Epigenetic Regulation of Tumor Cell Phenotype

Dissertation

for the award of the degree

"Doctor rerum naturalium (Dr. rer. nat.)"

Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

Submitted by

Vivek Kumar Mishra

born in

Rewa, India

Göttingen, 2016

Thesis supervisor

Prof. Dr. Steven A. Johnsen

Members of the Thesis Committee:

Prof. Dr. Steven A. Johnsen (Reviewer)

Department of General, Visceral and Pediatric Surgery

University of Göttingen Medical School, Göttingen

Prof. Dr. Matthias Dobbelstein (Reviewer)

Department of Molecular Oncology

University of Göttingen Medical School, Göttingen

Prof. Dr. Michael Zeisberg

Department of Nephrology and Rheumatology

University of Göttingen Medical School, Göttingen

Date of oral examination: June 8, 2016

Affidavit

I hereby declare that the PhD thesis entitled "Epigenetic Regulation of Tumor Cell Phenotype" has been written independently and with no other sources and aids than quoted.

_____

Vivek Kumar Mishra

April, 2016

Göttingen

"My parents have been my all – they bore me, raised me, supported me, taught me, and loved me. Their support, encouragement, quiet patience and unwavering love are undeniably the bedrock upon which the years of my life have been built. It is to them that I dedicate this thesis".

Table of contents

Acknowledgements

"No one can whistle a symphony. It takes an orchestra to play it."
-**H.E. Luccock**

While enthusiastically writing these lines and excitingly awaiting the day of submission, I realized that the mountain that I have climbed was not just a solitary achievement. Many people have helped me during this journey. Though I will stand alone on the cover of this thesis I will try to make them part of it.

With deep sense of gratitude and immense pleasure, I express my heartfelt thanks to my revered supervisor Prof. Dr. Steven A. Johnsen for his implacable support, constant motivation and an excellent research atmosphere in the lab. It is under your constructive guidance that I grew as a scientist and explored new dimensions of research. I shall never forget to do MORE, do BETTER and do FASTER.

I would like to extend my sincere thanks to Prof. Dr. Hans Will for his inspirational thoughts and ever cheerful attitude. You are truly a wonderful human being and great personality.

I sincerely thank my thesis committee members, Prof. Dr. Matthias Dobbelstein and Prof. Dr. Michael Zeisberg for their constructive discussions and helpful suggestions on my thesis project.

I am deeply grateful to Prof. John R. Hawse and Prof. Malayannan Subramaniam at the Mayo Clinic, Rochester, USA for giving me the opportunity to conduct an important part of my thesis project in their laboratory. It was a wonderful experience.

I extend my sincere thanks to Dr. Vijayalakshmi Kari and Dr. Florian Wegwitz for their expert advice and constructive guidance during the course of my thesis project.

On a special note, I would like to thank Dr. Upasana Bedi, Zeynab Najafova and Wanhua Xie for helpful discussions and all the fun and laughter.

I would like to convey heartfelt thanks to the members of AG Johnsen. Thank you for being a part of this incredible journey and being wonderful colleagues. Special word of thanks to Robyn and Frederike for their great help with mice experiments.

Back home, I wish to extend thanks to my two special friends Shruti and Divneet for sharing a wonderful bond of friendship with me.

Heartfelt thanks to Brigette, Hartmut, Roselinde, Martin and Mareike for all the love and support.

My final words, and the most important ones, go to my family. My parents always encouraged and supported me to pursue my ambitions and dreams. My elder brothers Rajesh and Rakesh and sisters Pratibha and Sangeeta for always being there to guide me through tough times and never letting me feel short of love and care.

Lastly and more importantly, I want to thank Katja for bringing so much positive energy and influence into my life. With you around, I never felt short of inspiration and motivation.

Abbreviations

α	alpha
B	beta
γ	gamma
μ	micro (10^{-6})
°C	degree Celsius / centigrade
Å	angstrom
APS	ammonium persulfate
α-SMA	alpha-smooth muscle actin
BAM	binary alignment map
BGP	β-glycerophosphate
bp	base pair
BRCA1	breast cancer 1
BSA	bovine serum albumin
cDNA	complimentary DNA
CEAS	cis-regulatory element annotation system
ChIP	chromatin immunoprecipitation
ChIP-seq	ChIP coupled with high throughput sequencing
cm	centimetre
CO_2	carbon dioxide
CSC	cancer stem cell
Da	dalton
DAPI	4',6-diamidino-2-phenylindole
DEPC	diethylpyrocarbonate
ddH_2O	double distilled water
DMBA	7,12-Dimethylbenz(a)anthracene
DMEM	dulbecco's Modified Eagle's Medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNMT	DNA methyl transferase
dNTP	deoxyribonucleotide
DTC	disseminated tumor cell
E-cad	e-cadherin
EDTA	ethylenediaminetetraacetic acid

et al.	et alii = and others
EtOH	ethanol
EMT	epithelial-to-mesenchymal transition
EMT-TF	EMT-transcription factor
ENCODE	encyclopedia of DNA elements
F	forward
FAD	Flavin adenine dinucleotide
FBS	fetal bovine serum
FC	fold change
FDR	false Discovery Rate
Fig.	figure
Fwd	forward
g	gravity (9.81 m/s^2)
hr.	hour
H1	histone 1
H2A	histone 2A
H2B	histone 2B
H3	histone 3
H3K27ac	histone 3 acetylated at lysine 27
H3K9ac	histone 3 acetylated at lysine 9
H3K4me1	histone 3 monomethylated at lysine 4
H4	histone 4
HAT	histone acetyltransferase
HDAC1	histone deacetylase 1
HDAC2	histone deacetylase 2
hg19	human genome project version 19
HMT	histone methyltransferase
HRP	horse radish peroxidase
hs	homo sapiens
HSC70	heat shock 70KDa protein
IAA	iodacetamide
IF	immunofluorescence
IgG	immunoglobulin G
IGV	integrative genomics viewer
Kb	kilo base pairs

KDa	kilo Dalton
kg	kilogram
KLF	Krüppel-like factor
KLF10	Krüppel-like transcription factor 10
LiCl	lithium chloride
log	logarithm
LSD1	lysine-specific demethylase 1
m	milli (10^{-3})
M	molar, mol/L
MACS	model-based analysis of ChIP-seq
MEM	minimum Essential Medium
MET	mesenchymal-to-epithelial transition
min	minute
mRNA	messenger RNA
MMP2	matrix metalloproteinase 2
MMP7	matrix metalloproteinase 7
n	nano (10^{-9})
N-cad	n-cadherin
NEM	N-ethylmaleimide
NP-40	Nonidet P40
n.s.	non-significant
NSCLC	non-small cell lung cancer
NuRD	nucleosome remodelling deacetylase
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with Tween-20
PCA	principle component analysis
PCR	polymerase chain reaction
pH	*Potential hydrogenii* = potential of hydrogen
qPCR or qRT-PCR	quantitative real-time PCR
R	reverse
RI	TGFβ receptor type-1
RII	TGFβ receptor type-2
RNA	ribonucleic acid
RNA-Pol II	RNA Polymerase II

RNA-seq	sequencing of rt-transcribed RNA
RPKM	reads per kilo base per million mappedreads
R-SMAD	receptor regulated SMAD
RT	room temperature
RT-PCR	reverse transcription PCR
SAM	sequence alignment map
s.d.	standard deviation
SDS	sodium dodecylsulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
sec	second
siRNA	small interfering RNA
Taq	*Thermus aquaticus*
TEMED	N,N,N',N'-Tetramethylethylenediamine
TF	transcription factor
TGFβ	transforming growth factor-beta
TIEG1	TGFβ-inducible early gene 1
Tris	tris(hydroxymethyl)aminomethane
TR	transcribed region
TSS	transcriptional start site
U	unit (enzyme activity)
ub	ubiquitin
up	upregulated
V	voltage
VIM	vimentin
vs.	versus
v/v	volume per volume
w/v	weight per volume
wt	wild type

List of Figures

Abstract

The Transforming Growth Factor-β (TGFβ)/SMAD signaling pathway can function as either a tumor suppressor or metastasis promoter during tumor progression. In normal epithelial cells and early stages of epithelial tumorigenesis TGFβ functions as a tumor suppressor to decrease cell proliferation or induce apoptosis. However, during malignant progression tumor cells no longer respond to the anti-proliferative effects of TGFβ, but instead undergo an epithelial-to-mesenchymal transition (EMT) whereby cells acquire a migratory and invasive phenotype which promotes tumor metastasis. Resolution of the dichotomy in TGFβ function and a further understanding of its tumor suppressor and metastasis promoting functions may uncover new strategies for the treatment of epithelial cancers. Previous studies have demonstrated an important role of the TGFβ-Inducible Early Gene-1 (TIEG1)/Krüppel-like Factor-10 (KLF10) as a central regulator of TGFβ/SMAD signaling and the anti-proliferative functions of TGFβ. In this study we examined the role of KLF10 in controlling the TGFβ-induced EMT and show that depletion of KLF10 results in a more pronounced induction of EMT. Moreover, chromatin immunoprecipitation (ChIP) and chromatin immunoprecipitation-sequencing (ChIP-seq) analysis shows that KLF10 directly binds to GC-rich sequences in the promoter region of the EMT-promoting transcription factor SLUG/SNAI2 to repress its transcription. Consistent with these findings, an analysis of KLF10 in lung cancer revealed that KLF10 levels are decreased in lung cancer vs. normal samples. Furthermore, *in vivo* study revealed a significantly increased tumor incidence and tumor size in *Klf10$^{-/-}$* mice compared to the wild type mice. Additional ChIP studies showed that KLF10 recruits HDAC1 to the SNAI2 promoter and is required for the removal of activating histone acetylation marks. These findings reveal a previously unknown function of KLF10 in suppressing TGFβ-induced EMT and

represent a significant advancement in the understanding the molecular dichotomy of TGFβ function during tumor progression.

In a more global approach, we have utilized a dual LSD1/HDAC inhibitor 4SC-202 to study the effect on tumor cell phenotype. We have shown that combined inhibition of LSD1 and HDACs significantly block the TGFβ-induced EMT. Immunohistochemical staining of LSD1 in pancreatic cancer samples revealed that LSD1 is highly expressed in a subset of tumors. Consistent with this finding, in our xenograft study we have shown that 4SC-202 significantly decreases the tumor size. Together these findings revealed the potential role of small molecule inhibitors against epigenetic modifiers in targeted anticancer therapy.

1. Introduction

Cancer is considered to be a disease of accumulation of genetic abnormalities like mutation, amplification, deletion or translocation. At the cellular level, the process of conversion of a normal cell to a cancer cell (tumor initiation) starts with the accumulation of genetic alterations in systems that regulate the cell behavior. This allows a single normal cell to break open the barrier of controlled cell division and start proliferating abnormally in an uncontrolled manner (Hanahan and Weinberg, 2011). Furthermore, these abnormalities help cancer cells to survive and proliferate in the local microenvironment in the initial stage and help them to metastasize in the later stage. During the course of progression from a pre-malignant to a metastatic tumor the gene expression pattern plays an important role, which includes the expression of a specific subset of genes and the repression of others. Epigenetic has emerged as a central player in the regulation of gene expression. In the recent past, various studies have uncovered the underlying epigenetic changes involved in cancer development and progression. Therefore, cancer is no longer considered as a disease of only genetic mutations but rather a disease of genetic and epigenetic abnormalities (Baylin and Jones, 2011; Esteller, 2007; Sandoval and Esteller, 2012). Moreover, it is believed that the aberrations in epigenetic regulators and genetic mutations cooperatively fuel tumor development and progression (Hitchins et al., 2011; Schepers and Clevers, 2012; Sharma et al., 2010).

1.1 Chromatin structure

The literal meaning of 'epigenetics' is 'outside genetics'. However, the term epigenetics is used to describe the heritable change in the cellular phenotype that occurs without any change in the genome (Berger et al., 2009). In eukaryotic cells, chromatin is composed of DNA, RNA and proteins. Within the nucleus, DNA is

wrapped around an octamer of the four core histone proteins forming a structure called the nucleosome, the basic unit of chromatin structure (Kornberg, 1974). Each nucleosome contains approximately 147 base pairs of DNA and two each of the histones H2A, H2B, H3 and H4 (Fig.1) (Dawson and Kouzarides, 2012; Kouzarides, 2007). Additionally, histone H1 binds to the DNA wrapped around the nucleosome and linker DNA between two nucleosomes and helps in maintaining the chromatin structure (Laybourn and Kadonaga, 1991). Histones are basic proteins that contain a globular domain and a charged amino terminal "tail" that protrudes out from the nucleosome (Kornberg, 1974; Luger et al., 1997). Histone tails are prone to undergo post-translational modifications at specific amino acid residues that lead to alteration in chromatin structure (Campos and Reinberg, 2009).
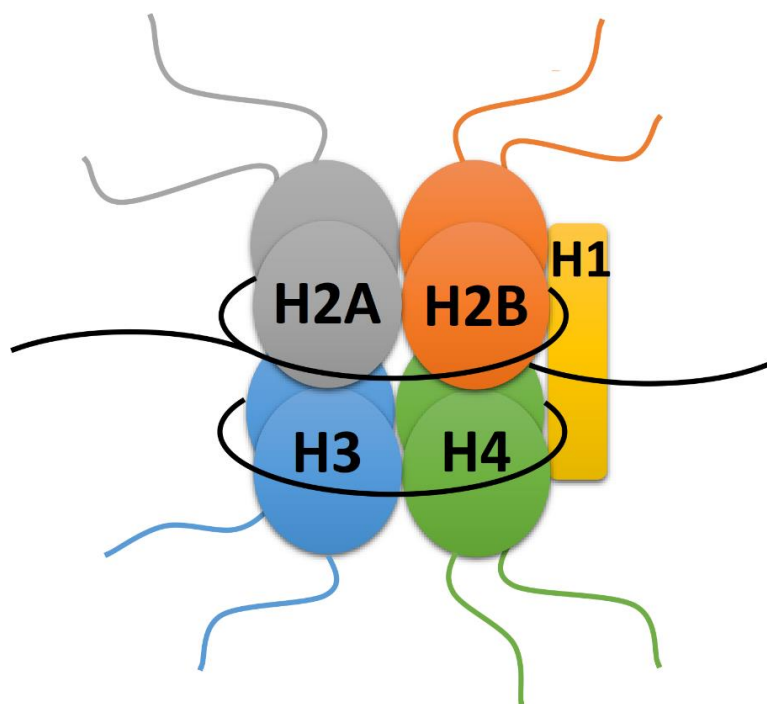
Figure 1: Structure of nucleosome. The nucleosome is a basic structure of chromatin. Each nucleosome is composed of an octamer of histones H2A, H2B, H3 and H4 (each of the histones are present as dimer). Approximately 147 base pairs DNA are wrapped around the histone octamer. In addition, histone H1 holds the end of the nucleosomal DNA and also holds together two nucleosomes thus, helps to maintain the chromatin structure. Dynamic post-translational histone modifications take place at the tail regions of the histones.

Based on the degree of compaction and ease of accessibility, chromatin has been divided into two major forms: "heterochromatin" is the highly condensed state of the chromatin and considered to contain the transcriptionally inactive region of the genome, whereas "euchromatin" is the more open and easily accessible form and comprises the more actively transcribed parts of the genome (Fig.2). Euchromatin is associated with key cellular processes like replication and transcription that require direct access to the DNA. In contrast, heterochromatin regions are repetitive elements and primarily contain inactive genes (Cheung and Lau, 2005; Li, 2002; Weintraub and Groudine, 1976).
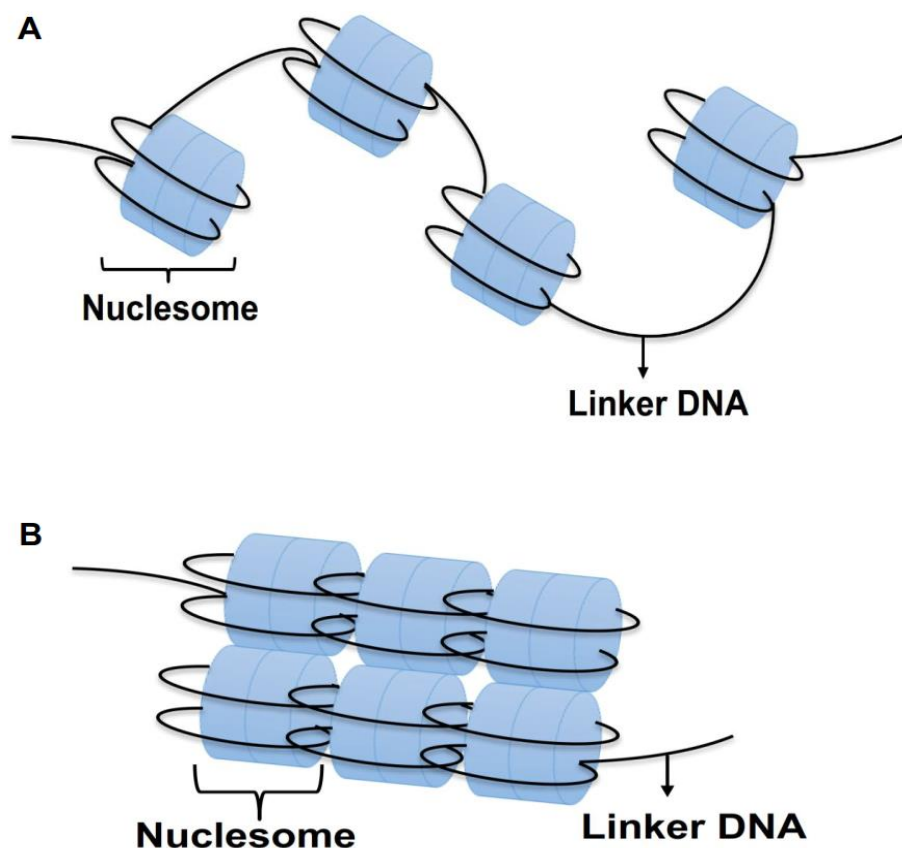
Figure 2: Schematics showing euchromatin and heterochromatin. (A) Euchromatin is an open state of chromatin which is easily accessible to the binding partners like transcription factors. It is mainly associated with actively transcribed region of the genome. (B) Heterochromatin is a highly condensed state of chromatin and is considered to be the transcriptionally inactive part of the genome. It contains repetitive elements that are associated with centromere and telomere.

1.1.1 Histone modifications

A variety of post-translational epigenetic modifications occur at the level of DNA and histone proteins, thereby altering chromatin structure and modulating gene expression by controlling the accessibility of DNA to transcription factors or by directly recruiting transcriptional co-factors. Post-translational histone modifications include methylation, acetylation, phosphorylation, ubiquitination and sumoylation and are carried out in a highly specific manner by chromatin-modifying enzymes (Fig.3). Many of these histone modifications take place on the N- and C-terminals of the tail of histone proteins which are protruding from the nucleosome. The enzymes responsible for carrying out histone modifications are highly specific and act on defined amino acid residues.

Depending on the signaling conditions in the cell, different types or combinations of these modifications are present on different genes and provide a plethora of possible combinations of chromatin modifications. The combinations of modifications or "histone code" can lead to the suppression (gene silencing) or expression (gene activation) of particular genes and thereby determine cell fate, for example, by directing cell proliferation and differentiation. The "histone code" hypothesis predicts that the pattern of histone modifications present on the histone tails forms a code that can be "read" by proteins or protein complexes that then positively or negatively direct DNA-associated processes such as transcription, co-transcriptional RNA processing, DNA replication and DNA repair (Feinberg and Tycko, 2004; Jenuwein and Allis, 2001). Post-translational histone modifications are dynamic and reversible in nature, allowing a high level of epigenetic plasticity in response to extrinsic and intrinsic stimuli. Whereas some of these modifications appear to be exclusively associated with active genes and others with inactive genes, some exceptions have been found.

For example, the so-called "bivalent domains" are genomic regions frequently located near the promoters of selected developmental genes in embryonic stem cells. These genes exhibit both activating (H3K4me3) and repressive (H3K27me3) histone modifications on the same gene at the same time (Bannister et al., 2002; Bernstein et al., 2006; Rice and Allis, 2001; Strahl and Allis, 2000; Zhang and Reinberg, 2001).
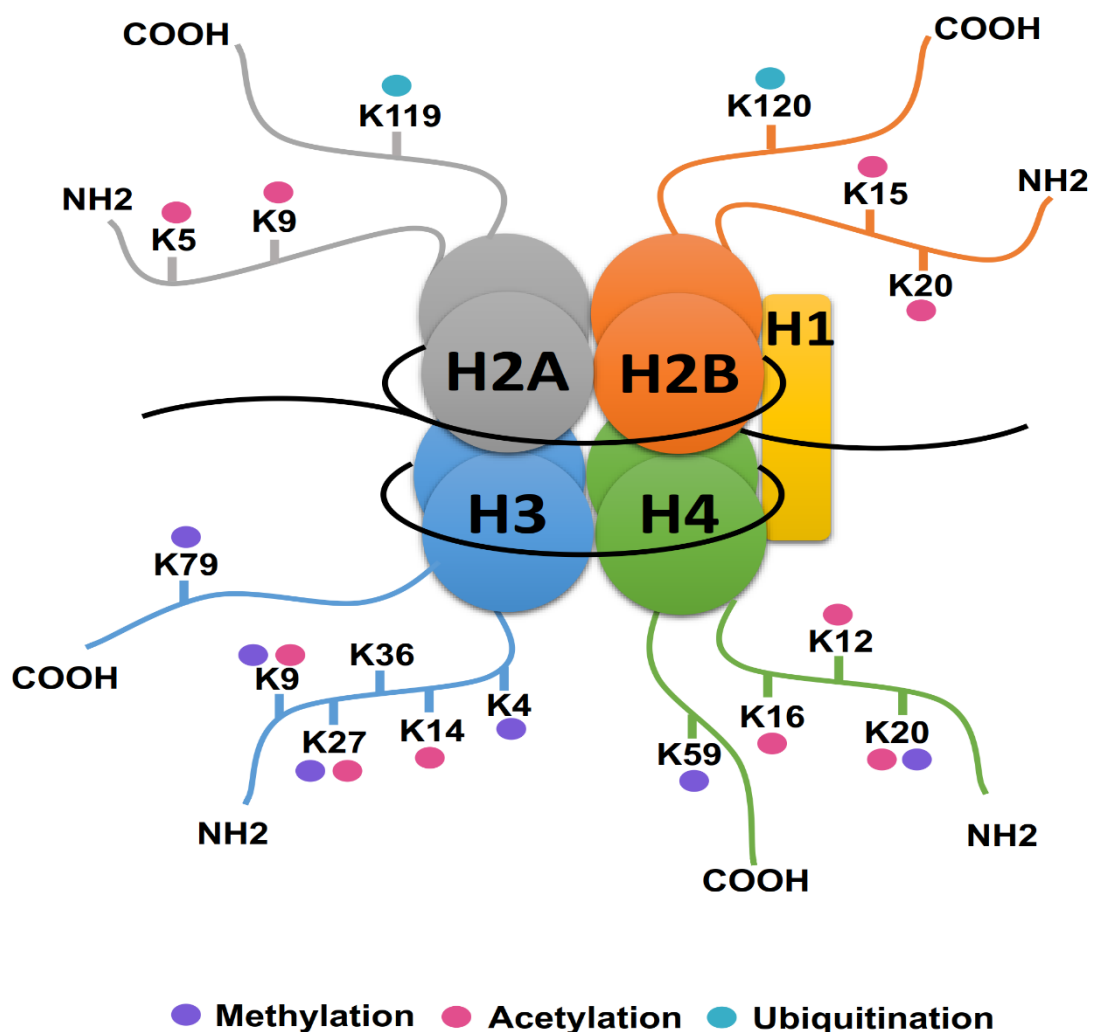
Figure 3: Schematics of different post-translational histone modifications. Different types of histone modifications take place at the NH$_2$- and COOH- terminal end of the histone tail. These include acetylation, methylation, ubiquitination, phosphorylation and sumoylation. Histone modifications are carried out by epigenetic regulators in a very specific manner at certain amino residues on the histone tails. Some of the most well studied ones are shown here.

1.1.2 Epigenetic regulators and their interplay

In a multicellular organism different cell types have different functions, which is mainly attributed to a defined and specific subset of genes that they express. Dynamic covalent modifications of nucleosomal DNA and histones by epigenetic regulators lead to changes in chromatin architecture and remodelling in a way that allows the cell specific gene expression and silencing. Moreover, these epigenetic regulators either act alone or get recruited in a large complexes to modify the chromatin structure and regulate its accessibility to the transcriptional machinery (Hayes, 2002; Horn and Peterson, 2002; Narlikar et al., 2002).

Epigenetic regulators can be broadly classified into four types: "epigenetic writers" are the enzymes responsible for modifying histone substrates by adding chemical marks such as methyl or acetyl groups (e.g., histone lysine and arginine methyltransferases, histone acetyltransferases); "epigenetic readers" are the proteins that recognize specific modifications or combinations of modifications that have been placed on the histone proteins (e.g., bromo- and chromodomain-containing proteins); "epigenetic erasers" are the enzymes that catalyse the removal of the histone modifications (e.g., histone deacetylases and histone demethylases); and finally, chromatin remodelling enzymes and histone chaperones physically alter chromatin structure by moving, removing, adding, or replacing nucleosomes or specific histones within the chromatin (Kouzarides, 2007; Lee et al., 2010; Strahl and Allis, 2000; Taverna et al., 2007; Wilson and Roberts, 2011). Thus, the modulation of the expression or activity of any of these classes of epigenetic regulators can have wide-ranging effects on the cellular transcriptional profile and might thereby impact cell fate determination and important tumor relevant processes such as proliferation and EMT.

1.2 Epithelial-to-mesenchymal transition

Tumor metastasis requires a cascade of biological processes that enables cancer cells to move from the primary tumor site to distant organs, to become acclimatized to the foreign tissue microenvironment and to begin to proliferate again, thus giving rise to secondary tumor. Metastasis occurs through a cascade of steps that involves dissemination form primary tumor, primary invasion, intravasation, survival during circulation, extravasation, formation of micrometastasis and finally colonization to form macrometastasis (Fidler, 2003a; Mishra and Johnsen, 2014; Scheel and Weinberg, 2012; Valastyan and Weinberg, 2011).

Epithelial-to-mesenchymal transition (EMT) is a phenomenon including changes in the cellular phenotype allowing epithelial cells to convert into mesenchymal cells (Fig.4). During EMT, epithelial cells lose their characteristic features like polarity and cell-cell adhesion through the dissolution of tight junction (claudins and occludins) and adherens junction (E-cadherin and cytokeratins) and, in contrast display increased expression of mesenchymal markers (e.g., N-cadherin, Vimentin, Fibronectin, and alpha-smooth muscle actin) and become migratory and invasive (Kang and Massagué, 2004; Scheel and Weinberg, 2012; Sleeman et al., 2012; Tiwari et al., 2012).

EMT is an evolutionary conserved process that plays an important role in normal embryonic development (e.g., EMT is indispensable during gastrulation and neural crest formation) during which cells need to migrate over long distances in order to give rise to various adult tissues and organs (Nieto, 2013). Once the embryonic cells have migrated to their final destination, they frequently undergo a reverse process of EMT known as the mesenchymal-to-epithelial transition (MET), which facilitates their differentiation into multiple cell lineages (Craene and Berx, 2013). Similarly,

disseminated tumor cells (DTCs) frequently revert to an epithelial phenotype by undergoing MET in order to colonize at distant metastatic sites and to give rise to secondary tumors.

A stringent balance between EMT and MET is essential for maintaining tissue homeostasis. The ability of cells to switch between epithelial and mesenchymal phenotypes is called cellular plasticity and is also an important characteristic of aggressive metastatic cancer cells (Hugo et al., 2007; Polyak and Weinberg, 2009).
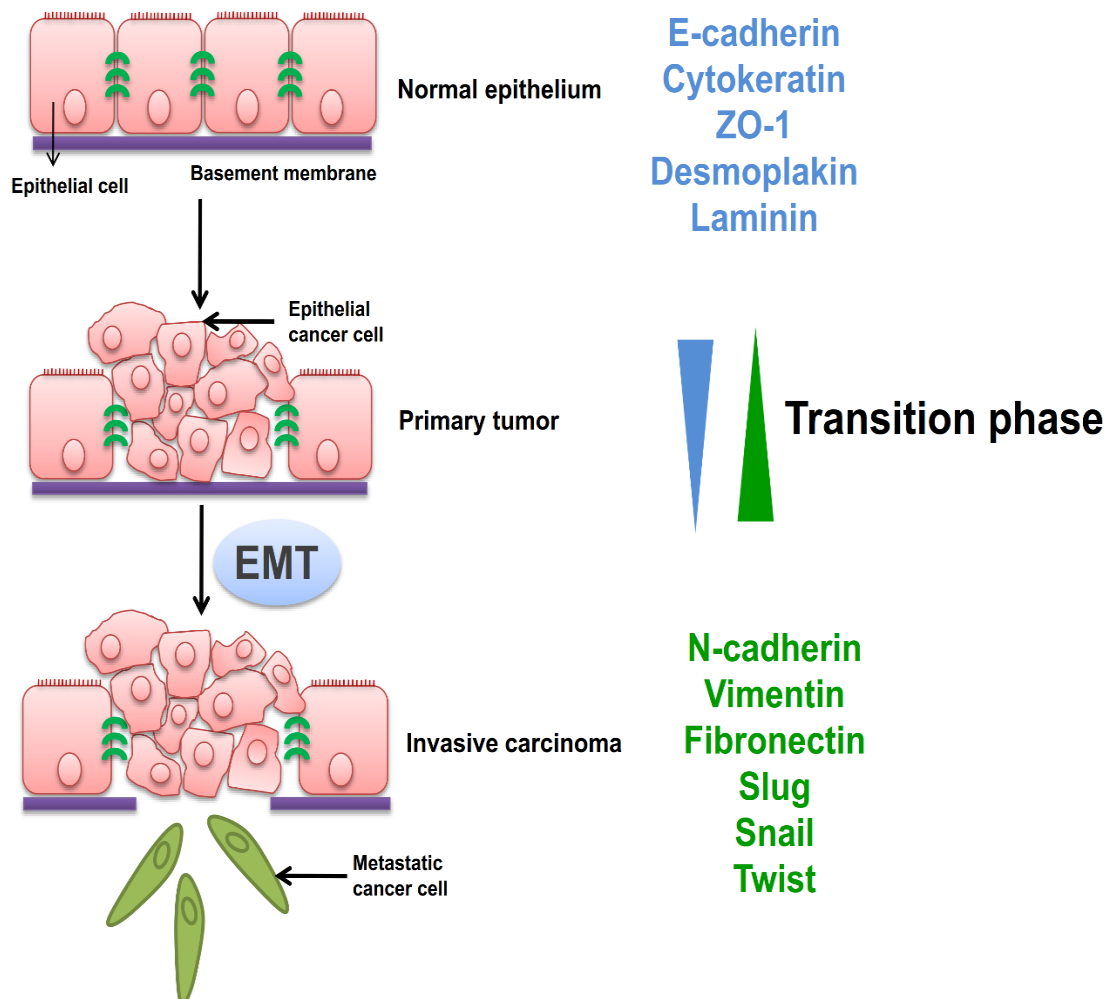
Figure 4: The process of EMT. EMT involves a dynamic change in the cellular phenotype and conversion from epithelial to mesenchymal type. The expression of epithelial markers like E-cadherin, cytokeratin are downregulated and mesenchymal markers like N-cadherin, Vimentin are upregulated. During the transition, there is an intermediate phase where cells express moderate levels of both epithelial as well as mesenchymal markers (Modified from (Mishra and Johnsen, 2014).

1.2.1 Master regulators of EMT

Certain key transcription factors orchestrate the induction of EMT during normal embryonic development and cancer progression and also act, at least in part, as transcriptional repressors of E-cadherin (*CDH1*) gene expression. Loss of E-cadherin is considered to be the fundamental event in EMT during normal development and cancer progression, its expression being down-regulated in advanced stage tumors. Many different extracellular signals have been shown to induce EMT including fibroblast growth factor, transforming growth factor-β (TGFβ), Notch and Wnt signaling (Bailey et al., 2007; Shipitsin et al., 2007; Vincan and Barker, 2008; Vogelstein and Kinzler, 2004; Wang et al., 2006; Yang and Weinberg, 2008a). Most of these extracellular factors carry out their action by regulating the transcription factors that repress epithelial genes such as *CDH1* and promote the transcription of genes that impart cells an invasive phenotype (Peinado et al., 2004a, 2007; Thiery and Sleeman, 2006).

Transcription factors involved in orchestrating EMT (EMT-TFs) include SNAIL1 (SNAI1), SLUG (SNAI2), ZEB1 (TCF8), ZEB2 (SIP1) and the basic helix-loop-helix factors (bHLH) E47 (TCF3) and TWIST1 (Craene and Berx, 2013; Moreno-Bueno et al., 2008; Peinado et al., 2007; Thiery and Sleeman, 2006). SNAIL1 suppresses the transcription of the *CDH1* gene by binding to E-boxes present in the promoter region of the gene where it recruits co-repressors to down-regulate its expression and promote EMT (Batlle et al., 2000; Cano et al., 2000). ZEB1 and ZEB2 has also been shown to directly bind to the *CDH1* promoter to inhibit its transcription, thereby decreasing E-cadherin expression (Comijn et al., 2001; Eger et al., 2005). The bHLH protein E47 has also been shown to repress the E-cadherin expression by directly binding to the *CDH1* promoter (Pérez-Moreno et al., 2001). Similarly, Twist1 also

decreases E-cadherin expression and cooperates with the epigenetic regulator and stem cell marker BMI1 (Vesuna et al., 2008; Yang et al., 2010). Thus, these EMT-TFs all repress the epithelial phenotype and promote a mesenchymal phenotype by regulating the expression of central genes (e.g., *CDH1*) involved in EMT to promote changes in cell morphology, cell adhesion, proliferation, cell survival, migration and invasion.

1.2.2 Mesenchymal-to-epithelial transition in tumor metastasis

Recent reports have highlighted the necessity for disseminated tumor cells to revert to epithelial phenotype in order to successfully establish macrometastasis at distant sites (Peinado et al., 2011; Zheng and Kang, 2014). It is believed that tumor cells at the metastatic site revert back to epithelial phenotype that allows them to proliferate and populate the secondary tumor by undergoing MET (Fig. 5) (Brabletz, 2012; Chaffer et al., 2007). Dynamic state of EMT and occurrence of MET is further strengthened by the fact that DTCs show characteristic feature of EMT whereas resulting secondary metastatic cells are largely epithelial (Bonnomet et al., 2012; Chao et al., 2010; Zheng and Kang, 2014). However, it is not well understood why DTCs need to revert back to an epithelial character to form macrometastasis. Previously it has been shown that EMT regulators put a block on cell growth and division (Chaffer et al., 2006; Ocaña et al., 2012; Tsai et al., 2012; Vega et al., 2004). This implies that in order to form macrometastasis DTCs need to proliferate immediately after colonization therefore, they need to revert back to an epithelial state. Furthermore, the E-cadherin gene (*CDH1)* has been shown to be differently methylated in primary (hypermethylated) and metastatic (demethylated) tumors strengthening the occurrence of MET (Graff et al., 2000; Nass et al., 2000).
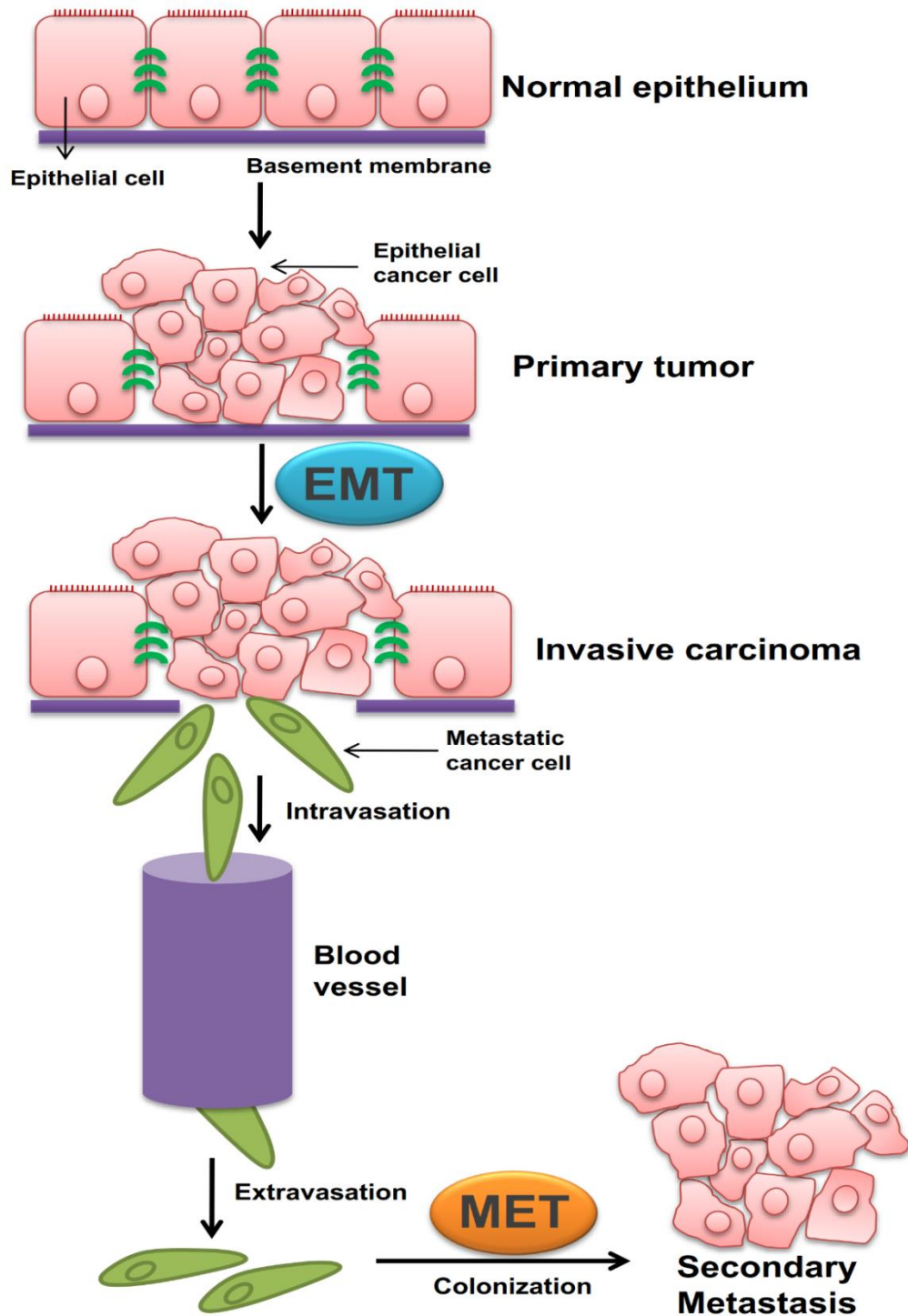
Figure 5: Tumorigenesis and metastasis requires EMT and MET. After dissemination from the primary tumor via EMT, disseminated cells have to enter into and travel through the blood vessels (intravasation). After reaching the distant site, cells exit the blood vessels (extravasation) and start to colonize. To be able to seed the secondary tumor, the cells need to revert back to an epithelial phenotype by undergoing MET. Modified from (Mishra and Johnsen, 2014).

1.2.3 EMT and cancer stem-like cells

Cancer stem-like cells (CSCs) are a small sub-population of cells within the tumor that have the ability of self-renewal and give rise to new tumor when injected into an immunocompromised mice model. The first study to show the existence of subpopulation of tumor cells that has the ability to seed new tumor was carried out in human acute myeloid leukaemia (AML) (Lapidot et al., 1994). In subsequent study these leukaemia initiating cells were isolated based on cell surface markers and were shown to have higher capacity to form a tumor compared to bulk of the tumor cells (Bonnet and Dick, 1997). Since then, CSCs have been isolated in different types of cancers including breast (Al-Hajj et al., 2003; Ginestier et al., 2007), lung (Eramo et al., 2007), prostate (Collins et al., 2005), pancreatic (Hermann et al., 2007) and colon (Ricci-Vitiani et al., 2007).

Since the discovery of CSCs many possibilities have been proposed to understand their evolution. One of the breakthrough study showed that EMT can generate the epithelial cells with stem cell-like properties (Mani et al., 2008). Furthermore, these CSCs expressed a cell surface marker CD44^{high}CD24^{low} compared to the rest of the population which was CD24^{high}CD44^{low}. Later, another study showed that EMT-TF ZEB1 is involved in generation and enhanced tumorigenic potential of CSCs (Chaffer et al., 2013). Another stud in pancreatic cancer showed that dissemination of tumor cells occurred in the early stage which was associated with EMT and disseminated cells exhibited the stem cell properties (Rhim et al., 2012).

Conventional chemotherapeutics target the actively dividing cells. However, targeting CSCs with the conventional therapy is not an option as these cells remain in quiescent state and are not actively dividing which to certain extent is attributable to their mesenchymal properties (Pece et al., 2010; Roesch et al., 2010). Furthermore, CSCs

are resistant to chemo- and radiotherapeutics because of increased expression of multi-drug resistance pump (like ABCG2) that can readily efflux the drugs out of the cells (Zhou et al., 2001). This implies that conventional therapies that targets the bulk of the tumor cells but is ineffective against the CSCs will result in relapse of the tumor after certain period (Fig. 6).
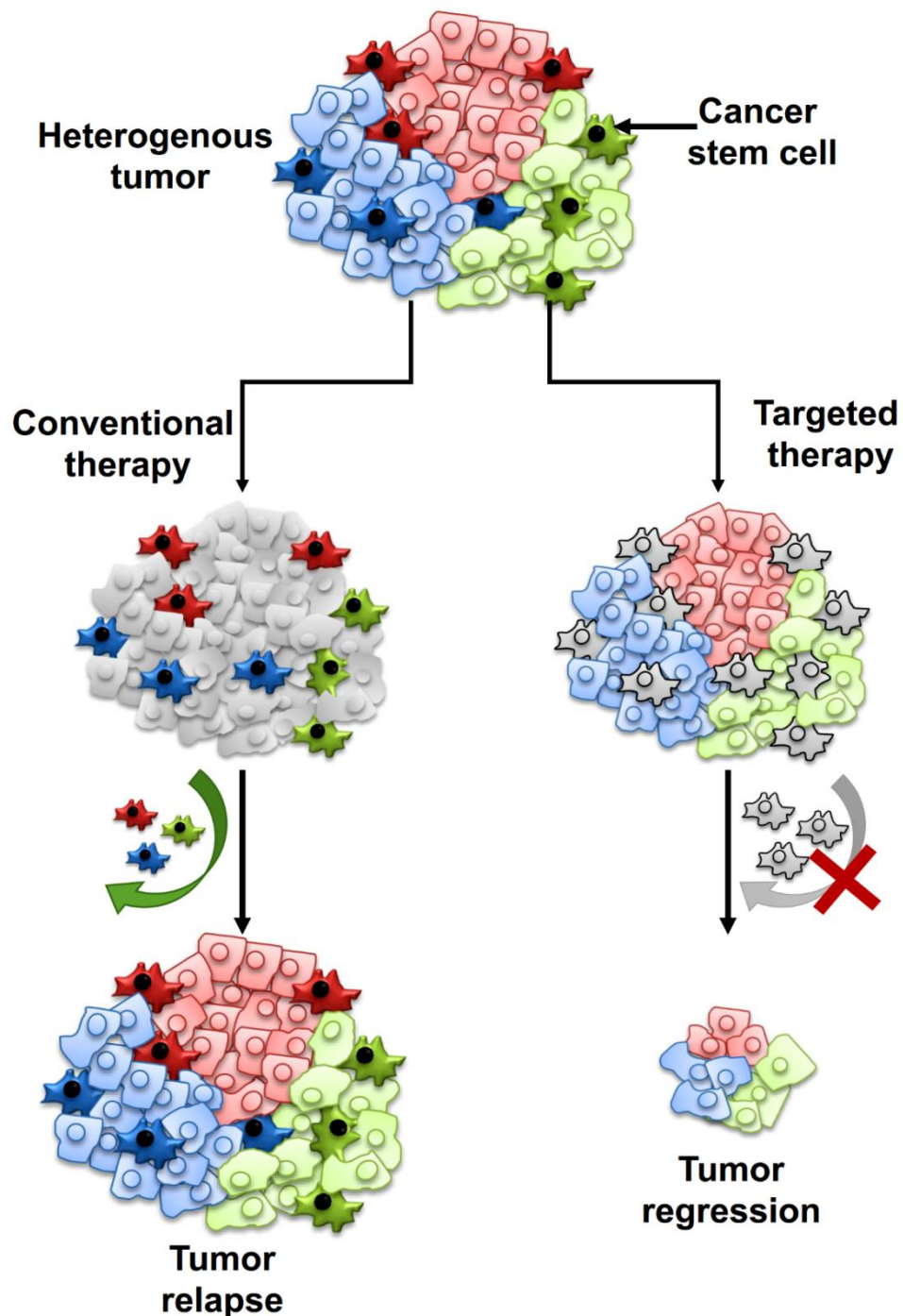
Figure 6: Conventional versus targeted cancer therapy against CSCs. Conventional therapy targets actively dividing cells. However, CSCs are usually in quiescent state and are not actively dividing. Therefore, bulk of the tumor cells are killed but the CSCs are spared which can repopulate and seed secondary tumor resulting in relapse. On the other hand, therapy specifically targeting CSCs will result in regression of tumor because bulk of the tumor cells do not have the self-renewal capacity hence, cannot seed the formation of new tumor.

1.3 Epigenetic regulation of EMT

The plasticity and reversibility of the epithelial and mesenchymal phenotypes in tumor cells and the dynamic activation and repression of genes involved in the EMT and MET processes is an outcome of increased binding of EMT-TFs to the enhancer or promoter regions and the epigenetic state of the target genes. Furthermore, post-translational histone modifications play an important role in determining the accessibility of transcription factors and epigenetic regulators to chromatin. Several studies have uncovered the underlying mechanism of interplay of epigenetic modifiers and other transcription factors in the regulation of EMT.

1.3.1 DNA methylation in EMT

DNA methylation is of the best characterized and most extensively studied chromatin modifications that plays an important role in maintaining a stable and heritable transcriptional repression of gene expression. DNA methylation is carried out by a family of DNA methyltransferase (DNMT) enzymes that covalently transfer a methyl group to the cytosine residues of the DNA, most commonly (more than 98 %) occurring at the CpG dinucleotides (Bird, 2002; Cedar and Bergman, 2009). CpG islands that are present around promoter regions of the genes are abnormally hypermethylated during malignant transformation, thereby leading to transcriptional repression and the silencing of important cell regulatory genes such as tumor suppressor genes (Robertson, 2005).

