to inhibit invasion and metastasis in lung cancers.* Oncogene, 2015.
292. Weber, B., et al., *Methylation of human microRNA genes in normal and neoplastic cells.* Cell Cycle, 2007. **6**(9): p. 1001-5.
293. Lujambio, A., et al., *A microRNA DNA methylation signature for human cancer metastasis.* Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13556-61.
294. Calin, G.A., et al., *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.* Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15524-9.
295. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* Nature, 2005. **435**(7043): p. 834-8.
296. Volinia, S., et al., *A microRNA expression signature of human solid tumors defines cancer gene targets.* Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2257-61.
297. Esquela-Kerscher, A. and F.J. Slack, *Oncomirs - microRNAs with a role in cancer.* Nat Rev Cancer, 2006. **6**(4): p. 259-69.
298. Calin, G.A., et al., *Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers.* Proc Natl Acad Sci U S A, 2004. **101**(9): p. 2999-3004.
299. He, L., et al., *A microRNA component of the p53 tumour suppressor network.* Nature, 2007. **447**(7148): p. 1130-4.
300. Chang, T.C., et al., *Widespread microRNA repression by Myc contributes to tumorigenesis.* Nat Genet, 2008. **40**(1): p. 43-50.
301. Hatley, M.E., et al., *Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21.* Cancer Cell, 2010. **18**(3): p. 282-93.
302. O'Donnell, K.A., et al., *c-Myc-regulated microRNAs modulate E2F1 expression.* Nature, 2005. **435**(7043): p. 839-43.
303. Saito, Y., et al., *Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells.* Cancer Cell, 2006. **9**(6): p. 435-43.
304. Kumar, M.S., et al., *Impaired microRNA processing enhances cellular transformation and tumorigenesis.* Nat Genet, 2007. **39**(5): p. 673-7.
305. Yu, Z., et al., *Aberrant allele frequencies of the SNPs located in microRNA target sites are potentially associated with human cancers.* Nucleic Acids Res, 2007. **35**(13): p. 4535-41.
306. Ziebarth, J.D., A. Bhattacharya, and Y. Cui, *Integrative analysis of somatic mutations altering microRNA targeting in cancer genomes.* PLoS One, 2012. **7**(10): p. e47137.
307. Mayr, C. and D.P. Bartel, *Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells.* Cell, 2009. **138**(4): p. 673-84.
308. Poliseno, L., et al., *A coding-independent function of gene and pseudogene mRNAs regulates tumour biology.* Nature, 2010. **465**(7301): p. 1033-8.
309. Tay, Y., J. Rinn, and P.P. Pandolfi, *The multilayered complexity of ceRNA crosstalk and competition.* Nature, 2014. **505**(7483): p. 344-52.
310. Iorio, M.V., et al., *MicroRNA gene expression deregulation in human breast cancer.* Cancer Res, 2005. **65**(16): p. 7065-70.
311. Blenkiron, C., et al., *MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype.* Genome Biol, 2007. **8**(10): p. R214.
312. Farazi, T.A., et al., *MicroRNA sequence and expression analysis in breast tumors by deep sequencing.* Cancer Res, 2011. **71**(13): p. 4443-53.
313. Volinia, S., et al., *Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA.* Proc Natl Acad Sci U S A, 2012. **109**(8): p. 3024-9.
314. Dvinge, H., et al., *The shaping and functional consequences of the microRNA landscape in breast cancer.* Nature, 2013. **497**(7449): p. 378-82.
315. Tavazoie, S.F., et al., *Endogenous human microRNAs that suppress breast cancer metastasis.* Nature, 2008. **451**(7175): p. 147-52.
316. Png, K.J., et al., *MicroRNA-335 inhibits tumor reinitiation and is silenced through genetic and epigenetic mechanisms in human breast cancer.* Genes Dev, 2011. **25**(3): p. 226-31.
317. Yu, F., et al., *let-7 regulates self renewal and tumorigenicity of breast cancer cells.* Cell, 2007. **131**(6): p. 1109-23.
318. Zhu, S., et al., *MicroRNA-21 targets tumor suppressor genes in invasion and metastasis.* Cell Res, 2008. **18**(3): p. 350-9.
319. Herschkowitz, J.I., et al., *Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells.* Proc Natl Acad Sci U S A, 2012. **109**(8): p. 2778-83.
320. Knezevic, J., et al., *Expression of miR-200c in claudin-low breast cancer alters stem cell functionality, enhances chemosensitivity and reduces metastatic potential.* Oncogene, 2015.