Loss of E-cadherin is the hallmark of EMT and aberrant promoter hypermethylation of the *CDH1* gene has been reported in many types of epithelial cancers (Chang et al., 2002; Chen et al., 2003; Lombaerts et al., 2006; Yoshiura et al., 1995). Recently, dimethylation of histone H3 on lysine 9 (H3K9me2) has been shown to be required for the DNA methylation of the CDH1 promoter (Dong et al., 2013). Similarly, direct interaction between DNMT1 and SNAIL has been shown to be responsible for the repression of *CDH1* gene (Espada et al., 2011). DNA methylation has been show to act cooperatively with histone modifications to repress target genes. One of the studies have reported that concomitant occurrence of DNA methylation, histone methylation and deacetylation at the CpG island within the *CDH1* promoter resulted in gene repression (Koizume et al., 2002).

DNA methylation has been shown to be involved in generation of cancer stem-like cells. A detailed comparison of DNA methylation patterns between CD44^{high}CD24^{low} and CD44^{low}CD24^{high} cells have demonstrated that the CD44^{high}CD24l^{ow} -enriched population of cells displays higher expression and hypomethylation of the genes encoding EMT-TFs, implying that changes in DNA methylation play an important role during the induction of EMT (Bloushtain-Qimron et al., 2008).

1.3.2 Role of various histone modifications in EMT

Epigenetic changes like histone methylation and acetylation have been frequently shown to be associated with numerous disorders including different types of cancer. Post-translational histone modifications alter the chromatin structure in a way that results in aberrant gene expression and altered cell phenotype which are common features of cancer (Fraga et al., 2005; Seligson et al., 2005).

1.3.2.1 Histone methylation

The methylation of histone proteins usually occurs at the amino acid side chains of arginine and lysine residues and is carried out by histone methyltransferases (HMT). Depending on the substrate that they act upon, they can be further classified as lysine (KMT) or arginine (PRMT) methyltransferases (Kouzarides, 2007). Arginine residues can be mono-, asymmetrically, or symmetrically dimethylated, whereas lysine residues can be mono-, di-, or trimethylated. Histone methylations can be activating or repressive depending on which position of amino acid that has been modified. Examples of lysine methylation events that are well characterized include methylation at K4, K9, K27, K36 and K79 of histone H3 and at K20 of histone H4. Trimethylations of K4 (H3K4me3), K36 (H3K36me3) and K79 (H3K79me3) are generally considered to be "active marks" and are frequently associated with expressed genes, whereas transcriptionally silenced genes are marked with repressive marks including H3K9me2, H3K9me3 and H3K27me3 (Kouzarides, 2007).

Histone methylation has been shown to be involved in EMT. A recent study has shown that EMT-TF SNAIL interacts with G9a (histone methyltransferase) and helps in the recruitment of G9a and DNMT to the CDH1 gene promoter leading to its repression (Dong et al., 2012). Another study have uncovered an interaction of SNAIL1 with another H3K9 methyltransferase SUV39H1, which also methylates H3K9 and promotes DNA methylation and the repression of the CDH1 promoter (Dong et al., 2013). Furthermore, the methyltransferase SET8, which catalyzes the repressive mark H4K20me1, has been shown to promotes EMT and breast cancer metastasis by interacting with the EMT-TF Twist at the CDH1 promoter (Yang et al., 2012).

Additionally, members of the polycomb repressive complex (PRC1 and PRC2) have been shown to be involved in promoting EMT. PRC1 contains the ubiquitin ligases

BMI1 and RING1, whereas PRC2 contains the HMT EZH2 and the additional subunits EED and SUZ12. EZH2 works together with SUZ12 and EED to catalyze the repressive trimethylation of H3K27 on the promoters of target genes leading to their repression (Sparmann and van Lohuizen, 2006). SNAIL has been shown to recruit PRC2 subunits EZH2 and SUZ12 to the CDH1 promoter, which in turn catalyzes H3K27me3 and leads to gene repression (Herranz et al., 2008).

1.3.2.2 Histone demethylation

Histone methylation was initially believed to be a highly stable and irreversible modification. However, many recent studies have disproven this supposition through the identification of histone demethylases and have provided important new insights into the nature of the dynamic regulation of histone methylation. Histone demethylases act as molecular "erasers" to remove methyl groups from lysine side chains (Shi and Whetstine, 2007; Trojer and Reinberg, 2006)**.** The first histone demethylase identified was LSD1 (KDM1A), which was initially shown to remove mono- or dimethyl groups from H3K4, leading to transcriptional repression (Shi et al., 2004)**.** The second class of histone demethylases belongs to the Jumonji-domain family of proteins, which contain a conserved Jumonji C (JmjC) domain and, in contrast to LSD1, can frequently fully demethylate trimethylated lysine residues by successively removing all three methyl group (Klose et al., 2006).

Histone demethylation has been shown to be linked with promoting EMT. A recent study has reported that physical interaction between SNAIL and LSD1 lead to recruitment of LSD1 to the promoter of epithelial genes and subsequent removal of H3K4me2 causing transcriptional repression (Lin et al., 2010). Another histone demethylase, KDM6B (JMJD3) has also been shown to promote EMT in breast cancer. Chromatin immunoprecipitation (ChIP) analyses have revealed that KDM6B

enhances the expression of SNAIL1 by removing the repressive mark H3K27me3 from its promoter, with KDM6B-induced EMT being dependent upon SNAIL1 expression (Ramadoss et al., 2012)**.** Another study have highlighted the role of KDM4B (JMJD2B) in increasing the expression of vimentin by interacting with β-catenin and demethylating repressive H3K9 mark on *VIM* gene promoter (Zhao et al., 2013).

1.3.2.3 Histone acetylation

Histone acetylation is probably the best characterized reversible histone modification. Histone proteins contain numerous lysine residues that can be acetylated by various histone acetyltransferases under various conditions. Most frequently, histone acetylation is associated with active transcription (You and Jones, 2012). Histone acetylation involves the transfer of an acetyl group from acetyl coenzyme-A to the side chain of lysine residues of histone proteins, thereby neutralizing the positive charge of the lysine residue and potentially loosening the chromatin structure because of the decreased interaction between the DNA and nucleosome. Recognition of acetylated lysine residues is performed by the bromodomain-containing family of proteins, which are components of several chromatin remodelling and transcriptional coactivator complexes, including some histone acetyltransferase and methyltransferase complexes (Dawson and Kouzarides, 2012). All core histone proteins are subject to modification. For instance, histones H3 and H4 both have a large number of lysine residues that can be acetylated (e.g., H3K9, H3K14, H3K18 and H4K16).

Some studies indicate a role of histone acetylation in promoting EMT. Acetyltransferase activity of p300/CBP and the subsequent hyperacetylation of Smad2 and Smad3 has been reported to be enhanced during TGFβ-induced EMT (Ko et al., 2013). However, a more detailed study of the role of histone acetylation in tumor

progression and the EMT program might open new avenues for epigenetic drug-based anticancer therapy.

1.3.2.4 Histone deacetylation

The reversal of histone acetylation is carried out by histone deacetylases (HDACs). In contrast to histone acetylation, histone deacetylation is associated with gene repression. HDACs restore the positive charge on the lysine residues of histone tails, preventing the recognition by bromodomain epigenetic readers and compacting the chromatin structure. Therefore, chromatin is less accessible for the binding of transcription factors and the recruitment of the transcriptional machinery to target genes (Glozak and Seto, 2007; Ropero and Esteller, 2007). HDACs are generally present in multi-subunit complexes in association with other proteins such as Mi2/NuRD, Sin3A and Co-REST (Dawson and Kouzarides, 2012). So far, 18 human HDACs have been identified and have been divided into four major classes based on sequence homology: Class I (HDACs 1, 2, 3, and 8), Class II (HDACs 4, 5, 6, 7, 9, and 10), Class III (Sirtuins including SIRT 1–7) and Class IV (HDAC11). Class I, II and IV HDACs require a Zn2+ ion for their catalytic action in contrast to Class III HDACs, which require NAD+ as a cofactor (Marks et al., 2001).

Several studies have reported the involvement of HDACs in cancer progression and EMT. In two different studies, the HDAC1/2 complex has been demonstrated to be recruited to and to repress the expression of the CDH1 promoter by either SNAIL (Peinado et al., 2004b) or ZEB1 (Aghdassi et al., 2012a). Furthermore, a Class III HDAC, SIRT1 has been shown to promote EMT and the metastatic growth of prostate cancer cells. SIRT1 directly represses the CDH1 promoter via its interaction with ZEB1, whereby it reduces RNA Polymerase-II recruitment and leads to transcriptional repression (Byles et al., 2012). Additionally, SIRT2 has been shown to be involved in

promoting EMT in hepatocellular carcinoma by regulating GSK3β/β-catenin signaling (Chen et al., 2013a). Together these findings highlight the role of HDACs in EMT and metastasis thus suggesting a potential target for anti-metastasis therapy.

1.4   Epigenetic therapy against EMT

Epigenetic mechanisms play a crucial role in the regulation of gene expression in healthy cells and aberrant changes in these mechanisms have been linked to the onset and progression of cancer. Because of the reversible nature of the epigenetic modifications and the amenability of epigenetic modifiers and readers to small molecule inhibitors, they might serve as ideal targets for therapeutic intervention. Indeed, efforts are being made to develop drugs that can restore the normal epigenetic state in cancer cells by inhibiting the enzymes that add or remove epigenetic modifications and the proteins that recognize them. Indeed, inhibitors for many of the epigenetic modifiers discussed in this review have previously been shown to be successful in blocking EMT and tumor metastasis (Fig. 7).

Hypermethylation of the CDH1 gene promoter has been shown to be one of the reasons for its silencing during EMT, however, treatment of E-cadherin-negative cancer cells with the DNMT inhibitor 5-azacytidine results in the reversal of gene repression and the restoration of the cell morphology to a more epithelial-like state (Yoshiura et al., 1995). However, other studies suggest that DNMT inhibition actually promotes tumor invasion and EMT in breast cancer. One of the studies involving the treatment of breast cancer cells with 5-aza-2′-deoxycytidine resulted in increased invasiveness and metastatic capacity (Ateeq et al., 2008). Thus, more data are required to fully assess the potential of DNMTs as therapeutic targets against cancer metastasis and EMT.
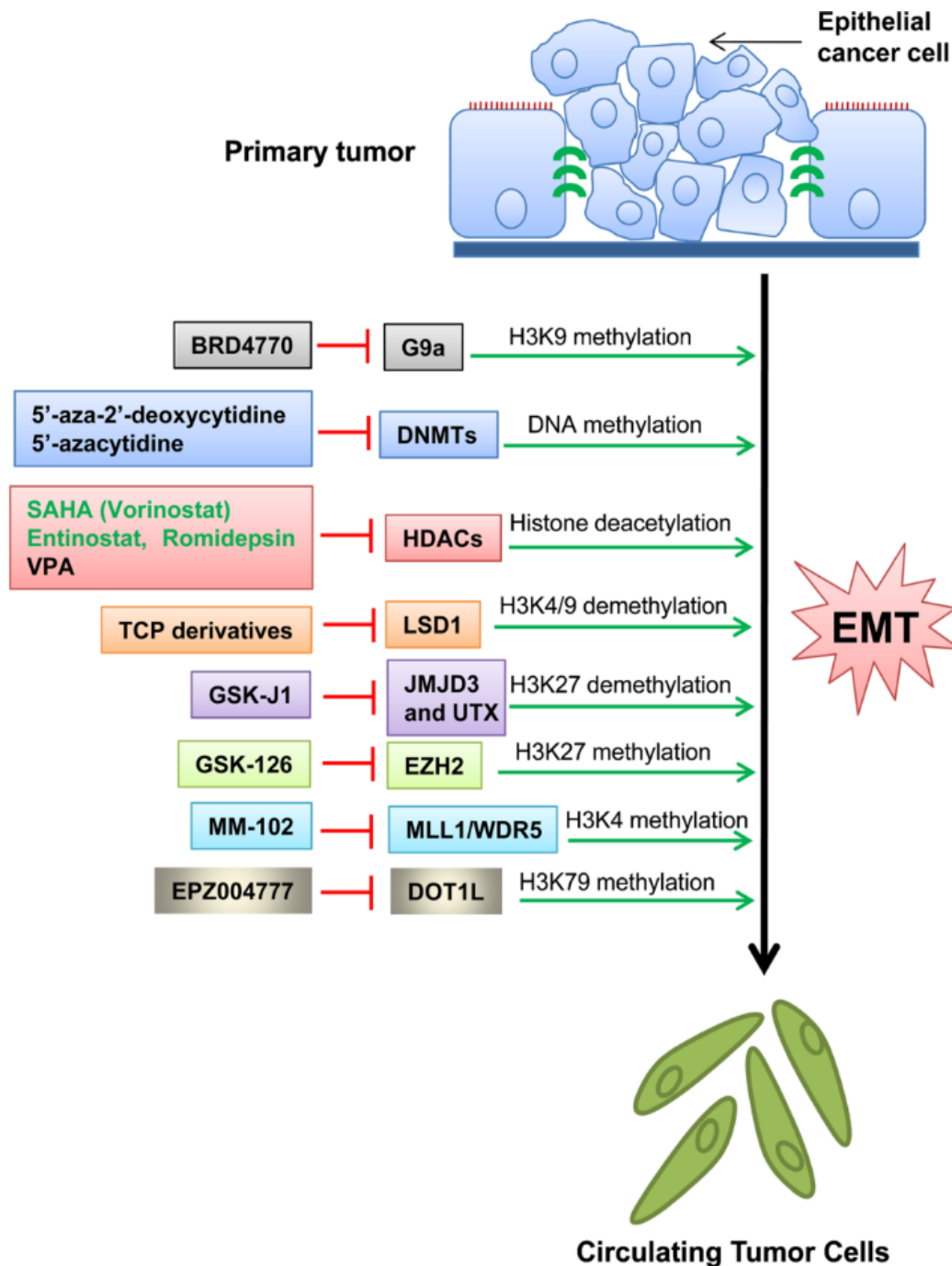
Figure 7: Targeting EMT by different small molecule inhibitors against epigenetic modifiers. Various epigenetic modifiers (writers and erasers) have been shown to be involved in regulating EMT. Targeting these epigenetic modifiers by small molecule inhibitors might be an effective treatment option to block cancer metastasis. Several such inhibitors have been recently discovered and tested: BRD4770 (Yuan et al., 2012), Entinostat (Kummar et al., 2007), Romidepsin (Saijo et al., 2012), TCP derivatives (Harris et al., 2012; Schenk et al., 2012), GSK-J1 (Kruidenier et al., 2012), GSK-126 (McCabe et al., 2012), MM-102 (Senisterra et al., 2013) and EPZ004777 (Daigle et al., 2011). Figure is adapted from (Mishra and Johnsen, 2014).

Most histone modifications are highly dynamic in nature. Because of the reversibility and plasticity of these modifications, histone-modifying enzymes represent promising potential therapeutic targets for the prevention or treatment of cancer metastasis and EMT. So far, a major focus has been directed to the potential utility of histone deacetylase inhibitors (HDACi) in cancer treatment. The HDACi trichostatin A (TSA) has been shown to effectively suppress TGFβ-induced EMT (Chen et al., 2013b; Kaimori et al., 2010; Lei et al., 2010; Witta et al., 2006; Yoshikawa et al., 2007). However, there are reports showing that HDACi can enhance the effect of TGFβ-induced EMT (Ji et al., 2015; Kong et al., 2012). Therefore, additional *in vivo* data and clinic I trials for HDACi will be required to address the concerns over their use in targeting EMT. Furthermore, combinatorial therapy of HDACi with other epigenetic modifier can be a potential option (O'Connor et al., 2006).
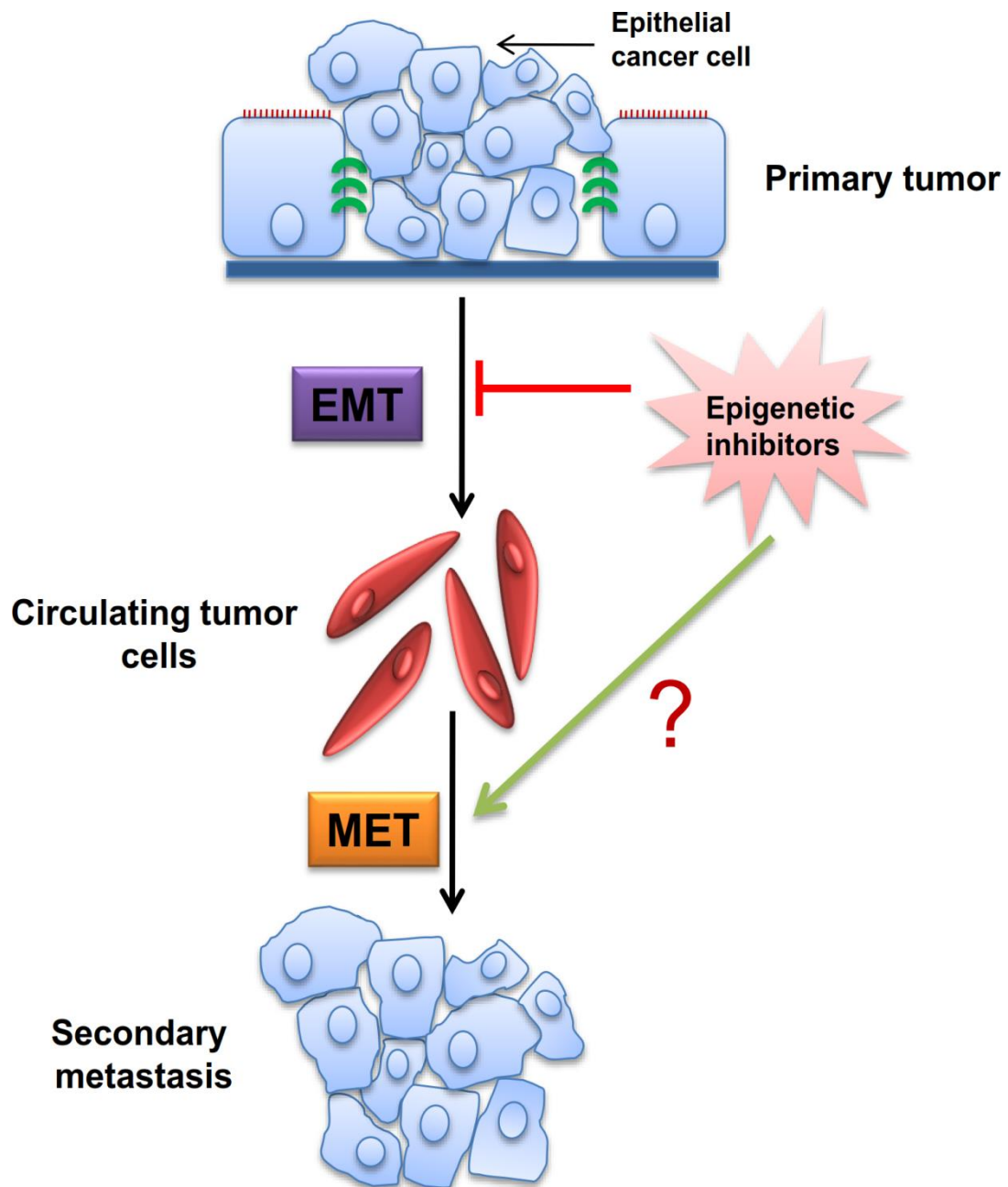
Figure 8: Inhibiting EMT can enhance the cancer metastasis. Inhibition of EMT in some cancer types can result in increased metastasis. Dissemination of tumor cells in some cancer types occur quite early. Inhibiting EMT in these types of cancer may result in MET hence, enhanced metastasis.

1.5 Transforming Growth Factor-β Signaling

Transforming growth factor-β (TGFβ) signaling is an evolutionary conserved ubiquitous signaling pathway which is crucial for the development and homeostasis of multicellular organisms. The importance of TGFβ signaling in multicellular organisms is evident from its effects on key cellular processes like proliferation, differentiation, migration and extra-cellular matrix production (Massague, 2000; Massagué et al., 2000). Furthermore, TGFβ signaling is dispensable in embryonic development where it plays an important role in organ morphogenesis, tissue homeostasis and immune system modulation. TGFβ is widely known for inducing epithelial-to-mesenchymal transition (EMT) that allows epithelial cells to lose their polarity and attain a mesenchymal shape. During embryonic development, EMT is required for cells to migrate from the site of their origin and participate in tissue and organ formation at distant sites. Further, numerous cell types in the body respond to the effects of TGFβ however, these effects are highly cell specific and context dependent (Massagué, 2012). Perturbation in TGFβ signaling has been implicated in a number of pathological diseases including developmental disorders, organ fibrosis and cancer, which is to some extent attributed to its pleiotropic effects in different cell types.

1.5.1 Molecular mechanism of TGFβ signaling

TGFβ signaling occurs through two different transmembrane receptors: TGFβ type-I (RI) and type-2 (RII) receptors (Fig. 9). Both RI and RII have serine/threonine kinase activity and are present as inactive dimers. While RI is required for phosphorylating the R-SMADs which then carry out the downstream signaling, the only known function of RII is to activate RI. The active dimer TGFβ ligand binds to the extracellular domain of RII which recruits and facilitates the formation of a hetero-tetrameric complex between RI and RII. In this complex, the activated RII kinase phosphorylates the

serine/threonine residues of the GS domain in RI resulting in its activation. Activated RI initiates the intracellular signaling by phosphorylating the receptor regulated R-SMADS SMAD2 and SMAD3, resulting in their increased affinity for the common-mediator SMAD (Co-SMAD or SMAD4), which is required for the further downstream signaling. R-SMADs and Co-SMADs form a complex and translocate to the nucleus. SMAD complex together with DNA binding cofactors and co-activators or co-repressors regulate the target gene expression. As different cell types express specific interactors of the SMAD complex, TGFβ signaling targets different genes in a cell-type specific manner (Derynck and Zhang, 2003; Massagué, 2000; Shi and Massagué, 2003).

Negative regulation of TGFβ signaling is carried out by inhibitory SMADs like SMAD7 and SMURFs (SMURF1 and 2) which provide a negative feedback loop. SMAD7 binds to the type-1 receptor, blocks the phosphorylation of R-SMAD and inhibits the complex formation between R-SMAD and Co-SMAD (Nakao et al., 1997). Additionally, SMURFs are involved in ubiquitination followed by proteasome-mediated degradation of active SMAD2 in the nucleus. TGFβ receptors are also targeted for proteasomal degradation by SMAD7 in cooperation with SMURF1 and SMURF2. SMAD7 forms a complex with SMURF1 and 2 in the nucleus and translocates to the plasma membrane upon TGFβ stimulation, it recruits ubiquitin ligases to the active TGFβ type-1 receptor which leads to its proteasomal degradation (Ebisawa et al., 2001; Kavsak et al., 2000; Suzuki et al., 2002).
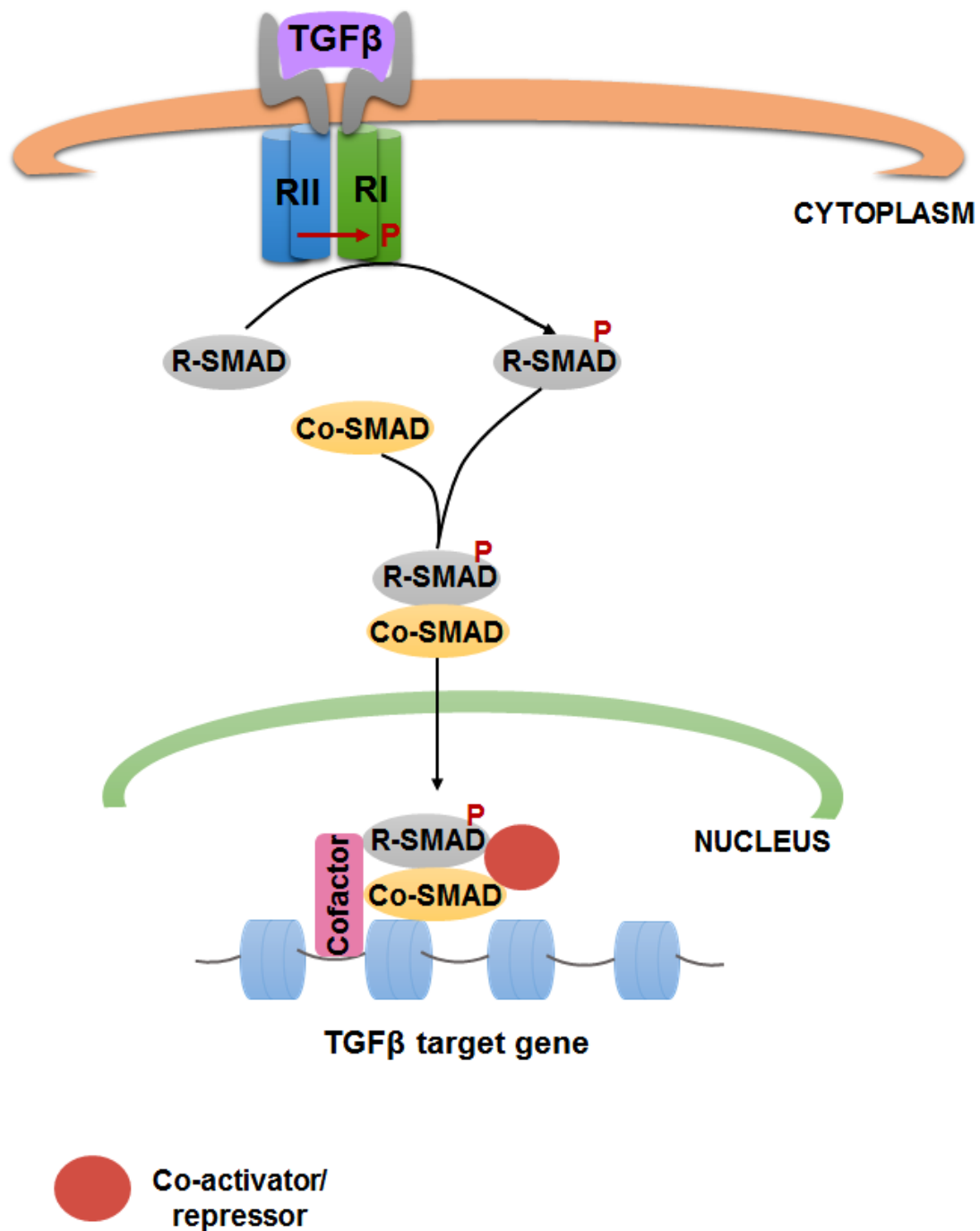
Figure 9: The Transforming Growth Factor-β signaling pathway. TGFβ signaling occurs through the association of type-1 (RI) and type-2 (RII) transmembrane receptors which are present as dimers. Both RI and RII contain a serine/threonine kinase domain in their cytoplasmic domains. TGFβ ligand binding to the transmembrane receptors (RI and RII) activates the RII which in turn phosphorylates the GS domain (inactive in the basal state) of RI leading to its activation. Activated RI then phosphorylate the R-SMAD (receptor activated SMADs) which binds to Co-SMAD and form a complex. The complex then translocates to the nucleus and, in cooperation with the DNA binding cofactors and co-activators/repressors regulates the target gene expression.

1.5.2 TGFβ signaling in cancer

TGFβ regulates a plethora of cellular processes like proliferation, apoptosis and differentiation in nearly all types of cells. Further, it also maintains tissue homeostasis and regulates cellular microenvironment thus, abnormality in TGFβ signaling often results in tumorigenesis. Components of TGFβ are often found to be mutated in malignancies especially in metastatic cancers (Antony et al., 2009; Caestecker et al., 2000; Chung et al., 1996; Levy and Hill, 2006; Lu et al., 1996; Myeroff et al., 1995). Cancer cells become resistant to the tumor suppressive effects of TGFβ and in later stages cancer cells use TGFβ signaling to migrate, invade and metastasize (Dalal et al., 1993; Kingsley et al., 2007; Massagué, 2008). TGFβ is like a double edged sword: it can act as a tumor suppressor as well as an oncogene in a cancer stage specific manner.

1.5.2.1 Tumor suppressive role of TGFβ

Cell proliferation and apoptosis are pivotal to control tumor formation and abnormality in these processes can be damaging. Downregulation of TGFβ receptor expression or impairment in their availability at the extracellular surface in tumor cells makes them resistant to growth inhibitory effects of TGFβ. Additionally, mutation in genes encoding SMAD proteins (SMAD2 and SMAD4) have been implicated in tumor development, implying that SMADs can be important regulators of tumor suppressive function of TGFβ (Coffey et al., 1988; Laiho et al., 1990; Moses et al., 1990).

TGFβ keeps a check on cell proliferation by activating CDK (cyclin-dependent kinase) inhibitors p15 and p21 and blocking the passage of cells to the G1 phase of the cell cycle. TGFβ induced p21 interacts with and inhibits the activity of CDK2-cyclinE/A complex thus, causes cell cycle arrest. Similarly, p15 interacts with and inhibits the activity of CDK4/6-cyclinD complex (Datto et al., 1995; Hannon and Beach, 1994;

Reynisdóttir et al., 1995; Sandhu et al., 1997). Furthermore, mutant p53 has been shown to confer resistance to cancer cells against growth arrest by TGFβ by hindering the CDK4 inhibition (Ewen et al., 1995). Additionally, TGFβ induces cell growth arrest by deactivating c-Myc, a potent growth-inducing transcription factor. Elevated levels of c-Myc are inhibitory to anti-proliferative effects of TGFβ thus making the cells resistant (Alexandrow et al., 1995; Staller et al., 2001; Warner et al., 1999).

In addition to its anti-proliferative effect, TGFβ exert an anti-apoptotic effect, which is cell type and context dependent. The exact mechanism behind the anti-apoptotic effects of TGFβ remains elusive, however, the Daxx adaptor protein, which is involved in JNK mediated apoptotic event has been shown to interact with the TGFβ type-II receptor during TGFβ induced apoptosis (Perlman et al., 2001). Moreover, activation of caspase 3 and 8 and deactivation of Bcl-xL (B-cell lymphoma-extra) large has been also implicated in the pro-apoptotic effects of TGFβ (Rotello et al., 1991; Saltzman et al., 1998; Selvakumaran et al., 1994).

1.5.2.2 Tumor promoting effects of TGFβ

Although TGFβ has an anti-tumor activity, tumor cells in the later stage become refractory to its growth inhibitory effects. One of the proposed mechanism is that tumor cells that become unresponsive to tumor suppressive functions of TGFβ express elevated levels of proto-oncogenes like c-Myc (Akiyoshi et al., 1999; Chen et al., 2001; Luo et al., 1999). Tumor cells that become unresponsive to the growth inhibitory effects of TGFβ but retain its functional signaling activity exhibit enhanced capacity of migration and invasion (Moustakas and Heldin, 2007). Furthermore, tumor- derived TGFβ is utilized by the tumor cells to escape the immune surveillance. Tumor-secreted TGFβ allows them to escape the host immune system by inhibiting the growth of key immune cells like macrophages, natural-killer cells, T-lymphocytes and neutrophils

(Ashcroft, 1999; Letterio and Roberts, 1998). Additionally, tumor-derived TGFβ also promotes the secretion of angiogenic factors that sustains the formation of a metastatic tumor (Noboru Ueki et al., 1992; Schwarte-Waldhoff et al., 2000; Stearns et al., 1999; Wikström et al., 1998).

One of the most widely investigated outcomes of deregulated TGFβ signaling is epithelial-to-mesenchymal transition (EMT) (Derynck and Akhurst, 2007). TGFβ induced EMT has been frequently shown to be associated with metastatic cancer. EMT helps the cells to disseminate and migrate to distant sites such that they may give rise to metastatic tumor (Kingsley et al., 2007; Oft et al., 1998; Thiery and Chopin, 1999). TGFβ is a potent inducer of EMT and TGFβ induced EMT has been shown to generate cells with the capacity to propagate new tumor, termed as cancer stem-like cells or tumor-initiating cells (Mani et al., 2008; Shipitsin et al., 2007). TGFβ stimulation in normal and transformed mammary epithelial cells resulted in the acquisition of stem cell-like properties and the capacity to form mammospheres. Furthermore, TGFβ transformed mammary epithelial cells exhibited high tumorigenic potential *in vivo*.

TGFβ is involved in promoting distal metastasis of tumor cells. Previously it has been shown that TGFβ promotes breast cancer cells to metastasize to the lung by activating the expression of the angiopoietin-like 4 (ANGPTL4) gene in disseminated tumor cells. ANGPTL4 helps the tumor cells to invade through the pulmonary walls and colonize (Padua et al., 2008).

1.6 Krüppel-like Transcription Factors

Krüppel-like factors (KLFs) are a highly conserved family of zinc-finger containing DNA-binding transcription factors. KLFs are expressed in a wide variety of human tissues and play a role in regulating a diverse array of cellular processes. The wide range of cellular processes that KLFs have been linked with includes growth and development, cell proliferation, differentiation, apoptosis and pluripotency (Bieker, 2001). KLFs share a homology with transcription factor Sp1, which also comprises C_2H_2-type zinc-fingers, therefore, they are broadly classified as Sp1/KLF family (Kaczynski et al., 2003). The first member of the KLF family was discovered in 1993 and was named EKLF (Erythroid Krüppel-like factor or KLF1) based on its homology to the Krüppel gene of *Drosophila melanogaster* (Miller and Bieker, 1993). The KLF family comprises 17 known members which are grouped according to their structural and functional similarities and are named KLF1-17 in order of their discovery. KLFs have varied tissue expression, some of them are ubiquitously expressed (KLF6, KLF10 and KLF11) while others are specific to certain tissue types (like KLF1 in erythroid cells, KLF2 in lung), thus, they have mutually exclusive as well as similar functions (Armstrong et al., 1998; Black et al., 2001; Turner and Crossley, 1999).

KLFs share a structural homology in their carboxy-terminal end where three zinc-finger domains are located. The first two zinc-fingers contain 25 amino acids each, while the third zinc-finger contains 23 amino acids and each of them can bind to three base pairs of the target DNA sequence. However, the functional diversity among the KLFs results from the differences in their functional amino terminal end (containing an activation or a repression domain) that allows them to recruit different binding partners like co-activators or co-repressors (Dang et al., 2000; Suske et al., 2005). Based on their functional divergence KLFs have been placed into three categories. KLFs 3, 8

and 12 (Group 1) interact with carboxy-terminal binding protein (CtBP), and thus, mainly function as transcriptional co-repressors. KLFs 1, 2, 4, 5, 6 and 7 (Group 2) interact with histone acetyltransferases and function as transcriptional activators. KLFs 9, 10, 11, 13, 14 and 16 (Group 3) interact with the transcriptional co-repressor Sin3A and thus, have repressor function (Kaczynski et al., 2003; Zhang and Bieker, 1998). KLFs 15 and 17 however, do not possess a specific motif sequence for interacting partners.

1.6.1 Krüppel-like Factors in cancer

KLFs have been implicated in tissue development and physiology of different organ systems. Perturbation in the functions of KLFs is associated with a wide variety of disorders including obesity, respiratory and cardiovascular disorders, inflammatory conditions and cancer. KLFs have an altered function in different types of cancer and they have regulatory effects on various processes like cancer cell proliferation, apoptosis and metastasis. Since different KLFs have varied expression in wide range of tissues and cancers they can either act as tumor suppressors or oncogenes in context dependent manner (Limame et al., 2010; Tetreault et al., 2013).

KLFs have been implicated in regulating cell growth and proliferation in normal and cancer cells mainly by targeting important cell cycle regulators like cyclin dependent kinases (CDKs) and CDK inhibitors like CDKN1A and CDKN1B. KLF4 has been shown to inhibit tumor cell proliferation in various cancer types including pancreatic, colorectal, lung and cervical cancer (Hu et al., 2009; Li et al., 2012; Zammarchi et al., 2011). However, the anti-proliferative effects of KLF4 are prone to switch into an oncogenic function depending on the expression of CDKN1A. Inactivation of CDKN1A results in a blockage of antitumor function of KLF4, thus switching it into an oncoprotein, indicating that the outcome of KLF4 function is dependent on CDKN1A

expression (Rowland and Peeper, 2006). Similarly, KLF6 generally inhibits cancer cell proliferation in different cancer types like NSCLS, HCC and prostate cancer through CDKN1A and CDKN1B (Narla et al., 2001). However, mutant forms or spliced isoforms of KLF6 promote cancer cell proliferation and tumorigenesis.

KLFs have been also implicated in regulating apoptosis, a mechanism that cancer cells have to circumvent. KLF4 promotes apoptosis in myeloid leukaemia cells by enhancing the expression of the pro-apoptotic BAX gene and repressing the anti-apoptotic BCL2 gene (Li et al., 2010). Similarly, KLF6 promotes apoptosis and inhibits tumorigenesis in melanoma cells by deactivating ERK2 and cyclin-D1 (Huh et al., 2010). However, KLF6 and other KLFs have also been shown to exert anti-apoptotic effects through their spliced isoforms. Downregulation of KLF6 in NSCLC (non-small cell lung cancer) and HCC (hepatocellular carcinoma) cell lines induces apoptosis, implying that KLF6 has an anti-apoptotic function in these cancers (Narla et al., 2005).

Additionally, KLFs exert their anti-tumorigenic role by modulating various signaling pathways involved in promoting cancer. KLF4 inhibits Wnt/β-catenin signaling by abrogating the TCF4 and β-catenin binding, thus, exerts a negative effect on cell proliferation (Evans et al., 2010). Furthermore, KLFs have also been implicated in modulating the estrogen (ER) and NOTCH signaling pathways. KLF5 exerts contrasting effects on breast cancer cells depending on their ER status. It has anti-proliferative effects on ER$^+$ breast cancer cells, however, it does not affect the proliferation of ER$^-$ breast cancer cells (Guo et al., 2010). This finding shows that the cellular context has a great impact on KLF function and gives insight into the divergent roles of KLFs.

1.6.2 Krüppel-like Factor-10

Krüppel-Like Factor-10 (KLF10) is a member of the zinc-finger containing Krüppel-like family of transcription factors and was originally identified as an early response gene for TGFβ in human fetal osteoblasts hence also termed as TGFβ-Inducible Early Gene-1 or TIEG1 (Subramaniam et al., 1995). Remarkably, the KLF10 expression level was increased within 30 min of TGFβ treatment and was increased up to 10-fold more than the basal expression 2 hrs post-treatment. Since the discovery of KLF10, numerous studies have reported a role in various cellular processes and altered function in various disorders including cancer.

Similar to other KLF family members, KLF10 contains three zinc-finger domains on its carboxy-terminal end and shares a homology with the Sp1 transcription factor. KLF10 has three repression domains (R1, R2 and R3) which are spread across the gene and preferentially bind to GC rich regions in the DNA to regulate target gene transcription. Additionally, KLF10 contains Src homology-3 (SH3) binding domains at the C-terminal end (Fig. 10). The KLF10 gene contains 5 coding exons and encodes a 480 amino acid protein (Subramaniam et al., 2007). Interestingly, KLF10 shares great homology with the EGR-alpha (early growth response-alpha) gene and they are transcribed from alternative promoters of the same gene located on chromosome 8q22.2 (Subramaniam et al., 1998). However, KLF10 and EGR-alpha proteins differ by 12 amino acids at their amino terminal end and KLF10 expression is higher compared to EGR-alpha in most tissues (Fautsch et al., 1998). Furthermore, the unique N-terminal end of the KLF10 protein distinguishes it from EGR-alpha and rest of the KLF family members (Blok et al., 1995).
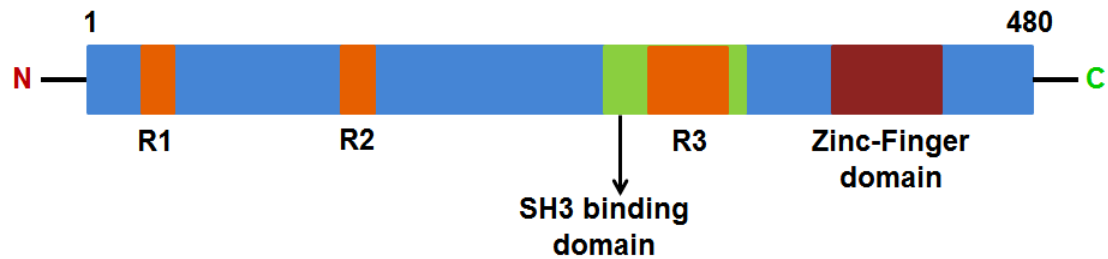
Figure 10: Protein structure of KLF10. KLF10 encodes a 480 amino acid protein. At the C-terminal end it has 3 zinc-finger domains which allow it to bind to the target DNA. Additionally, it contains 3 unique repression domains R1, R2 and R3. Several proline-rich SH3 (Src homology-3) domains are present at the C-terminal end (modified from Subramaniam et al., 2007).

1.6.3 The role of KLF10 in TGFβ Signaling

KLF10 was identified as an early response gene to TGFβ and thereafter, various studies have focussed on elucidating the role of KLF10 in regulating the TGFβ signaling pathway. Overexpression of KLF10 has been shown to imitate some of the general effects of TGFβ signaling in different cell types. For instance, overexpression of KLF10 in osteosarcoma cells resulted in enhanced alkaline phosphatase activity and decreased cell proliferation in a manner similar to TGFβ (Hefferan et al., 2000). Furthermore, KLF10 overexpression has been shown to mimic anti-proliferative effects of TGFβ in various cell lines (Chalaux et al., 1999; Ribeiro et al., 1999; Tachibana et al., 1997). Additionally, KLF10 overexpression has been shown to elicit the anti-proliferative effects of TGFβ by activating p21 expression (Johnsen et al., 2004)

TGFβ signaling is governed via SMAD proteins R-SMAD (SMAD2 and 3) and co-SMAD (SMAD4) which play an activating role in the while, the inhibitory SMAD7 is responsible for the negative feedback loop (Fig. 11). KLF10 has been shown to positively regulate TGFβ signaling by enhancing the expression of SMAD2 and repressing SMAD7 gene. KLF10 binds to a specific element in the proximal promoter region of the SMAD7 gene which represses SMAD7 expression, thus releasing the

negative feedback loop of TGFβ signaling (Johnsen et al., 2002a). However, even in the absence of SMAD7, KLF10 is still capable of enhancing TGFβ signaling, implying that KLF10 may have another target. In another study, it has been shown that in addition to repressing inhibitory SMAD7 gene, KLF10 activates the transcription of SMAD2 gene thus, a dual mechanism has been proposed through which KLF10 enhances the TGFβ signaling (Johnsen et al., 2002b).
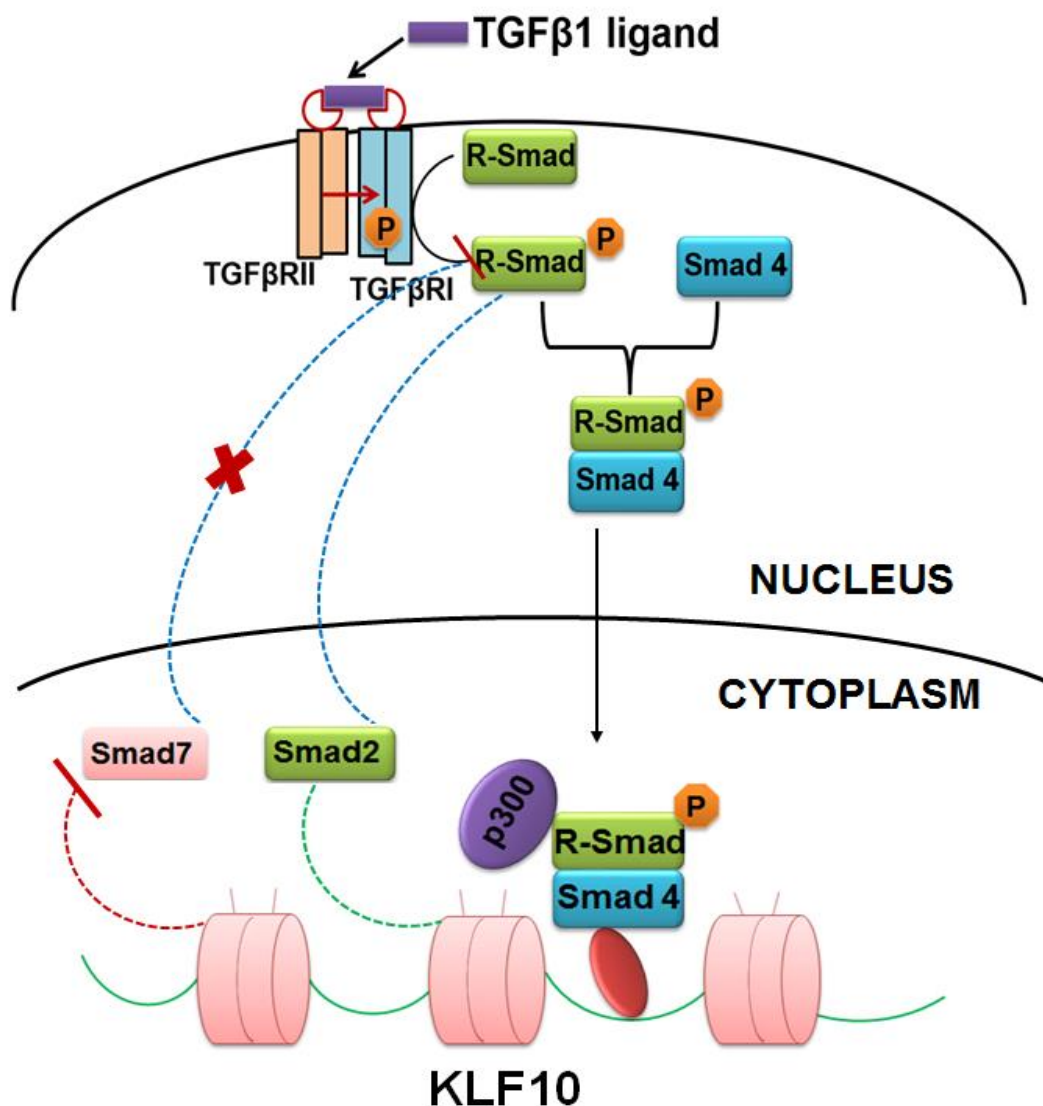
Figure 11: Role of KLF10 in TGFβ signaling. KLF10 enhances TGFβ signaling through a dual mechanism. KLF10 activates SMAD2 expression, thus, providing a positive feedback loop. Additionally, KLF10 binds to the promoter region of the inhibitory SMAD7 gene and inhibits its expression which can no longer inhibit the phosphorylation of the R-SMADs (receptor-activated SMADs) and thus, releases the negative feedback loop.

1.7 Lysine-Specific Demethylase 1

Histone methylation which usually occurs at the lysine and arginine amino acid residue of the histone tail, was considered to be stable and irreversible. However, discovery of the first histone demethylase, lysine-specific demethylase 1 (LSD1; also known as KDM1A or BHC110 or AOF2) changed this perception (Shi et al., 2004) and histone methylation is considered as more dynamic in nature.