321. Gibbons, D.L., et al., *Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression.* Genes Dev, 2009. **23**(18): p. 2140-51.
322. Korpal, M., et al., *Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization.* Nat Med, 2011. **17**(9): p. 1101-8.
323. Dykxhoorn, D.M., et al., *miR-200 enhances mouse breast cancer cell colonization to form distant metastases.* PLoS One, 2009. **4**(9): p. e7181.
324. Martello, G., et al., *A MicroRNA targeting dicer for metastasis control.* Cell, 2010. **141**(7): p. 1195-207.
325. Kumar, M.S., et al., *Dicer1 functions as a haploinsufficient tumor suppressor.* Genes Dev, 2009. **23**(23): p. 2700-4.
326. Lambertz, I., et al., *Monoallelic but not biallelic loss of Dicer1 promotes tumorigenesis in vivo.* Cell Death Differ, 2010. **17**(4): p. 633-41.
327. Zhang, B., et al., *microRNAs as oncogenes and tumor suppressors.* Dev Biol, 2007. **302**(1): p. 1-12.
328. Ma, L., et al., *Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model.* Nat Biotechnol, 2010. **28**(4): p. 341-7.
329. Bouchie, A., *First microRNA mimic enters clinic.* Nat Biotechnol, 2013. **31**(7): p. 577.
330. Hayes, J., P.P. Peruzzi, and S. Lawler, *MicroRNAs in cancer: biomarkers, functions and therapy.* Trends Mol Med, 2014. **20**(8): p. 460-9.
331. Cheng, J., et al., *piRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells.* Clin Chim Acta, 2011. **412**(17-18): p. 1621-5.
332. Huang, G., et al., *Altered expression of piRNAs and their relation with clinicopathologic features of breast cancer.* Clin Transl Oncol, 2013. **15**(7): p. 563-8.
333. Law, P.T., et al., *Deep sequencing of small RNA transcriptome reveals novel non-coding RNAs in hepatocellular carcinoma.* J Hepatol, 2013. **58**(6): p. 1165-73.
334. Mercer, T.R., M.E. Dinger, and J.S. Mattick, *Long non-coding RNAs: insights into functions.* Nat Rev Genet, 2009. **10**(3): p. 155-9.
335. Zhao, Y., et al., *Role of long non-coding RNA HULC in cell proliferation, apoptosis and tumor metastasis of gastric cancer: a clinical and in vitro investigation.* Oncol Rep, 2014. **31**(1): p. 358-64.
336. Fan, Y., et al., *TGF-beta-induced upregulation of malat1 promotes bladder cancer metastasis by associating with suz12.* Clin Cancer Res, 2014. **20**(6): p. 1531-41.
337. Xu, Z.Y., et al., *Knockdown of long non-coding RNA HOTAIR suppresses tumor invasion and reverses epithelial-mesenchymal transition in gastric cancer.* Int J Biol Sci, 2013. **9**(6): p. 587-97.
338. Ying, L., et al., *Upregulated MALAT-1 contributes to bladder cancer cell migration by inducing epithelial-to-mesenchymal transition.* Mol Biosyst, 2012. **8**(9): p. 2289-94.
339. Luo, M., et al., *Long non-coding RNA H19 increases bladder cancer metastasis by associating with EZH2 and inhibiting E-cadherin expression.* Cancer Lett, 2013. **333**(2): p. 213-21.
340. Gupta, R.A., et al., *Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis.* Nature, 2010. **464**(7291): p. 1071-6.
341. Yuan, J.H., et al., *A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma.* Cancer Cell, 2014. **25**(5): p. 666-81.
342. Richards, E.J., et al., *Long Non-coding RNAs (LncRNA) Regulated by Transforming Growth Factor (TGF) beta: LncRNA-HIT-MEDIATED TGFbeta-INDUCED EPITHELIAL TO MESENCHYMAL TRANSITION IN MAMMARY EPITHELIA.* J Biol Chem, 2015. **290**(11): p. 6857-67.
343. Nigro, J.M., et al., *Scrambled exons.* Cell, 1991. **64**(3): p. 607-13.
344. Guo, J.U., et al., *Expanded identification and characterization of mammalian circular RNAs.* Genome Biol, 2014. **15**(7): p. 409.
345. Jeck, W.R. and N.E. Sharpless, *Detecting and characterizing circular RNAs.* Nat Biotechnol, 2014. **32**(5): p. 453-61.
346. Conn, S.J., et al., *The RNA Binding Protein Quaking Regulates Formation of circRNAs.* Cell, 2015. **160**(6): p. 1125-34.
347. Hansen, T.B., et al., *Natural RNA circles function as efficient microRNA sponges.* Nature, 2013. **495**(7441): p. 384-8.
348. Memczak, S., et al., *Circular RNAs are a large class of animal RNAs with regulatory potency.* Nature, 2013. **495**(7441): p. 333-8.
349. Nieto, M.A., *The ins and outs of the epithelial to mesenchymal transition in health and disease.* Annu Rev Cell Dev Biol, 2011. **27**: p. 347-76.
350. Polyak, K. and R.A. Weinberg, *Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits.* Nat Rev Cancer, 2009. **9**(4): p. 265-73.
351. Magee, J.A., E. Piskounova, and S.J. Morrison, *Cancer stem cells: impact, heterogeneity, and uncertainty.* Cancer Cell, 2012. **21**(3): p. 283-96.
352. Scheel, C. and R.A. Weinberg, *Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links.* Semin Cancer Biol, 2012. **22**(5-6): p. 396-403.
353. Moreno-Bueno, G., F. Portillo, and A. Cano, *Transcriptional regulation of cell polarity in EMT and cancer.* Oncogene, 2008. **27**(55): p. 6958-69.
354. Hong, W. and K.L. Guan, *The YAP and TAZ transcription co-activators: key downstream effectors of the mammalian Hippo pathway.* Semin Cell Dev Biol, 2012. **23**(7): p. 785-93.