1.7.1 Structure of LSD1

The LSD1 recombinant protein contains three functional domains: SWIRM domain (SWI3, RSC8,Moira), OXIDASE domain (also known as AOL or amine oxidase-like) and TOWER domain (Fig. 12) (Chen et al., 2006; Stavropoulos et al., 2006). N-terminal region of LSD1 is unstructured and is dispensable for its catalytic activity (Forneris et al., 2005).
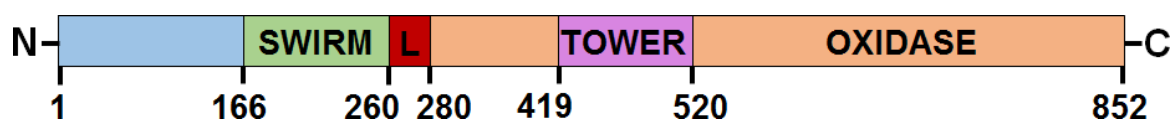
Figure 12: Protein structure of LSD1. Recombinant LSD1 protein is composed of 852 amino acid residues and three functional domains. Towards the C-terminal end it has OXIDASE domain and on N-terminal end SWIRM domain. In between SWIRM and OXIDASE domain is TOWER domain. Modified from (Chen et al., 2006).

To further understand the mechanism behind demethylase activity of LSD1, crystal structure was elucidated (Chen et al., 2006; Stavropoulos et al., 2006). Crystal structure revealed that LSD1 is composed of three different structural domains which are arranged in a form of highly asymmetric molecule ($\sim$60 Å wide and $\sim$140 Å high) (Stavropoulos et al., 2006). N-terminal SWIRM domain and C-terminal OXIDASE domain comprise the core of the protein that binds non-covalently to the FAD and functions as the enzymatic domain. While the TOWER domain protrudes from the core

and provides the platform for the interacting partners. The SWIRM domain is comprised of six-helical bundle structure and is packed against the OXIDASE domain, which contains two functional lobes, an FAD-binding lobe and substrate-binding lobe (Stavropoulos et al., 2006). The catalytic site of the LSD1 is located within the substrate-binding domain (Fig. 13).
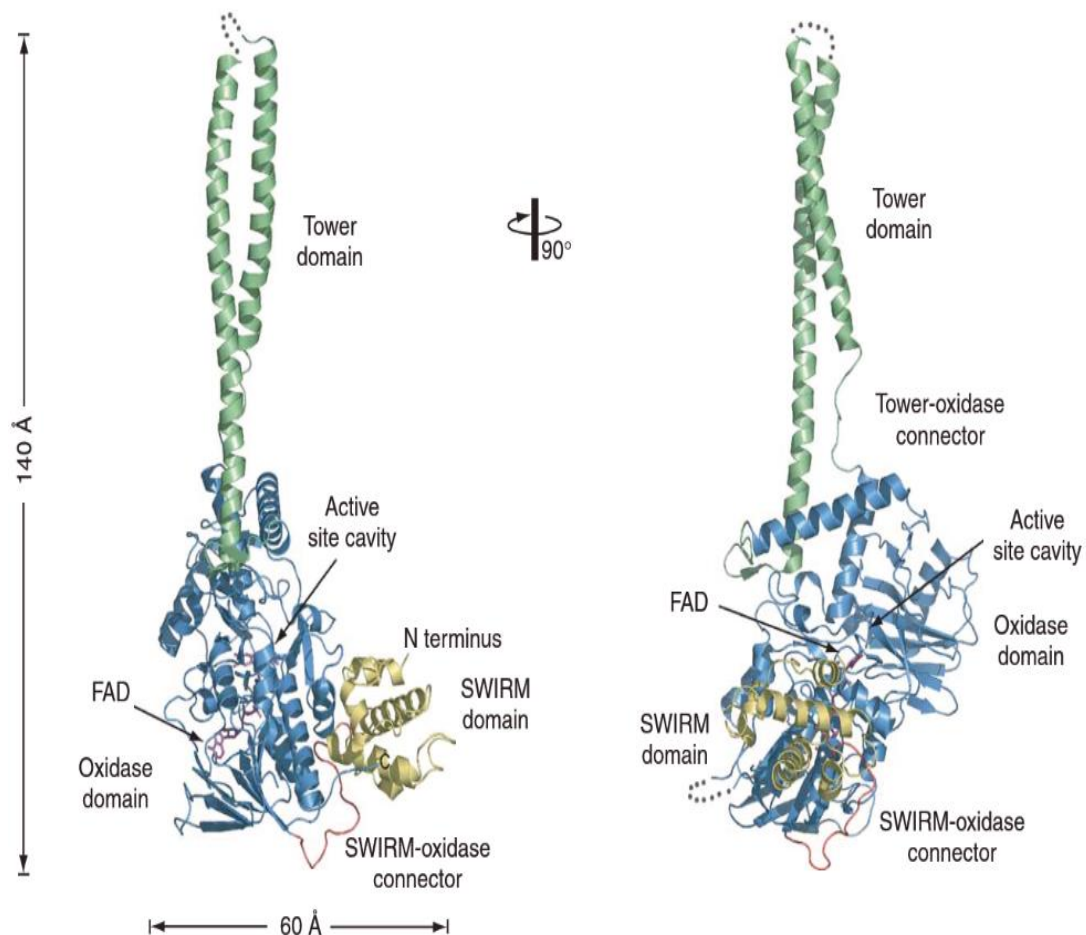
Figure 13: Structure of LSD1 in ribbon representation. Crystal structure of LSD1 depicting various domains (SWIRM, TOWER and OXIDASE) and active catalytic site. Adapted from (Stavropoulos et al., 2006).

1.7.2 Histone demethylation activity of LSD1

LSD1 was discovered as a histone demethylase which specifically demethylates lysine 4 residue of histone 3 and hence act as a transcriptional repressor (Shi et al., 2004). LSD1 is a component of different types of transcriptional repressor complexes including HDAC1, HDAC2 and CoREST. LSD1 catalyses demethylation of mono- or

dimethylated H3K4 via FAD-dependent oxidative reaction (Hakimi et al., 2003a; Humphrey et al., 2001; Shi et al., 2004). Furthermore, LSD1 has also been shown to act as transcriptional activator by demethylating repressive H3K9 methylation marks (Kahl et al., 2006). Interaction of LSD1 with androgen (AR) or estrogen (ER) nuclear hormone receptors have been implicated in its switch for substrate specificity to H3K9me1/me2 (Metzger et al., 2005; Perillo et al., 2008). Therefore based on its substrate specificity LSD1 can act as transcriptional repressor as well as activator (Fig. 14). These contrasting functions of LSD1 could be due to different interacting partners that can modulate its substrate specificity.
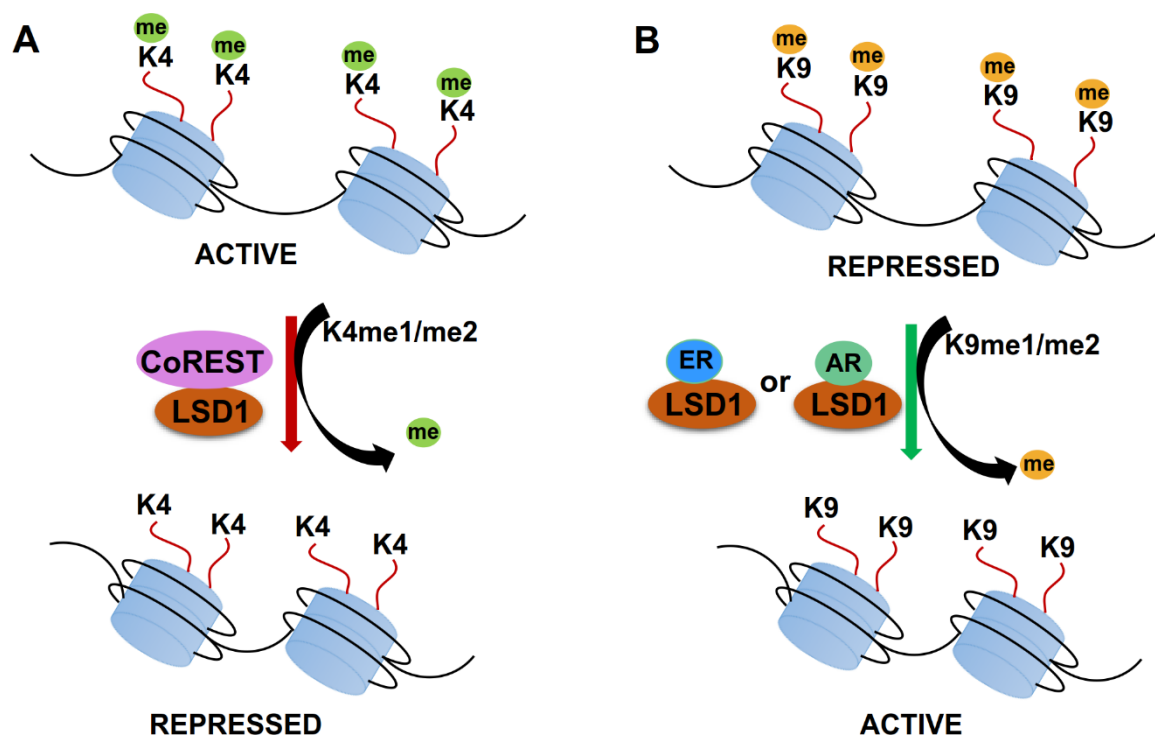
Figure 14: Dual function of LSD1. LSD1 can exert its demethylase activity on target gene to cause repression or activation. (A) LSD1 in complex with co-repressors like CoREST and HDACs, demethylate the active histone marks H3K4me1/me2 and causes transcriptional repression. (B) In presence of AR or ER, LSD1 changes the substrate specificity and acts as a transcriptional coactivator by demethylating mono- and dimethylated H3K9 repressive marks.

1.8   Aims of the study

The Transforming Growth Factor-β (TGFβ)/SMAD signaling pathway can function as either a tumor suppressor or metastasis promoter during tumor progression. In normal epithelial cells and early stages of epithelial tumorigenesis TGFβ functions as a tumor suppressor to decrease cell proliferation or induce apoptosis. However, during malignant progression tumor cells no longer respond to the anti-proliferative effects of TGFβ, but instead undergo an epithelial-to-mesenchymal transition (EMT) whereby cells acquire a migratory and invasive phenotype which promotes tumor metastasis. Resolution of the dichotomy in TGFβ function and a further understanding of its tumor suppressor and metastasis promoting functions has been an elusion for decades. Previous studies have demonstrated an important role of the TGFβ-Inducible Early Gene-1 (TIEG1)/Krüppel-like Factor-10 (KLF10) as a central regulator of TGFβ/SMAD signaling and the anti-proliferative functions of TGFβ.

Therefore, one of the aims of this study was to investigate the potential role of KLF10 in TGFβ-induced EMT. To elucidate the molecular mechanism we have peroformed various cell culture based assays, mice model and high-throughput ChIP-seq and RNA-seq.

In a more global approach, we have investigated the efficacy of small molecule inhibitors against the epigenetic modifiers in regulating the tumor cell phenotype. For this purpose we have utilized various small molecule inhibitors which are either in clinical trials or clinical development. To unravel the mechanism we have performed high-throughput ChIP-seq and RNA-seq, xenograft study in mice and cell culture based assays.

2. Materials

2.1 Technical equipment

EQUIPMENT	COMPANY
Agarose gel chamber	Harnischmacher Labortechnik, Kassel
Balance	Sartorius AG, Goettingen
Bandelin Sonoplus Sonicator	Bandelin electr. GmbH & Co. KG, Berlin
Bioruptor® Plus Sonicator	Diagenode SA, Liege, Belgium
Biological safety Cabinet "Safe 2020"	Thermo Fisher Scientific, Waltham, USA
CFX96™ Optical Reaction Module	Bio-Rad Laboratories GmbH, Muenchen
C1000™ Thermal Cycler	Bio-Rad Laboratories GmbH, Muenchen
Centrifuge 4 °C (Fesco 21)	Thermo Fisher Scientific, Waltham, USA
Centrifuge (Magefuge 1.OR)	Thermo Fisher Scientific, Waltham, USA
Centrifuge 4 °C (5417R)	Eppendorf AG, Hamburg
Counting chamber (Neubauer)	Brand GmbH & Co. KG, Wertheim
DynaMag™ 96 side	Life Technology, Carlsbad, USA
DynaMag™ 2	Life Technology, Carlsbad, USA
Eclipse TS100	Nikon, Tokyo, Japan
Electrophoresis & Electrotransfer Unit	GE healthcare Europe GmbH, Muenchen
Gel iX Imager	Intas Science Imaging GmbH, Goettingen
HERAcell 150i CO_2 Incubator	Thermo Scientific, Waltham, USA
Isotemp® water bath	Thermo Fisher Scientific, Waltham, USA
Invert Microscope "Axiovert 40 CFL"	Carl Zeiss MicroImaging GmbH, Goettingen
Magnet stirrer "MR3001"	Heidolph GmbH & Co. KG, Schwabach
Microwave	Clatronic International GmbH, Kempen
Mini Trans-Blot™ Cell	Bio-Rad Laboratories, Hercules, USA
Mini-PROTEAN Tetra Cell	Bio-Rad Laboratories, Hercules, USA
Mr. Frosty® cryo Freezer	Thermo Fisher Scientific, Waltham, USA
Nano Drop® ND-1000	Peqlab Biotechnology GmbH, Erlangen
pH meter inoLab®	WTW GmbH, Weilheim
Pipette Aid® portable XP	Drummond Scientific Co., Broomall, USA

Pipettes "Research" series	Eppendorf AG, Hamburg
Power supply Power pack P25T	Biometra GmbH, Goettingen
PowerPac^{TM} Basic Power Supply	Bio-Rad Laboratories,Hercules, USA
PowerPac^{TM} HC Power Supply	Bio-Rad Laboratories,Hercules, USA
Qubit® 2.0 Fluorometer	Invitrogen GmbH, Karlsruhe
Repeat pipette	Gilson Inc., Middleton, USA
Refrigerator	Liebherr GmbH, Biberach
Scanner Epson V700 Photo	Seiko Epson, Suwa, Japan
Shaker "Rocky"	Schutt Labortechnik GmbH, Goettingen
Test tube rotator	Schutt Labortechnik GmbH, Goettingen
Thermo mixer C	Eppendorf AG, Wessling-Berzdorf
Vortex-Genie 2	Electro Scientific Industr. Inc., Portland, USA
X-ray cassettes	Rego X-ray GmbH, Augsburg
-20 °C Freezer	Liebherr GmbH, Biberach
-150 °C Freezer (MDF-C2156VAN)	Panasonic, Kadoma, Japan
-80 °C Freezer "Hera freeze"	Thermo Fisher Scientific, Waltham, USA
2100 Bioanalyzer	Agilent Technology, Santa Clara, USA

2.2   Consumable materials

MATERIAL	COMPANY
Cellstar tissue culture dish 100x20 mm	Greiner Bio-One GmbH, Frickenhausen
Cellstar tissue culture dish 145x20 mm	Greiner Bio-One GmbH, Frickenhausen
Cellstar PP-tube 15 and 50ml	Greiner Bio-One GmbH, Frickenhausen
Cellstar 6- and 12-well cell culture plate	Greiner Bio-One GmbH, Frickenhausen
Cryo Tube Vial (1.8 ml)	Thermo Fisher Scientific, Waltham, USA
Cell scraper (16 cm)	Sartstedt AG & Co., Nümbrecht
DNA loBind Tube (0.5 and 1.5 ml)	Eppendorf AG, Wessling-Berzdorf
Gel blotting paper (Whatman paper)	Sartorius AG, Göttingen
Glass coverslip (18 mm)	Gebr. Rettberg GmbH, Göttingen

Hybond™-PVDF Transfer Membrane	GE Healthcare Europe GmbH, München
Microtube 1.5 ml, conical	VWR International GmbH, Darmstadt
Microtube 0.5 ml, 1.5 ml, 2 ml	Sarstedt AG & Co., Nümbrecht
NORM-JECT syringe	Henke Sass Wolf GmbH, Tuttlingen
PCR plate white(96-well Multiplate®)	Bio-Rad Laboratories GmbH, München
Parafilm® M	Pechiney Plastic Packaging, Chicago, USA
Petri dish 92x16 cm	Sarstedt AG & Co., Nümbrecht
PET track-etched cell culture inserts	BD Bioscience, Franklin Lakes, NJ, USA
Pipette tips	Greiner Bio-One GmbH, Frickenhausen
Pipette filter tips	Sarstedt AG & Co., Nümbrecht
Shandon coverplate	Thermo Fisher Scientific, Waltham, USA
Syringe filter, Ca-membrane	Sartorius AG, Göttingen
Ultra low attachment plates	Cornig Life Sciences, NY, USA
X-ray films "Super RX"	Fujifilm Corp, Tokyo, Japan

2.3   Chemicals

CHEMICAL	COMPANY
Acetic acid	Carl Roth GmbH & Co., KG, Karsruhe
Adefodur WB developing concentrate	Adefo-Chemie Gmbh, Dietzenbach
Adefodur WB fixinf concentrate	Adefo-Chemie Gmbh, Dietzenbach
Agarose	Biozym Scientific GmbH, Oldendorf
Agencourt® AMPure® XP Beads	Beckman Coulter Inc. Brea USA
Ammonium persulfate	Carl Roth GmbH & Co. KG, Karsruhe
Ammonium sulfate	Carl Roth GmbH & Co. KG, Karsruhe
Ampicillin	AppliChem GmbH, Darmstadt
Aprotinin	Carl Roth GmbH & Co. KG, Karsruhe
Bovine Serum Albumin (BSA)	Carl Roth GmbH & Co. KG, Karsruhe
Bromophenol blue	Sigma-Aldrich Co., St. Louis, USA
Calcium chloride	Carl Roth GmbH & Co. KG, Karsruhe
Chloroform	Carl Roth GmbH & Co. KG, Karsruhe

Co-precipitant Pink	Bioline, Luckenwalde
Glycine	Carl Roth GmbH & Co. KG, Karsruhe
Hydrochloric acid	Carl Roth GmbH & Co. KG, Karlsruhe
Immobilon™ Western HRP substrate	Merck Millipore KGaA, Darmstadt
Iodoacetamide	Sigma-Aldrich Co., St. Louis, USA
Isopropanol	Carl Roth GmbH & Co. KG, Karlsruhe
Leupeptin	Carl Roth GmbH & Co. KG, Karlsruhe
Linear Acrylamide	Thermo Fisher Scientific, Waltham, USA
Lithium Chloride (8 M)	Sigma-Aldrich Co., St. Louis, USA
Magnesium chloride	Carl Roth GmbH & Co. KG, Karlsruhe
Methanol	Carl Roth GmbH & Co. KG, Karlsruhe
Monopotassium phosphate	Carl Roth GmbH & Co. KG, Karlsruhe
N-ethylmaleimide	Sigma-Aldrich Co., St. Louis, USA
Nickel chloride ($NiCl_2$)	Sigma-Aldrich Co., St. Louis, USA
Nonidet™ P-40	Sigma-Aldrich Co., St. Louis, USA
Opti-MEM	GIBCO®, Invitrogen GmbH, Darmstadt
PBS tablets	Sigma-Aldrich Co., St. Louis, USA
Pefabloc SC	Carl Roth GmbH & Co. KG, Karlsruhe
Penicillin-Streptomycin solution	Sigma-Aldrich Co., St. Louis, USA
Peptone	Carl Roth GmbH & Co. KG, Karlsruhe
Potassium chloride	AppliChem GmbH, Darmstadt
Potassium dihydrogen phosphate	Carl Roth GmbH & Co. KG, Karlsruhe
Protein A Sepharose™ CL-4B	GE Healthcare, Uppsala, Sweden
RNase inhibitor	New England Biolabs, Frankfurt am Main
RNAiMAX	Invitrogen GmbH, Karlsruhe
Roti® Phenol	Carl Roth GmbH & Co. KG, Karlsruhe
Rotipherose® Gel 30	Carl Roth GmbH & Co. KG, Karlsruhe
Rotipuran® chloroform	Carl Roth GmbH & Co. KG, Karlsruhe
Rotipuran® isoamylalcohol	Carl Roth GmbH & Co. KG, Karlsruhe
Sepharose™ CL-4B	GE Healthcare, Uppsala, Sweden
Skim milk powder	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium acetate	Carl Roth GmbH & Co. KG, Karlsruhe

Sodium Azide	AppliChem GmbH, Darmstadt
Sodium chloride	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium deoxycholate	AppliChem GmbH, Darmstadt
Sodium dodecylsulfate (SDS)	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium Fluoride	AppliChem GmbH, Darmstadt
di-Sodium hydrogen phosphate	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium hydroxide	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium pyruvate	GIBCO®, Invitrogen GmbH, Darmstadt
SYBR Green I	Roche Diagnostics GmbH, Mannheim
TEMED	Carl Roth GmbH & Co. KG, Karlsruhe
Tris	Carl Roth GmbH & Co. KG, Karlsruhe
Triton X-100	AppliChem GmbH, Darmstadt
Trypsin-EDTA	Life Technology, Carlsbad, USA
TRIzol® Reagent	Invitrogen GmbH, Karlsruhe
Tween-20	AppliChem GmbH, Darmstadt
Trehalose	Pancreac AppliChem GmbH, Darmstadt
Xylene	Carl Roth GmbH & Co. KG, Karlsruhe

2.4   Kits and Reagents

KIT and REAGENT	COMPANY
Agilent High Sensitivity DNA Kit	Agilent Technology, Santa Clara, USA
Lipofectamine^{TM} 2000	Life technology, Carlsbad, USA
Lipofectamine^{TM} RNAiMAX	Life technology, Carlsbad, USA
Microplex Library Preparation^{TM} Kit	Diagenode SA, Liege, Belgium
Microplex Library Preparation^{TM} v2 Kit	Diagenode SA, Liege, Belgium
NEBNext® Ultra^{TM} Library Prep Kit	New England Biolabs, Ipswich, USA
Qubit dsDNA HS Assay	Life Technology, Carlsbad, USA
Immobilon Western Chemiluminiscent HRP substrate	Millipore, Billerica, USA
SuperSignal® West Femto Maximum	Thermo Fisher Scientific, Waltham, USA

2.5 Nucleic acids

2.5.1 siRNA Oligonucleotides

For preparing the transfection mix individual siRNAs against the respective genes were pooled in a 1:1:1:1 ratio.

siRNA	Target sequence (5'-3')	Source	Cat. No.
siKLF10 #1	CACCAGACCUGCCCAAUGA	Dharmacon	D-006566-01
siKLF10 #2	GAAGUGAGCAAGCUAAAUG	Dharmacon	D-006566-02
siKLF10 #3	GAUAAGGAGUCACAUCUGU	Dharmacon	D-006566-03
siKLF10 #4	GAAGAACCCACCUAAAUGU	Dharmacon	D-006566-04
Non-targeting siRNA-5	-	Dharmacon	D-001210-05
Luciferase GL2 duplex control siRNA	CGUACGCGGAAUACUUCGA	Dharmacon	
siLSD1 #1	UGAAUUAGCUGAAACACAA	Thermo Scientific	D-009223-01
siLSD1 #2	GACAAGCUGUUCCUAAAGA	Thermo Scientific	D-009223-02
siLSD1 #3	GUAAAGCCACCCAGAGAUA	Thermo Scientific	D-009223-03
siLSD1 #4	CUAUAAAGCUCCAAUACUG	Thermo Scientific	D-009223-04

2.5.2 RT-PCR Primers

Reverse transcription primers were purchased from Sigma-Aldrich, Germany.

qPCR Primers

qPCR primers were designed using the primer designing tool of the NCBI (http://www.ncbi.nlm.nih.gov/tools/primer-blast/).

Gene name	Primer sequence (5'-3')	Reference
E-cadherin F	CTTTGACGCCGAGAGCTACA	This study
E-cadherin R	AAATTCACTCTGCCCAGGACG	This study
N-cadherin F	GGGTCATCCCTCCAATCAAC	This study
N-cadherin R	ACCTGATCCTGACAAGCTCT	This study
MMP2 F	TGGATGATGCCTTTGCTCGT	This study
MMP2 R	TATCCATCGCCATGCTCCCA	This study
MMP7 F	CGGATGGTAGCAGTCTAGGGAT	This study
MMP7 R	TCAGAGGAATGTCCCATACCCA	This study
SNAI2 F	TCGGACCCACACATTACCTTG	This study
SNAI2 R	AAAAAGGCTTCTCCCCCGTG	This study
SNAI1 F	CCAGTGCCTCGACCACTATG	This study
SNAI1 R	CTGCTGGAAGGTAAACTCTGGAT	This study
TJP3 F	CAGAGCATGGAGGATCGTGG	This study
TJP3 R	TCAGGTTCTGGAATGGCACG	This study
ZEB1 F	GCGCAGAAAGCAGGCGAACCC	This study
ZEB1 R	CCCTTCCTTTCCTGTGTCATCCTCC	This study
CD24 F	GCTCCTACCCACGCAGATTT	This study
CD24 R	GAGACCACGAAGAGACTGGC	This study
HNRNPK F	ATCCGCCCCTGAACGCCCAT	Karpiuk et al., 2012
HNRNPK R	ACATACCGCTCGGGGCCACT	Karpiuk et al., 2012
RPLP0 F	GATTGGCTACCCAACTGTTG	Fritah et al., 2005
RPLP0 R	CAGGGGCAGCAGCCACAAA	Fritah et al., 2005

ChIP Primers

Gene name	Primer sequence (5'-3')	Reference
SNAI2 TSS F	CCAGTTCGCTGTAGTTTGGC	This study
SNAI2 TSS R	CAGACCCGCTGGCAAGAT	This study
SNAI2 TR F	AGTGATGGGGCTGTATGCTC	This study
SNAI2 TR R	CTCCCTCCCTTTTCTTTCCCAG	This study

2.6 Buffers

2.6.1 ChIP buffers

Nuclear preparation buffer

Component	Stock concentration [M]
NaCl	5
EDTA (pH 8.0)	0.5
Tris-HCl (pH 7.5)	1
NP-40 (v/v)	10 %
Triton-x 100 (v/v)	10 %
NaF	0.5

Dilution buffer

Component	Stock concentration [M]
NaCl	5
EDTA (pH 8.0)	0.5
Tris-HCl (pH 8.0)	1
NP-40 (v/v)	10 %
Sodium deoxycholate (w/v)	10 %
NaF	0.5

IP buffer

Component	Stock concentration [M]
NaCl	5
EDTA (pH 8.0)	0.5
Tris-HCl (pH 8.0)	1
NP-40 (v/v)	10 %
Sodium deoxycholate (w/v)	10 %
NaF	0.5
SDS (w/v)	10 %

Wash buffer

Component	Stock concentration [M]
LiCl	8
EDTA (pH 8.0)	0.5
Tris-HCl (pH 8.5)	1
NP-40 (v/v)	10 %
Sodium deoxycholate (w/v)	10 %
NaF	0.5

Sonication buffer-1

Component	Stock concentration [M]
EDTA (pH 8.0)	0.5
Tris-HCl (pH 8.0)	1
SDS (w/v)	1 %

Sonication buffer-2

Component	Stock concentration [M]
NaCl	5
EDTA (pH 8.0)	0.5
Tris-HCl (pH 8.0)	1
NP-40 (v/v)	10 %
NaF	0.5

TE buffer

Component	Stock concentration [M]
EDTA (pH 8.0)	0.5
Tris-HCl (pH 8.0)	1

Cross-linking buffer

Component	Stock concentration [M]
Formaldehyde	37 %
PBS	1X

Proteinase inhibitors

Inhibitor	Stock concentration
$NiCl_2$	1 mM
Pefabloc	1 mM
Aprotinin/Leupeptin	1 ng/µl
N-ethylmaleimide	1 mM
Indole acetamide	10 µM
Glycerol 2-phosphate disodium salt hydrate	10 mM

2.6.2 Western blot buffers

SDS separating gel (X %)

Acrylamide	X %
Tris-HCl (pH 8.8)	375 mM
SDS (w/v)	0.1 %
APS (v/v)	0.1 %
TEMED	0.04 %

Stacking gel (5%)

Acrylamide	X %
Tris-HCl (pH 6.8)	125.5 mM
SDS (w/v)	0.1 %
APS (v/v)	0.1 %
TEMED	0.1 %

Transfer buffer

Western salts (10X) (v/v)	10 %
Methanol (v/v)	15 %

6X Laemmli buffer

Tris-HCl (pH 6.8)	0.35 M
Glycerol	30 %
SDS (w/v)	10 %
DTT	9.3 %
Bromophenol blue	0.02 %

Western salts (10X)

Glycine	1.92 M
Tris-HCl (pH 8.3)	250 mM
SDS (w/v)	0.02 %

Running buffer

Glycine	200 mM
Tris	25 mM
SDS (w/v)	0.1 %

TBS

Component	Stock Concentration
NaCl	150 mM
KCl	2.68 mM
$Na_2HPO_4 x2H_2O$	4.29 mM
KH_2PO_4 (pH 7.4)	1.47 mM

TBS-T

TBS + 0.1 % (w/v) Tween-20

RIPA buffer

PBS	1 X
NP-40 (v/v)	1 %
Sodium deoxycholate (v/v)	0.5 %
SDS (w/v)	0.1 %

Blocking solution

TBST	1 X
Skimmed milk (w/v)	5 %

Proteinase inhibitors

Inhibitor	Stock concentration
Pefabloc	1 mM
Aprotinin/Leupeptin	1 ng/µl
Glycerol 2-phosphate disodium salt hydrate	10 mM

2.6.3 qRT-PCR buffer

Component	Concentration
Tris-HCl (pH 8.8)	75 mM
$(NH_4)_2SO_4$	20 mM
Tween-20	0.01 %
$MgCl_2$	3 mM
dNTPs	200 µM
Taq DNA polymerase	0.5 U/reaction
Triton X-100	0.25 %
Trehalose	300 mM
Random primer	30 nM
SYBR Green I	1:80,000

2.7 Proteins, enzymes and standards

2.7.1 Antibodies

Primary antibodies

Primary antibodies for ChIP, western blot and immunofluorescence were used in below mentioned concentration and dilutions. The antibody dilutions for western blot analysis were supplemented with 0.01% sodium azide.

Antibody	Source	Cat. No.	Clone	IF	WB	ChIP
Vimentin	Santa Cruz	sc6260	v9	1:400	1:2500	
E-cadherin	Cell Signaling	3195	24E10	1:200	1:1000	
N-cadherin	Cell Signaling	#13116	D4R1H		1:1000	
ZEB1	Sigma	HPA027524		1:200	1:500	
ZO-1	Cell Signaling	8193	D7D12	1:200		
SNAI2	Cell Signaling	9585	C19G7	1:200	1:1000	
SNAI1	Abcam	ab17732			1:1000	
KLF10	Abcam	ab184182	EPR12102(2)		1:1000	1 µg
CD24	Hans Peter		SWA11		1:2	
p21 (CDKN1A)	Cell Signaling	#2947	12D1		1:1000	
H3K9Ac	Diagenode	C15410004				1 µg
H3K27Ac	Diagenode	C15410196				1 µg
HDAC1	Diagenode	C15410053				2 µg
LSD1	Cell Signaling	#2184	C69G12			1 µg
H3K4me1	Diagenode	C15410194				1 µg
HDAC2	Santa Cruz	sc-7899	H54			1 µg
HSC-70	Santa Cruz	sc7298	B-6		1:10000	
IgG (non-specific)		ab46540				1 µg

Secondary antibodies

NAME	Source	Cat. No.	WB	IF
Anti-mouse (IgG)-HRP	Santa Cruz		1:10,000	
Anti-rabbit (IgG)-HRP	Santa Cruz		1:10,000	
Alexa Fluor® 594 Goat Anti-Mouse IgG (H+L)	Life Technologies	A11005		1:500
Alexa Fluor® 488 Goat Anti-Rabbit IgG (H+L)	Life Technologies	A11008		1:500

2.7.2 Enzymes

Enzyme	Company
Taq DNA Polymerase	Prime Tech, Minsk, Belarus
Reverse Transcriptase (M-MuLV)	New England Biolabs, FFM
Proteinase-K	Life Technology, Carlsbad, USA
RNase A	Qiagen GmbH, Hilden
RNase Inhibitor	New England Biolabs, FFM

2.7.3 Molecular weight standards

Standard	Company
Gene RulerTM DNA-Ladder	Fermentas GmbH, St. Leon-Rot
PageRulerTM Prestained Protein Ladder	Fermentas GmbH, St. Leon-Rot

2.8 Cell culture medium

DMEM cell culture medium

DMEM phenol red-free medium

10% fetal bovine serum

100 U/ml penicillin

100 µg/ml streptomycin

DMEM/F12 cell culture medium

DMEM phenol red-free medium (high glucose)

10% fetal bovine serum

100 U/ml penicillin

100 µg/ml streptomycin

MEM cell culture medium

MEM phenol red-free medium (high glucose)

10% fetal bovine serum

1% L-Glutamine

100 U/ml penicillin

100 µg/ml streptomycin

Cell freezing medium

DMEM - 42%

Fetal bovine serum - 50%

DMSO - 8%

PBS for cell culture

1 PBS tablet per 500 ml of distilled H_2O

2.9   Cell lines

Cell line	Species	Origin	Source
A549	Human	Lung adenocarcinoma	Prof. Ekkehard Dikomey, UKE, Hamburg
Panc1	Human	Pancreatic cancer	Dr. Elisabeth Heßmann, University Medical Center, Göttingen
L3.6	Human	Pancreatic cancer	Dr. Elisabeth Heßmann, University Medical Center, Göttingen
BxPC3	Human	Pancreatic adenocarcinoma	Dr. Elisabeth Heßmann, University Medical Center, Göttingen

2.10 Growth factors and inhibitors

NAME	Company
Recombinant Human TGF-β1	R&D Systems® GmbH, Germany
SP2509 (LSD1 inhibitor)	Selleckchem, USA
Vorinostat	Selleckchem, USA
Resminostat	4SC AG, Martinsreid, Germany
4SC-202	4SC AG, Martinsreid, Germany

2.11 Software and online tools

NAME	SOURCE
useGalaxy	https://usegalaxy.org/
Galaxy Cistrome	http://cistrome.org/ap/root
Galaxy Deeptools	http://deeptools.ie-freiburg.mpg.de/
REVIGO GO analysis	http://revigo.irb.hr/
DAVID GO analysis	https://david.ncifcrf.gov
Gene Set Enrichment Analysis	http://software.broadinstitute.org/gsea/index.jsp
R statistical software	https://www.r-project.org/
DESeq2 package	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
DiffBind package (version 3.2)	http://bioconductor.org/packages/release/bioc/html/DiffBind.html
GREAT analysis software	http://bejerano.stanford.edu/great/public/html/
ReMap online tool	http://tagc.univ-mrs.fr/remap/
Oncomine database	https://www.oncomine.org/resource/main.html
Kaplan-Meier plotter	http://kmplot.com/analysis/index.php?p=background
Zeiss ZEN lite software	http://www.zeiss.com/microscopy/en_de/products/microscope-software/zen-lite.html
Bio-Rad CFX Manager 3.1	Bio-Rad Laboratories, Hercules, USA
Image Lab Version 5.2 build 14	Bio-Rad Laboratories, Hercules, USA
Integrative Genome Viewer 2	https://www.broadinstitute.org/software/igv/download
Primer designing tool/NCBI primer-BLAST	www.ncbi.nlm.nih.gov/tools/primer-blast/

3. Methods

3.1 Cell culture

3.1.1 Culturing of adherent cells

A549 (lung adenocarcinoma) and Panc1 (pancreatic ductal carcinoma) cells were cultured in phenol-red free high-glucose Dulbecco's modified Eagle's medium (DMEM) whereas BxPC3 (pancreatic adenocarcinoma) and L3.6 (pancreatic cancer cell) were grown in minimum essential medium Eagle (MEM). Media was supplemented with 10% fetal bovine serum (FBS), 100 U/ml penicillin, 100 μg/ml streptomycin at 37 °C and 5% CO_2.

For TGFβ treatment, cells were washed once with 1X PBS and fresh medium was added. Cells were treated with 5 ng/ml TGFβ for either 90 min or 72 hours.

Cells were treated with different inhibitors for 12 hours at following concentrations: 4SC-202 (1 μM), SP2509 (500 nM), Resminostat (1 μM), Vorinostat (1 μM).

3.1.2 siRNA mediated reverse transfection

siRNA mediated reverse transfection was performed using LipofectamineTM RNAiMAX according to the manufacturer's instruction. For transfection in a 6-well plate, 30 pmol of used siRNA's were diluted in 500 μl of opti-MEM per well following which 5 μl of LipofectamineTM RNAimAX was added and mixed gently. Transfection mix was then incubated at RT for 20 min. While the transfection mix was incubating cells were washed twice with 1X PBS and trypsinized and diluted in antibiotics free medium. Cells were counted using Neubauer counting chamber and approximately 250,000 cells were added to each well containing transfection master mix. For transfections in 10 cm culture plates a four times higher volume of the reagents and cell number was used.

3.1.3 Migration assay

A transwell migration assay was performed to monitor the migration potential of the cells upon knock-down of a gene of interest as well as various treatment conditions. Post-transfection cells were trypsinized and approximately 25,000 cells were seeded into the upper chamber of the cell culture inserts (por size 8.0 µm) which were pre-equilibrated with serum-free medium for approximately 30 minutes. Cells were allowed to migrate through the membrane for 48 hours. Cells from the upper side of the membrane were scraped off using a Q-Tip without disturbing the cells on the other side. Migrated cells were then fixed with 100% methanol for 10 min. Following fixation cells were stained with crystal violet (0.1%) dissolved in 10% (v/v) formaldehyde for 10 min. For getting rid of excess staining inserts were rinsed in distilled water twice and allowed to dry. Migrated cells were visualized under the microscope.

3.1.4 Colony formation assay

Initially, approximately 250,000 cells were seeded in a 6-well plate and transfection was performed. 24 hour post-transfection cells were washed with PBS and the medium was replaced with medium containing antibiotics. 48 hour later cells were trypsinized and approximately 2,000-3,000 cells were seeded into each well of a 6-well plate and allowed to grow for 5-7 days and during this time period they were treated with inhibitors at the respective concentrations. Next, colonies were fixed with 100% methanol for 10 minutes and subsequently stained with 0.1% crystal violet for 20 minutes at RT. Cells were destained with tap water to remove excess stain and allowed to dry at RT. Plates were then scanned with a Scanner Epson V100 photo.

3.2   Molecular biology

3.2.1   RNA isolation

RNA isolation was performed using QIAzol® reagent according to the manufacturer's instructions. Briefly, medium was sucked off and cells were washed twice with PBS and then lysed by adding 500 µl QIAzol® reagent per well. Cells were gently scraped and collected into 1.5 ml tubes. RNA was isolated by chloroform extraction and isopropanol precipitation method. In short, 100 µl chloroform was added to the samples which were then vortexed for approximately 20 min and centrifuged at 10,000g for 20 min at 4 °C. The upper aqueous phase was collected and then samples were precipitated with isopropanol overnight at -20 °C. On the following day samples were centrifuged at maximal speed for 20 min at 4 °C, pellets were washed twice with 70% ethanol, dried and resuspended in 40 µl DEPC-treated water. The RNA concentration was measured using NanoDrop and used for cDNA synthesis or RNA-seq.

3.2.2   cDNA synthesis

For cDNA synthesis 1 µg of total RNA was mixed with 6 µl of master mix containing 2 µl of 15 µM random primers and 4µl of 2.5 mM dNTP mix, volume was brought up to 16 µl with DEPC-treated water and incubated for 5 min at 70 °C. Next, 4 µl of reverse transcription master mix containing 2 µl 10x reaction buffer, 10 units RNAse inhibitor, 25 units of M-MuLV reverse trascriptase and 1.625 µl DEPC-treated water were added to each sample. cDNA synthesis was performed at 42 °C for 1 h followed by enzymatic inactivation at 95 °C for 5 min. Samples were then diluted with DEPC-treated water to a volume of 50 µl.

3.2.3 Quantitative real-time PCR

Quantitative real-time PCR was performed in a reaction volume of 25 µl. For each reaction 1 µl of cDNA or ChIP DNA was mixed with 8.5 µl of ddH$_2$O, 1.5 µl of 5 µM primer mix (reverse and forward) and 14 µl of qRT-PCR mix. qRT-PCR mix was prepared with following the reagents: 75 mM Tris-HCL (pH 8.8), 20 mM (NH$_4$)$_2$SO$_4$, 0.01% Tween-20, 3 mM MgCl$_2$, 200 µM dNTs, 20 U/ml Taq polymerase, 0.25% Triton X-100, 1:80,000 SYBR Green I and 300 mM Trehalose.

The PCR was reaction was performed using the following two-step protocol

95 °C – 2 min

95 °C – 15 sec

60 °C – 1 min   40x

The PCR reaction was followed by a melting curve analysis from 60 °C to 95 °C with read every 0.5 °C.

A standard curve made from all the cDNA samples was used for the quantification. HNRNPK was used as an internal reference gene to normalize all the qRT-PCR samples following statistical analysis. The expression levels were displayed relative to the control sample and expressed as "relative mRNA levels". In case of ChIP qRT-PCR, quantification was performed using a standard curve made from the ChIP input DNA. ChIP samples were normalized to their corresponding input sample and displayed as "% of input".

3.2.4 Chromatin-immunoprecipitation (ChIP)
3.2.4.1 Cross-linking and sonication

For ChIP experiments cells were either grown in 10 cm or 15 cm plates. Cells were cross-linked with 1% formaldehyde in PBS for either 10 min (histone modifications

ChIP) or 15 min (KLF10 and HDAC1 ChIP) at RT. The formaldehyde cross-linking reaction was then quenched by adding 1.25 M glycine for 5 min. After that cells were washed twice with ice-cold PBS. 1 ml of nuclear preparation buffer (lysis buffer) containing proteinse inhibitor cocktail was added on to the cells which were then scraped and collected into a 1.5 ml tubes. Nuclei were then centrifuged at 12,000g for 1 min at 4 °C, the supernatant was discarded and the pellet was gently resuspended in 1 ml nuclear preparation buffer and again centrifuged. The final pellet was resuspended in 200 µl sonication buffer-1 (1% SDS) containing an proteinase inhibitor cocktail and incubated on a rotating wheel for 15 min at 4 °C. Next, the SDS content was diluted to 0.33% by adding 100 µl of sonication buffer-2 (no SDS) to each sample. Samples were then sonicated using Bioruptor® Pico (Diagenode) at high power with 30 sec on/off pulse for 25 or 30 cycles. Sonicated samples were then centrifuged at 12,000 at 8 °C for 10 min and if there was no visible pellet then proceeded with pre-clearing step.

3.2.4.2 Shearing check

To confirm that sonication was efficient and chromatin was sheared properly a shearing check was performed before proceeding with pre-clearing step. Briefly, 10 µl of the sonicated chromatin sample were taken in a separate tube and 100 µl of sonication buffer-1 and 1 µl of Proteinase-K (20mg/ml) was added and incubated overnight in a thermo-shaker at 65 °C (800 rpm). On the following day samples were briefly spin down and DNA isolation was performed by phenol/chloroform/isoamylic alcohol extraction as described above (for visualizing the pellet pink precipitant was used). DNA was dissolved in 15 µl Tris 10 mM (pH 8) containing 100 µg/ml RNAseA and incubated 1 h at 37 °C (700 rpm). DNA was then mixed with the loading dye and run on a 1.5% agarose gel at 100 V. The gel was analyzed on gel documentation and

shearing was considered efficient if most of the fragments were found to be concentrated around 150-300 bp range in all the samples.

3.2.4.3 Pre-clearing and immunoprecipitation

For pre-clearing 100 µl of 50% sepharose bead slurry was added to the samples and incubated for 1 h at 4 °C. Following pre-clearing, samples were centrifuged, supernatant was discarded and the pellet was resuspended in dilution buffer containing a proteinase inhibitor cocktail. The desired number of aliquots was made and samples were either snap frozen in liquid nitrogen and stored at -80 °C or proceeded with the immunoprecipitation step. 10 µl input sample (10% of ChIP extract) for the corresponding ChIP sample was taken in separate tube and snap frozen. For the immunoprecipitation step 100 µl of chromatin extract was diluted up to 500 µl using IP buffer containing proteinse inhibitor cocktail and incubated with the appropriate amount (refer to materials) of the respective antibody overnight on a rotating wheel at 4 °C. Chromatin complexes were then pulled down by adding 30 µl of Protein-A sepharose 50% slurry prepared in IP buffer and incubated for 2 h at 4 °C. Afterwards, samples were centrifuged at 2000g for 2 min at 4 °C following which washing steps with different buffers (ice-cold) were carried out in the following order: thrice with IP buffer, twice with wash buffer, twice with IP buffer and twice with TE buffer. Between the washing steps beads were gently mixed to make sure the beads are in suspension with the buffers. In the subsequent downstream sample processing steps input samples were also included.

3.2.4.4 DNA isolation

Immunprecipitated chromatin complexes were treated with 15 µl RNAse A (10µg) diluted in 10 mM Tris pH 8 (0.2 µg/µl) and incubated for 30 min at 37 °C. Next, any protein bound to DNA was removed by treating the samples with 1 µl of Proteinase-K

(20 mg/ml) and incubated overnight at 65 °C on a thermo-shaker. On next day, samples were centrifuged at 2,000g for 2 min at RT and the supernatant was transferred to a fresh tube. DNA was precipitated by adding 10 µl of 8 M LiCl and 4 µl colorless co-precipitant (Bioline). Subsequently, 200 µl of premixed phenol/chloroform/isoamylic alcohol (25:24:1) was added to each sample and vortexed for 30 sec and then centrifuged at maximal speed for 2 min at RT. The aqueous phase was collected in a fresh tube and back extraction was performed by adding 200 µL Tris pH 8 10 mM + 0.4 M LiCl and vortexed for 30 sec. Samples were centrifuged at maximal speed and again the aqueous phase was collected and pooled with the first one. Precipitation was performed by incubation with 1 ml 100% ethanol overnight at -80 °C. On the following day samples were centrifuged at 15,000g for 30 min at 4 °C. Pellets were washed with 70% ethanol and again centrifuged at 15,000g for 5 min. All the ethanol was removed and pellets were dried and dissolved in 40 µl of DEPC-treated water. 5 µl ChIP DNA was aliquot into a separate tube and diluted for analysis by quantitative real-time PCR to check for the efficiency of ChIP. ChIP with non-specific IgG antibody was performed to subtract the background binding in the main samples. ChIP samples were normalized to input DNA ad represented as "% of Input".

3.3   Protein biochemistry

3.3.1   Sample preparation and SDS-PAGE

For extracting the protein, the cells were washed with 1x PBS and scraped in ice-cold RIPA buffer containing proteinase inhibitor cocktail (1 mM Pefabloc, 1 ng/µl Aprotinin/Leupeptin, 10 mM BGP and 1 mM NEM). To shear the genomic DNA samples were sonicated for 10 cycles at 30 sec on/off pulse using a Bioruptor® Pico (Diagenode) at high power. Subsequently, cell lysates were diluted with Laemmli

buffer to 1x and boiled at 95 °C for 10 min. Protein samples of interest were separated using denaturing agent SDS (sodium dodesylsulfate) in a polyacrylamide gel upon electrophoresis (SDS-PAGE). Composition of resolving and stacking gels used are described in the Materials section. Polyacrylamide gels were run in SDS running buffer at 20 mA/gel.