355. Sawada, A., et al., *Tead proteins activate the Foxa2 enhancer in the node in cooperation with a second factor.* Development, 2005. **132**(21): p. 4719-29.
356. Mahoney, W.M., Jr., et al., *The transcriptional co-activator TAZ interacts differentially with transcriptional enhancer factor-1 (TEF-1) family members.* Biochem J, 2005. **388**(Pt 1): p. 217-25.
357. Maeda, M., K.R. Johnson, and M.J. Wheelock, *Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition.* J Cell Sci, 2005. **118**(Pt 5): p. 873-87.
358. Lehembre, F., et al., *NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin.* EMBO J, 2008. **27**(19): p. 2603-15.
359. Consortium, F., et al., *The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line.* Nat Genet, 2009. **41**(5): p. 553-62.
360. Waldmeier, L., et al., *Py2T murine breast cancer cells, a versatile model of TGFbeta-induced EMT in vitro and in vivo.* PLoS One, 2012. **7**(11): p. e48651.
361. Larkin, S.B., I.K. Farrance, and C.P. Ordahl, *Flanking sequences modulate the cell specificity of M-CAT elements.* Mol Cell Biol, 1996. **16**(7): p. 3742-55.
362. Lai, D., et al., *Taxol resistance in breast cancer cells is mediated by the hippo pathway component TAZ and its downstream transcriptional targets Cyr61 and CTGF.* Cancer Res, 2011. **71**(7): p. 2728-38.
363. Kitagawa, M., *A Sveinsson's chorioretinal atrophy-associated missense mutation in mouse Tead1 affects its interaction with the co-factors YAP and TAZ.* Biochem Biophys Res Commun, 2007. **361**(4): p. 1022-6.
364. Tian, W., et al., *Structural and functional analysis of the YAP-binding domain of human TEAD2.* Proc Natl Acad Sci U S A, 2010. **107**(16): p. 7293-8.
365. Li, Z., et al., *Structural insights into the YAP and TEAD complex.* Genes Dev, 2010. **24**(3): p. 235-40.
366. Chen, L., et al., *Structural basis of YAP recognition by TEAD4 in the hippo pathway.* Genes Dev, 2010. **24**(3): p. 290-300.
367. Ikeda, Y., et al., *Gastric cancer surgery for patients with liver cirrhosis.* World J Gastrointest Surg, 2009. **1**(1): p. 49-55.
368. Beckerle, M.C., *Zyxin: zinc fingers at sites of cell adhesion.* Bioessays, 1997. **19**(11): p. 949-57.
369. Hirata, H., H. Tatsumi, and M. Sokabe, *Zyxin emerges as a key player in the mechanotransduction at cell adhesive structures.* Commun Integr Biol, 2008. **1**(2): p. 192-5.
370. Smith, M.A., et al., *A zyxin-mediated mechanism for actin stress fiber maintenance and repair.* Dev Cell, 2010. **19**(3): p. 365-76.
371. Mori, M., et al., *Zyxin mediates actin fiber reorganization in epithelial-mesenchymal transition and contributes to endocardial morphogenesis.* Mol Biol Cell, 2009. **20**(13): p. 3115-24.
372. Pan, D., *The hippo signaling pathway in development and cancer.* Dev Cell, 2010. **19**(4): p. 491-505.
373. Sanchez-Tillo, E., et al., *EMT-activating transcription factors in cancer: beyond EMT and tumor invasiveness.* Cell Mol Life Sci, 2012. **69**(20): p. 3429-56.
374. Bhattaram, P., et al., *Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors.* Nat Commun, 2010. **1**: p. 9.
375. Zhao, B., Q.Y. Lei, and K.L. Guan, *The Hippo-YAP pathway: new connections between regulation of organ size and cancer.* Curr Opin Cell Biol, 2008. **20**(6): p. 638-46.
376. Anbanandam, A., et al., *Insights into transcription enhancer factor 1 (TEF-1) activity from the solution structure of the TEA domain.* Proc Natl Acad Sci U S A, 2006. **103**(46): p. 17225-30.
377. Weber, M., et al., *Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome.* Nat Genet, 2007. **39**(4): p. 457-66.
378. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
379. Stadler, M.B., et al., *DNA-binding factors shape the mouse methylome at distal regulatory regions.* Nature, 2011. **480**(7378): p. 490-5.
380. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
381. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.
382. Siddharthan, R., E.D. Siggia, and E. van Nimwegen, *PhyloGibbs: a Gibbs sampling motif finder that incorporates phylogeny.* PLoS Comput Biol, 2005. **1**(7): p. e67.
383. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression.* Nat Rev Cancer, 2002. **2**(6): p. 442-54.
384. Gunasinghe, N.P., et al., *Mesenchymal-epithelial transition (MET) as a mechanism for metastatic colonisation in breast cancer.* Cancer Metastasis Rev, 2012. **31**(3-4): p. 469-78.
385. Shibue, T. and R.A. Weinberg, *Metastatic colonization: settlement, adaptation and propagation of tumor cells in a foreign tissue environment.* Semin Cancer Biol, 2011. **21**(2): p. 99-106.
386. Tsai, J.H. and J. Yang, *Epithelial-mesenchymal plasticity in carcinoma metastasis.* Genes Dev, 2013. **27**(20): p. 2192-206.
387. Biswas, S., et al., *Inhibition of TGF-beta with neutralizing antibodies prevents radiation-induced acceleration of metastatic cancer progression.* J Clin Invest, 2007. **117**(5): p. 1305-13.
388. Li, X., et al., *MiR-200 can repress breast cancer metastasis through ZEB1-independent but moesin-dependent pathways.* Oncogene, 2014. **33**(31): p. 4077-88.

389. Iorio, M.V. and C.M. Croce, *MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review.* EMBO Mol Med, 2012. **4**(3): p. 143-59.
390. Li, J., et al., *MicroRNA-125b suppresses the migration and invasion of hepatocellular carcinoma cells by targeting transcriptional coactivator with PDZ-binding motif.* Oncol Lett, 2015. **9**(4): p. 1971-1975.
391. Wu, D., et al., *microRNA-125b inhibits cell migration and invasion by targeting matrix metallopeptidase 13 in bladder cancer.* Oncol Lett, 2013. **5**(3): p. 829-834.
392. Taube, J.H., et al., *Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes.* Proc Natl Acad Sci U S A, 2010. **107**(35): p. 15449-54.
393. Aslakson, C.J. and F.R. Miller, *Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor.* Cancer Res, 1992. **52**(6): p. 1399-405.
394. Dweep, H., et al., *miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes.* J Biomed Inform, 2011. **44**(5): p. 839-47.
395. Harazono, Y., et al., *miR-655 Is an EMT-suppressive microRNA targeting ZEB1 and TGFBR2.* PLoS One, 2013. **8**(5): p. e62757.
396. Li, X.L., et al., *A p21-ZEB1 complex inhibits epithelial-mesenchymal transition through the microRNA 183-96-182 cluster.* Mol Cell Biol, 2014. **34**(3): p. 533-50.
397. Suzuki, H., et al., *DNA methylation and microRNA dysregulation in cancer.* Mol Oncol, 2012. **6**(6): p. 567-78.
398. Eger, A., et al., *DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells.* Oncogene, 2005. **24**(14): p. 2375-85.
399. Comijn, J., et al., *The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion.* Mol Cell, 2001. **7**(6): p. 1267-78.
400. Ding, X., et al., *Signaling between transforming growth factor beta (TGF-beta) and transcription factor SNAI2 represses expression of microRNA miR-203 to promote epithelial-mesenchymal transition and tumor metastasis.* J Biol Chem, 2013. **288**(15): p. 10241-53.
401. Soule, H.D., et al., *Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10.* Cancer Res, 1990. **50**(18): p. 6075-86.