3.3.2 Western blot analysis

After separating the proteins by SDS-PAGE they were detected by western blot analysis (Towbin et al., 1979) using specific antibodies against protein of interest. Separated proteins were then transferred to a nitrocellulose membrane at 100 V for 2 h depending on the size of the protein following which the membranes were incubated with blocking solution (5% skimmed milk in TBS-T) for 1 h to prevent non-specific antibody binding. After that the membranes were incubated with the respective antibodies prepared in blocking solution (antibody dilution is described in Materials section) overnight at 4 °C. On the next day, membranes were washed three times with TBS-T buffer and incubated with horseradish peroxidase-conjugated secondary antibodies in blocking solution for 1 h at RT. Membranes were again washed three times with TBS-T buffer and HRP signals were detected using enhanced chemo luminescence and either exposed to X-ray films or western blot imager (Biorad).

3.3.3 Immunofluorescence

Cells were grown onto glass coverslip in 24-well plate. Cells were gently washed with 1x PBS twice and fixed with 4% paraformaldehyde for 10 minutes and then washed with 1x PBS. Permealization was achieved using 0.1% Triton X-100 (prepared in PBS) for 10 minutes following which cells were washed twice with 1x PBS. Cells were blocked in 10% FBS in 1x PBS for 20 minutes (to avoid non-specific antibody binding) and then incubated with the primary antibody (diluted in blocking solution) overnight

at 4°C. On the next day cells were washed three times with 1x PBS and then incubated with the respective Alexa-488 (A11008) or Alexa-594 (A11005) conjugated secondary antibody (Invitrogen) for 1 h at RT. Cells were washed twice with 1x PBS and then incubated with DAPI (Sigma, D9542) diluted 1:15,000 in 1x PBS for 5 minutes at room temperature. The coverslip was then mounted on to the glass slide using mounting medium (Dako, S3023). Images were captured using the AXIO Scope.A1 microscope from Zeiss and data was analyzed using the ZEN 2 lite software.

3.4 Next generation sequencing

3.4.1 Chromatin immunoprecipitation-sequencing (ChIP-seq)

After checking the efficiency of chromatin immunoprecipitation by quantitative real-time PCR isolated DNA was used for sequencing. To ensure that amount of DNA is high enough for the purpose of library preparation DNA concentration was measured using a Qubit dsDNS HS assay on a Qubit® 2.0 Fluorometer. ChIP sequencing was performed in triplicates.

3.4.1.1 Library preparation

2-10 ng DNA were used for the library preparation. DNA samples were first resonicated to get fragments of 200-300 bp size using Bioruptor® Pico (Diagenode). Fragmented DNA was then used to prepare the library with NEBNext Ultra DNA library preparation kit (New England Biolabs) (for KLF10 ChIP) or MicroPlex™ Library Preparation Kit v2 from Diagenode® (for the all other ChIPs) according to the manufacturer's protocol. Briefly, end prep reaction was performed by adding end prep enzyme mix and end prep reaction buffer (total reaction volume 65 µl) to each sample and completing the process on a thermal cycler programmed for 20 °C for 30 min and 65 °C for 30 min. Afterwards, adaptor ligation was performed where the following components were mixed to the samples: Blunt/TA ligase master mix (15 µl), NEBNext

adaptor for Illumina (2.5 µl) and ligation enhancer (1 µl) in a final volume of 83.5 µl and samples were incubated in a thermal cycler for 15 min at 20 °C. Upon completion 3 µl of USER enzyme were added and samples were incubated for 15 min at 37 °C. Size selection of the adaptor ligated DNA was performed using 0.9x AMPure XP beads on a magnetic stand with two washing steps with 80% alcohol. Beads were allowed to dry and then resuspended in 28 µl of 10 mM Tris pH 8.0. Subsequently, PCR amplification was performed with 23 µl of the DNA during which specific Index primers (barcode) were used for each DNA library to allow that after sequencing samples can be separated. Afterwards, PCR amplified DNA was cleaned using AMPure XP beads and the DNA was resuspended in 33 µl 10 mM Tris pH 8.0.

Each ChIP DNA library concentration was measured using Qubit® 2.0 Fluorometer and fragment sizes of the libraries were analyzed using Agilent Bioanalyzer 2100 (High Sensitivity DNA assay). 12 ChIP DNA libraries were pooled together to a final concentration of 10 nM which was later diluted to 2 nM. cDNA libraries were then sequenced using HiSeq 2500 (Illumina) sequencer at the Transcriptome Analysis Laboratory, Göttingen.

3.4.2 RNA sequencing

RNA sequencing was performed in either duplicates or triplicates. Before proceeding with the library preparation RNA integrity was checked on an agarose gel and by visualizing the bands under gel documentation. Once the RNA integrity was confirmed the samples were used for library preparation.

3.4.2.1 Library preparation

RNA library preparation was performed using the NEBNext® Ultra™Library Prep Kit according to the manufacturer's instructions. Briefly, 500 ng to 1 µg of total RNA were used for the further downstream processing. From the total RNA, mRNA enrichment

was performed using a polyadenylayed magnetic beads following which RNA was fragmented by incubating the samples at 94 °C for 15 min in a thermal cycler. Subsequently, fragmented and primed mRNA was used for the first and second strand cDNA synthesis. Double-stranded cDNA was then purified using Agencourt AMPure XP beads and afterwards end repair reaction (single base overhangs at the 5' end) which was immediately followed by adaptor ligation (sequencing adaptor) and purification steps. Purified adaptor ligated cDNA was then PCR amplified where one universal common primer was used for all the samples while for the purpose of barcoding a specific index primer was used for each sample that later allows the separation of individual sample loaded on the same lane of the sequencer.

3.5 Bioinformatic analysis of ChIP and RNA sequencing data

3.5.1 Analysis of ChIP-sequencing data

Raw ChIP sequencing data was processed and analyzed using tools available on publicly accessible servers (Galaxy, Galaxy/Cistrome and Galaxy/deepTools). Before proceeding with the downstream processing of the data all the files were checked for their quality by running FASTQ quality check (FastQC) command on Galaxy (S. Andrews Babraham Institute) to ensure that the quality of the raw data was good enough to be processed further.

3.5.2 Mapping and Peak Calling

After performing the quality check on the raw data the FASTQ files were used to align the sequenced reads to the human genome. For this purpose the Bowtie2 function on Galaxy was used and the reads were aligned to human reference genome (*hg19*, GRCh37) (Langmead and Salzberg, 2012). Final aligned files from the Bowtie2 tool were in SAM format (Sequence Alignment/Map) which were converted to BAM format (Binary Alignment/Map) using the SAMtools package on Galaxy (Li et al., 2009).

The BAM files (BAM files from the triplicate ChIP sample were merged) containing the mapped and aligned reads were used for the purpose of peak calling using Model-based Analysis of ChIP-seq 2 (MACS2) tool available on Galaxy (Zhang et al., 2008). Using the signal from the input sample as the background MACS2 provides the significantly enriched genomic regions in the ChIP sample calling them as a peak. Minimum FDR (q-value) cutoff for peak detection was set to 0.05 and.

3.5.3 Normalization and visualization of ChIP sequencing data

The BAM files were used for the purpose of normalization using the bamCoverage tool available on the public server Galaxy/deepTools (Ramírez et al., 2014). The tool divides the genome into bins of specific size (default setting '50' was used) and then it calculates the number of reads that overlaps with each of the bins in the whole genome. The total number of mapped reads was then normalized using the reads per kilobase per million (RPKM) option. Recently ENCODE Project Consortium has listed specific genomic loci which were described as artificial high signal regions which were excluded from the purpose of normalization. The bamCoverage output file 'bigwig' was then used to visualize the binding intensities at individual genomic regions using Integrative Genomics Viewer software (Robinson et al., 2011). To determine the enrichment at various genomic locations, Cis-regulatory Annotation System (CEAS) was used which gives the relative enrichment of the ChIP regions at specific genomic locations compared to the whole genome (Shin et al., 2009). Furthermore, bigwig files were used to generate heatmaps and aggregate profile plots over defined genomic regions (deepTools and Cistrome).

3.5.4 Differential Binding analysis

Differential binding analysis was performed for the H3K4me1 and H3K27ac binding sites using DiffBind (Ross-Innes et al., 2012) package on R. Genomic coordinate file

for all the genes was obtained using UCSC genome browser (Karolchik et al., 2004). Significantly enriched H3K4me1 bound regions were used to perform GREAT (Genomic Region Enrichment of Annotations Tool) analysis (McLean et al., 2010) to find the associated genes. ReMap was performed on the genomic regions from GREAT to find the enriched transcription factors at the given genomic regions (Griffon et al., 2014). Motif analysis was preformed using SeqPos motif tool on Cistrome/Galaxy (He et al., 2010).

3.5.5 Analysis of RNA-sequencing data

RNA-seq was performed in triplicates for all the experiments (duplicate for A549 cells). The quality of the raw data (FASTQ) was checked using the FastQC tool on Galaxy. Fastq files were mapped to the human genome (hg19) using the TopHat tool on Galaxy (settings were set to 'very sensitive') (Kim et al., 2013). BAM files were then coordinate sorted using SortSam (version 1.126.0) from Picard tools on Galaxy. Subsequent files were then used for read counting using the HTSeq tool (version 0.6.0) (Anders et al., 2015) and then htseqcount files were used for measuring differential gene expression using DESeq2 package on R (Bioconductor version 3.2) (Love et al., 2014). Gene Ontology (GO) analysis was performed using the DAVID (Database for Annotation, Visualization and Integrated Discovery) software (Huang et al., 2009). Significantly enriched GO categories were selected based on the FDR value ≤ 0.05. Gene Set Enrichment Analysis (GSEA) was performed with standard parameters (1000 permutations of gene sets, Signal2Noise ranking metric) and significantly enriched pathways (c5.all gene sets) were selected (Subramanian et al., 2005). Molecular Signatures Database (MSigDB) was used to compute the overlap between available gene set and the gene set from our RNA-seq data.

3.6 Xenograft study

For each animal, one million tumor cells were resuspended in 20 µl of a 1:1 mixture of DMEM medium and BD Matrigel Matrix High Concentration (HC), Growth Factor Reduced (GFR) (BD Bioscience) and kept on ice until transplantation. 8 to 16 weeks old virgin NMRI foxn1nu/nu mice (Janvier Labs) were anesthetized by Isofluran inhalation-narcosis (2-3 %, Forene). The cell suspensions were injected with a 0.3 ml Micro-Fine syringe (BD Bioscience) into left abdominal flank. The operation was performed under sterile conditions. After the mice has developed tumor, they were randomly divided into two groups (n=12 per group): control and treated. Mice were treated with either vehicle (methylcellulose) or 4SC-202 for days (twice per day). Mouse weight and size of growing tumors were measured every day.

4. Results

1. KLF10 and epithelial-to-mesenchymal transition

In this study the role of KLF10 in TGFβ-induced epithelial-to-mesenchymal transition (EMT) was investigated. The main focus of the study was to uncover the molecular mechanism through which KLF10 acts as a tumor suppressor and whether it can play a role in metastasis. Previous studies have well documented the role of KLF10 in cancer and also as a central regulator of TGFβ signaling, which is frequently perturbed in different types of cancer.

TGFβ signaling is one of the most important and potent drivers of cancer metastasis, which is the leading cause of cancer-related death. It was therefore, the goal of the study to find out if KLF10, a central regulator of TGFβ signaling, can block its pro-metastatic effects.

4.1 KLF10 as a tumor suppressor

4.1.1 KLF10 expression is downregulated in lung and breast cancer

In order to elucidate the role of KLF10 as a tumor suppressor it was important to check if KLF10 expression is perturbed in cancer compared to the normal tissues. For this purpose the publicly available Oncomine database was utilized. The Oncomine database contains a large collection of gene expression datasets from different types of cancer, allowing the users to check the expression of any particular gene in a wide range of independent datasets from different cancer types. Oncomine gene expression datasets for lung adenocarcinoma and breast carcinoma were utilized to check for the expression levels of KLF10. Consistent with a potential tumor suppressor

function, KLF10 gene expression was significantly downregulated in lung adenocarcinoma compared to normal lung (Fig. 15).
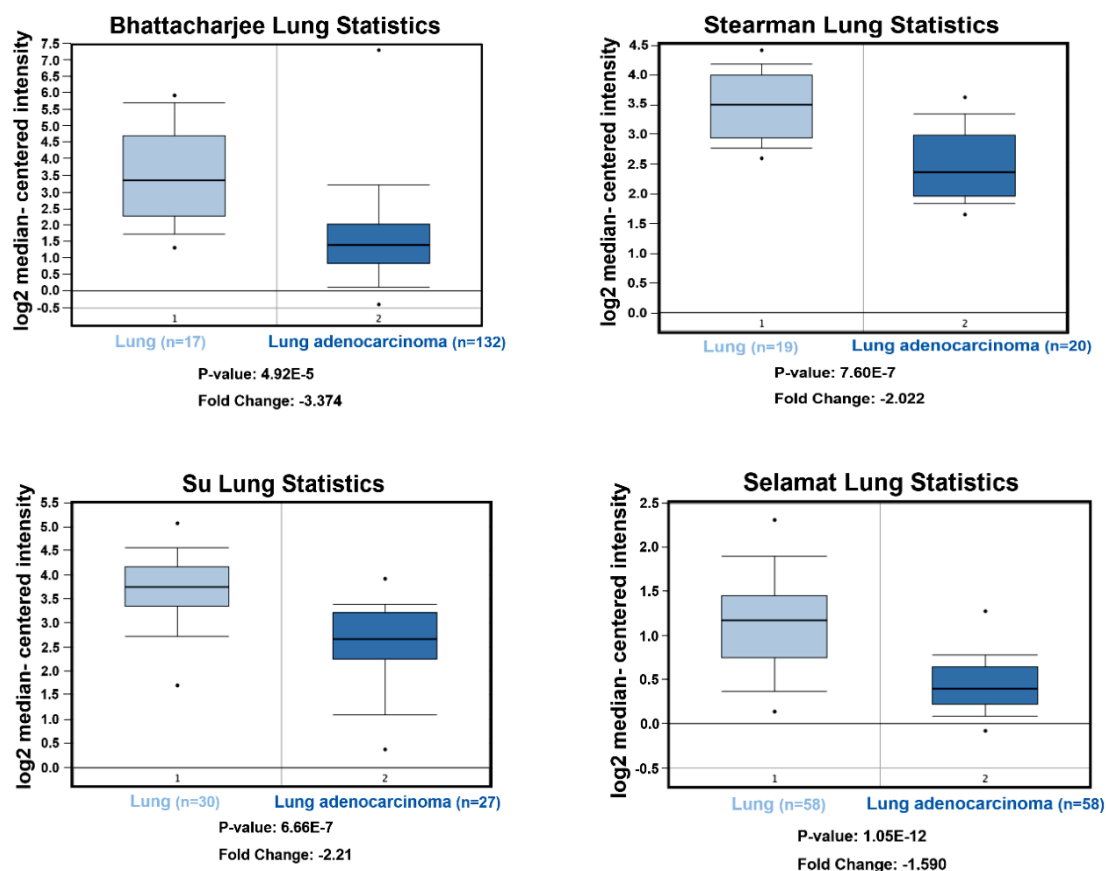
Figure 15: KLF10 expression in lung adenocarcinoma. KLF10 expression was found to be significantly downregulated in lung adenocarcinoma compared to the normal lung samples as shown in four independent datasets. Data was generated using publicly available Oncomine database.

Furthermore, given previous data indicating decreased expression of KLF10 in breast cancer (Reinholz et al., 2004; Subramaniam et al., 1998), we also examined the available datasets for breast carcinoma and found that KLF10 expression was significantly downregulated in invasive breast carcinoma compared to the normal breast (Fig. 16). Low KLF10 expression in invasive carcinoma supports the tumor suppressor role of KLF10 as well as a possible role in regulating metastasis. Previously it was shown that KLF10 expression in breast cancer is stage dependent highest being in normal breast and almost complete loss of expression in metastatic breast cancer (Subramaniam et al., 1998).
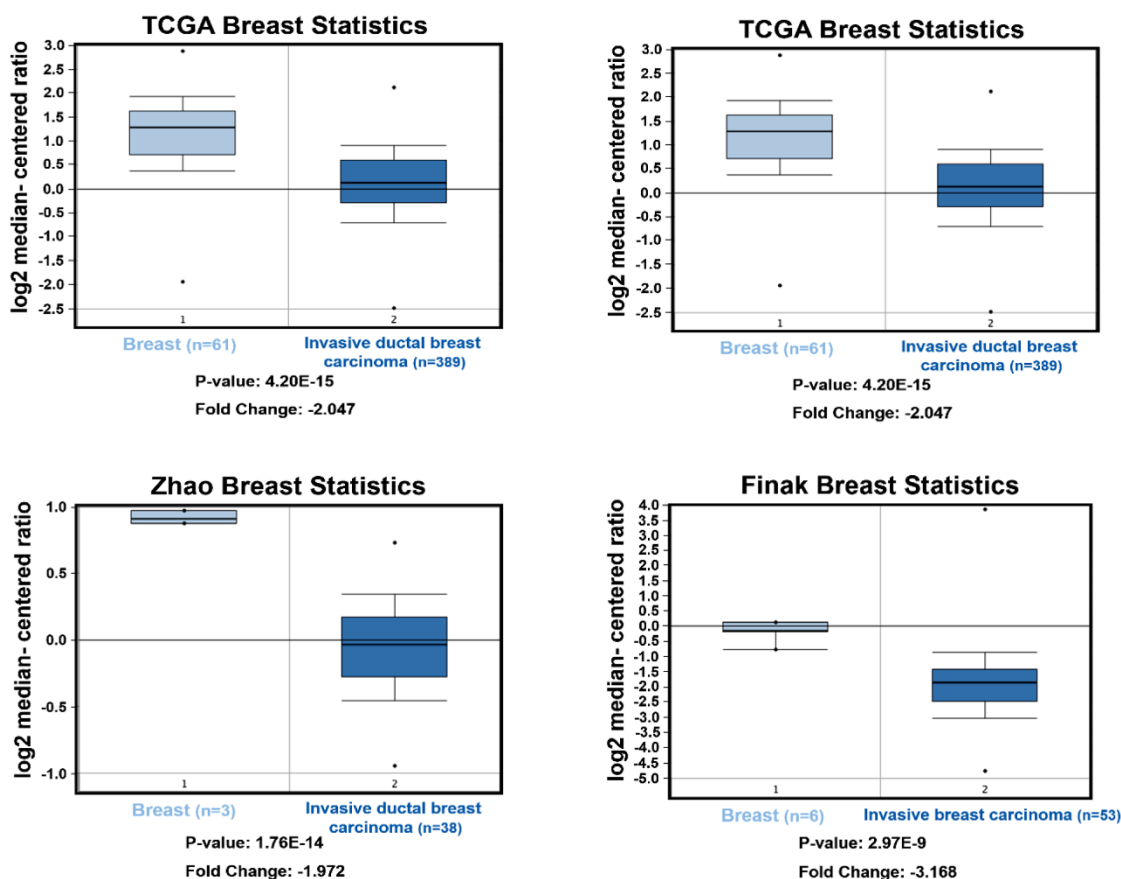
Figure 16: KLF10 expression in breast carcinoma. KLF10 expression was found to be significantly downregulated in invasive breast carcinoma compared to the normal breast samples as shown in four independent datasets. Data was generated using publicly available Oncomine database.

4.1.2 KLF10 expression correlates with disease outcome

In line with decreased expression of KLF10 in cancer samples we sought to investigate if KLF10 expression level can be a prognostic marker to predict the disease outcome in cancer patients. For this purpose we utilized an online survival analysis tool called Kaplan-Meier Plotter to determine the prognostic role of KLF10 in lung and breast cancer. We found that lung cancer patients with low KLF10 expression have poor overall survival rate as compared to the patients with high KLF10 expression (Fig. 17A). Furthermore Kaplan-Meier plotter for breast cancer (luminal-B subtype)

revealed that low KLF10 expression was related to poor disease and metastasis free survival (DMFS) compared to high KLF10 expression (Fig. 17B).
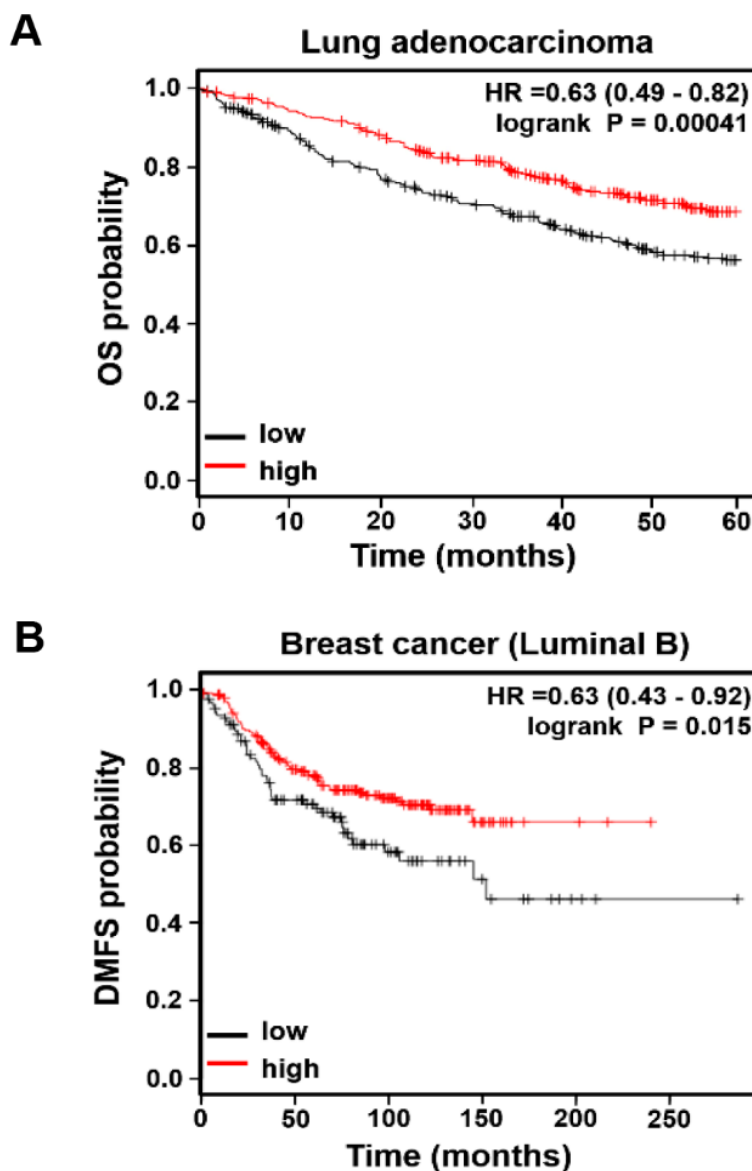
Figure 17: KLF10 expression level can predict disease outcome. (A) Kaplan-Meier plot for lung adenocarcinoma showing that patients with high KLF10 expression have significantly higher overall survival (OS) probability compared to patients with low KLF10 expression. (B) Kaplan-Meier plot for breast cancer (subtype 'luminal B') showing that patients with high KLF10 expression have significantly higher disease and metastasis free survival (DMFS) compared to the patients with low KLF10 expression.

4.2   KLF10 knock-out results in tumor formation *in vivo*

In order to gain further insight into the tumor suppressor role of KLF10 in an *in vivo* model system for lung cancer we carried out a study in *Klf10* knock-out mice. To induce lung tumor formation in *Klf10* knock-out mice and wild type mice we utilized the commonly used laboratory chemical carcinogen DMBA which has been previously reported to induce lung cancer with a high incidence (Duro de Oliveira et al., 2013). 2-5 day old mice were treated with DMBA and then allowed to develop tumors over a period of four months following which they were they were analyzed for lung tumor incidence and tumor size. Importantly neither wild type nor the *Klf10* knock-out mice had any other genetic alterations such as *KRAS or EGFR.*
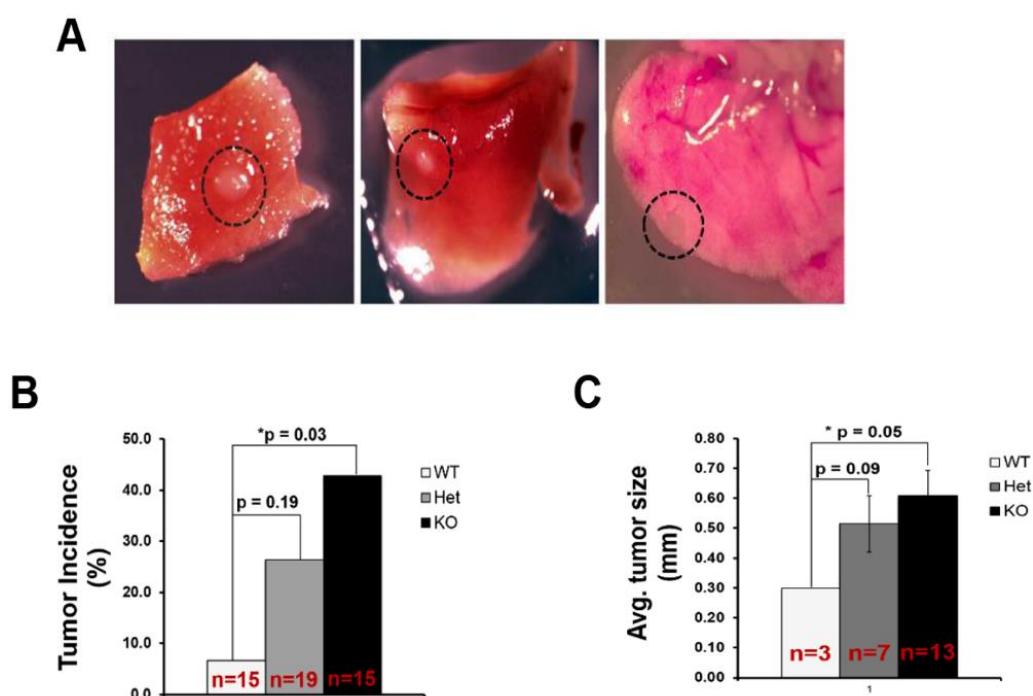
Figure 18: KLF10 mutation results in tumor formation. (A) Microscopic images of cancerous lesions in KLF10 knock-out mice lung. (B) Increased tumor incidence in KLF10 knock-out mice compared to the wild type mice, p-value = 0.03. (C) Increased tumor size in KLF10 knock-out mice compared to wild type mice, p-value = 0.05. (WT= wild type, Het= heterozygous, KO= knock-out).

After four months of DMBA treatment, wild type and *Klf10* knock-out mice were analyzed for tumor formation (Fig. 18A). Interestingly, *Klf10* knock-out mice had more lung tumors compared to the wild type mice implying that mutation in KLF10 resulted in significantly increased tumor incidence (Fig. 18B). Furthermore, we also observed significantly higher tumor size in *Klf10* knock-out mice compared to the wild type mice (Fig. 18C). These results from our *in vivo* study not only establish a tumor suppressor role for KLF10 in lung cancer but also strengthen the notion that loss of KLF10 expression can promote tumor formation.

4.3 KLF10 and TGFβ signaling

Our data supports that KLF10 acts as a tumor suppressor and low KLF10 expression is associated with poor overall survival rate in lung cancer patients. Previous data suggests that KLF10 is an important regulator of TGFβ signaling (Johnsen et al., 2002a) and overexpression of KLF10 has been shown to mimic the anti-proliferative function of TGFβ (Johnsen et al., 2004). Previous studies have well documented the pro-metastatic effects of TGFβ in late stages of cancer (Dalal et al., 1993; Friedman et al., 1995; Gorsch et al., 1992; Huang et al., 2003). One of the most important goals of this thesis work was to find out if KLF10, being a central player in TGFβ signaling, can fine tune the balance between anti-proliferative and pro-metastatic functions of TGFβ signaling. To investigate this we performed a transcriptome-wide study to determine the effect of KLF10 on TGFβ regulated genes.

4.3.1 The majority of TGFβ regulated genes are affected by KLF10

In order to study the transcriptome-wide effects of KLF10 on TGFβ regulated genes we performed high throuput RNA sequencing. Since KLF10 expression was perturbed in lung adenocarcinoma patients we chose A549 (lung adenocarcinoma) cell line. In parallel we also performed the same study in another cell line, Panc1

(pancreatic epithelial cancer cell line) where overexpression of KLF10 in these cells has been shown to induce TGFβ like anti-proliferative effects and apoptosis. Furthermore, A549 and Panc1 cell lines are TGFβ responsive. Initially cells were transfected with either non-targeting control siRNA or siRNA targeting KLF10 and 24 hours post-transfection cells were either treated with TGFβ or vehicle. Total RNA was isolated 72 hours post treatment and was used to perform mRNA sequencing. Significantly (padj ≤ 0.05) up (log2fc ≥ 1.5) or down (log2fc ≤ -1.5) regulated genes were selected for further analysis. In the heatmap, "TGFβ vs Control" represents significantly regulated genes in TGFβ-treated cells compared to the vehicle (control) treated cells and "siKLF10+TGFβ vs TGFβ" represents significantly regulated genes in KLF10-depleted TGFβ-treated cells compared to the cells treated with TGFβ alone. Remarkably a significant fraction of TGFβ-regulated genes was affected by KLF10 depletion both in A549 as well as Panc1 cells (Fig. 19). However, there was a certain number of genes that remained unaffected by KLF10 depletion, implying that these genes do not require KLF10 for their regulation. Interestingly some TGFβ-upregulated genes were further upregulated, whereas some downregulated genes were further downregulated following KLF10 depletion whereas a number of other genes were differently regulated by KLF10 depletion. Strikingly, we obtained similar results from transcriptome-wide studies in two different cancer cell lines which could mean that the effect of KLF10 on TGFβ signaling is a general phenomenon in different cancer types.
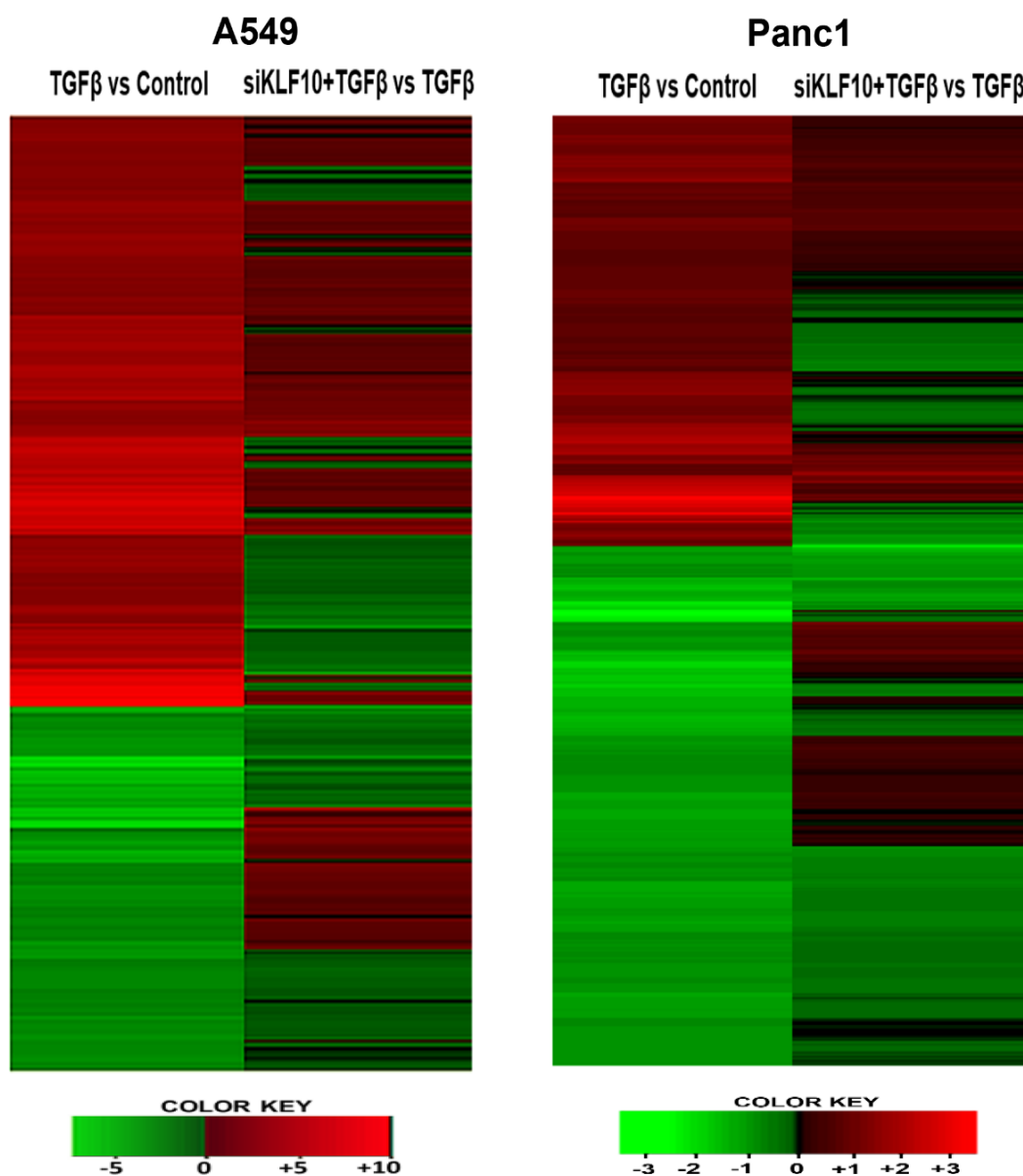
Figure 19: Transcriptome wide effects of KLF10 on TGFβ regulated/targeted genes. Heatmap from RNA-seq data in A549 and Panc1 cells showing significant number of TGFβ regulated/targeted genes were affected by KLF10. A549 and Panc1 cells were transfected with non-targeting control siRNA or siRNA targeting KLF10. 24 hours post-transfection cells were treated with or without TGFβ for 72 hours following which RNA was harvested and used for high-throughput RNA sequencing. Heat map was generated using statistically significant up and down regulated genes in the TGFβ vs control condition (padj-value ≤ 0.05, cutoff of ± 1.5 log2fold change, red color represents upregulated and green color represents downregulated genes).

4.3.2 KLF10 affects pathways related to EMT and metastasis

In order to investigate the biological relevance of the TGFβ regulated genes which were affected by KLF10, we selected the significantly regulated genes from 'TGFβ vs Control' and 'siKLF10+TGFβ vs TGFβ' conditions with a cutoff of log2fc ±1 and padj value ≤0.05. Initially, we identified the genes which were overlapping between two conditions (TGFβ affected or following KLF10 depletion) for which we made a Venn diagram using the list of significantly regulated genes. As shown in the Venn diagram (Fig. 20A & B), in A549 cells 448 (upregulated) and 206 (downregulated) genes whereas in Panc1 cells 114 (upregulated) and 147 (downregulated) genes were found to be overlapping. Further we pooled the overlapping set of up and down regulated genes and performed gene ontology (GO) analysis using DAVID (Database for Annotation, Visualization and Integrated Discovery) online tool. Interestingly, all the top enriched pathways were related to cell adhesion, extracellular matrix, cell motility and cell migration in both A549 and Panc1 cells (Fig. 20C & D). One of the important characteristics of differentiated cells is that they adhere to the extracellular matrix and to neighboring cells. However, during metastasis cells lose their cell-cell contact and overcome the extracellular matrix barrier enabling migration to distant sites. As gene ontology analysis revealed that several pathways related to EMT and metastasis were enriched following KLF10 depletion, we hypothesized that KLF10 can inhibit cancer cells from entering into EMT and thereby suppress metastasis. The rationale behind this hypothesis was that KLF10 differentially regulates subsets of genes controlled by TGFβ signaling. Notably TGFβ is one of the best characterized growth factors that induces EMT in cancer cells (Xu et al., 2009).

Figure 20: KLF10 regulates pathways pertaining to EMT. Significantly up or down regulated genes from the RNA-seq data were merged and used to plot Venn diagram for A549 (A) and Panc1 (B) cells. Significant number of genes regulated by TGFβ and KLF10 were found to be overlapping. Overlapping up and down regulated genes were merged to perform DAVID gene ontology analysis. Top enriched pathways in DAVID analysis were related to EMT in both A549 (C) as well as Panc1 (D) cells.

4.3.3 KLF10 regulates EMT and metastasis signature

To gain further insight into the transcriptome-wide effects of KLF10 we investigated the pathways affected following KLF10 depletion. GSEA (Gene Set Enrichment Analysis) is a tool that can be used to identify significantly enriched pathways in cancer that are predefined for a particular set of genes. For our GSEA analysis we used 'c2.all' curated gene set and analyzed significantly enriched pathways. Interestingly, significantly enriched pathways were related to poorly differentiated metastatic cancer with EMT-like phenotype. GSEA results support the hypothesis that perturbation of KLF10 expression results in an EMT-like phenotype leading to metastatic cancer (Fig. 21A & C).

Patients diagnosed with metastatic cancer often have a poor life expectancy. In our Kaplan-Meier plot analysis we observed that low KLF10 expression is associated with poor survival of lung cancer patients. In support of that finding, GSEA revealed a significant enrichment of pathways associated with poor survival in lung cancer. Further, we used Molecular Signatures Database (MSigDB) in GSEA to look for pathways associated with significantly regulated genes from our transcriptome data. Remarkably various pathways related to EMT and metastasis were significantly enriched. Notably, the gene set KRAS signaling was significantly enriched. KRAS signaling is quite frequently mutated and is one of the most important driver mutations in lung cancer. These types of analyses were also performed for Panc1 transcriptome data which also showed enrichment of several EMT and metastasis related pathways. Again the most striking finding was the enrichment of gene set related to KRAS signaling which is also quite frequently mutated in pancreatic cancer (Fig. 21B & D).

A

A549 GSEA

Enrichment plot:
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP

NES = -2.1375823
FDR = 0

Veh_TGFβ (positively correlated)

Zero cross at 9548

siKLF10+TGFβ (negatively correlated)

Enrichment plot:
SHEDDEN_LUNG_CANCER_POOR_SURVIVAL_A6

NES = -2.1944106
FDR = 0

Veh_TGFβ (positively correlated)

Zero cross at 9548

siKLF10+TGFβ (negatively correlated)

Enrichment plot:
RHODES_UNDIFFERENTIATED_CANCER

NES = -2.3627214
FDR = 0

Veh_TGFβ (positively correlated)

Zero cross at 9548

siKLF10+TGFβ (negatively correlated)

Enrichment plot:
WINNEPENNINCKX_MELANOMA_METASTASIS_UP

NES = -2.1712918
FDR = 0

Veh_TGFβ (positively correlated)

Zero cross at 9548

siKLF10+TGFβ (negatively correlated)

Enrichment plot:
POOLA_INVASIVE_BREAST_CANCER_UP

NES = -1.753971
FDR = 0.014588

Veh_TGFβ (positively correlated)

Zero cross at 9548

siKLF10+TGFβ (negatively correlated)

Enrichment plot:
VANTVEER_BREAST_CANCER_METASTASIS_DN

NES = -1.840753
FDR = 0.00564

Veh_TGFβ (positively correlated)

Zero cross at 9548

siKLF10+TGFβ (negatively correlated)

B

GSEA_MSigDB__Gene Set Name	p-value	FDR q-value
Hallmark_Epithelial_Mesenchymal_Transition	$3.87\ e^{-43}$	$1.93\ e^{-41}$
Hallmark_KRAS_SIGNALING_UP	$1.48\ e^{-23}$	$3.7\ e^{-22}$
Hallmark_HYPOXIA	$3.24\ e^{-17}$	$2.4\ e^{-16}$
TGFB_UP.V1_UP	$7.85\ e^{-21}$	$1.48\ e^{-19}$
NABA_MATRISOME	$2.74\ e^{-67}$	$1.29\ e^{-63}$

C **Panc1 GSEA**

Enrichment plot:
CHARAFE_BREAST_CANCER_BASAL_VS_
MESENCHYMAL_UP
NES = -1.5014148
FDR = 0
Veh_TGFβ (positively correlated)
Zero cross at 9009
siKLF10+TGFβ (negatively correlated)

Enrichment plot:
SWEET_KRAS_ONCOGENIC_SIGNATURE
NES = -1.4933548
FDR = 0
Veh_TGFβ (positively correlated)
Zero cross at 9009
siKLF10+TGFβ (negatively correlated)

Enrichment plot:
HATTMAN_B_CLL_POOR_SURVIVAL_UP
NES = -1.5723573
FDR = 0
Veh_TGFβ (positively correlated)
Zero cross at 9009
siKLF10+TGFβ (negatively correlated)

Enrichment plot:
AGUIRRE_PANCREATIC_CANCER_COPY_NUM
BER_UP
NES = -1.4992714
FDR = 0
Veh_TGFβ (positively correlated)
Zero cross at 9009
siKLF10+TGFβ (negatively correlated)

Enrichment plot:
VERHAAK_GLIOBLASTOMA_MESENCHYMAL
NES = -1.5892941
FDR = 0
Veh_TGFβ (positively correlated)
Zero cross at 9009
siKLF10+TGFβ (negatively correlated)

Enrichment plot:
RICKMAN_METASTASIS_DN
NES = -1.85016
FDR = 0
Veh_TGFβ (positively correlated)
Zero cross at 9009
siKLF10+TGFβ (negatively correlated)

D

GSEA_MSigDB__Gene Set Name	p-value	FDR q-value
Hallmark_Epithelial_Mesenchymal_Transition	$2.21\ e^{-38}$	$1.11\ e^{-36}$
Hallmark_HYPOXIA	$2.21\ e^{-21}$	$7.27\ e^{-20}$
Hallmark_Apoptosis	$2.7\ e^{-12}$	$1.69\ e^{-11}$

Figure 21: KLF10 regulates EMT and metastasis signature. For GSEA TGFβ-treated sample was compared with siKLF10+TGFβ sample. Veh_TGFβ represents sample treated with siControl and TGFβ. GSEA in A549 (A) and Panc1 (C) cells revealed enrichment of gene sets related to poorly differentiated and metastatic cancers in siKLF10+TGFβ-treated sample compared to the siControl+TGFβ-treated sample. Molecular signature database further confirmed the enrichment of EMT like phenotype and driver mutations (B) and (D).

4.4   Genome-wide occupancy of KLF10

Our results strongly suggested that KLF10, can play an important role in blocking metastasis by fine-tuning the balance between anti-proliferative and pro-metastatic functions of TGFβ signaling. This finding was supported by transcriptome data. In order to uncover direct mechanisms of action and identify key target genes of KLF10 we performed genome-wide occupancy studies. For this purpose we performed chromatin-immunoprecipitation followed by deep sequencing (ChIP-seq).

4.4.1   KLF10 is enriched on promoter region

We examined the enrichment of KLF10 binding at specific genomic locations like promoter, coding exons, introns using a tool called Cis-regulatory Annotation System (CEAS). CEAS gives the relative enrichment of ChIP binding regions at specific genomic locations compared to the whole genome. Interestingly, CEAS revealed that as compared to the genome the majority of the KLF10 binding sites (42.6%) were confined to the promoter region of genes in comparison to the whole genome (1.1%) (Fig. 22A & B)). CEAS analysis also revealed that a significant portion (29.8%) of KLF10 binding sites were confined to coding exons whereas a very small percentage of them were confined to introns (12.8%) and distal intergenic regions (4.3%). Additionally we also performed aggregate plot analysis to look for the KLF10 signals around the transcriptional start sites (TSS) of the KLF10 bound genes. Consistent with the results of CEAS analysis, we observed that highest KLF10 signal intensity near the TSS region of the gene whereas very low signal was observed 5Kb upstream or downstream of the TSS (Fig. 22C). Taken together these results corroborate that the transcription factor KLF10 is mostly recruited to the promoter region and coding exons of target genes genome-wide. Furthermore, as compared to the genome, very few KLF10 binding sites were found to be enriched on introns and distal intergenic regions.

A

GENOME

- Promoter (<=1000 bp): 1.1 %
- Promoter (1000−2000 bp): 0.7 %
- Promoter (2000−3000 bp): 0.6 %
- Downstream (<=1000 bp): 0.9 %
- Downstream (1000−2000 bp): 0.7 %
- Downstream (2000−3000 bp): 0.6 %
- 5'UTR: 0.4 %
- 3'UTR: 1.4 %
- Coding exon: 1.9 %
- Intron: 41.6 %
- Distal intergenic: 50.0 %

ChIP

- Promoter (<=1000 bp): 42.6 %
- Promoter (1000−2000 bp): 0.000 %
- Promoter (2000−3000 bp): 2.1 %
- Downstream (<=1000 bp): 2.1 %
- Downstream (1000−2000 bp): 0.000 %
- Downstream (2000−3000 bp): 0.000 %
- 5'UTR: 6.4 %
- 3'UTR: 0.000 %
- Coding exon: 29.8 %
- Intron: 12.8 %
- Distal intergenic: 4.3 %

B

KLF10 binding on promoter region

KLF10 binding on gene

Figure 22: KLF10 binding sites are confined to the promoter region. (A) Pie chart depicting the relative enrichment of KLF10 binding regions at various genomic locations compared to the whole genome. (B) Percentage enrichment of KLF10 binding regions around the promoter region and various locations in the gene compared to the genome. (C) Aggregate plot depicting average KLF10 signal ±5 Kb around the transcriptional start site (TSS).

4.5   KLF10 targets the EMT transcription factor SNAI2

EMT is coordinated by key transcription factors which are induced in response to external stimuli or growth factors like TGFβ. SNAI2 is one of the most important EMT transcription factors (EMT-TF) which has been shown to directly repress E-cadherin gene expression, thus initiating a key step in EMT. In our transcriptome data we found that expression of SNAI2 (but not the other EMT-TFs) was significantly upregulated following KLF10 depletion. Based on this observation we next investigated whether SNAI2 may be a central target of KLF10 in controlling metastasis.