6 Acknowledgments

Though only my name appears on the cover of this thesis, many people have directly or indirectly contributed to its production. I would like to thank all those people who have made this dissertation possible and of whom my graduate experience in Basel has been one that I will cherish forever.

First of all, I would like to express my deepest gratitude to Gerhard Christofori. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered. The inspiring and friendly work atmosphere he created in his lab was a great source of motivation for me and I admire his expertise and appreciated his conceptual inputs and patience. At this point, I would like to thank the other members of my PhD committee, Mohamed Bentires-Alj and Thomas Brabletz, for their assistance and conceptual inputs for my research projects.

Special thanks goes to Lorenz Waldmeier, my former master thesis supervisor and later Tead2 project partner during my PhD studies. I have learned a lot from working with him and I am deeply grateful for long and productive project discussions – in the lab or outside over a beer.

I would also like to thank all former and present members of the Christofori group who supported my professional life and shared their knowledge, skills and gave helpful advices. Particularly, I have to thank Dana Ronen, Dorothea Gruber and Ruben Bill for being great labmates – I greatly value their research suggestions, reading and commenting on this thesis and most importantly their friendship. I am really happy that I was part of the DBM PhD student club and get to know Carlos Mayer, Frederic Laurent and Fabrizio Botindari. I am thankful for a great time organizing several student events, which opened the door to “easily” meet and get to know other people from the DBM.

Most importantly, none of this work would have been possible without my family who supported me through my entire life. Long, uplifting telephone calls and relaxing weekends at home gave me the energy to keep on going. I am also deeply grateful to Arne whose love and encouragement supported me the past years – THANK YOU!