4.5.1   KLF10 depletion significantly enhances SNAI2 expression

To validate our finding from the transcriptome data examined changes in the expression level of SNAI2 mRNA and protein levels by qRT-PCR, western blotting and immunofluorescence. Briefly, A549 cells were transfected with control siRNA or siRNA

targeting KLF10. 24 hours post-transfection cells were treated with TGFβ to examine the immediate gene activation (90 minutes) as well as sustained gene expression changes (72 hours). After treatment both RNA and protein samples were harvested. For comparison we performed similar experiments in Panc1 and MDA-MB-231 cells. In Panc1 cells were only treated for 72 hours and since MDA-MB-231 cells display high basal levels of TGFβ signaling and mesenchymal phenotype we did not treat the cells with TGFβ. Interestingly, qRT-PCR results revealed that there was a significant increase in gene expression of SNAI2 following KLF10 depletion both in A549 as well as Panc1 cells (Fig. 23A & B). Furthermore, at protein level we observed similar results in A549 cells as well as Panc1 cells (Fig. 23C). Strikingly, despite having a strong mesenchymal phenotype knock-down of KLF10 in MDA-MB-231 cells resulted in further elevation of SNAI2 protein levels (Fig. 23C). Furthermore, in immunofluorescence staining we observed a strong nuclear staining of SNAI2 in KLF10-depleted A549 cells treated with TGFβ which further supports the finding at gene expression and protein level (Fig. 23D). Taken together these results revealed that KLF10 depletion led to significant increase of SNAI2 expression implicating that SNAI2 could be an important target gene of KLF10
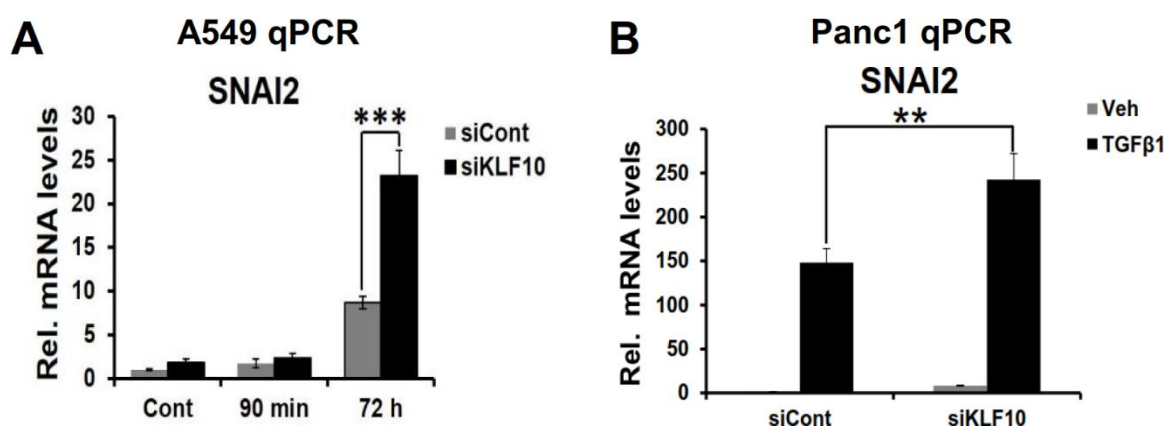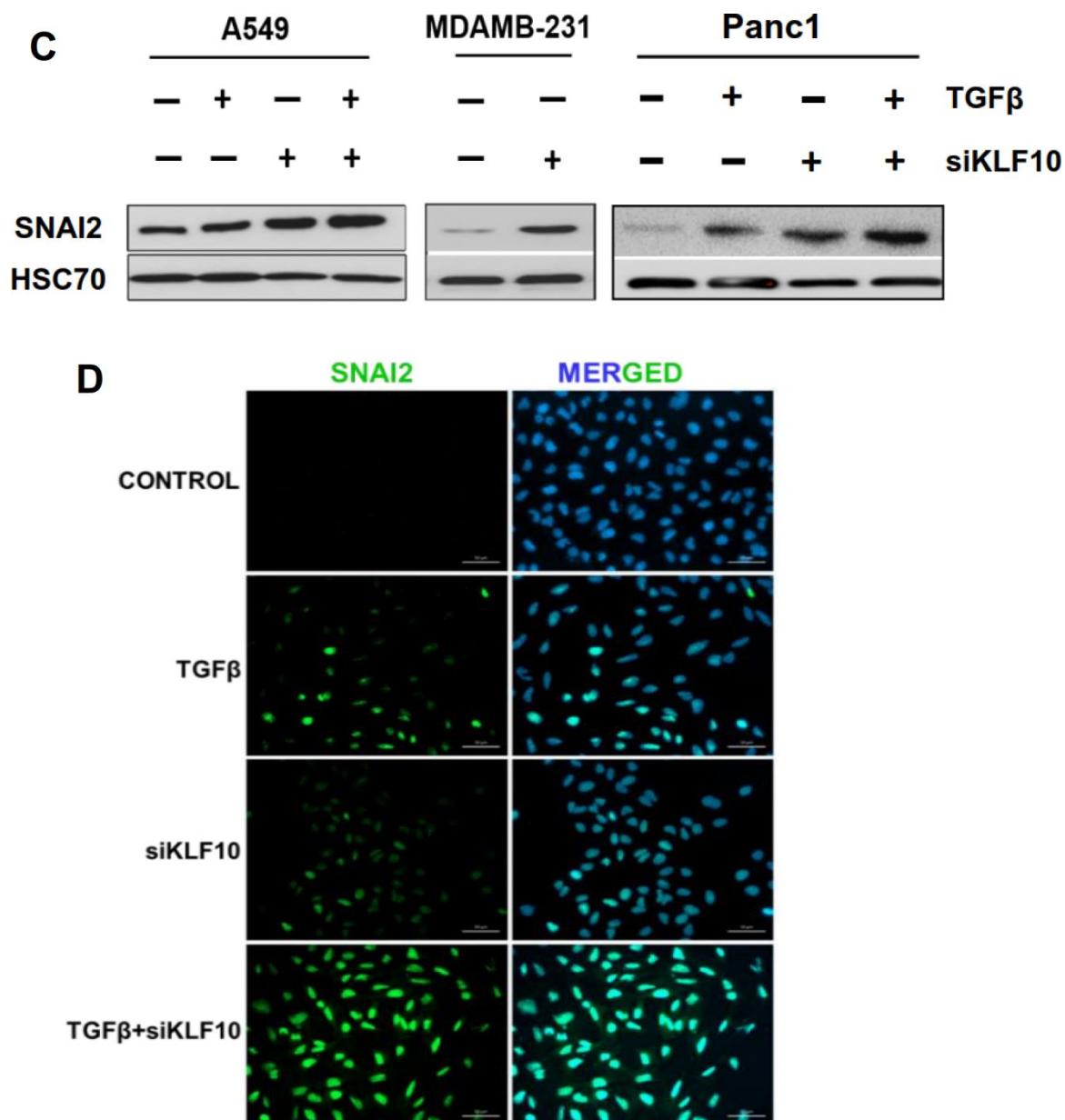
Figure 23: KLF10 depletion resulted in increased SNAI2 expression. (A) & (B) qRT-PCR analysis for SNAI2 in A549 and Panc1 cells respectively. Significantly increased expression of SNAi2 was observed following KLF10 depletion in TGFβ-treated cells compared to the control cells (Veh). Data is represented as mean ± SD. n=3. (C) Western blot for SNAI2 in A549, MDA-MB-231 and Panc1 cells. Protein level of SNAI2 was increased following KLF10 depletion. (D) Immunofluorescence staining for SNAI2 in A549 cells. Strong nuclear staining of SNAI2 was observed in KLF10-depleted cells treated with TGFβ. Nuclei were stained with DAPI. Scale bar represents 50 μm.

4.5.2 KLF10 directly binds to the SNAI2 promoter

After confirming the change in SNAI2 gene expression following perturbation of KLF10 expression we investigated if KLF10 directly binds to SNAI2 gene and represses it. To address this question we went back to KLF10 ChIP-seq data to check if there is an enrichment of KLF10 binding on SNAI2 gene. Remarkably we observed a prominent KLF10 peak around the promoter region of the SNAI2 gene which was consistent with the transcriptome wide data where we observed significantly enhanced expression of SNAI2 upon KLF10 depletion (Fig. 24A). To confirm that KLF10 indeed directly binds to SNAI2 we designed primers across the transcriptional start site (TSS) of the SNAI2 gene where a KLF10 peak was visualized. For a control we also designed primers within the transcribed region (TR) of the gene where KLF10 occupancy was not observed and used it as a negative site. Consistent with the ChIP-seq results, qRT-PCR analysis of ChIP samples showed a significant enrichment of KLF10 binding on the TSS region of the SNAI2 gene, while the TR displayed background levels of KLF10 enrichment (Fig. 24B). Similar ChIP experiments in Panc1 and MDA-MB-231 cells confirmed these results (Fig. 24C & D). Taken together, these results show that: first, we identified SNAi2 as a target gene of KLF10 and second it further strengthened the role of KLF10 in blocking metastasis. Furthermore, similar results obtained in three cell lines representing different epithelial cancer types and strongly support the finding that SNAI2 is a general target gene for KLF10.
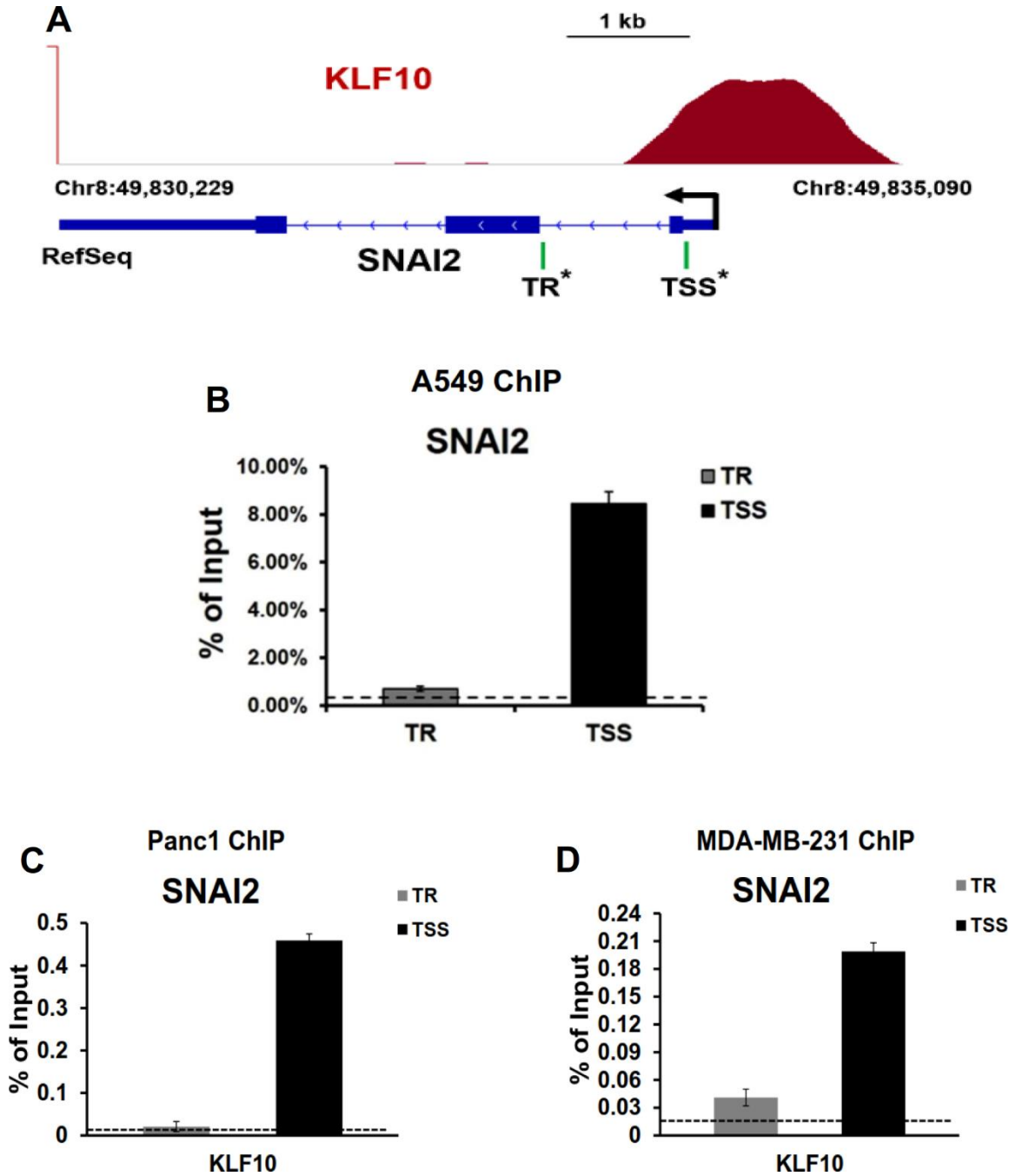
Figure 24: KLF10 directly bind to SNAI2 promoter. (A) ChIP-seq profile of KLF10 on SNAI2 gene. Significant peak was observed around TSS of SNAI2 but no peak was observed within the gene. Transcription direction is indicated by black arrow mark. Scale bar represents 1 kilobase. (B) ChIP analysis of KLF10 occupancy on transcriptional start site (TSS) and transcribed region (TR) of SNAI2 gene in A549 cells. Immunoprecipitated DNA is compared to input and shown as percentage. IgG antibody was used as a negative control to subtract the background level and is shown as black dotted line. Data is represented as mean ± SD. n=3. (C) & (D) ChIP analysis of KLF10 occupancy on TSS and TR of SNAI2 gene in Panc1 and MDA-MB-231 cells respectively.

4.6 KLF10 represses SNAI2 transcription by an epigenetic mechanism

After finding out the target gene for KLF10 we sought to elucidate the mechanism by which KLF10 is repressing SNAI2. Since KLF10 is a transcription factor and does not have an enzymatic activity it cannot repress or activate a gene on its own which means there must be an interacting partner. Previous studies reported that KLF10 can interact and form complex with co-repressors like histone deacetylases (HDACs) (Jin et al., 2012) and demethylases like KDM6B (Kim et al., 2010) and recruit them on to the target gene to cause transcriptional repression.

4.6.1 KLF10 is required for recruitment of HDAC1 to the SNAI2 gene

Based on the previous reports we performed ChIP experiments for HDAC1 to check for its recruitment to the SNAI2 gene promoter and if there is any change in the recruitment upon KLF10 depletion. We used the primers designed on TSS of SNAI2 to check for the enrichment of HDAC1 at the KLF10 binding site. We observed that HDAC1 was recruited on to the SNAI2 TSS in the presence of KLF10. However, there was a significant decrease in the recruitment of HDAC1 after KLF10 depletion implying that HDAC1 is involved in the repression of SNAI2 by KLF10 (Fig. 25A). Knock-down of KLF10 was confirmed by western blotting (Fig. 25B). Interestingly, we observed that even though HDAC1 recruitment was decreased after TGFβ treatment alone, it was significantly stronger decreased after KLF10 depletion, which further supports the potential role of KLF10 in regulating the pro-metastatic function of TGFβ signaling.
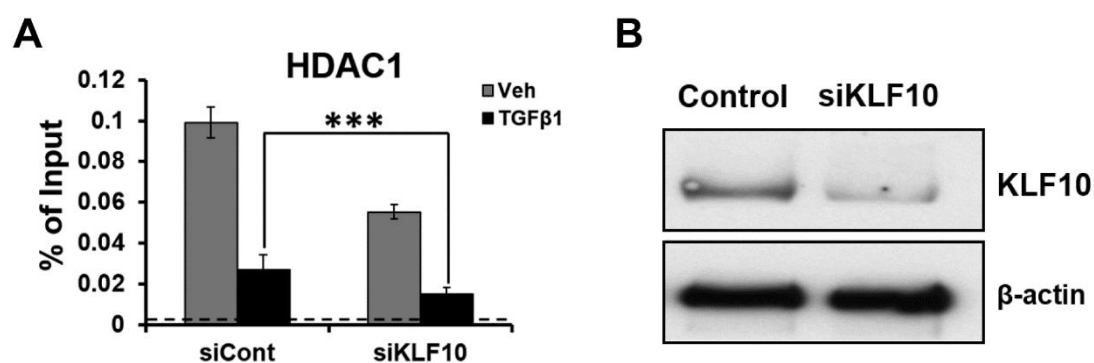
Figure 25: KLF10 form co-repressor complex with HDAC1 to repress SNAI2. (A) ChIP qPCR results showing that HDAC1 is recruited by KLF10 on to the SNAI2 TSS which significantly decreased upon KLF10 depletion. Immunoprecipitated DNA is compared to input and shown as percentage. IgG antibody was used as a negative control to subtract the background level and is shown as black dotted line. Data is represented as mean ± SD. n=3. (B) Western blot results confirming the knock-down of KLF10. β-actin was used as a loading control.

4.6.2 KLF10 depletion leads to enhanced acetylation of the SNAI2 gene

Histone deacetylases are epigenetic "erasers" which remove the activating acetylation marks from histones, thus attenuating gene expression. Therefore we next investigated if KLF10 depletion results in alterations in the acetylation status of the SNAI2 gene. To address this question we performed a ChIP-seq experiment for active acetylation marks H3K9ac and H3K27ac with the same treatment condition as described above for HDAC1 ChIP. Interestingly, ChIP-seq data revealed that SNAI2 gene was already marked with some amount of H3K9ac and H3K27ac marks which could mean that the gene is poised for activation. However, we observed significantly enhanced H3K9ac and H3K27ac around the TSS of the SNAI2 gene following KLF10 depletion compared to the control TGFβ condition (Fig. 26A & C). Furthermore, we also confirmed this finding by performing qRT-PCR from the ChIP samples to check for the enrichment of acetylation marks around TSS region of SNAI2. As expected qRT-PCR results showed a significant increase in H3K9ac and H3K27ac marks at the SNAI2 TSS upon KLF10 depletion (Fig. 26B & D). Taken together these results

confirm cooperative activity between KLF10 and HDAC1 where KLF10 serves to recruit HDAC1 to the SNAI2 promoter, resulting in reduced acetylation of the gene and subsequent gene repression.
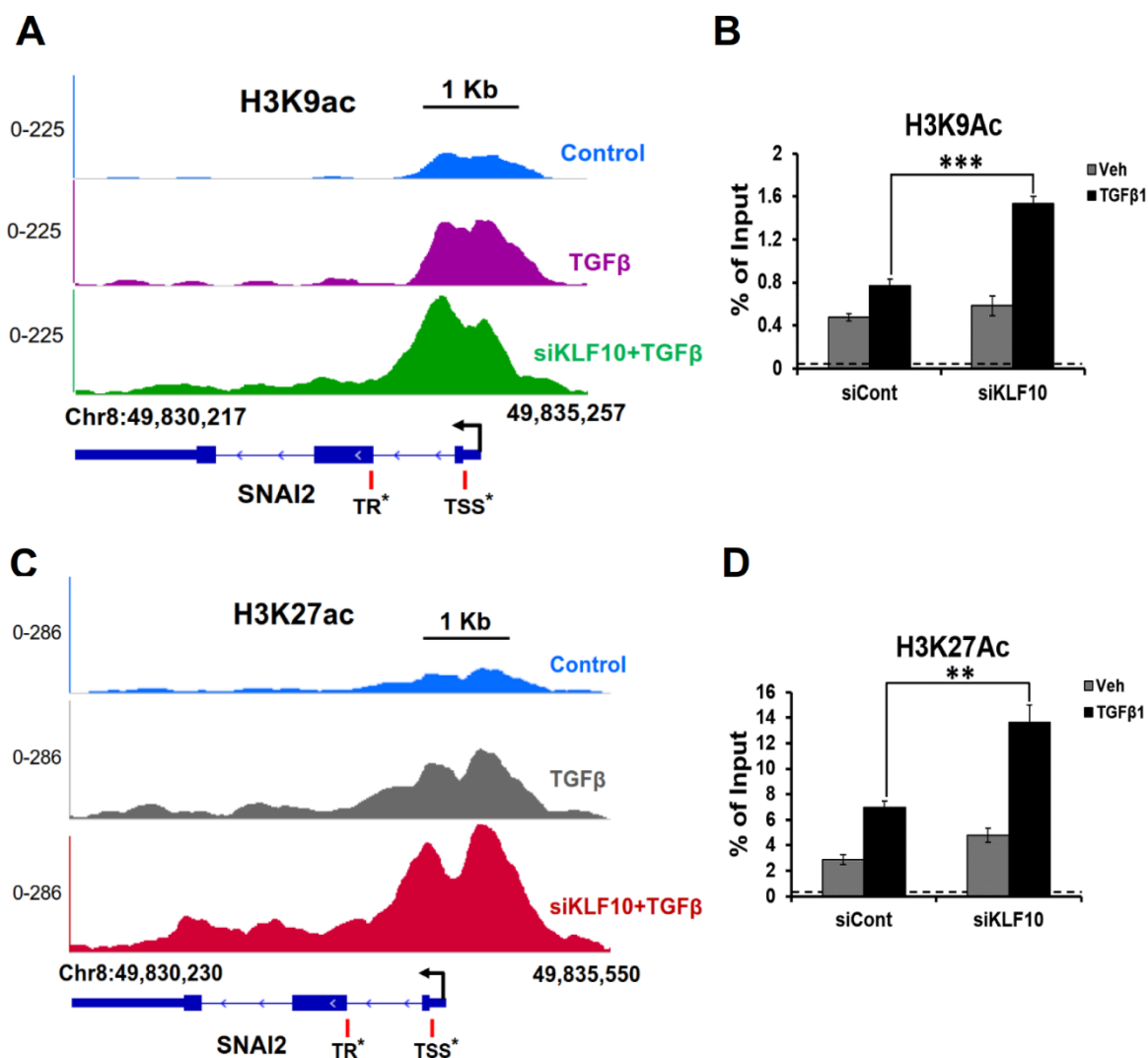
Figure 26: KLF10 depletion results in enhanced acetylation on the SNAI2 gene. (A) & (C) ChIP-seq profile of H3K9ac and H3K27ac respectively in untreated (control), TGFβ-treated (TGFβ) and TGFβ-treated and KLF10 depleted (siKLF10+TGFβ) conditions on SNAI2. Increased acetylation was observed upon KLF10 depletion. Scale bar is represented in kilobase. (C) & (D) ChIP qRT-PCR analysis for H3K9ac and H3K27ac respectively confirming the significantly increased acetylation on SNAI2 upon KLF10 depletion. Immunoprecipitated DNA is compared to input DNA and shown as percentage. IgG antibody was used as a negative control to subtract the background level and is shown as black dotted line. Black arrow indicates the direction of the transcription and red lines mark the sites where primers were designed. Data is represented as mean ± SD. n=3.

4.6.3 KLF10 depletion is associated with enhanced acetylation

To further elaborate the finding that KLF10 promotes a repressed state of the target gene, we examined whether deacetylation generally requires KLF10 at its target genes. To address this question we checked for the average H3K9ac and H3K27ac signals (±5Kb) around the TSS region of KLF10-bound genes genome-wide in an aggregate plot for all the KLF10-bound genes. Interestingly, we observed a similar trend in the H3K9ac and H3K27ac acetylation marks across the KLF0 bound genes. Notably, the highest average signal for H3K9ac was observed upon KLF10 depletion and was even higher than the signal obtained after TGFβ treatment (Fig. 27A). Furthermore, as compared to H3K9ac a slightly different trend was observed for H3K27ac. Average H3K27ac signal for TGFβ condition was higher than the signal obtained after KLF10 depletion. However, the overall highest signal was observed following TGFβ treatment and KLF10 depletion (Fig. 27B). In general, KLF10 depletion leads to an overall increase in the H3K9ac and H3K27ac marks on its target genes. Additionally, these results indicate that deacetylation could be a general mechanism of action of KLF10 to repress target gene expression. In summary, these data suggest that KLF10 acts together with HDAC1 to inhibit the acetylation of the target gene, making the chromatin inaccessible to the transcriptional machinery, thus ultimately leading to inhibition of gene transcription.

A

H3K9ac

Control
TGFβ
siKLF10
siKLF10+TGFβ

Average Signal

-5 Kb   TSS   +5 Kb

Distance from the TSS

B

H3K27ac

Control
TGFβ
siKLF10
siKLF10+TGFβ

Average Signal

-5 Kb   TSS   +5 Kb

Distance from the TSS
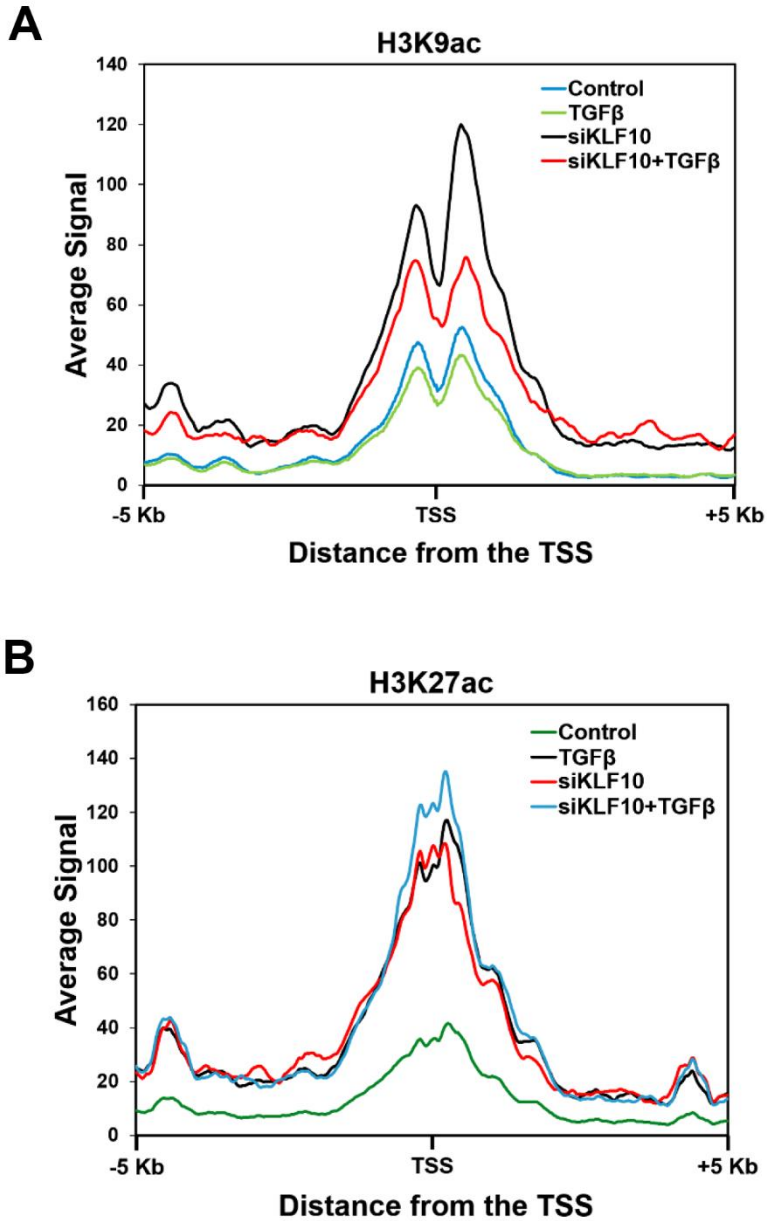
Figure 27: KLF10 depletion is associated with enhanced acetylation marks on its target genes. (A) & (B) Aggregate plot analysis of the ChIP-seq data for H3K9ac and H3K27ac respectively, on the KLF10-bound genes genome-wide. Analysis was performed at ±5 Kb around the TSS of the genes. Overall increased signal for H3K9ac and H3K27ac was observed on the genes upon KLF10 depletion.

4.7 KLF10 regulates EMT and metastasis

Results from the previous experiments pointed towards a potential role of KLF10 in regulating the pro-metastatic function of TGFβ by directly repressing the expression of SNAI2, one of the important EMT-TFs regulating EMT. Based on these findings we examined if perturbation of KLF10 expression indeed has an impact on TGFβ-induced EMT.

4.7.1 KLF10 depletion enhances TGFβ-induced EMT

To test the effect of perturbing KLF10 expression on TGFβ-induced EMT we performed several cell culture based assays in A549 cells. During EMT cells undergo morphological changes, lose cell-cell adhesion and attain an elongated mesenchymal shape. When treated with TGFβ cells depleted for KLF10 became more mesenchymally shaped compared to cells treated with TGFβ alone but having intact KLF10 (Fig. 28A). Initiation of EMT is marked by the downregulation of epithelial markers and upregulation of mesenchymal markers. To test this in our model system we performed gene expression analysis, western blotting and immunofluorescence staining for classical EMT markers. As shown in Fig. 28B, epithelial markers (E-cadherin and MMP7) were significantly downregulated and mesenchymal markers (N-cadherin and MMP2) were significantly upregulated following KLF10 depletion in TGFβ-treated cells compared to the cells treated with TGFβ alone. Additionally, we also observed the same trend at the protein level. As depicted in Fig. 28D, there was a complete loss of E-cadherin expression upon KLF10 depletion. In parallel we also performed similar experiments in Panc1 cells which showed similar results (downregulation of epithelial markers and upregulation of mesenchymal markers) at both gene expression and protein levels (Fig. 28C & E). Moreover, we also performed immunofluorescence staining for epithelial and mesenchymal markers. As shown in

Fig. 28F, epithelial markers ZO-1 and E-cadherin were completely lost whereas the mesenchymal marker Vimentin was significantly upregulated upon KLF10 depletion in TGFβ-treated cells. Taken together these results confirm that KLF10 indeed suppresses the TGFβ-induced EMT.
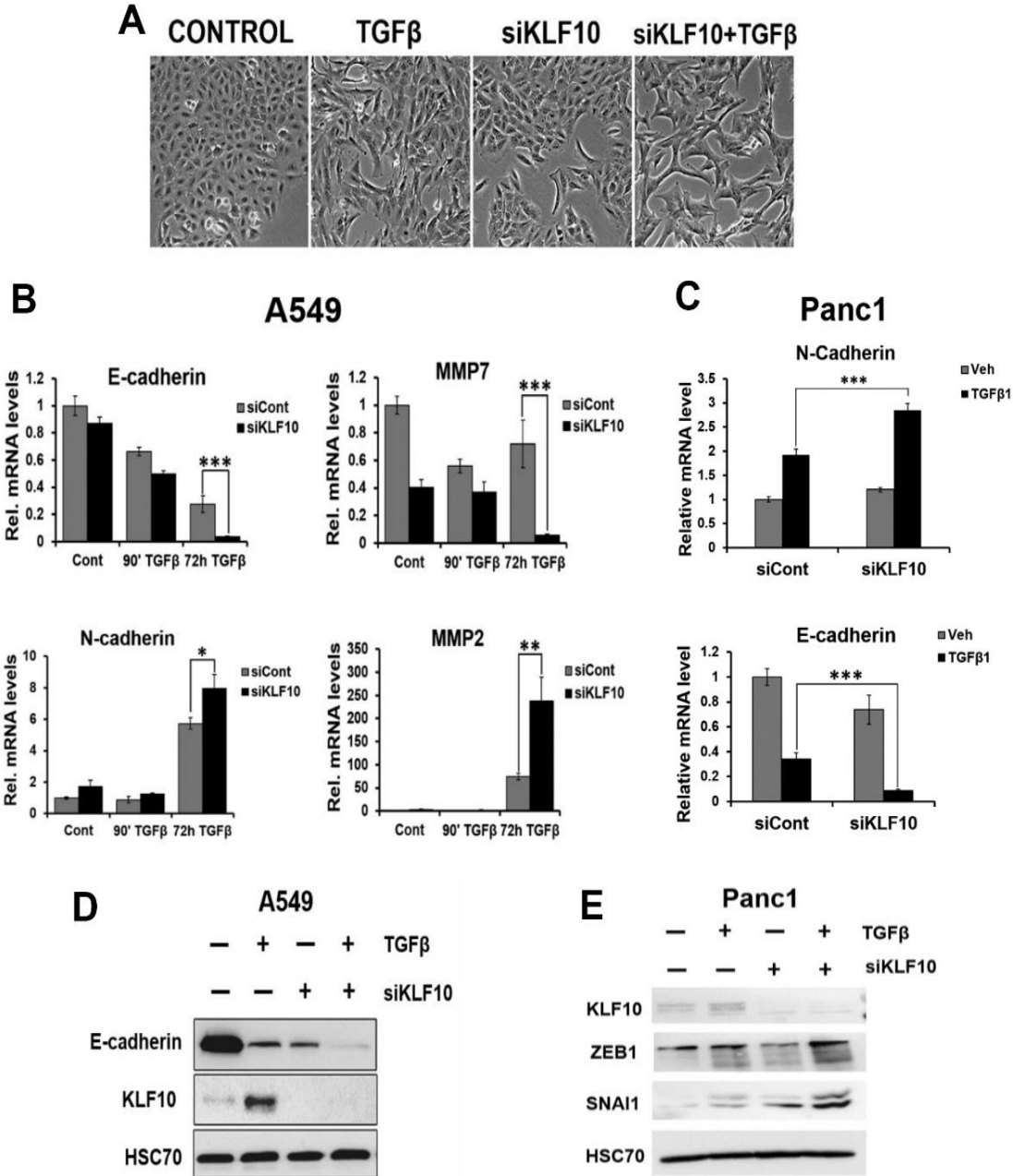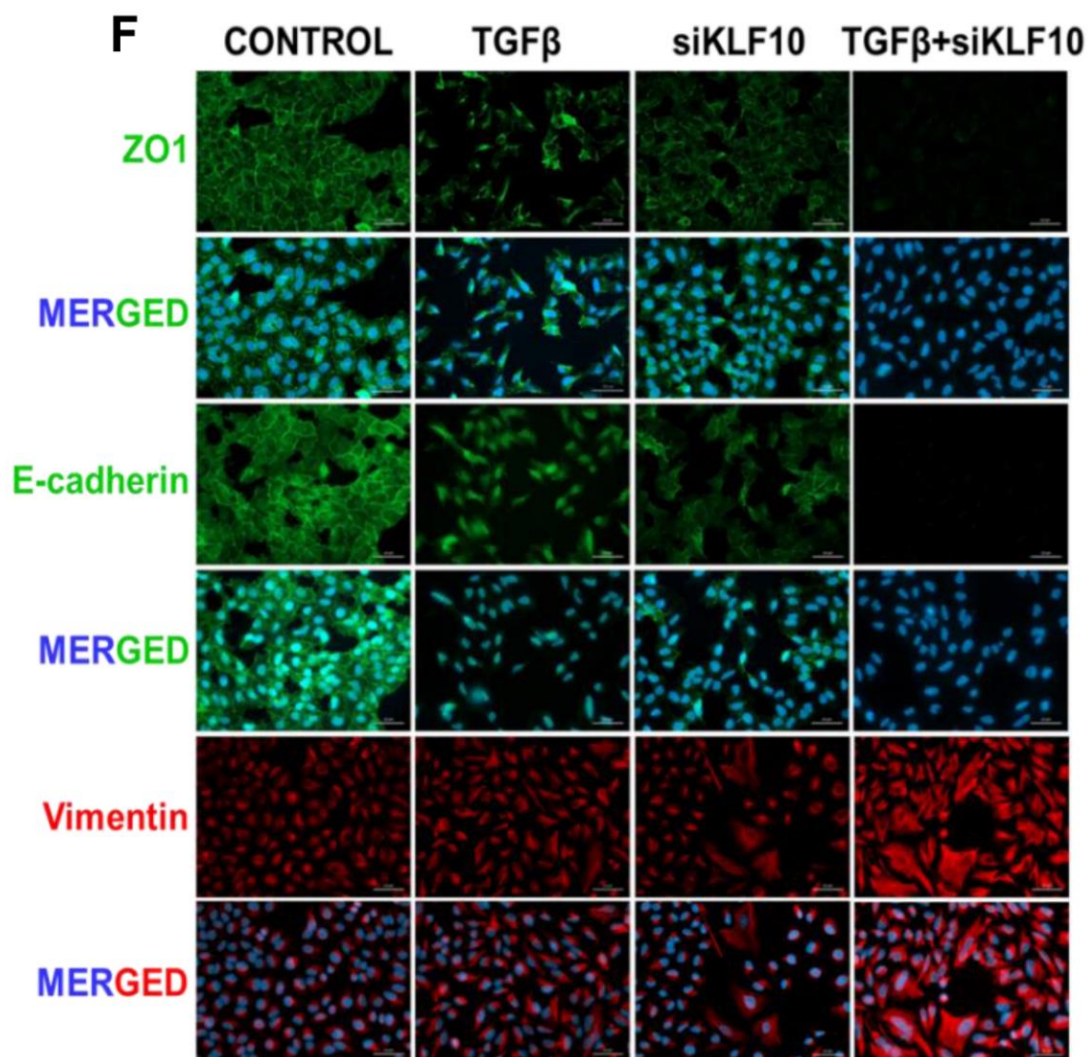
Figure 28: KLF10 depletion enhances TGFβ-induced EMT. (A) Phase contrast images showing the morphological changes in the cells upon undergoing EMT. Magnification 10x. (B) In A549 cells gene expression level of epithelial (E-cadherin and MMP7) and mesenchymal (N-cadherin and MMP2) markers were analyzed by qRT-PCR and shown as "relative mRNA levels" as compared to HNRNPK expression level. Data are represented as mean ± SD. n=3. ***$p ≤ 0.005$, **$p ≤ 0.01$, *$p ≤ 0.05$. (C) In Panc1 cells gene expression level of epithelial (E-cadherin) and mesenchymal (N-cadherin) marker was analyzed by qRT-PCR and shown as "relative mRNA levels" as compared to HNRNPK expression level. Data are represented as mean ± SD. n=3. ***$p ≤ 0.005$, **$p ≤ 0.01$, *$p ≤ 0.05$. (D) Western blot analysis of whole cell protein lysates from A549 cells showing significant loss of E-cadherin upon TGFβ treatment in KLF10 depleted cells. KLF10 blot shows upregulation of KLF10 protein level upon TGFβ treatment and complete loss following siRNA mediated knockdown. HSC70 was used as a loading control. (E) Western blot analysis of whole cell lysates from Panc1 cells showing loss of epithelial marker and upregulation of mesenchymal markers upon KLF10 depletion. HSC70 was used as a loading control. (F) Immunofluorescence staining showing enhanced EMT induction upon TGFβ treatment in KLF10 depleted cells. Cells were stained for epithelial markers (ZO1 and E-cadherin) and mesenchymal marker (Vimentin). Nuclei were stained with DAPI. Scale bar represents 50 μm.

4.7.2 KLF10 depletion results in enhanced migratory potential

One of the hallmark features of EMT is that it imparts the cells with potential to migrate to distant sites thereby promoting metastasis. As we have demonstrated in the earlier experiments that low expression of KLF10 can result in an enhanced EMT-like phenotype we speculated that KLF10 depletion might lead to increased cell migration. To address this question we performed transwell migration assays with A549 and MDA-MB-231 cells. Cells were transfected with control or KLF10 siRNAs and were allowed to grow for 48 hours. Post-transfection cells were seeded into 8.0 μm PET track-etched membrane cell culture inserts and were treated with or without TGFβ for 48 hours. Migration assay results revealed that KLF10 depletion leads to significantly enhanced migratory potential in cells treated with TGFβ compared to the cells treated with TGFβ alone (Fig. 29). Strikingly, even MDA-MB-231 cells, which are highly metastatic, showed increased migration capacity upon KLF10 depletion.
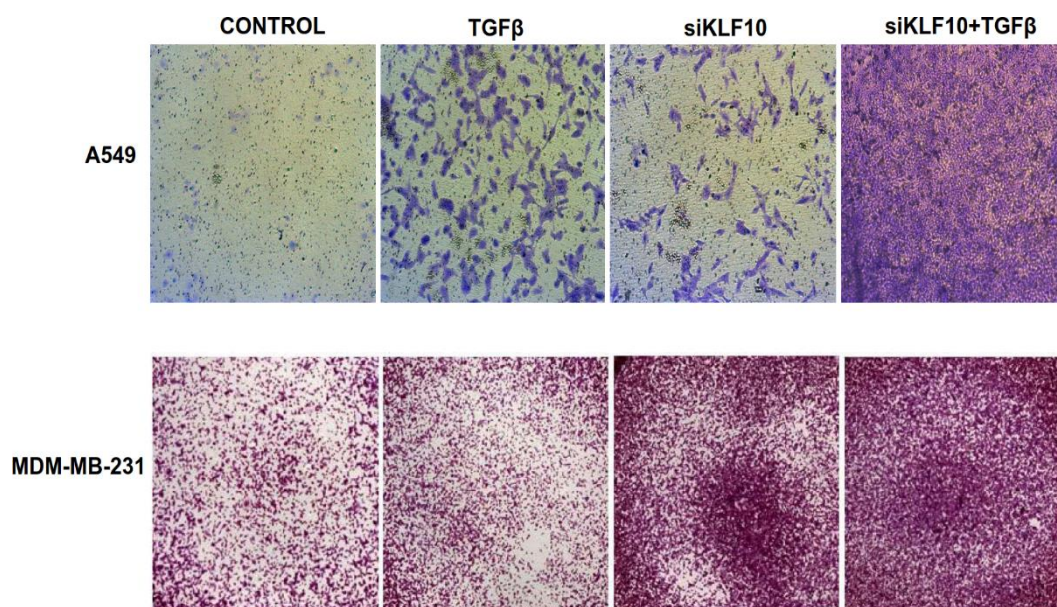
Figure 29: KLF10 depletion results in enhanced migratory potential. Transwell migration assay was performed in A549 and MDA-MB-231 cells using Boyden chamber inserts (8.0 μm). Cells were transfected with control or KLF10 siRNA and after 48 hours were split and seeded into the upper chamber of the inserts (25,000 cells). Cells were treated with or without TGFβ for additional 48 hours following fixation with 100% methanol and staining with crystal violet. Increased migration was observed upon KLF10 depletion in both A549 as well as MDA-MB-231 cells. Image represents 10 x magnifications.

2. Pharmacological inhibition of epigenetic regulators using small molecule inhibitors

After investigating the potential role of transcription factor KLF10 in TGFβ-induced EMT, we wanted to take a more global approach and study the effects of pharmacological inhibition of epigenetic modifiers using small molecule inhibitors which are either in clinical trials or clinical development. Epigenetic regulators controlling DNA and histone modifications have been shown to be frequently deregulated in cancer and are major focus of targeted therapies. Several small molecule inhibitors targeting epigenetic regulators have been developed and either in clinical development or clinical trials. LSD1 is frequently overexpressed in many cancer types and its high expression in prostate cancer was correlated with tumor relapse (Kahl et al., 2006; Metzger et al., 2005). Given the potential role of LSD1 and HDACs in tumorigenesis and tumor progression we have utilized small molecule inhibitors against LSD1 (SP2509) and HDACs (Vorinostat and Resminostat). To investigate the potential of combined inhibition of LSD1 and HDACs as a targeted therapy we have used dual LSD1/HDAC inhibitor 4SC-202 which is currently under clinical development (http://www.4sc.com/product-pipeline/clinical/4sc-202/).

4.8 4SC-202 blocks-TGFβ induced EMT and drives the cells towards differentiation

4.8.1 Transcriptome wide effect of 4SC-202 on TGFβ regulated genes

4SC-202 is an epigenetic drug candidate currently under clinical development. According to the phase-I clinical data it targets the important signaling pathways involved in cancer development and progression. Since EMT is one of the key steps in the formation of metastatic cancer we sought to investigate if 4SC-202 can inhibit EMT. Initially, we wanted to determine the transcriptome wide effect of 4SC-202 on

the TGFβ regulated genes by a RNA sequencing study. We chose Panc1 cells as a model system to induce EMT via TGFβ since our previous study showed that Panc1 cells respond to TGFβ stimulation and undergo EMT. We treated the cells with TGFβ (5 ng/ml) and 4SC-202 (1 µg/ml) alone or in combination for 12 hours. Control cells were treated with DMSO (vehicle). After 12 hours post-treatment RNA samples were harvested and used for library preparation followed by RNA sequencing. For further analysis we selected the significantly (padj ≤ 0.05) up- (log2fc ≥ 1.5) or down- (log2fc ≤ -1.5) regulated genes and TGFβ-regulated genes were compared to differentially regulated genes upon 4SC-202 treatment. As depicted in the heatmap Fig. 30, the majority of the TGFβ upregulated genes (shown in red) were downregulated after 4SC-202 treatment whereas TGFβ downregulated genes (shown in green) were upregulated in the 4SC-202 sample. Interestingly, the effect of TGFβ was essentially totally blocked by 4SC-202 which indicates that 4SC-202 can block the TGFβ-induced EMT. LSD1 has been shown to be frequently overexpressed in different types of metastatic cancer with an aggressive phenotype. Furthermore, LSD1 which is inhibited by 4SC-202, has been shown to promote EMT and impaired LSD1 expression has been found to be associated with blocked EMT (Fig. 30). Our findings from the transcriptome study correlate well with the inhibition of EMT by reduced LSD1 expression as treatment with 4SC-202 completely reversed the effect of TGFβ on its target genes many of which are involved in inducing EMT. Taken together these results revealed a potential function of 4SC-202 in inhibiting EMT.
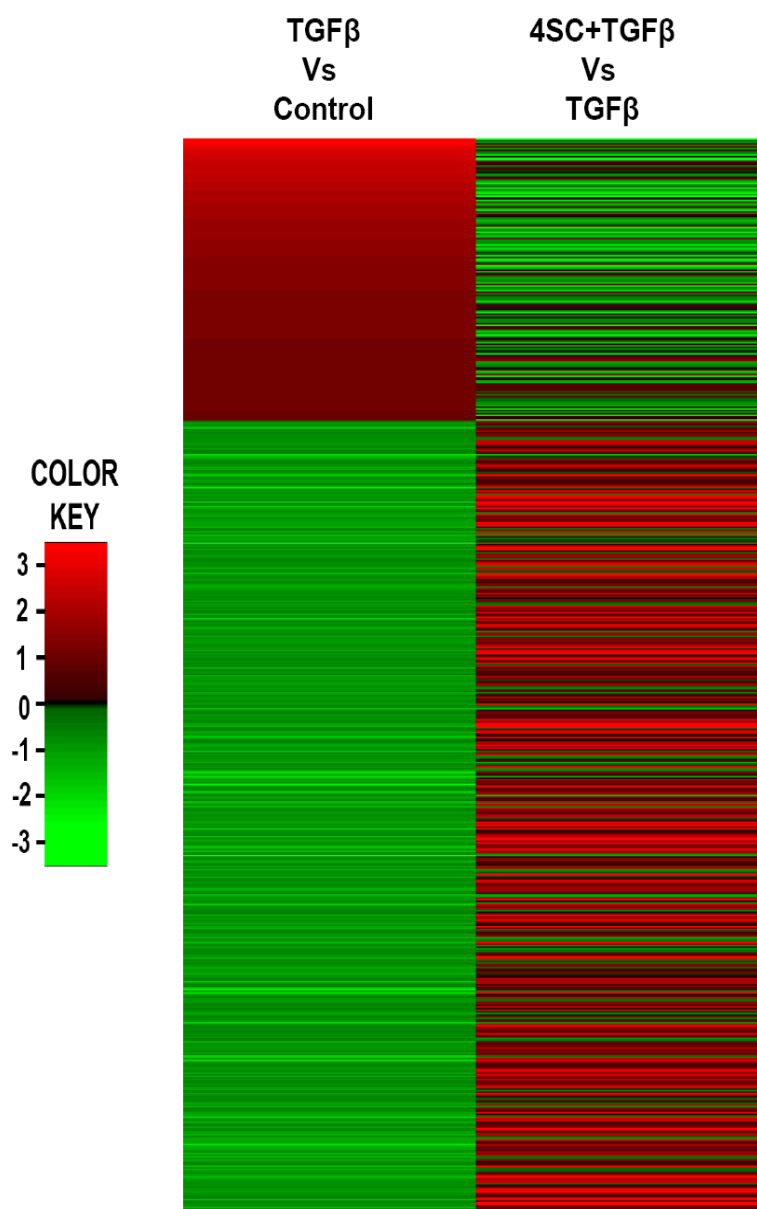
Figure 30: Transcriptome wide effects of 4SC-202 on TGFβ regulated genes. Heatmap of RNA-seq data from Panc1 cells. Significant numbers of TGFβ regulated/targeted genes were affected by 4SC-202. Panc1 cells were treated with or without TGFβ (5 ng/ml) for 72 hours to induce EMT and thereafter 4SC-202 was added to the cells for 12 hours. Heat map was generated using statistically significant (padj-value ≤ 0.05) up (red) and down (green) regulated genes (cutoff of ± 1.5 log2fold change). As depicted in the heatmap, TGFβ upregulated genes were downregulated and downregulated genes were reactivated upon 4SC-202 treatment.

4.8.2 4SC-202 regulates pathways related to cellular homeostasis and maintaining cell identity

To gain additional insight to the effects of 4SC-202 action we investigated the signature pathways and genes differentially regulated by 4SC-202. Initially, we performed Gene Set Enrichment Analysis (GSEA) to identify the significantly enriched pathways. Interestingly, GSEA results revealed that pathways pertaining to cellular homeostasis and organization of the cellular architecture were highly enriched in 4SC-202 treated cells compared to the cells that had undergone TGFβ-induced EMT (Fig. 31A). This could imply that, in the presence of 4SC-202 cells are pushed back to normal state or the epithelial state. Furthermore, we performed GO analysis using the statistically significantly regulated genes and then selected the top 100 GO terms to perform further analysis using online tool REVIGO (REduce VIsualize Gene Ontology). REVIGO takes into account a list of Gene Ontology terms and summarizes them into a cluster of meaningful non-redundant GO terms. As depicted in Fig. 31B, significantly enriched GO terms were all related to maintenance of cellular homeostasis by controlling cell division, organization of cellular components and regulation of gene expression. Interestingly several pathways involved in the organization or remodeling of chromatin and regulation of gene expression were also among the significantly enriched GO terms. The enrichment of GO terms like regulation of transcription by RNA Polymerase II, gene expression and chromatin silencing indeed point towards a cellular state where changes in chromatin state and gene expression status occur. These findings coincide with a recent report where the authors have shown that TGFβ-induced EMT affects the chromatin morphology and also induces a global reprogramming of the chromatin state (McDonald et al., 2011). To further strengthen our findings we selected the list of significantly up- and down-regulated genes and used them to perform Gene Ontology (GO) analysis. GO analysis results revealed

that, most highly enriched GO terms were related to cell cycle, chromosome condensation, cytoskeleton organization and cellular response to stress (Fig. 31C). Taken together these results revealed that 4SC-202 affected genes are associated with pathways regulating cellular homeostasis and maintaining the cell identity.
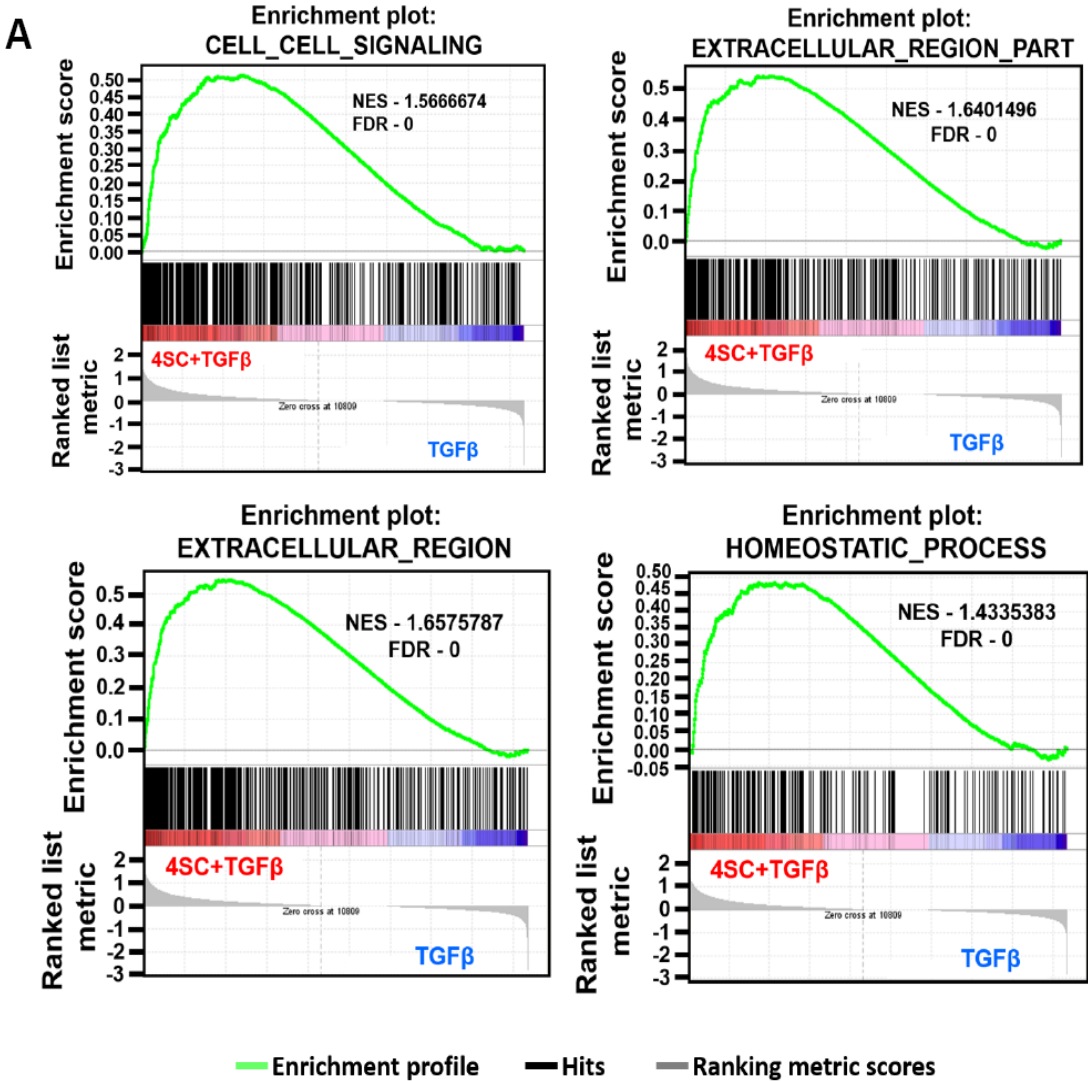
B
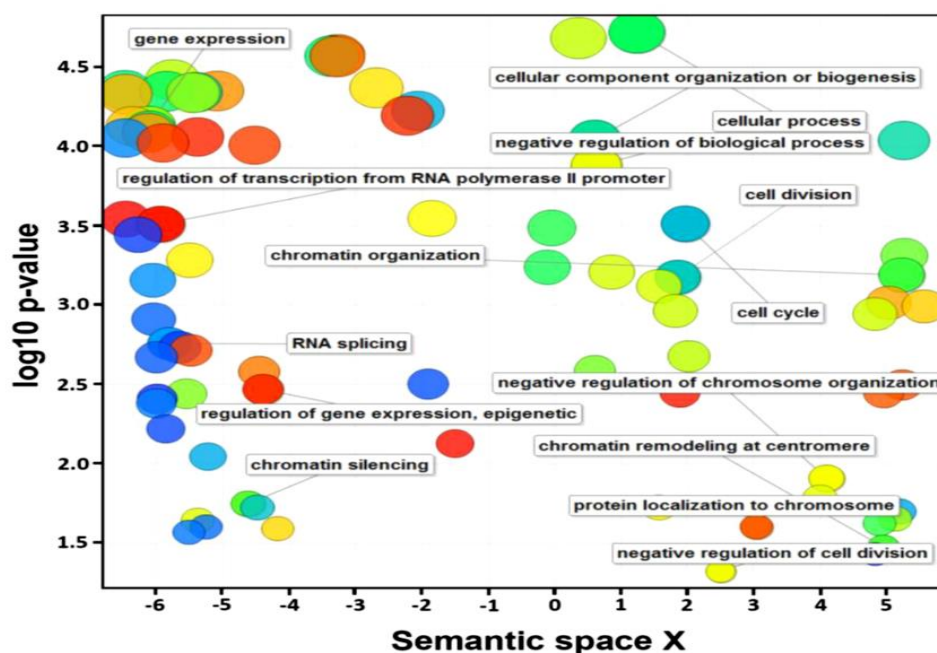
C

Annotation cluster 1					
GO term	Enriched pathway	Enrichment score	Count	p-value	FDR
SP_PIR_KEYWORDS	Cell cycle	21.38	157	5.0E-30	7.5E-27
GOTERM_BP_FAT	Cell cycle phase	21.38	149	2.1E-28	4.0E-25
GOTERM_BP_FAT	Cell cycle	21.38	226	1.4E-27	2.7E-24
GOTERM_BP_FAT	M phase	21.38	127	1.7E-27	3.2E-24
GOTERM_BP_FAT	Cell cycle process	21.38	176	4.8E-25	9.1E-22
GOTERM_BP_FAT	Mitotic cell cycle	21.38	122	1.0E-19	1.9E-16
Annotation cluster 2					
GO term	Enriched pathway	Enrichment score	Count	p-value	FDR
GOTERM_CC_FAT	Condensed chromosome	15.01	130	3.4E-16	4.9E-13
GOTERM_CC_FAT	Chromosomal part	15.01	111	1.9E-14	2.8E-11
Annotation cluster 3					
GO term	Enriched pathway	Enrichment score	Count	p-value	FDR
GOTERM_BP_FAT	Microtubule-based process	8.15	76	2.7E-10	5.0E-7
GOTERM_BP_FAT	Cytoskeleton organization	8.15	103	2.6E-7	4.9E-4
Annotation cluster 4					
GO term	Enriched pathway	Enrichment score	Count	p-value	FDR
GOTERM_BP_FAT	DNA metabolic process	7.65	134	7.7E-13	1.4E-9
GOTERM_BP_FAT	Cellular response to stress	7.65	137	3.9E-10	7.3E-7
GOTERM_BP_FAT	Response to DNA damage stimulus	7.65	95	1.9E-8	3.5E-5

Figure 31: 4SC-202 regulates pathways related to cellular homeostasis and maintaining cell identity. (A) GSEA comparing control versus 4SC-202 condition. Significantly enriched pathways are shown. (B) Results from REVIGO clustered the significantly enriched GO terms into meaningful non-redundant clusters showing the pathways related to gene expression and chromatin organization being highly represented. (C) Significantly (padj ≤ 0.05) up or downregulated genes from the RNA-seq data were used to perform GO analysis using the DAVID online tool. Significantly enriched annotated cluster (selected based on p-value and FDR) with related GO term are shown.

4.8.3 4SC-202 blocks TGFβ-induced EMT

Since EMT is one of the key steps in the formation of metastatic cancer we investigated whether 4SC-202 can inhibit or reverse the occurrence of EMT. To test this hypothesis we treated Panc1 cells with TGFβ (5 ng/ml) and 4SC-202 (1 µg/ml) alone or in combination for 12 hours. Control cells were treated with DMSO (vehicle). After 12 hours post-treatment RNA and protein samples were harvested. We performed gene expression analysis and western blot to check for the regulation of key EMT markers and investigate the effects of 4SC-202 treatment. Interestingly, in our gene expression results we observed that classical epithelial markers like E-cadherin, TJP3 and MMP2 which were downregulated by TGFβ were significantly blocked in downregulation upon co-treatment with 4SC-202. Similarly, classical EMT markers like N-cadherin, SNAI1, ZEB1 and MMP7 which were upregulated upon TGFβ treatment were found to be significantly blocked in their regulation upon co-treatment with 4SC-202 (Fig. 32A). Furthermore, western blot results also showed similar effects of 4SC-202 on epithelial and mesenchymal markers (Fig. 32B). Next, we performed immunofluorescence staining for epithelial and mesenchymal marker and interestingly observed that expression of E-cadherin, which was lost following TGFβ treatment, was maintained with 4SC-202 co-treatment. Likewise, the mesenchymal marker ZEB1, which was significantly upregulated by TGFβ stimulation, was significantly blocked in its upregulation by 4SC-202 co-treatment (Fig. 32C). Taken together, these results revealed that 4SC-202 can indeed block TGFβ-induced EMT.

A
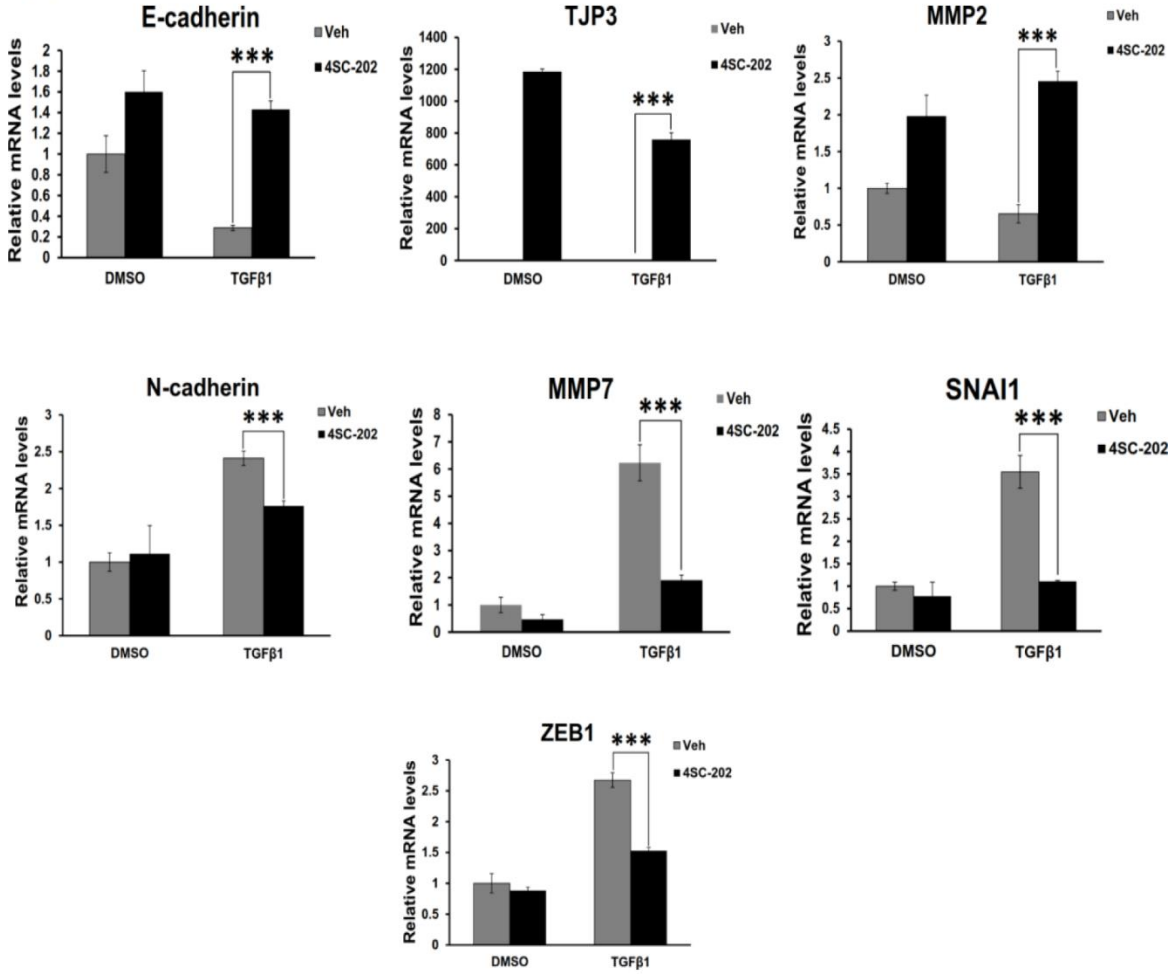
B
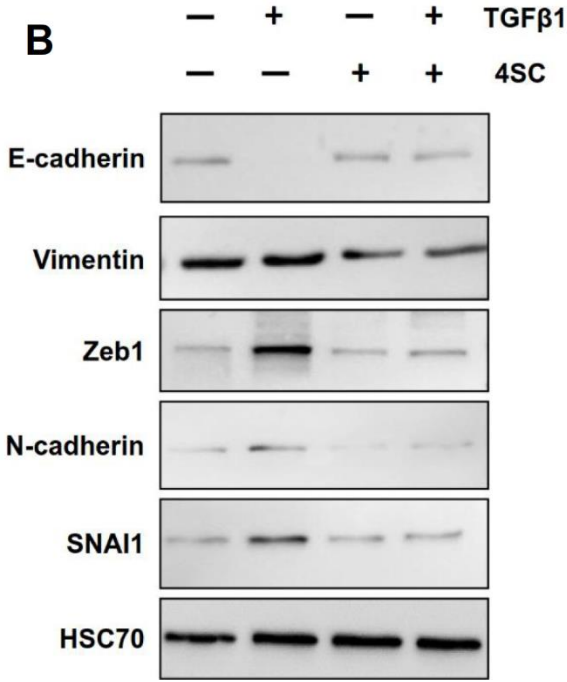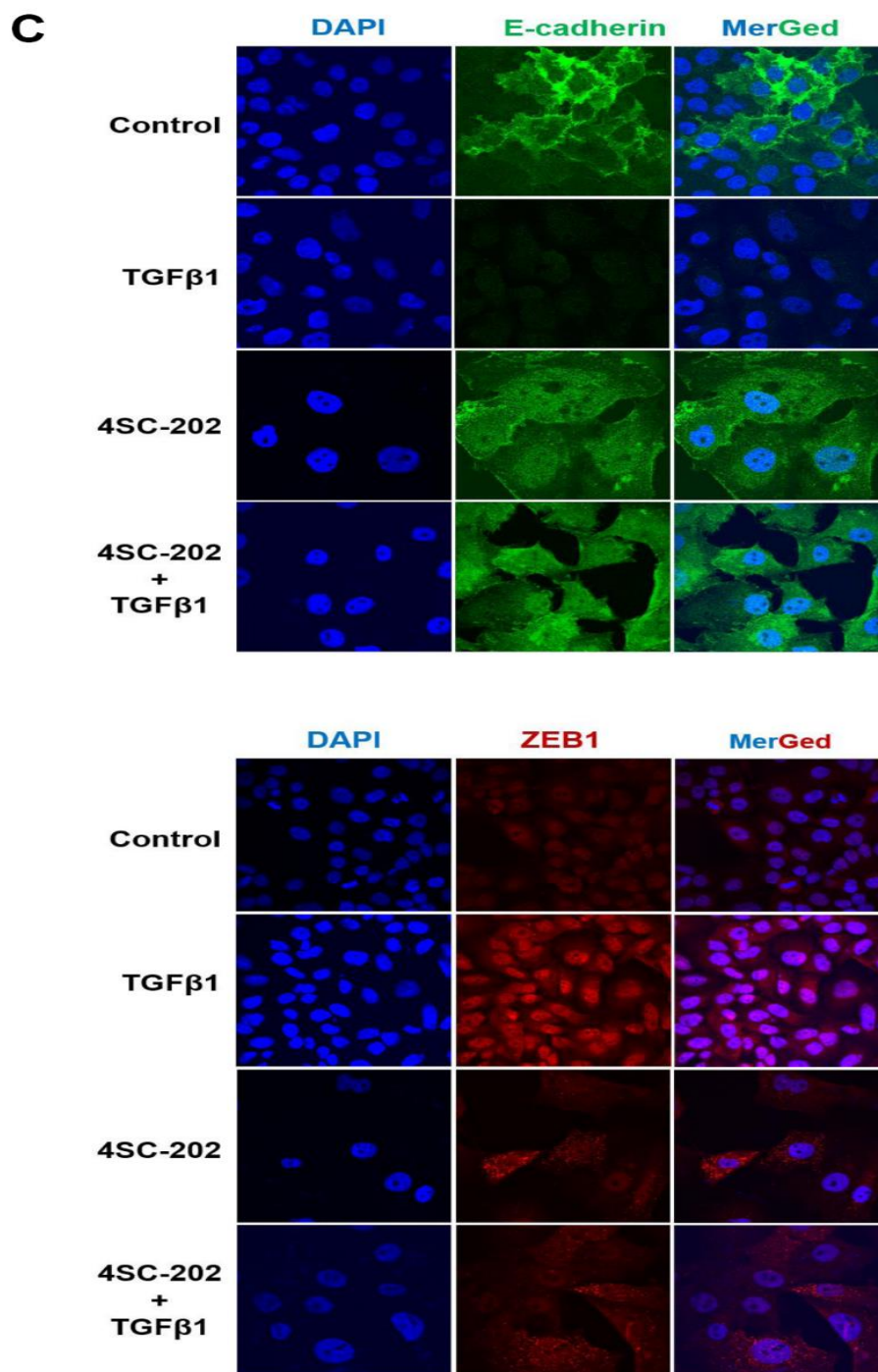
Figure 32: 4SC-202 blocks TGFβ-induced EMT. (A) qRT-PCR results showing that epithelial markers (E-cadherin, TJP3, MMP2) were re-expressed and mesenchymal markers (N-cadherin, ZEB1, SNAI1, MMP7) were downregulated following 4SC-202 treatment in TGFβ stimulated cells. Data is shown as "relative mRNA levels" as compared to RPLP0 expression level. Data are represented as mean ± SD. n=3. ***p ≤ 0.005, **p ≤ 0.01, *p ≤ 0.05. (B) Western blot results showing the upregulation of epithelial marker protein (E-cadherin) and downregulation of mesenchymal marker proteins (Vimentin, N-cadherin, SNAI1, ZEB1) upon 4SC-202 treatment in TGFβ treated cells. HSC70 was used as a loading control. (C) Immunofluorescence staining showing enhanced EMT induction upon TGFβ treatment in Panc1 cells. However, EMT induction was reversed upon treatment with 4SC-202. Cells were stained for the epithelial marker (E-cadherin) and mesenchymal marker (ZEB1). Nuclei were stained with DAPI. Scale bar represents 50 μm.

4.8.4 4SC-202 promotes a differentiated phenotype

Having verified that 4SC-202 can prevent TGFβ-induced EMT next we investigated whether 4SC-202 can promote differentiation. We used the RNA and protein samples from the previous experiment and checked for the expression of CD24, a marker associated with a differentiated phenotype. Previous reports have shown that cancer stem-like cells are negative for CD24, implying that CD24-positive cells possess a differentiated phenotype. Interestingly, our gene expression analysis and western blot results revealed a significant upregulation of CD24 expression upon treatment with 4SC-202 (Fig. 33A & B). To further confirm the association of 4SC-202 with a more differentiated phenotype we investigated whether it can inhibit the cell migration as cells that have undergone EMT have an enhanced migratory potential. To check this we performed transwell migration assay where cells were either treated with TGFβ or 4SC-202 alone and also in combination. Interestingly migration assay results revealed that 4SC-202 completely blocked the TGFβ-induced migration (Fig. 33C). Taken together, these results validated that 4SC-202 can inhibit TGFβ-induced EMT and significantly block migration. However, its potential role in promoting differentiation will need to be validated further.
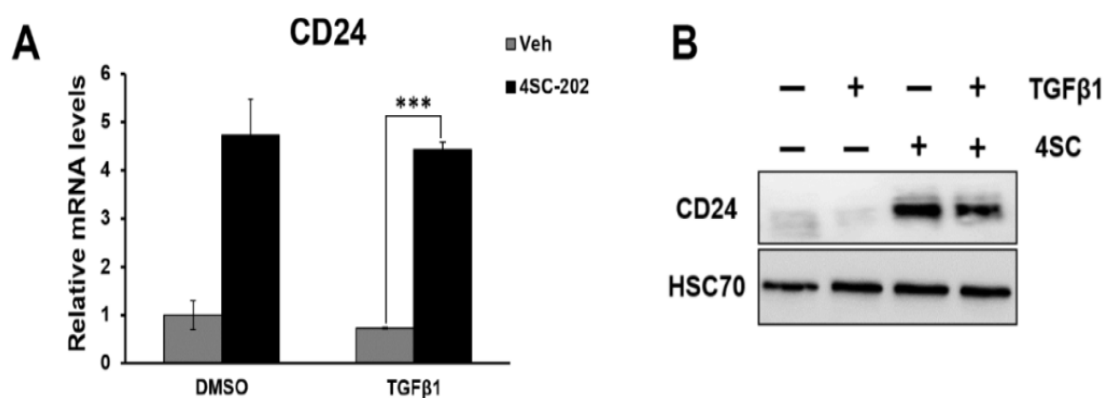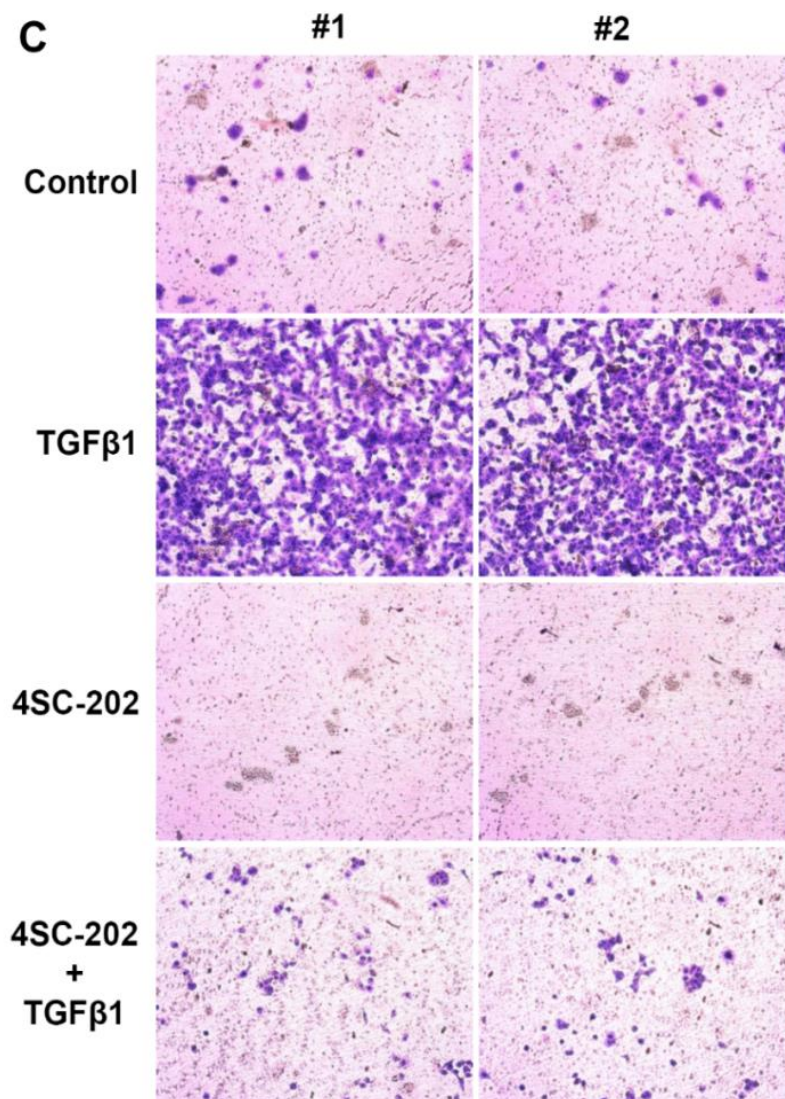
Figure 33: 4SC-202 treatment drives the cells towards differentiation. (A) qRT-PCR results showing that CD24, a marker associated with a differentiated cell phenotype was significantly upregulated following 4SC-202 treatment in TGFβ-stimulated cells. Data is shown as "relative mRNA levels" as compared to RPLP0 expression levels. Data is represented as mean ± SD. n=3. ***p ≤ 0.005, **p ≤ 0.01, *p ≤ 0.05. (B) Western blotting results showing significantly enhanced expression of CD24 protein levels following 4SC-202 treatment in TGFβ-stimulated and non-stimulated cells. HSC70 was used as a loading control (C) Transwell migration assay was performed in Panc1 cells treated with TGFβ or 4SC-202 and their co-treatment. Panc1 cells displayed enhanced migration capacity upon TGFβ-stimulation. Migration potential of the cells was essentially blocked by 4SC-202 treatment. Cells were fixed with 100% methanol and stained with crystal violet. Images were taken using light microscope at 10x magnification.

4.9 Target specificity of 4SC-202

4.9.1 Similar gene expression pattern was observed in three different pancreatic cancer cell lines

To further elucidate the mechanism of action of 4SC-202 we investigated whether or not its effects are specific. For this we performed transcriptome-wide studies across different pancreatic cancer cell lines. We chose three pancreatic cancer cell lines namely L3.6, BxPC3 and Panc1 and treated them with 4SC-202 for 12 hours while the control cells were treated with DMSO (vehicle). Before proceeding with further analysis of the transcriptome data we performed quality check and PCA (principle component analysis) plot to ensure that all the replicates from a given condition correlate with one another. As depicted in Fig. 34A, one of the replicates from control condition (vehicle) in L3.6 cells had a high variance compared to the other two replicates and hence, was not included in further analyses. Importantly, the PCA plot from vehicle-treated and 4SC-202 treated samples revealed differences between two conditions within the same cell line (Fig. 34B). After confirming the quality of the data and any possible variability between the samples, we proceeded with further analyses of the RNA-seq data and compared the differentially regulated genes between the 4SC-202 and vehicle treated conditions in all three different cell lines. Statistically significant (padj ≤ 0.05) up- (log2fold ≥ 1.5) or down- (log2fold ≤ 1.5) regulated genes in L3.6 cells were selected for the heatmap. As shown in the (Fig. 34C), the heatmap showed that the effect of 4SC-202 on the transcriptome of all the three cell lines was very similar. Having observed a similar pattern of gene regulation in the different pancreatic cancer cell lines following 4SC-202 treatment we investigated the common pathways that were enriched due to the affected genes. For this purpose we utilized the samples from all three cell lines and performed gene set enrichment analysis. Interestingly, several pathways related to the extracellular matrix and tissue development were

enriched which further suggests that 4SC-202 affects differentiation-related genes. Furthermore we also observed an enrichment of several pathways that were related to growth inhibition of different types of cancers like breast, colorectal and gastric (Fig. 34D). Taken together these results revealed that 4SC-202 has a similar gene regulation pattern across three different pancreatic cancer cell lines and differentially regulated genes were found to associated with growth inhibitory effects on different types of cancers.
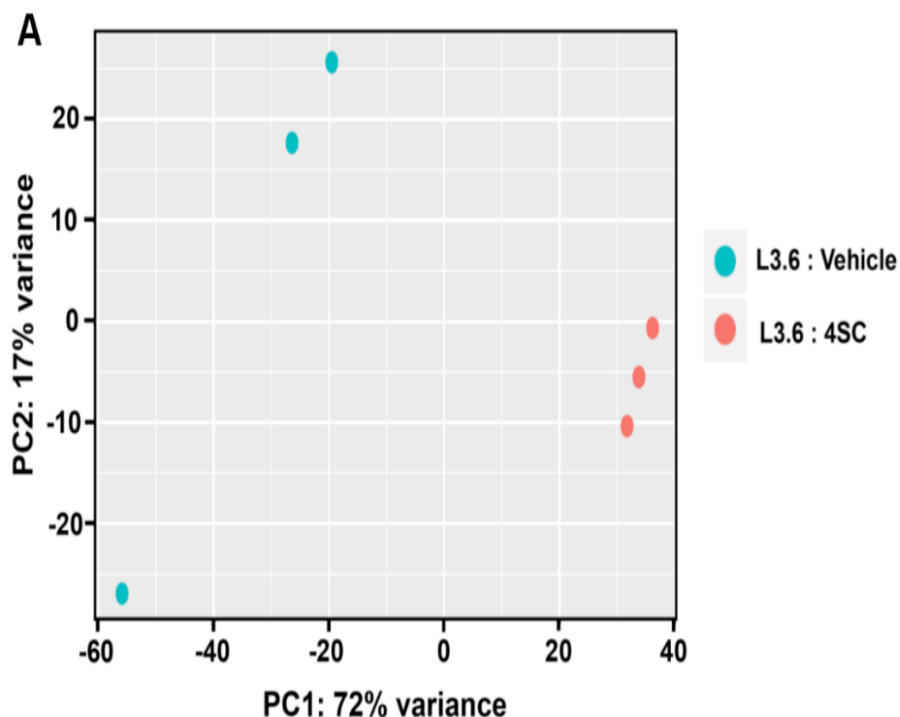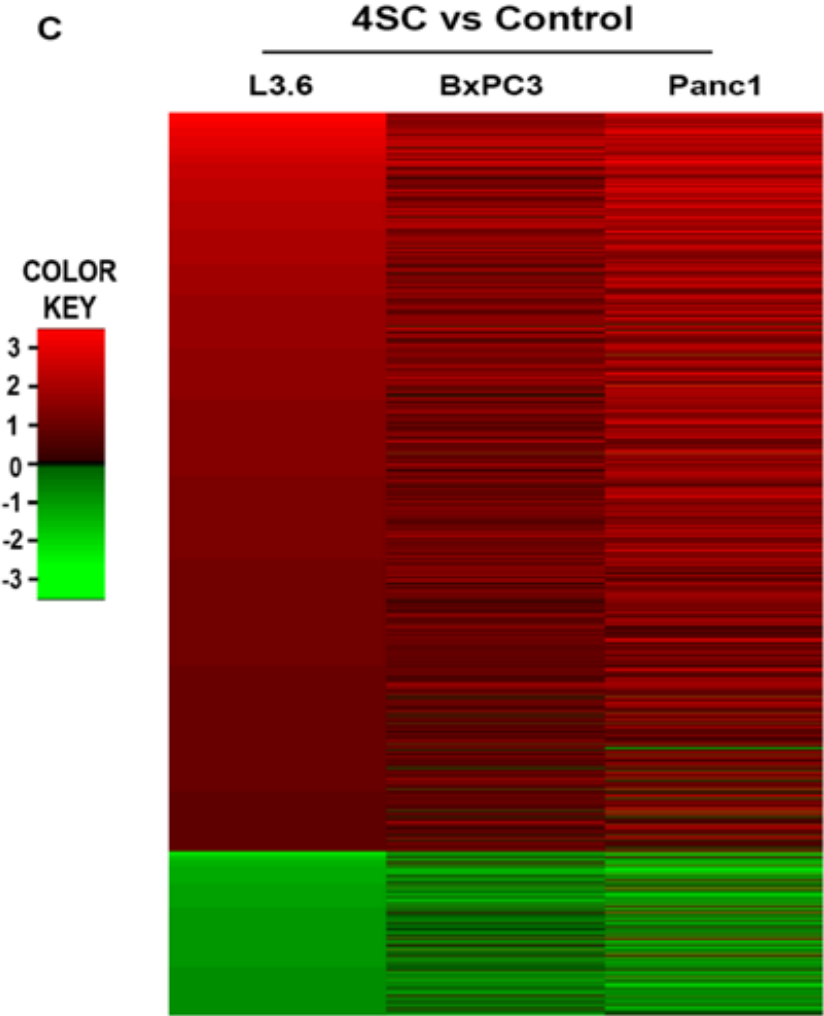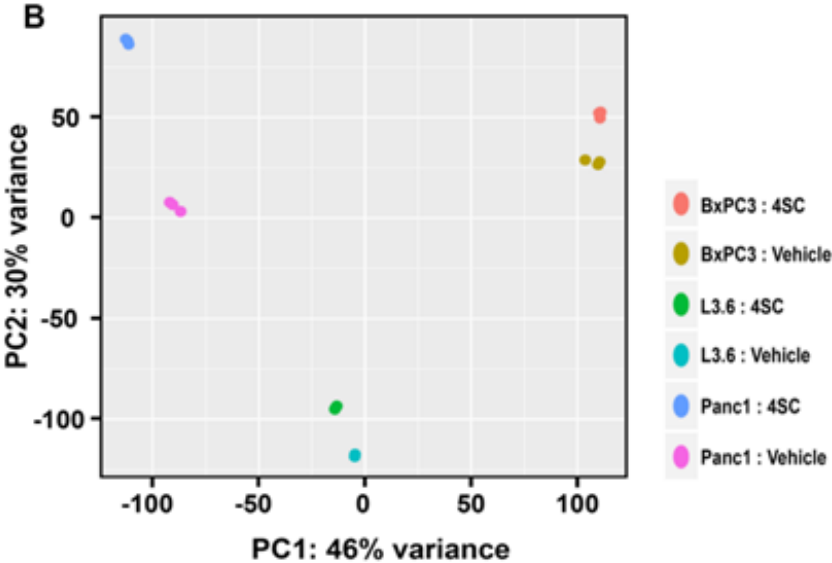
D
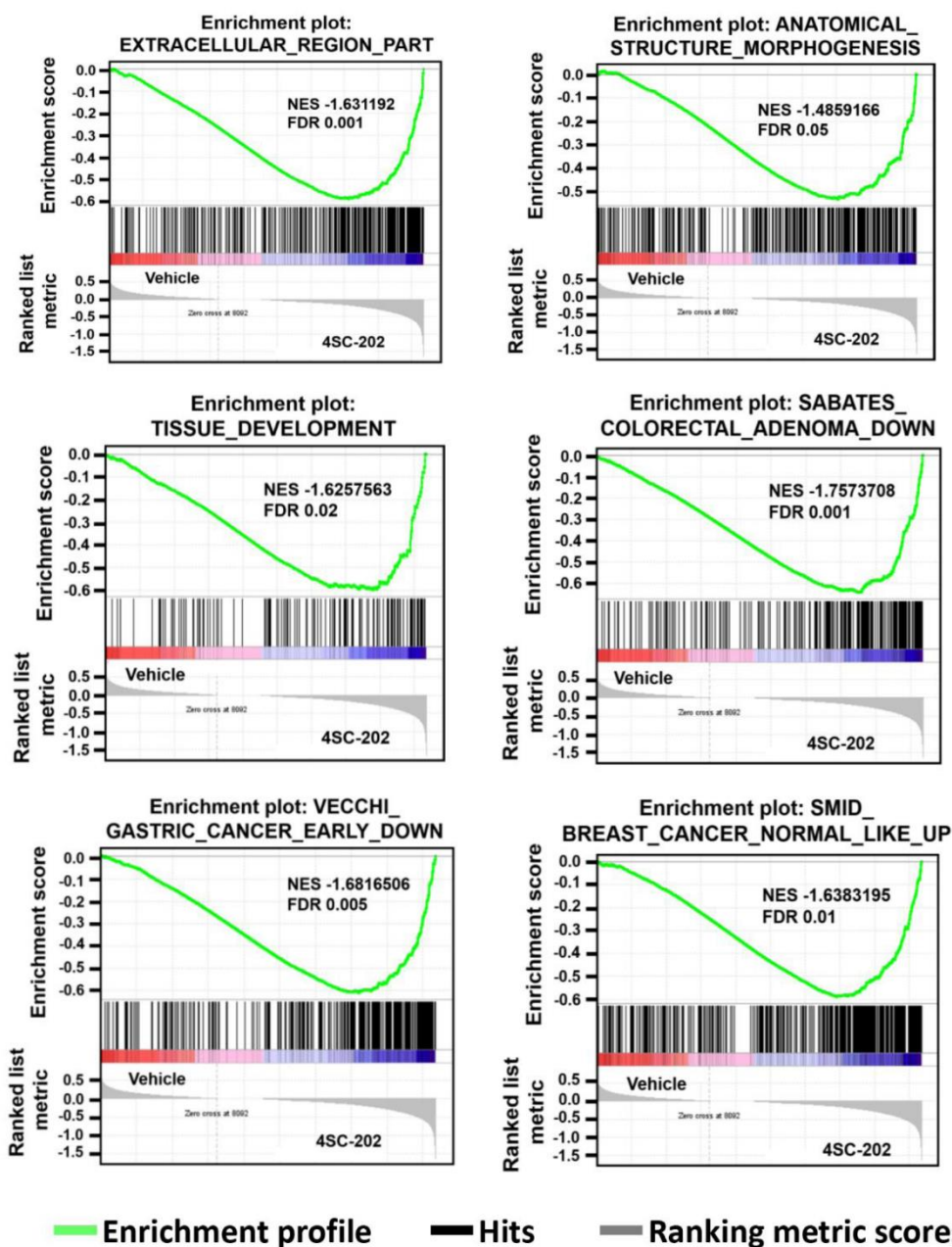
Figure 34: Similar gene expression pattern was observed in three different pancreatic cancer cell lines. (A) PCA plot from RNA-seq samples in L3.6 cell line showing that there was a large variance between one of the vehicle treated samples compare to the other two samples. This sample was not included in further analyses. Also noteworthy is the large variance between the vehicle and 4SC-202 treated samples. (B) PCA plot from RNA-seq samples in L3.6, BxPC3 and Panc1 cell lines showing the variance between vehicle and 4SC-202 treated samples. (C) Heatmap showing similar pattern of gene expression in L3.6, BxPC3 and Panc1 cell line following 4SS-202 treatment. Statistically significant (padj ≤ 0.05) up (log2fold ≥ 1.5) or down (log2fold ≤ 1.5) regulated genes were selected for the heatmap. Upregulated genes are shown in red while downregulated genes are shown in green. (D) GSEA results showing the enrichment of pathways pertaining to extracellular matrix, tissue development and inhibition of cancer growth and development.

4.10 Comparative study of 4SC-202 with other HDAC inhibitors

Since 4SC-202 has an inhibitory activity against HDACs (class I HDACs including HDAC1, 2 and 3) we performed a comparative study of 4SC-202 with two other potent HDAC inhibitors Resminostat (1 µM) and Vorinostat (1 µM) (also known as (suberoylanilide hydroxamic acid or SAHA). Resminostat is a potent and selective inhibitor of HDAC1, 3 and 6 while having a comparatively smaller potency against HDAC8. Vorinostat is a reversible pan-HDAC inhibitor with inhibitory activity against all 11 known human HDACs including both class-I and class-II HDACs. Both Vorinostat and Resminostat alter the chromatin structure of the transformed cells resulting in cell cycle arrest thus causing apoptosis and differentiation (Munshi et al., 2006; Richon, 2006). We performed transcriptome analysis on L3.6 cells treated with either 4SC-202, Resminostat, Vorinostat or DMSO (vehicle). Further analysis of the RNA-seq data revealed that there was a high degree of similarity between the samples treated with Resminostat or Vorinostat as depicted by principle component analysis (PCA) plot (Fig. 35A). Interestingly, there was a significant difference between samples treated with 4SC-202 and samples treated with Resminostat or Vorinostat, pointing towards a difference in the mode of action of 4SC-202 compared to the other two HDAC inhibitors. To investigate the possible similarity or differences in the regulation of gene expression caused by these inhibitors we selected the significantly (padj ≤ 0.05) up or down regulated genes in any of the treatment condition and used these genes for a heatmap. Due to differences in the magnitude of effects different cut-off values (log2fc) were used for different inhibitor treatments to achieve an equal number of up or down regulated genes in each conditions. As depicted in the heatmap (Fig. 35B), and similar to the PCA plot, Resminostat and Vorinostat had quite similar effects and were clustered together while the effects of 4SC-202 was weaker but it still

showed some similarity in the observed effects. To further validate the differences in the pattern of gene regulation between these inhibitors we selected significantly regulated genes from all three treatment conditions and used those to plot a Venn diagram and check for potential overlap between the regulated genes. As shown in the Venn diagram (Fig. 35C), a significant portion of the Resminostat and Vorinostat regulated genes were found to overlap with each other. In contrast, the overlap of 4SC-202 regulated genes with either Resminostat or Vorinostat was very low which again correlates well with the PCA and heatmap results. Interestingly despite notable differences in the overlapping set of regulated genes between the different inhibitors we observed that slightly over one thousand genes displayed overlap. We, thus performed Gene Ontology analysis to determine what common pathways may be regulated by all three substances. As shown in Fig. 35D, the Gene Ontology results revealed that apart from GO terms related to transcription regulation there were several enriched pathways related to or involved in cell morphogenesis and development of cell projections to facilitate cell-cell adhesion. Furthermore, GO terms related to cell differentiation and specifically neuronal differentiation was also enriched which implicates a potential role of 4SC-202 in promoting differentiation. Taken together these results revealed similarity in the action of 4SC-202, Resminostat and Vorinostat.

C
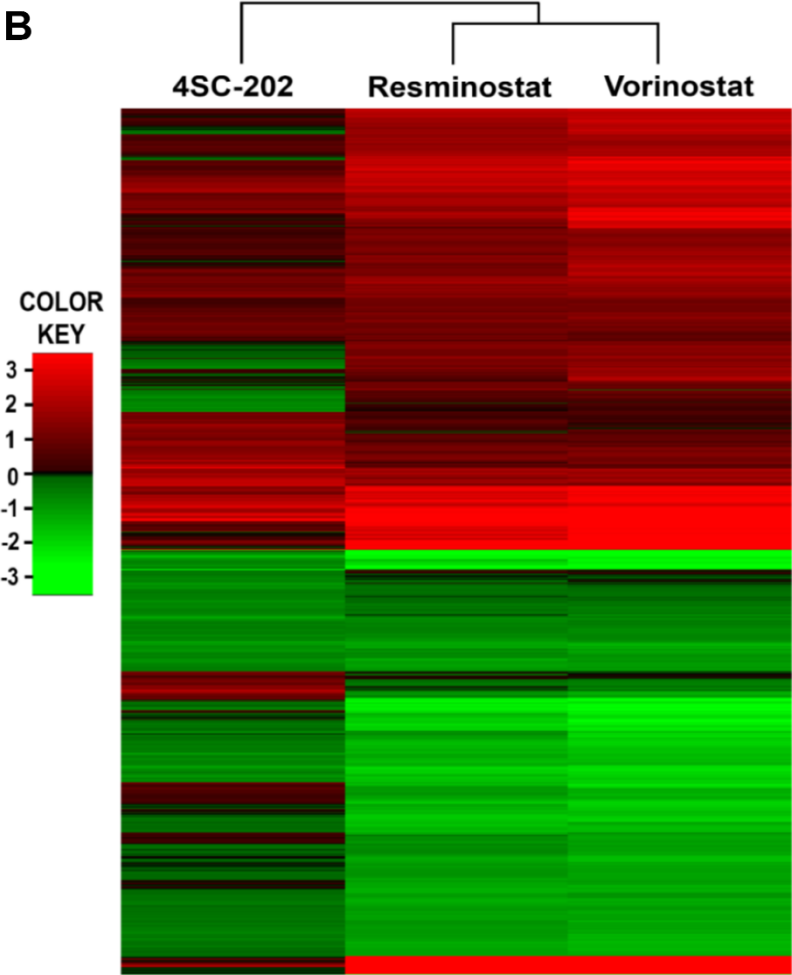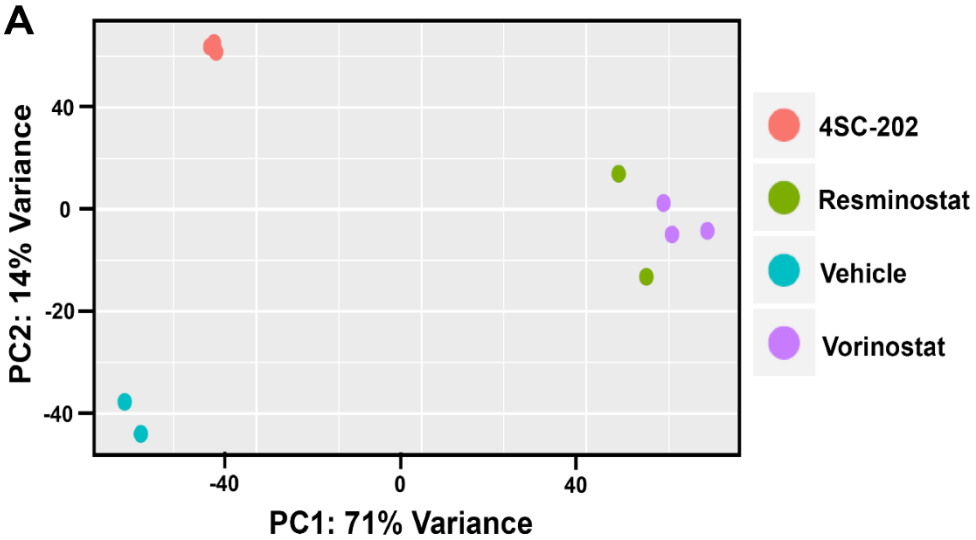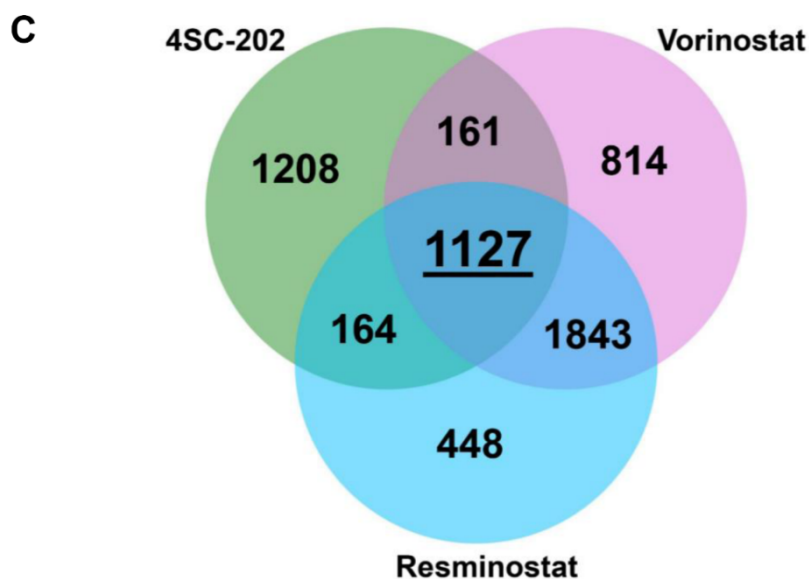
D

Annotation cluster 1					
GO term	**Enriched pathway**	**Enrichment score**	**Count**	**p-value**	**FDR**
GOTERM_BP_FAT	Cell projection organization	6.87	51	3.7E-10	2.9E-5
GOTERM_BP_FAT	Neuron differentiation	6.87	57	2.0E-8	3.6E-5
GOTERM_BP_FAT	Neuron development	6.87	49	8.5E-9	8.9E-7
GOTERM_BP_FAT	Cellular component morphogenesis	6.87	47	5.7E-6	9.2E-4
Annotation cluster 2					
GO term	**Enriched pathway**	**Enrichment score**	**Count**	**p-value**	**FDR**
GOTERM_BP_FAT	Vasculature development	3.98	33	2.4E-5	2.3E-3
GOTERM_BP_FAT	Blood vessel developement	3.98	32	3.7E-5	6.6E-2
Annotation cluster 3					
GO term	**Enriched pathway**	**Enrichment score**	**Count**	**p-value**	**FDR**
GOTERM_BP_FAT	Positive regulation of developmental process	3.46	39	7.6E-7	1.4E-3
GOTERM_BP_FAT	Positive regulation of cell differentiation	3.46	32	9.6E-6	1.7E-2
GOTERM_BP_FAT	Regulation of cell development	3.46	29	2.2E-5	3.9E-3
Annotation cluster 9					
GO term	**Enriched pathway**	**Enrichment score**	**Count**	**p-value**	**FDR**
GOTERM_MF_FAT	Transcription factor activity	2.59	86	1.0E-4	8.5E-2
SP_PIR_KEYWORDS	DNA_binding	2.59	132	1.5E-3	5.9E-2
SP_PIR_KEYWORDS	Transcription regulation	2.59	140	2.5E-3	3.9E-3

Figure 35: Comparative study of 4SC-202 with other HDAC inhibitors. (A) Principle component analysis (PCA) plot showing the variance and overlap within the replicates of each sample and also between two different samples. (B) Heatmap from the RNA-seq data in L3.6 cells treated with 4SC-202, resminostat and vorinostat showing that gene regulation pattern was similar for resminostat and vorinostat and they clustered together. However effects of 4Sc-202 were weaker in comparison to the other two inhibitors. (C) Venn diagram showing

that a significant portion of genes were overlapping between resminostat and vorinostat but in contrast overlap of any of them with 4SC-202 was very less. Over one thousand genes were overlapping between the three inhibitors. (D) DAVID Gene Ontology analysis was performed on the genes which were found to be overlapping between 4SC-202, resminostat and vorinostat. Enriched GO terms were associated with cell differentiation, angiogenesis, regulation of gene expression and transcriptional regulation.

4.11 Transcriptome wide effects of 4SC-202 differ from LSD1 and HDAC inhibition alone or their co-treatment

To further investigate the mechanism of action of 4SC-202 we performed a transcriptome wide study in L3.6 cells with LSD1 and HDAC1 inhibition alone or their combined inhibition and compared their effects with the effects of 4SC-202. We expected since 4SC-202 has an inhibitory effect against LSD1 and class I HDACs it will have effects on the gene expression similar to combined inhibition of LSD1 and HDACs. To inhibit the LSD1 activity we treated the cells with either siRNA against LSD1 (siLSD1) thus decreasing LSD1 gene expression or treated the cells with a specific LSD1 inhibitor SP2509 (LSD1i) (500 nM) which specifically inhibits the demethylase enzymatic activity of LSD1. Furthermore, for HDAC inhibition we again utilized the pan-HDAC inhibitor Vorinostat at the concentration used in the previous experiment. Cells were treated with the respective inhibitors at defined concentrations for 12 hours after which RNA was harvested and used for high throughput RNA sequencing. Interestingly as depicted in the PCA plot (Fig. 36A), siLSD1 and LSD1i had similar effect and correlated with each other. Furthermore, Vorinostat, siLSD1+Vorinostat and LSD1i+Vorinostat also correlated well with each other meaning that in the combined inhibition Vorinostat had stronger effects than LSD1i or siLSD1. Interestingly, we observed that 4SC-202 samples did not resemble siLSD1/LSD1i or Vorinostat or LSD1i+Vorinostat/siLSD1+Vorinostat combine treatments. To further validate the findings we used a heatmap analysis out of

significantly regulated genes in the different treatment conditions. We applied different cut-off (log2fold) values to select differentially regulated genes in different treatment conditions to have an equal number of regulated genes. We plotted the heatmap using the significantly regulated genes following treatment with 4SC-202 (Fig. 36B), Vorinostat (Fig. 36C) and LSD1i (Fig. 36D) and found that Vorinostat had a significantly stronger effect on gene expression compared to the siLSD1, LSD1i or 4SC-202. Taken together, these results revealed that 4SC-202, a dual inhibitor of LSD1 and class I HDACs acts in a different manner compared to the either LSD1i/siLSD1 or HDAC alone or their co-treatment.
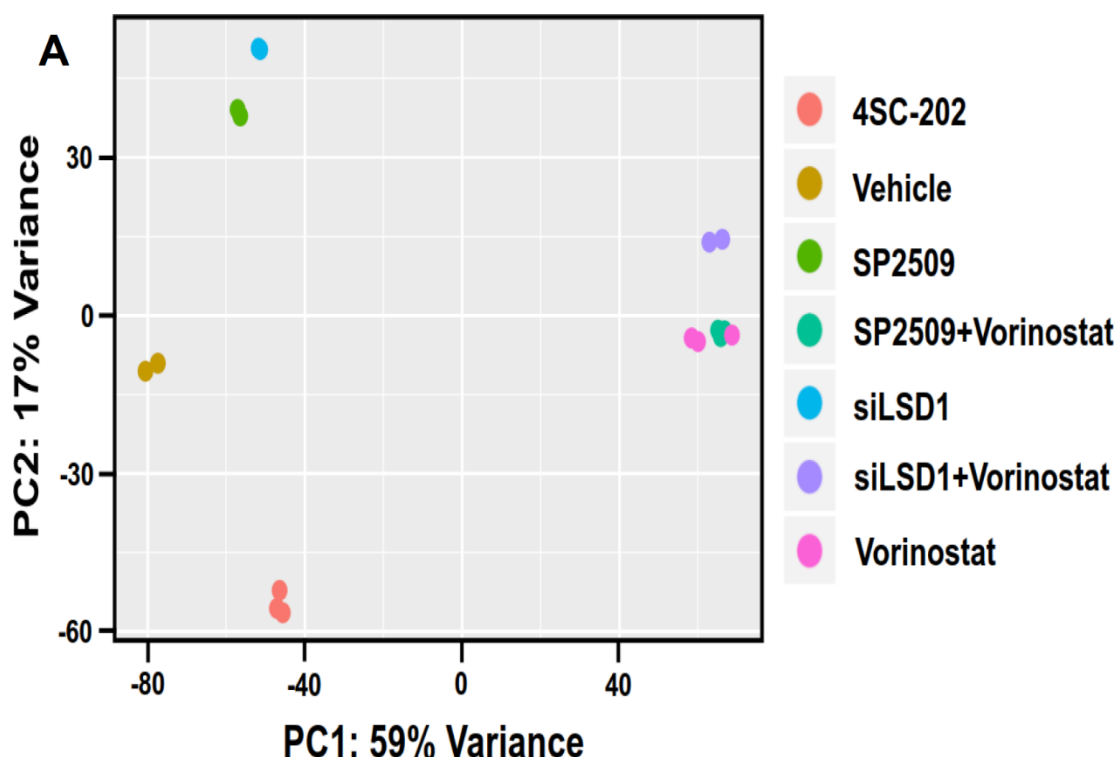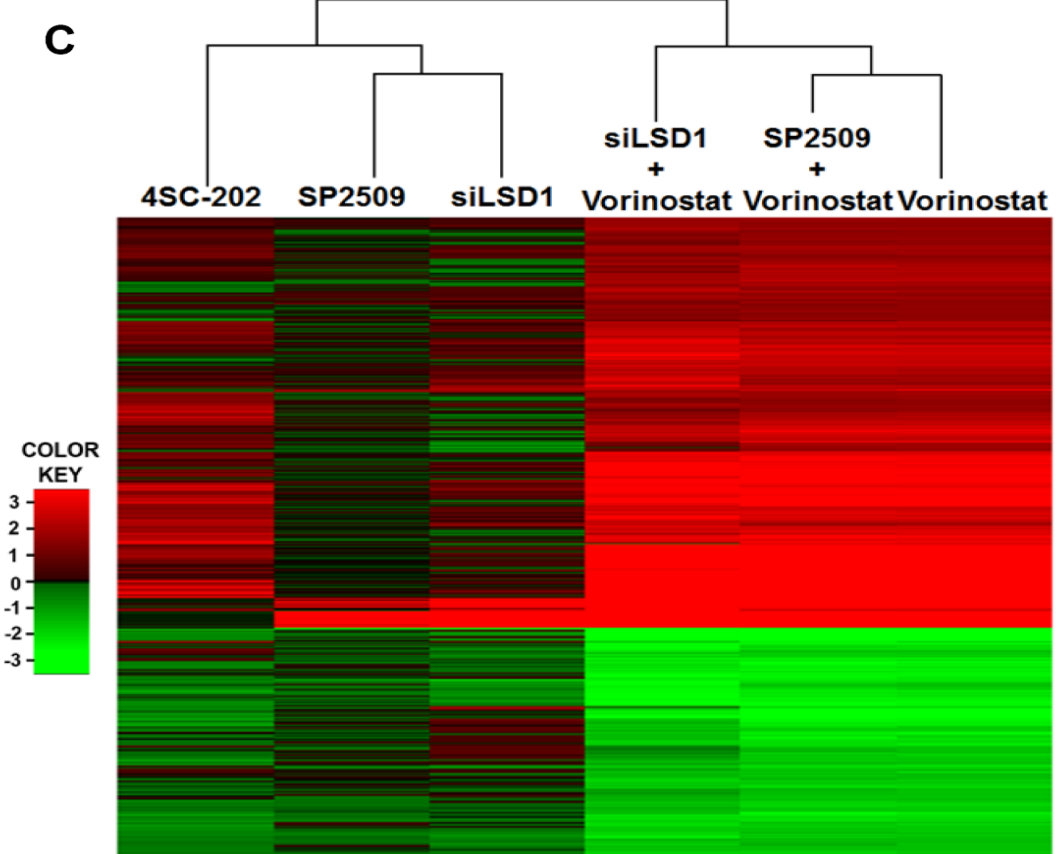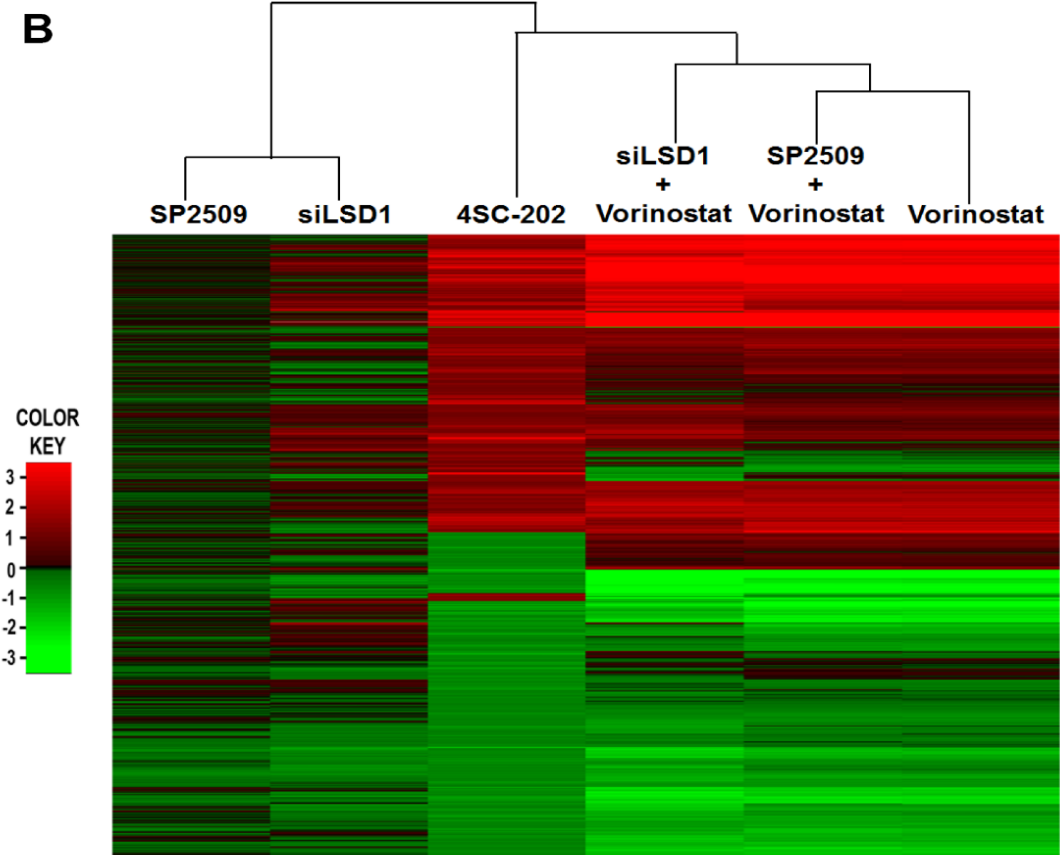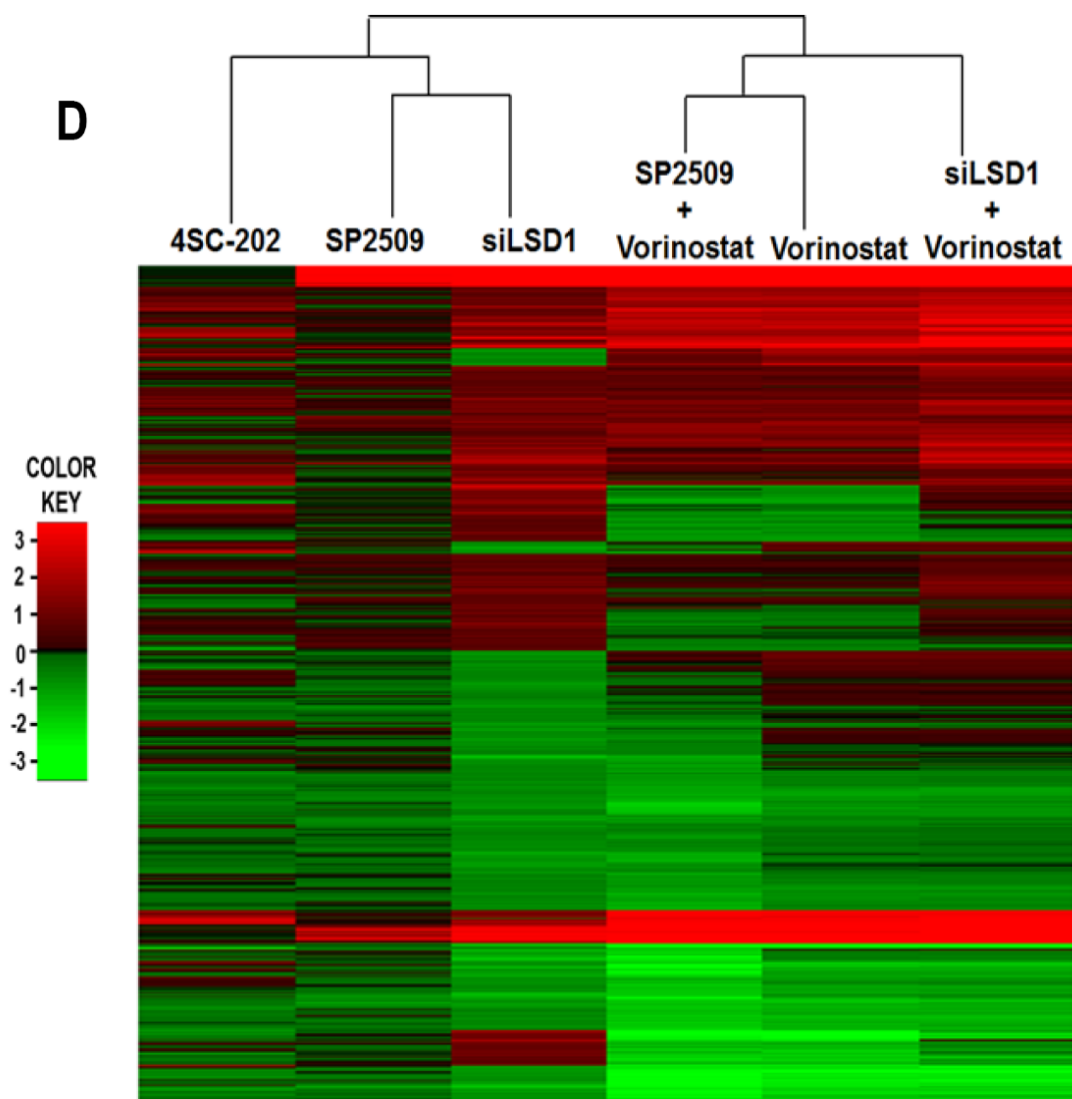
Figure 36: Transcriptome wide effects of 4SC-202 are different compared to LSD1 and HDAC inhibition alone or their co-treatment. (A) PCA plot depicting variability and correlation within and between the samples from the RNA-seq data in L3.6 cells treated with 4SC-202 (1 µM), SP2509 (500 nM), siLSD1 and Vorinostat (1 µM). (B) Significantly (padj ≤ 0.05) up (log2fc ≥ 1) or down (log2fc ≤ -0.7) regulated genes following 4SC-202 treatment were selected and then were compared for their regulation following treatment with Vorinostat, siLSD1, LSD1i, siLSD+Vorinostat and LSD1i+Vorinostat via a heatmap. Clustering was allowed between the samples and also between the genes. (C) Similar to the heatmap in B. Here the significantly (padj ≤ 0.05) up (log2fc ≥ 1.5) or down (log2fc ≤ -1.5) regulated genes were selected based on Vorinostat and then compared with the other treatments. (D) Similar to the heatmap in B or C. Here the significantly (padj ≤ 0.05) up (log2fc ≥ 0.65) or down (log2fc ≤ -0.65) regulated genes were selected based on SP2509 and then compared with the other treatments.

4.12 4SC-202 blocks the colony forming ability of cells in vitro

LSD1 has been reported to be highly expressed in poorly differentiated tumors and its inhibition has been shown to be associated with activation of all-trans-retinoic acid differentiation pathway in leukemia cells (Schenk et al., 2012; Schulte et al., 2009). Additionally, it has been reported that in human embryonic stem cells LSD1 maintains a balance between self-renewal and differentiation (Adamo et al., 2011). We observed that 4SC-202 upregulated the expression of CD24, a marker of differentiated cells and also inhibits TGFβ-induced migration in cells. We were further interested to investigate the effect of 4SC-202 on proliferation. We therefore examined the protein levels of the cell cycle regulator gene p21 (CDKN1A) following 4SC-202 treatment and observed significantly higher p21 levels (Fig. 37A). Elevated level of p21 protein was indicative of an anti-proliferative effect of 4SC-202. Therefore, we next investigated its effect on the colony forming ability of these cells. For this we performed a colony formation assay where we treated the cells with 4SC-202 and allowed them to grow for 7 days. In parallel we also treated the cells with either SP2509 or Vorinostat alone or in co-treatment to check for their independent or cooperative effects and compared them with the effects of 4SC-202. Interestingly, 4SC-202 significantly impaired proliferation of the cells (Fig. 37B). Furthermore, SP2509 or Vorinostat alone had mild effects on the cells but their co-treatment significantly affected the colony forming capacity. Notably, the effect of the co-treatment was similar to the effects of 4SC-202. Taken together these results revealed that 4SC-202 has strong effects on the proliferation of cancer cells which could in part be due to the elevated levels of p21 expression.
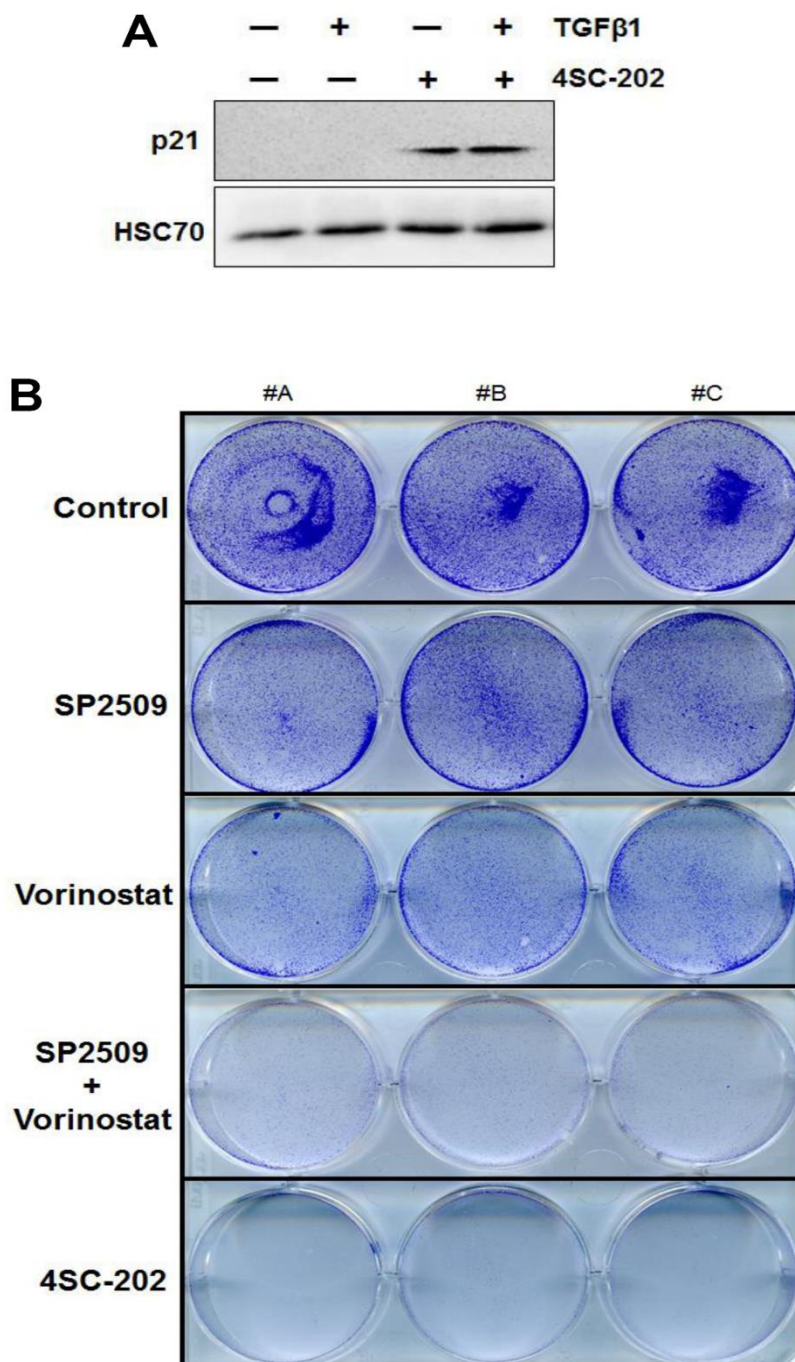
Figure 37: 4SC-202 blocks the colony forming ability of cells in vitro. (A) Western blot results showing the elevated level of p21 (CDKN1A) protein level following 4SC-202 treatment as compared to the control samples (treated with DMSO). HSC70 was used as a loading control. (B) Proliefration assay was performed by seeding approx. 2500 cells per well in 6-well plates and allowing them to grow for 7 days. Cells were then fixed with 100% methanol for 10 minutes at RT and then stained with 0.1% crystal violet for 20 minutes. Cells were treated with either DMSO, SP2509 (500 nM), Vorinostat (1 µM) or 4SC-202 (1 µM). 4SC-202 completely blocked the colony formation in cells which was in synergy with the co-treatment of SP2509 and Vorinostat.

4.13 4SC-202 regresses the tumor growth *in vivo*

To evaluate the efficacy of 4SC-202 as an anticancer drug we investigated its effect on tumor growth and development in a xenograft model using L3.6 cells. Cells were implanted into immune-deficient mice and were allowed to grow to a certain size. After the development of tumors mice were randomly divided into two groups (n=12 per group) and either treated with methylcellulose (vehicle) or 4SC-202 (120 mg/kg) for 4 days (twice per day). All the mice which were implanted with L3.6 cells developed tumors within a time frame of two weeks and treatment was started after the tumor size reached 100 mm^3 (Fig. 38A & B). Tumor growth and size were analyzed for both the vehicle and 4SC-202 treated mice. Since LSD1 has been shown to be overexpressed in many different types of cancer, inhibition of LSD1 would be expected to have deleterious effects on the tumor growth. Consistently we observed a significant reduction in tumor size in 4SC-202-treated mice as compared to the vehicle-treated mice (Fig. 38C). However, further analysis of the dissected tumor tissue samples from the vehicle and 4SC-202 treated mice will be performed to examine the differentiation status of the tumor and expression of other tumor markers.

We also performed immunohistochemistry analysis to investigate the expression status of LSD1 in human pancreatic cancer samples using tissue microarray (TMA). We observed a heterogeneous expression of LSD1 in different tumor samples from moderate to very high level. Representative images of the LSD1 staining in some of the pancreatic cancer samples are shown (Fig. 38D). Taken together from our xenograft study we have determined that 4SC-202 can inhibit the tumor growth *in vivo*. Furthermore, we have also shown that LSD1 is overexpressed in a subset of pancreatic cancers which further supports the hypothesis that targeting LSD1 can be a potential antitumor therapy.

A

5x10⁵ L3.6-cells
(1:1 Matrigel-DMEM mixture)

Xenograft

Mice were allowed
to develop tumor

Tumor

Mice were divided
into two groups

Control
n=12

Treatment
n=12

Methocel
(Vehicle)

4SC-202
120 mg/kg

Mice were treated with 4SC-202
or Vehicle for 4 days (twice
per day) and thereafter tumor
size was measured daily
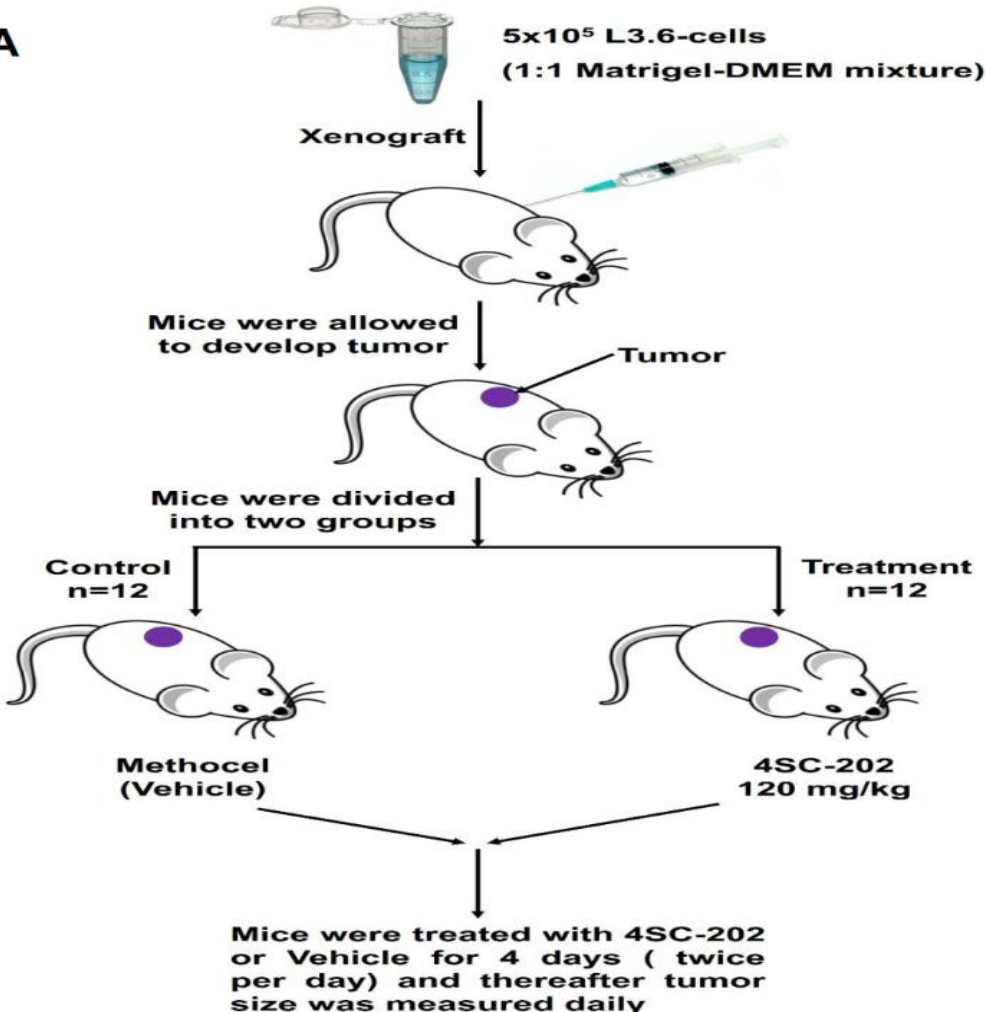
B **Representative images of the tumor-bearing mice**
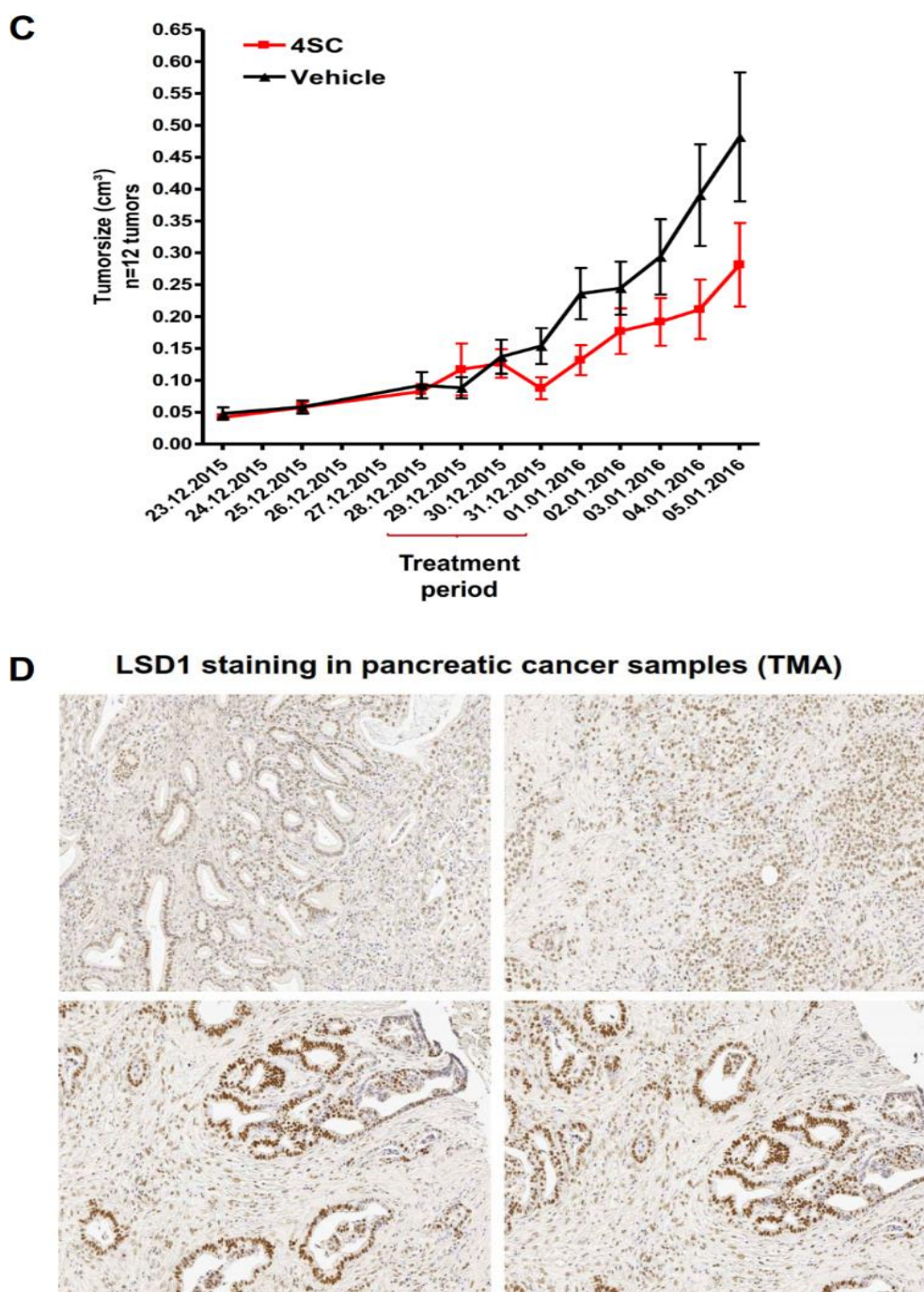
Figure 38: Xenograft study revealed anti-tumor activity of 4SC-202. (A) Schematics of the experimental set-up for xenograft study. L3.6 cells were implanted into the mice and allowed to develop tumor. Mice were either treated with 4SC-202 (120 mg/kg) or vehicle (methylcellulose) for 4 days (twice daily). (B) Representative images of the tumor-bearing mice. (C) Graph showing the tumor size in the treated and control mice over the period of study. Significant reduction in the tumor size was observed in the 4SC-202 treated mice. (D) Representative images of the immunohistochemistry staining for LSD1. Tissue microarray (TMA) containing pancreatic cancer samples from various patients was used to examine the expression status of LSD1.

4.14 4SC-202 leads to a genome-wide enrichment of H3K4me1 and H3K27ac marks

Since LSD1 and HDACs are associated with demethylation (H3K4me1 and H3K4me2) and deacetylation (like H3K27ac) activity respectively, we investigated the effect of their combined inhibition on global levels of H3K27ac and H3K4me1. We performed ChIP-seq for H3K27ac and H3K4me1 with and without 4SC-202 treatment and analyzed their genome-wide distribution. Initially, we checked for the change in levels of H3K27ac and H3K4me1 in western blot and observed a significant increase in the histone marks following 4SC-202 treatment (Fig. 39A & D). After confirming the inhibitory effect of 4SC-202 on HDACs and LSD1 we performed a ChIP-seq experiment and investigated the genome-wide enrichment of the above mentioned histone modifications around the TSS of all genes. As depicted in the aggregate plot around the TSS region (± 5Kb), we observed a significant increase in the genome-wide signal for H3K4me1 and H3K27ac histone marks in 4SC-202-treated condition compared to the control (Fig. 39B & E). These findings were also confirmed by the heatmap profile for the H3K4me1 and H3K27ac marks around the TSS region (± 3Kb) of the genes genome-wide. Interestingly, significantly enhanced signals for H3K4me1 and H3K27ac were observed around the TSS of genes (Fig. 39C & F). Taken together these findings confirmed the inhibitory action of 4SC-202 on HDACs and LSD1 as their respective inhibition would result in enhanced acetylation and methylation of target histone residues. Since LSD1 and HDACs are often found in complexes where they cooperatively lead to the formation of repressive chromatin structure it is likely that they have a common set of target genes, therefore, further analysis of the genome-wide data was performed on the genomic regions that are regulated by both LSD1 and HDAC based on H3K4me1 and H3K27ac ChIP-seq data.
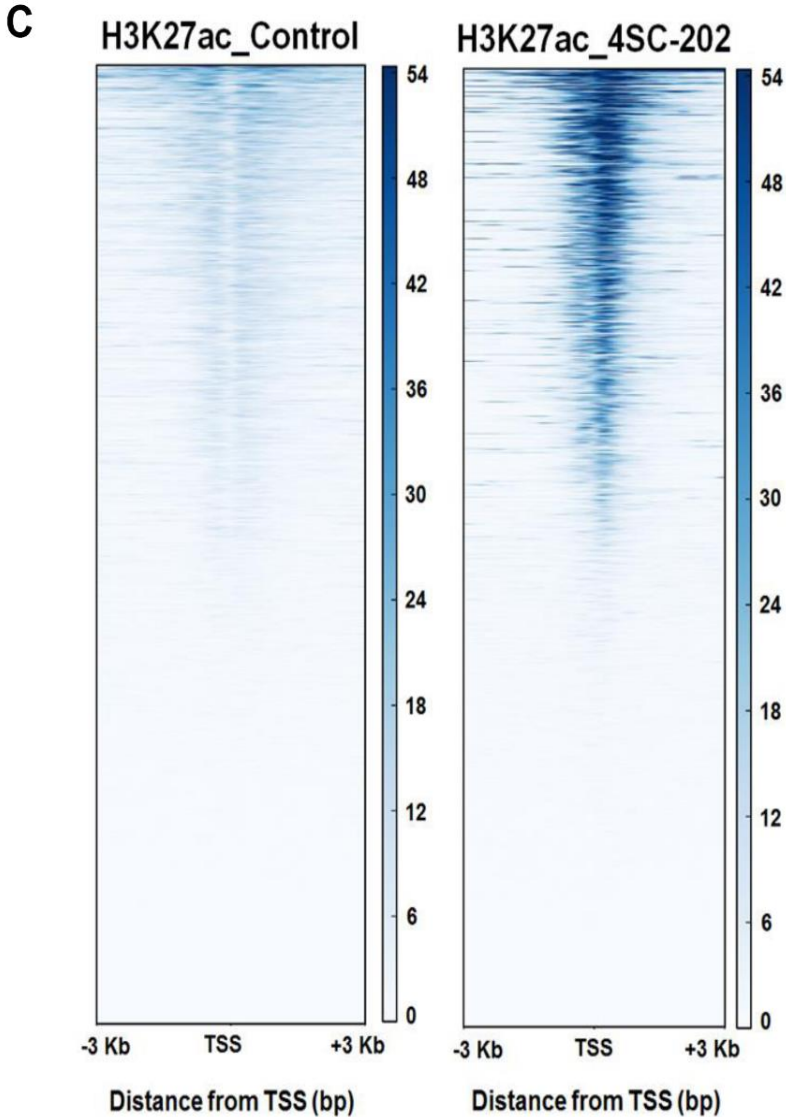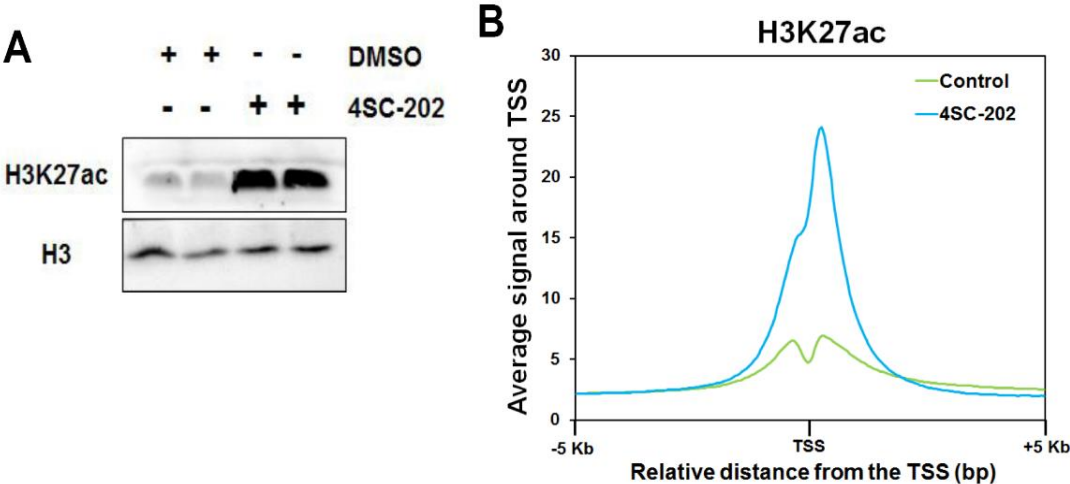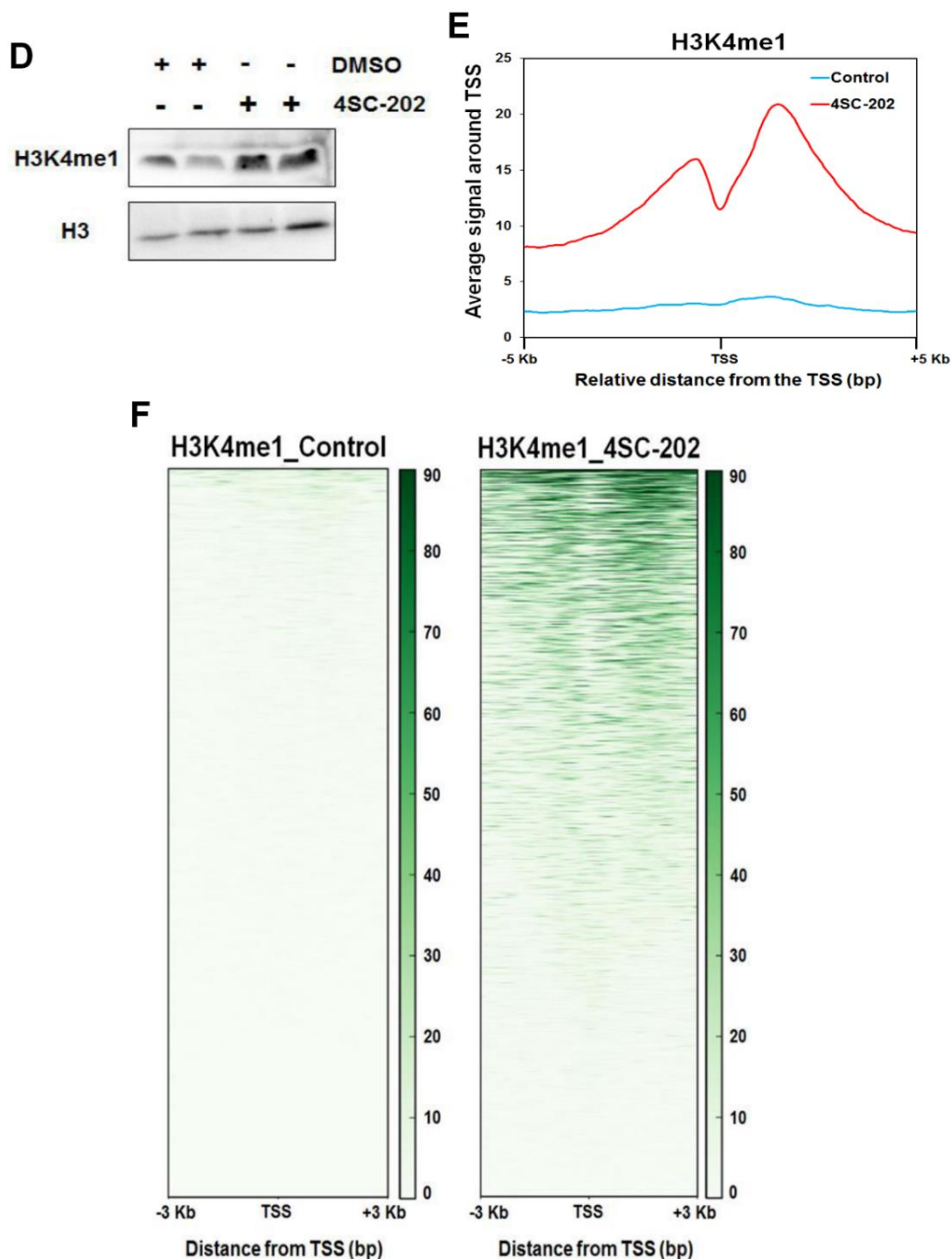
Figure 39: 4SC-202 leads to genome-wide enrichment of H3K4me1 and H3K27ac marks: (A) & (D) L3.6 cells were treated with DMSO or 4SC-202 for 12 hours. Western blotting was performed to check for the protein levels of H3K27ac and H3K4me1 respectively. Samples were loaded in duplicate (samples are from technical replicates of the experiment). Significant enrichment of H3K27ac and H3K4me1 was observed in samples treated with 4SC-202. Total H3 was used as a loading control. (B) & (E) Aggregate profile plot for H3K27ac and H3K4me1 respectively around the TSS (± 5Kb) of the all the genes genome-wide. The enrichment of both histone marks was significantly higher in 4SC-202 treated cells compared to the control (DMSO treated) cells. (C) & (F) Heatmap profile for H3K27ac and H3K4me1 respectively around the TSS (± 3Kb) region of the genes genome-wide. For plotting the heatmap the same maximum intensity was kept for the two treatment conditions for a better comparison. Significantly higher signal was observed for both histone modifications following 4SC-202 treatment.

4.15 H3K4me1 enriched sites are mainly associated with distal intergenic regions

Since H3K4me1 and H3K27ac histone marks have been shown to be associated with active enhancers and promoters we sought to investigate the enrichment of these histone modifications at various genomic locations genome-wide. We performed CEAS for H3K4me1 and H3K27ac histone marks and as depicted in the pie chart plot, fraction of H3K4me1 (33%) and H3K27ac (27%) enriched signals were associated with distal intergenic regions (Fig. 40A). Further, we performed DiffBind (differential binding analysis of ChIP-seq peak data) analysis which enables the determination of the differentially bound genomic regions between different datasets. We performed DiffBind analysis on H3K4me1 ChIP-seq data in comparison to control and 4SC-202 treated samples to check for the regions that have the highest increase in H3K4me1 signals and performed further analysis on those regions. As shown in the correlation plot (Fig. 40B), control samples correlated well with each other and similarly 4SC-202 treated samples correlated with each other but there was no cross-correlation between different conditions indicating that control and treated samples had substantial differences. One of the replicates from control samples was not included in the analysis because the number of reads was significantly low. Furthermore, binding affinity plot (Fig. 40C) revealed that 4SC-202 treated samples had significantly higher differentially bound sites compared to the control samples. In addition we took the differentially bound sites enriched in the 4SC-treated samples compared to the control samples and visualized those genomic regions on the IGV tool. The peak intensity around individual differentially-bound genomic regions was significantly higher in 4SC-202 treated samples compared to control samples (Fig. 40D).

A

H3K4me1_ChIP

4.9 %
3.1 %
1.8 %
1.5 %
1.1 %
0.9 %
1.1 %
1.6 %
2.6 %
33.1 %
48.4 %

■ Promoter (<=1000 bp): 4.9 %
■ Promoter (1000−2000 bp): 3.1 %
■ Promoter (2000−3000 bp): 1.8 %
■ Downstream (<=1000 bp): 1.5 %
■ Downstream (1000−2000 bp): 1.1 %
■ Downstream (2000−3000 bp): 0.9 %
■ 5'UTR: 1.1 %
■ 3'UTR: 1.6 %
■ Coding exon: 2.6 %
■ Intron: 48.4 %
■ Distal intergenic: 33.1 %

H3K27ac_ChIP

5.1 %
2.6 %
1.3 %
1.4 %
1.1 %
0.9 %
1.4 %
1.7 %
2.2 %
27.0 %
55.2 %

■ Promoter (<=1000 bp): 5.1 %
■ Promoter (1000−2000 bp): 2.6 %
■ Promoter (2000−3000 bp): 1.3 %
■ Downstream (<=1000 bp): 1.4 %
■ Downstream (1000−2000 bp): 1.1 %
■ Downstream (2000−3000 bp): 0.9 %
■ 5'UTR: 1.4 %
■ 3'UTR: 1.7 %
■ Coding exon: 2.2 %
■ Intron: 55.2 %
■ Distal intergenic: 27.0 %

B

Condition replicate

Color key

Correlation

Group

Control B
Control C
4SC-202 C
4SC-202 A
4SC-202 B

Control B
Control C
4SC-202 C
4SC-202 A
4SC-202 B

Figure 40: H3K4me1 enriched sites are associated with distal intergenic regions. (A) CEAS was performed on H3K4me1 and H3K27ac ChIP-seq data to find the relative enrichment of ChIP binding regions at various genomic locations. As depicted in the pie chart, 33% of H3K4me1 and 27% of H3K27ac binding regions were confined to the distal intergenic regions compared to the whole genome. (B) Correlation plot showing the appropriate correlation between the replicates of the control and 4SC-202 treated samples. (C) Binding affinity plot showing the statistically significant (FDR ≤ 0.05) enrichment of differentially bound sites in the H3K4me1 ChIP-seq data from 4SC-202 treated samples compared to the control samples. (D) Differentially bound regions obtained from the DiffBind analysis which were enriched for H3K4me1 signals were visualized under the IGV. ChIP-seq profile for a few representative genes are shown in the figure. Significantly enhanced signal intensity and peak height was observed in 4SC-202 treated samples compared to the control samples.

4.16 H3K4me1 enriched distal intergenic elements are associated with key cellular processes

In the recent past many studies have reported the importance of *cis*-regulatory elements that have been shown to play important role in the gene expression. Active enhancers display an enrichment of H3K4me1 and H3K27ac and can be located at a long distance (from ten to hundreds of Kb) from the promoter or TSS region of the genes. From DiffBind analysis we identified several differentially bound H3K4me1 enriched regions and we wanted to know if the differentially bound regions are associated with enhancer elements. For this we selected statistically significantly (FDR ≤ 0.05 and log2fc ≥ 2) enriched DiffBind regions and performed GREAT (Genomic Regions Enrichment of Annotation Tool) analysis to find out the regions that are located ± 5Kb around TSS (basal regulatory domain) plus up to 300Kb distal to TSS. As shown in the bar chart, the majority of the distal intergenic elements were located ± 50Kb to 500Kb away from the TSS region (Fig. 41A). Furthermore GREAT analysis also revealed that the distal intergenic elements were associated with the key biological processes like apoptosis, regulation of cell adhesion (positive) and migration (negative) etc. as shown in the Fig. 41B. Interestingly we also observed that the associated GO cellular component terms were associated with a differentiated cell phenotype like cell-cell adhesion, tight junction and extracellular matrix (Fig. 41C). Enhancers are frequently regions that have a cluster of binding sites (specific DNA sequences or motifs) which provide a platform for recruiting different transcription factors and together regulate the expression of a specific set of genes. We therefore investigated which transcription factors are enriched or associated with the H3K4me1 enriched regions using an online tool ReMap (Griffon et al., 2014). ReMap is a tool that has a broad collection of about 8 million TF binding sites generated from over 200

different TFs using publicly available and ENCODE ChIP-seq datasets. ReMap revealed that 99.31% of H3K4me1 enriched differentially bound sites overlapped with Remap sites (Fig. 41D). Furthermore, ReMap revealed that several important TFs like JUN and FOS that are well-known player in the regulation of key biological processes like cell proliferation and apoptosis were found to be significantly enriched at the given genomic regions (Fig. 41E). To further investigate the enrichment of TFs associated with the H3K4me1 enriched genomic regions we performed motif analysis using the SeqPos motif tool and found a significant enrichment of the motif for the members of AP1 complex of transcription factors (FOS, FOSB, FOSL2, JUN, JUNB, JUND), thus validating the enrichment of these TFs at the H3K4me1 associated genomic regions (Fig. 41F). Taken together these results revealed that 4SC-202 leads to the enrichment of H3K4me1 mark at the distal intergenic regions accompanying putative enhancers associated with various key TFs. Furthermore, the TFs that have been found to be enriched at the given sites may play a role in regulating subset of genes involved in promoting differentiation, apoptosis and regulating cell proliferation.

D

Overlapping regions

0.69 %

Number of peaks occupied

0 200 400 600 800 1000 1200 1400 1600 1800 2000 2200 2400 2600

■ Unmapped ■ Mapped

E

Enriched TFs in intersection

FOS	1 077.21
JUN	777.043
STAT3	743.55
FOSL2	588.532
JUND	523.364
MYC	449.692
CEBPB	393.593
NFKPB	288.676
NKX2-1	254.326
CDX2	238.883

0 100 200 300 400 500 600 700 800 900 1000 1100 1200 1300

-log10(E-value)

F

AP1/Jun/JunB/JunD/Fos/FosB/Fosl1

Information content

2

1.5

1

0.5

0

TGAcTcA

Z-score: -7.897
-10log(pval): 341.798

1 2 3 4 5 6 7 8 9

Position

Figure 41: H3K4me1 enriched distal intergenic elements were associated with key cellular processes. (A) GREAT (Genomic Regions Enrichment of Annotations Tool) analysis of H3K4me1 enriched genomic regions. Bar graph is showing that majority of the H3K4me1 enriched distal intergenic regions were located 5 to 500 Kb away from the TSS region of the genes. (B) Significantly enriched (-log10 binomial p-value) GO Biological Processes associated with the distal intergenic regions are shown. Pathways regulating cell apoptosis, organelle organization, immune response and defense processes were found to be enriched. (C) Significantly enriched (-log10 binomial p-value) GO cellular components associated with the distal intergenic regions are shown. GO terms related to maintenance of extracellular matrix, cell-cell adhesion and cell junction were found to be enriched. (D) Overview of the overlapping regions between the H3K4me1 enriched genomic regions form the ChIP-seq data and genomic regions from the ReMap tool. As depicted in the plot, 99.31% of the regions were found to be overlapping (marked in blue) while 0.69% regions did not map (marked in red). (E) Top-10 significantly enriched TFs associated with the H3K4me1 enriched genomic regions are shown in the bar graph. The enrichment was based on the overlap between the provided genomic regions and the genomic sites from the ReMap (based on the publicly available and ENCODE datasets). (F) Motif analysis was performed on the H3K4me1 enriched genomic regions using SeqPos motif tool (version 1.0.0) of the Galaxy/Cistrome. Significant enrichment of the AP1-TF (TF-complex containing JUN and FOS TFs) was observed.

5. Discussion

DISCUSSION-I

Tumor metastasis involves the accumulation of mutations that allow cancer cells to survive during uncontrolled proliferation, migration and colonization at distant sites to ultimately give rise to metastatic secondary tumors (Chambers et al., 2002; Fidler, 2003b; Gupta and Massagué, 2006). One of the most important aspects of formation of the metastatic tumor is the ability of cancer cells to lose cell-cell adhesion i.e. break-open the barrier of the extracellular matrix (Kalluri and Weinberg, 2009; Thiery and Sleeman, 2006). Cancer cells display cellular plasticity that allows them to undergo phenotypic changes through a process termed epithelial-to-mesenchymal transition (EMT) that can be induced through various external stimuli including growth factors and cytokines like TGFβ (Chaffer and Weinberg, 2011; Thiery, 2002). However, to be able to successfully colonize at distant sites and give rise to a secondary tumor cancer cells must revert back to an epithelial phenotype by undergoing mesenchymal-to-epithelial transition (MET) (Hugo et al., 2007; Yao et al., 2011). TGFβ signaling is one of the important regulators of EMT and has been often found to be overexpressed in metastatic cancers (Derynck et al., 2001; Massagué et al., 2000). Many factors play a role in mediating the anti-proliferative and pro-metastatic roles of TGFβ but there is no concrete evidence that could dissect the dichotomy of the dual role of TGFβ under normal and cancer conditions (Massagué, 2008). Furthermore, in recent years epigenetics has been shown to play a major role during normal as well as cancer development, therefore it is indispensable to study the epigenetic factors in the regulation of EMT and MET during cancer progression and metastasis (KIESSLICH et al., 2013; Tam and Weinberg, 2013; Wu et al., 2012a). Moreover, understanding

the mechanisms behind the cancer metastasis may open new avenues for cancer treatment and therapeutics.

5.1 KLF10 in tumor development and disease prognosis

One of the causative factors in the transformation of a normal cell to a cancer cell is the loss of function of tumor suppressor genes. Most of the tumor suppressor genes encode proteins which are responsible to keep a regulatory brake on the cell proliferation. Understandably there is a loss of function or inactivation of tumor suppressor genes during tumor development thus allowing them to undergo uncontrolled cell proliferation. Among various classes of tumor suppressor genes are genes encoding regulators of apoptosis. KLF10, a well-known tumor suppressor, has been shown to induce apoptosis in pancreatic cancer (Panc1), hepatoma cell line (Hep3B) and oligodendroglia cells (OLI-neu) (Bender et al., 2004; Chalaux et al., 1999; Ribeiro et al., 1999; Tachibana et al., 1997). In our study we explored the expression status of KLF10 primarily in lung and breast cancer. Interestingly, in correlation with its tumor suppressor role we found that KLF10 expression levels were significantly downregulated across different datasets in lung cancer as well as breast cancer samples in comparison to the control sample. This finding fits well with the previous reports showing that KLF10 expression in breast cancer is inversely correlated with breast cancer stage with highest expression in normal breast tissue and minimum or complete loss of expression in advanced stage and invasive breast cancer (Reinholz et al., 2004; Subramaniam et al., 1998).

It is a well-known fact that, in the path of conversion of a normal cell into a cancer cell, tumor suppressor genes play an inhibitory role unlike proto-oncogenes. Tumor suppressor genes generally operate by keeping a check on cell proliferation and induce apoptosis to maintain the tissue homeostasis. However, a loss of function

mutation or inactivation of tumor suppressor genes results in uncontrolled proliferation which is considered to be one of the most important hallmarks of cancer (Hanahan and Weinberg, 2011). To address the tumor suppressor function of KLF10 in an *in vivo* system we have utilized *Klf10* knock-out mice carrying a loss of function mutation in the KLF10 gene. Previously it has been shown that the commonly used laboratory chemical carcinogen DMBA induces lung tumor formation in mice (Duro de Oliveira et al., 2013). In our study a significantly higher incidence of lung tumor was observed in *Klf10* knock-out mice as compared to the wild type mice upon exposure to DMBA. Notably, *Klf10* knock-out mice did not carry any other mutational background like KRAS or EGFR.

5.2   KLF10 and dichotomy of TGFβ signaling

Impairment of various signaling pathways regulating cellular functions like cell proliferation, differentiation and apoptosis is a frequent occurrence during malignant transformation. One such signaling pathway is TGFβ signaling, which is governed by various factors and perturbations in TGFβ signaling during tumorigenesis and tumor progression has been repeatedly reported (Yang and Weinberg, 2008b). The effects of TGFβ signaling are more dependent on cellular contexts and this paradigm has eluded the scientific world for many years (Massagué, 2012). How cells read and respond to the effects of TGFβ signaling in different manners has been a mystery for more than half a century and it still is a contradiction. The scientific community has been trying to find the factors that can dissect the dichotomy of the dual role of the TGFβ signaling during cancer development and progression.

KLF10 has been previously reported to be a central player in TGFβ signaling where it either suppresses the inhibitory SMAD7 gene or upregulates the expression of the activating SMAD2 gene, thus enhancing the TGFβ signaling (Johnsen et al., 2002a,

2002b). We have performed a transcriptome-wide study in two different cell lines to determine if the effects of KLF10 on TGFβ regulated genes is a general phenomenon or is cell or cancer-type dependent. Transcriptome wide results in cell lines representing different cancer types showed the same result implying that effect of KLF10 on TGFβ signaling is a general phenomenon and not cancer-type specific. KLF10 is a tumor suppressor while TGFβ has tumor suppressing as well as tumor promoting functions. Previous studies have also shown that overexpression of KLF10 can mimic the anti-proliferative function of TGFβ (Johnsen et al., 2004) while the loss of KLF10 resulted in a pro-proliferative effects of TGFβ in mouse embryonic fibroblasts. Therefore, it can be speculated that KLF10 can be a central factor that could dissect the dual role of TGFβ.

5.3   KLF10 responsive genes are associated with differentiation

In the transcriptome data we have shown that a significant portion of the TGFβ regulated genes were overlapping with KLF10 affected genes indicating that KLF10 indeed targets genes involved in carrying out TGFβ functions. However, we were interested to elucidate whether the KLF10 responsive genes were associated with tumor suppressing or tumor promoting pathways. Cellular pathways like cell adhesion, extracellular matrix and cell migration were found to be significantly enriched GO terms in the Gene Ontology analysis. Since these pathways are associated with a cell that is rather in a differentiated state and these GO terms were enriched upon KLF10 depletion it points towards the association of KLF10 with a differentiated cellular state. Furthermore, we also observed the enrichment of gene sets relevant for the formation of poorly differentiated and metastatic cancer upon KLF10 depletion. Enrichment of pathways related to cell differentiation on the one hand versus pathways related to metastatic cancer on the other hand speaks for the role of KLF10 in affecting or

regulating both the tumor suppressing as well as tumor promoting functions of the TGFβ.

5.4 Role of KLF10 in cancer metastasis

To be able to invade the surrounding tissues or metastasize to distant sites cancer cells must lose their polarity, cell-cell contacts and become more mesenchymal by undergoing EMT. TGFβ is a potent inducer of EMT and thereby exerts its tumor promoting function in later stages of cancer. In our study, various results indicated that KLF10 could play a role in inhibiting the formation of metastatic cancer by regulating the genes associated with TGFβ-induced EMT.

5.4.1 KLF10 inhibits TGFβ-induced EMT

One of the hallmarks of initiation of EMT is the loss of expression of E-cadherin. Loss of E-cadherin marks the onset of dissolution of cell-cell adhesion and extracellular matrix thus allowing the cells to break-free and be able to migrate. We observed that KLF10 depletion itself was able to cause a change in cellular morphology making the cells more elongated and mesenchymal-like. However, KLF10 depletion in TGFβ-induced cells resulted in a complete loss of cell-cell contact and attainment of an elongated mesenchymal-like phenotype. Additionally, KLF10 depletion resulted in a significant loss of E-cadherin mRNA and protein expression which was coupled with elevated expression of genes associated with a mesenchymal phenotype.

Matrix metalloproteinases (MMPs) are one of the important classes of proteolytic enzymes and their function of protein degradation regulates various cellular processes. Apart from their involvement in maintaining tissue homeostasis, regulating cell growth and differentiation they have also been implicated in invasion and metastasis of cancer (Egeblad and Werb, 2002; Kessenbrock et al., 2010). We have

noted significantly elevated level of MMP2 and downregulation of MMP7 following KLF10 depletion. Interestingly, various studies have reported the significance of MMP2 in the context of cancer development and progression. MMP2 is indispensable for angiogenesis and depletion of MMP2 has been shown to be associated with decreased angiogenesis in chicken chorioallantoic membrane model and also in *Mmp2* deficient mice (Itoh et al., 1998). Furthermore, *Mmp2* deficient mice have been shown to form less colonies or metastasis in lung compared to the wild type mice (Fang et al., 2000). Consistent with our findings, our transcriptome data suggest that KLF10 is indeed able to block TGFβ-induced EMT, which further strengthens the possibility that KLF10 could be a critical factor that can dissect the dual role (tumor suppressing and tumor promoting) of TGFβ signaling in cancer.

5.4.2 KLF10 regulates EMT by targeting SNAI2

Since our transcriptome-wide data indicated a potential role for KLF10 in regulating TGFβ-induced EMT it was important to identify the genome-wide binding sites of KLF10 in order to find potential target genes. Interestingly, in our genome-wide study for KLF10 we determined that most of the KLF10 binding sites were confined to the promoter region and coding exons of genes while very few were at the distal intergenic regions. One of the interesting and most important findings from our genome-wide study was that SNAI2 emerged as a direct KLF10 target gene. A significant KLF10 peak was observed around the promoter region of the SNAI2 gene. SNAI2 is one of the most important transcription factors involved in the initiation of EMT by repressing the hallmark epithelial marker gene E-cadherin (Bolós et al., 2003; Hajra et al., 2002; Naber et al., 2013). SNAI2 has been shown to be frequently overexpressed in different types of cancer, mainly metastatic or advanced stage cancer. Especially in metastatic breast cancer, SNAI2 expression was found to be inversely correlated with the E-

cadherin expression and is related to poor prognosis (Côme et al., 2006; Jethwa et al., 2008; Pérez-Mancera et al., 2005). SNAI2 has also been shown to promote tumor cell migration and invasion in lung adenocarcinoma and to generate cells with properties of cancer stem-like cells in breast cancer (Bhat-Nakshatri et al., 2010; Shih, 2005).

To rule out the possibility of SNAI2 being a target of KLF10 specifically in lung cancer we have performed a ChIP analysis in two other cancer cell lines. We consistently observed the enrichment of KLF10 binding on the TSS region of the SNAI2 gene in A549, Panc1 and MDA-MB-231 cells, whereas the binding in the transcribed region was equivalent to the background. Furthermore, there was a significant upregulation of SNAI2 mRNA and protein levels upon KLF10 depletion which further strengthened SNAI2 as the KLF10 target gene in its quest to block TGFβ-induced EMT.

5.4.3 KLF10 represses SNAI2 by an epigenetic mechanism

EMT involves tightly coordinated and reversible changes in the expression of epithelial and mesenchymal marker genes and this plasticity has recently been credited to a large extent to epigenetic changes (Serrano-Gomez et al., 2016; Wang and Shang, 2013). Different classes of epigenetic regulators (readers, writers or erasers) are recruited by EMT regulators (EMT-TFs) to cause gene repression (of epithelial markers) or gene activation (of mesenchymal markers). Histone deacetylation is an epigenetic event that causes gene de-activation and is carried out by histone deacetylases (HDACs). HDACs remove the acetyl groups from the lysine residues of the histones thus making the chromatin more compact and limiting access for DNA binding transcription factors as well as transcriptional machinery to access the chromatin. HDAC1 is often found in complex with mSin3A together in a co-repressor complex and previous studies have reported that KLF10 also forms a co-repressor

complex with mSin3A to recruit HDAC1 in order to repress its target genes (Jin et al., 2012; Zhang et al., 2001). We have shown that KLF10 recruits HDAC1 to the SNAI2 promoter which leads to the transcriptional repression of the gene. Significantly reduced occupancy of HDAC1 on the SNAI2 promoter was observed upon KLF10 depletion uncovering a mechanism behind the transcriptional regulation of SNAI2 by KLF10. Furthermore, previous studies have reported that in response to external stimuli like TGFβ and other growth factors, SNAI2 and other EMT-TFs like SNAI1 (SNAIL) and ZEB1 recruit HDACs to the epithelial genes (especially E-cadherin) to repress them to initiate the EMT (Aghdassi et al., 2012b; von Burstin et al., 2009; Peinado et al., 2004a). HDACs (especially HDAC1 and HDAC2) have been shown to be overexpressed in different types of cancer and the most commonly targeted gene is *p21* which plays an important role in regulating processes like cell proliferation, differentiation and apoptosis (Hrzenjak et al., 2006; Huang and Guo, 2006; Huang et al., 2005; Sambucetti et al., 1999; Song et al., 2005).

The switch between histone acetylation and deacetylation is tightly coupled with active and repressed chromatin respectively. Active chromatin is marked by certain types of histone acetylation marks like H3K9ac and H3K27ac which contribute to an open and uncondensed chromatin structure accessible to the transcriptional machinery. We found that depletion of KLF10 leads to significantly enhanced H3K9 and H3K27 acetylation on the promoter of SNAI2 and correlates with the decreased recruitment of HDAC1. Consistent with the finding that KLF10 forms a co-repressor complex with HDAC1, our genome-wide data for H3K9 and H3K27 acetylation marks show an overall increase in these acetylation marks on KLF10-bound genes genome-wide following KLF10 depletion. It has been reported that KLF10 recruits HDAC1 to the target genes and causing their repression through altering the chromatin structure due

to removal of acetylation marks by HDAC1 (Jin et al., 2012). Furthermore, this is the first genome-wide study of KLF10. Together these results uncovered the epigenetic mechanism by which KLF10 targets the important EMT-TF SNAI2 to block the TGFβ-induced EMT.

To summarize, we propose that in the absence or low expression of KLF10, following TGFβ ligand binding to its receptors, SMAD complex binds to and activate the SNAI2 gene expression which in turn initiates the EMT program (Fig. 42).
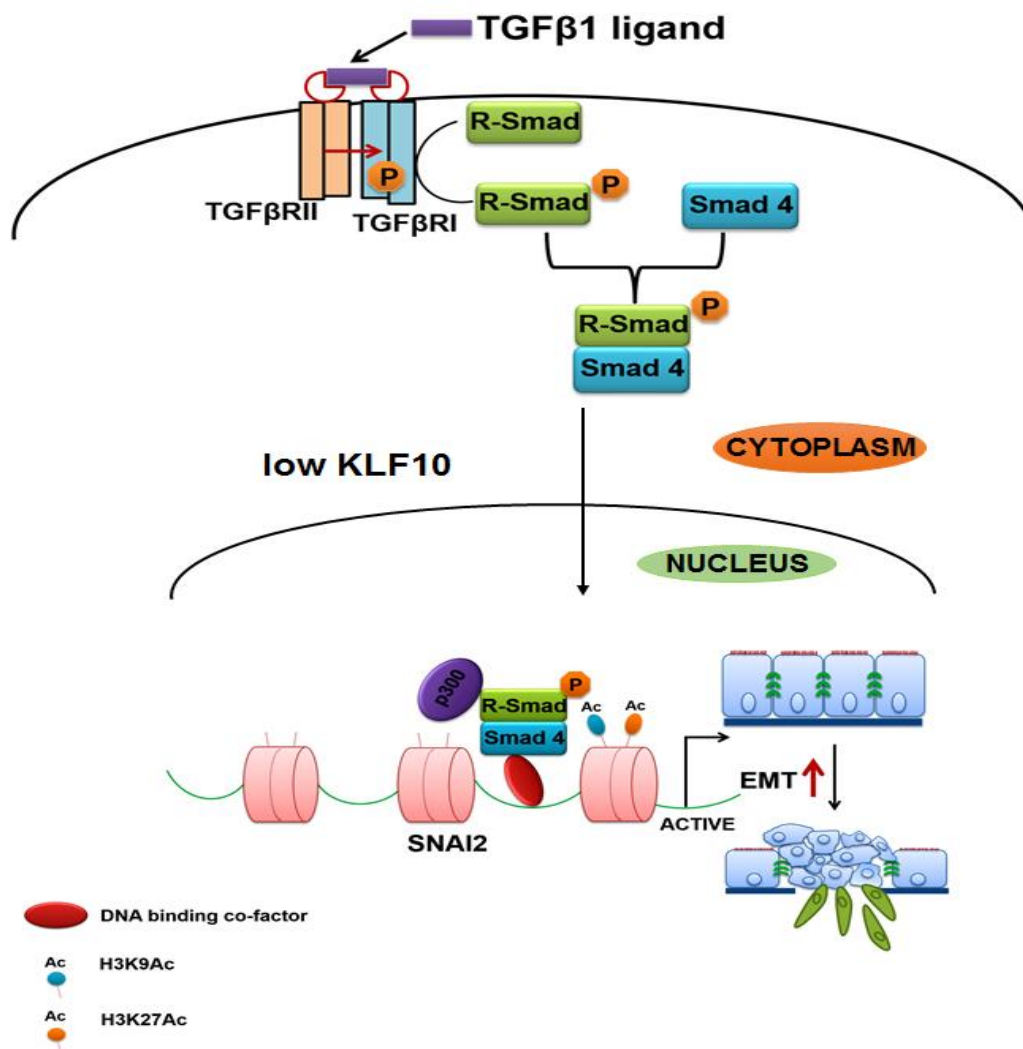
Figure 42: SNAI2 is activated in the absence of KLF10. In the absence of KLF10, in response to TGFβ stimulation cancer cells have elevated expression of SNAI2 which in turn then initiates the EMT program.

In the presence of KLF10, following TGFβ stimulation the SMAD complex activates KLF10 gene expression which in turn binds to the SNAI2 promoter and recruits HDAC1 leading to deacetylation and consequently transcriptional repression. Repressed SNAI2 cannot initiate the EMT program. Thus, cancer cells would no longer be able to metastasize (Fig. 43).
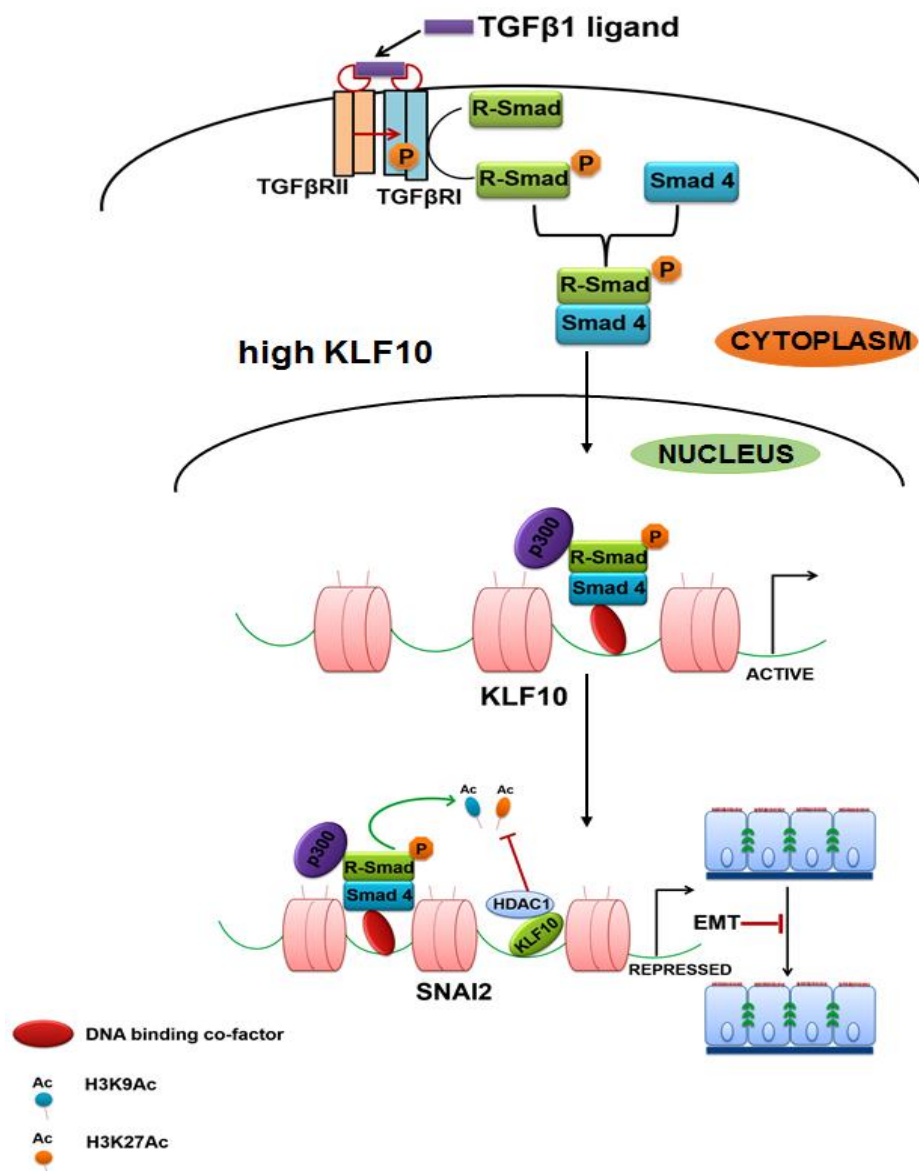
Figure 43: KLF10 activation by TGFβ results in repression of SNAI2. Following TGFβ stimulation the SMAD complex binds to and activates the transcription of KLF10. High KLF10 expression results in inhibitory effects on SNAI2 expression. KLF10 binds to and recruits HDAC1 to the SNAI2 promoter leading to deacetylation and transcriptional repression. As a consequence cancer cells can no longer undergo EMT and cannot metastasize.

5.5 KLF10 specifically targets certain TGFβ regulated genes

Previously it has been shown that KLF10 enhances TGFβ signaling through regulation of SMAD2 and SMAD7 expression (Johnsen et al., 2002a, 2002b). KLF10 has been shown to mimic the anti-proliferative effect of TGFβ (Johnsen et al., 2004) but nothing is known about role of KLF10 in pro-metastatic effects of TGFβ. In our transcriptome-wide study we found that KLF10 specifically upregulates some and downregulates certain TGFβ regulated genes. Additionally, we also found that significant fraction of TGFβ regulated genes were unaffected by KLF10 depletion implying that they are not regulated by KLF10. This finding could support the hypothesis that KLF10 upregulates the genes involved in anti-proliferative effects of TGFβ whereas downregulates those involved in pro-metastatic effects of TGFβ, thus fine tune the balance between its contrasting functions. Furthermore, it is relevant to know what kind of pathways are regulated by KLF10 affected genes because that could further shed a light on mechanism behind the tumor suppressor action of KLF10 and importantly if KLF0 expression level correlates with metastatic cancer.

5.6 KLF10 and cancer stem cell theory

Cancer stem cells (CSCs) are defined as a small subset of cells within a heterogeneous tumor population that have the capacity for self-renewal and sustaining the tumor growth (Clarke et al., 2006). These CSCs can be isolated based on specific cell surface marker expression and the most striking feature that distinguishes CSCs from the rest of the tumor cells is that they can repopulate themselves to generate parent tumors with similar heterogeneity (Eramo et al., 2007; Hermann et al., 2007; O'Brien et al., 2007; Singh et al., 2004). Cells undergoing EMT have been shown to acquire a characteristic similar to CSCs and were shown to generate tumors when implanted into a mouse model (Mani et al., 2008). We have found that KLF10 not only

inhibits the TGFβ-induced EMT but also blocks the migration capacity of the cells stimulated with TGFβ, which is one of the fundamental steps towards the formation of a metastatic tumor. Since CSCs are indispensable for repopulating the metastatic tumor, blocking their migratory potential may hinder them from reaching the distant site, consequently they will not be able to form metastases. Whether KLF10 depletion affects the enrichment of the CSC pool or not needs to be addressed in future studies.

5.7 Therapeutic relevance of KLF10 for cancer treatment and prognosis

Prognostic markers play an important role in cancer treatment by providing a way to accurately classify the disease events in a patient and whether or not a patient may survive the disease. We have elucidated that KLF10 expression may help to predict the disease outcome in lung and breast cancer (subtype Luminal B) patients. Low KLF10 expression was found to be associated with poor overall survival in lung cancer patients and poor disease and metastasis free survival (DMFS) in breast cancer patients. Most of the cancer related deaths, especially for lung cancer, are due to metastatic or advanced stage cancer, which is to a large extent based on a lack of proper molecular or prognostic markers that allow the early diagnosis of the cancers likely to metastasize. Based on our findings it can be speculated that KLF10 can indeed be established as a prognostic marker especially for lung cancer. Therefore, an in-depth investigation into this direction will be of significant therapeutic relevance.

Furthermore, EMT has been shown to generate cells with stem-like characteristics, implying that these cells can stay dormant for a long time without undergoing division. Most chemotherapeutics target the actively dividing cell, and hence are not as effective against CSCs because they do not divide actively and are thus chemo-resistant and can repopulate later to give rise to metastatic tumors. Since we have shown that KLF10 can block EMT we hypothesize that analyzing the KLF10 expression level can

be useful in categorizing tumors as therapy responder and non-responder. Tumors with low or no KLF10 expression will most likely have an enriched pool of CSCs and may be more likely to progress to metastatic disease whereas tumors with high KLF10 expression will likely be in a differentiated state with less probability of metastasizing. For this reason further in-depth research in this direction is required.

Taken together, we have uncovered the previously unknown function of KLF10 in inhibiting TGFβ-induced EMT. Furthermore, in our transcriptome-wide study in two different cell lines we have shown that KLF10 affects a significant number of TGFβ regulated genes and most of the overlapping genes (KLF10 and TGFβ responsive) were found to be relevant for pathways regulating EMT and metastasis. Further, we have establish SNAI2 as a direct target gene for KLF10 through genome-wide ChIP-seq (in A549 cells) and ChIP (in A549, Panc1 and MDA-MB-231 cells) studies. Mechanistically our study provides evidence that KLF10 forms a corepressor complex with HDAC1 and recruits it to the SNAI2 promoter to cause transcriptional repression. Furthermore, genome-wide analysis of the active histone marks H3K9ac and H3K27ac revealed significant increases in these two acetylation marks on the SNAI2 promoter following KLF10 depletion consistent with the recruitment of HDAC1 by KLF10 to repress SNAI2 expression. As KLF10 expression was found to be inversely correlated to the disease outcome in lung and breast cancer patients with low KLF10 expression resulting in poor overall survival, our study holds a clinical relevance. Further investigation might help to establish KLF10 as a prognostic marker.

DISCUSSION-II

Epigenetic changes are pivotal for the regulation of gene expression and recent studies have highlighted the perturbation of various epigenetic modifiers during cancer development and progression, thus pointing towards cooperation between genetic and epigenetic events during malignant transformation (You and Jones, 2012). Cancer is now regarded as a disease of genetic and epigenetic abnormalities but, in contrast to genetic mutations, epigenetic alterations are reversible in nature. The reversible nature of epigenetic modifications offers a great possibility to design targeted therapies against specific modifiers and recently numerous such epigenetic drugs have been designed and are undergoing clinical testing as a targeted drug therapy (Dawson and Kouzarides, 2012; Yoo and Jones, 2006). Many anticancer drugs target the bulk of the cancer cells, but a small subset of cells (termed as cancer stem-like cells) are often resistant to these drugs and overtime develop resistance to other drugs as well and thus cannot be targeted anymore (Bozic et al., 2013; Diaz Jr et al., 2012; Komarova and Wodarz, 2005). Combination therapy has been proposed to be an answer to overcome drug resistance in cancer cells considering the likelihood of a cell being resistant to all the drugs used in combination (two or more) is smaller and may offer a better chance for the success of the treatment (Komarova and Boland, 2013).

5.8 4SC-202 promotes differentiation

One of the important reasons of the failure of therapeutics against cancer is the stem cell characteristic of a small subset of cancer cells that possess resistance against drugs (Dean et al., 2005). Therefore, it is essential to design drugs that can target these cancer stem-like cells by promoting a differentiated phenotype that no longer has the ability to repopulate the tumor. We found that 4SC-202 blocks the induction of EMT in cells stimulated with TGFβ and forces them towards an epithelial phenotype

as confirmed by enhanced expression of epithelial genes. Simultaneous inhibition of the EMT-TFs further supports a role of 4SC-202 in blocking the EMT-promoting effects of TGFβ and in turn promoting an epithelial cell phenotype. LSD1 has previously been shown to be required for the SNAI1 mediated epithelial-to-mesenchymal transition (Lin et al., 2010). Another study has reported that LSD1 inhibition attenuates the SNAI1-mediated EMT, while combined inhibition of LSD1 and HDAC1 completely blocks it (Javaid et al., 2013). Furthermore, LSD1 has been shown to be responsible for causing chromatin reprogramming during TGFβ-induced EMT, which favors the transcriptional repression of epithelial genes (McDonald et al., 2011). Moreover, EMT has also been implicated in promoting migration of cancer cells and in our study we have shown that 4SC-202 completely blocks the TGFβ-induced migration in cells. It has also been reported that inhibiting the interaction between LSD1 and SNAI1 abrogated the invasive potential (Ferrari-Amorotti et al., 2013), whereas overexpression of LSD1 results in enhanced migration and invasion of cancer cells (Lv et al., 2012). Furthermore, 4SC-202 strongly enhances the expression of the CD24 gene, a well-regarded cell surface marker for differentiation which has been reported to be downregulated or absent in cancer stem cells (Jiang et al., 2011; Petkova et al., 2013). Additionally, LSD1 was shown to play an important role in maintaining a balance between self-renewal and differentiation potential of human embryonic stem cells by targeting key developmental genes involved in differentiation through its demethylase activity (Adamo et al., 2011). In the same study loss of LSD1 has been shown to cause activation of key differentiation genes while its overexpression results in rescue of the effect implying that LSD1 is essential for maintaining an undifferentiated state.

5.9   4SC-202 and cell cycle regulation via p21

Any anticancer drug would be expected to put a brake on the most important driving force of cancer development which is the uncontrolled cell proliferation with resistance to apoptosis. One of the important genes that has been shown to play a role in controlling cell proliferation is *p21* (CDKN1A), which is a cyclin-dependent kinase inhibitor exerting its effect by regulating cell cycle arrest (Brugarolas et al., 1995; Deng et al., 1995; Wade Harper et al., 1993). In our study we have found that 4SC-202 dramatically enhanced the expression of *p21* protein levels within 12 hours of drug treatment, even in the cells stimulated with TGFβ. Previous studies have reported that loss of LSD1 has a direct effect on *p21* expression (Lim et al., 2010). Furthermore, in our colony formation assay we have shown that 4SC-202 completely attenuated the colony forming ability of cancer cells. We also tested the effects of the LSD1 inhibitor SP2509 and the HDAC inhibitor Vorinostat either alone or in combinatorial treatment and demonstrated that effects of the combinatorial treatment were consistent with the 4SC-202 effects. Previously it has been reported that in *in vivo* and *in vitro* model systems for acute myeloid leukemia (AML), LSD1 and HDAC inhibitors show better effects when used in combination than individually (Fiskus et al., 2014).

5.10 4SC-202 decreases tumor growth *in vivo*

LSD1 and HDACs are often overexpressed in different types of human cancer suggesting a potential tumorigenic role (Kahl et al., 2006; Lv et al., 2012; Müller et al., 2013; Serce et al., 2012). Further, LSD1 has been largely found in association with HDAC1 and HDAC2 (Hakimi et al., 2002, 2003b). HDACs fuel the demethylase activity of LSD1 by creating hypoacetylated nucleosome thus facilitating LSD1 binding (Lee et al., 2006; Shi et al., 2005). We investigated the expression of LSD1 in pancreatic cancer and have determined that its expression level was in the range moderate to

very high, implying that it is overexpressed in a subset of pancreatic cancers. Furthermore, we explored the effects of inhibiting LSD1 and class-1 HDACs on the tumor growth in a xenograft mice model using the dual LSD1/HDAC inhibitor 4SC-202. Remarkably, combined inhibition of LSD1 and HDAC lead to significant decrease in tumor growth. LSD1 has been shown to interact with and inhibit p53 induced-apoptosis by repressing the transcriptional activation of p53 regulated genes (Huang et al., 2007). Tumor suppressors act to maintain cellular homeostasis mainly by controlling aberrant proliferation of the cells where p53 represents one of the most extensively studied and commonly inactivated tumor suppressors in cancer. Therefore, perturbation in the activity of p53 can result in aberrant proliferation and survival of damaged cells that can lead to tumor development (Parant and Lozano, 2003; Symonds et al., 1994). Furthermore, HDACs have also been shown to modulate p53 activity and alter its inhibitory activity on cell proliferation and apoptosis. HDAC2 has been shown to inhibit p53 activity by modulating p53-DNA binding (Harms and Chen, 2007). Furthermore, HDAC1 has bene shown to inhibit the p53 dependent activation of cyclin-dependent kinase inhibitor p21 implying that HDAC1 is an antagonist to p53 (Lagger et al., 2003). Since both LSD1 and HDACs have been shown to modulate p53 function, it can be speculated that their inhibition can have antitumor effects in a p53 dependent manner.

5.11 Epigenetic therapies targeting EMT may result in increased metastasis

Initiation of EMT is the first step towards cancer cell metastasis and it helps the cells to migrate from the primary site and invade to the surrounding tissues. For a cancer cell to be able to colonize at distant site and give rise to secondary tumor, it is important to revert back to epithelial phenotype by undergoing mesenchymal-to-epithelial transition (MET) (Nieto, 2013; Yao et al., 2011). We have shown that inhibition of LSD1

and HDACs blocked EMT. It has been reported that, disseminated breast cancer cells re-express epithelial markers like E-cadherin by undergoing MET at the distant site to form secondary tumor (Chao et al., 2010). Furthermore, a study conducted in a spontaneous mouse model for squamous cell carcinoma has shown that the EMT-TF Twist1 was able to promote EMT in the tumor cells. However, deactivation of Twist1 was required in the circulating tumor cells to undergo MET and form secondary metastasis (Tsai et al., 2012). In another study it has been shown that reversal of EMT at distant metastatic site allows the cells to acquire stem cell properties and enhanced proliferation capacity that allows them to colonize and form metastasis (Ocaña et al., 2012). We have shown that 4SC-202 treatment led to alterations in the expression of a subset of genes and pathways involved in differentiation, which was accompanied with inhibition of TGFβ-induced EMT. Inhibiting EMT has been considered as a therapeutic target to tackle metastatic spread of cancer. However, considering the recent data highlighting the importance of MET in metastatic colonization at distant site, inhibiting EMT cannot be a therapeutic strategy for all cancer types. In some cases like pancreatic and breast cancer, EMT and dissemination of cancer cells occur at a relatively early stage (Hüsemann et al., 2008; Rhim et al., 2012), thus inhibiting EMT in these types of cancer may promote metastasis. Therefore, an in-depth study is required to elucidate the epigenetic signature associated with EMT-TFs and their target genes. Furthermore, a better understanding of epigenetic mechanisms involved in the regulation of EMT will lead us to novel therapeutic targets which will be highly specific with minimal side effects (Bedi et al., 2014; Mishra and Johnsen, 2014).

5.12 LSD1 can modulate tumor cell phenotype through enhancer regulation

Recent studies have highlighted the importance of enhancer elements in regulating cell identity and maintaining tissue specificity during development and disease.

Enhancers are non-coding DNA (few hundred base pairs) elements that contain clusters of binding sites (6-10 bp long motif) which can be recognized by transcription factors in a sequence- and tissue-specific manner (Shlyueva et al., 2014). Enhancers play an important role in defining cell lineage by controlling the expression of specific sets of genes(Rada-Iglesias et al., 2011; Whyte et al., 2013). Interestingly, enhancers have distinct and unique chromatin landscape which is characterized by the presence of H3K4me1 (Heintzman et al., 2009). Additionally, the presence of active mark H3K27ac distinguishes active from "poised" enhancers (Creyghton et al., 2010).

Recently various studies have highlighted that mutations in enhancer-associated factors can result in cancer development (Gröschel et al., 2014; Yamazaki et al., 2014; Zhang et al., 2016). Furthermore, another study carried out in colon cancer has shown that changes in the epigenetic landscape of enhancers can lead to perturbation in gene expression in a manner that can result in colon cancer (Akhtar-Zaidi et al., 2012). Since LSD1 demethylate H3K4me1, which is a marker for enhancers, it can be speculated that overexpression of LSD1 can result in deregulation of a subset of genes in an enhancer dependent manner that can result in tumorigenesis.

Perturbation of cell fate commitment and acquisition of stem cell characteristics are important characteristics of cancer (Ben-Porath et al., 2008). Another study has shown that LSD1-mediated suppression of hematopoietic stem cell associated enhancers is required for their differentiation (Kerenyi et al., 2013). Since LSD1 is involved in governing the pluripotent and differentiation states of the cells, it can be inferred that inhibition of LSD1 activity can repress the differentiation-associated enhancers resulting in enrichment of cancer stem-like cells and increased metastasis.

5.13 LSD1 expression as a predictive biomarker for responsiveness to targeted therapy

Several studies have highlighted the overexpression of LSD1 in poorly differentiated and aggressive form of cancers and has been shown to correlate with poor outcome (Jie et al., 2013; Lim et al., 2010; Lv et al., 2012; Yu et al., 2013). Furthermore, overexpression of LSD1 in NSCLC has been shown to be associated with cell proliferation, migration and invasive phenotype (Lv et al., 2012). In our study we have shown that LSD1 is highly expressed in a subset of pancreatic tumors. However, we did not observe any significant correlation to the patient survival. LSD1 has been shown to be involved in silencing of tumor suppressor gene *BRCA1* (Wu et al., 2012b) and has been found to be inversely correlated with *BRCA1* expression in triple-negative breast cancer (Nagasawa et al., 2015). *BRAC1* mutant tumors have been shown to be sensitive to PARP inhibitors (Turner et al., 2008) therefore, expression status of LSD1 in breast cancer has been proposed as biomarker for patients that will respond to PARP inhibition based therapies (Nagasawa et al., 2015). Since we have observed higher expression of LSD1 in a fraction of pancreatic tumors, it can be proposed that LSD1 expression may stratify the patients that will respond to targeted therapy.

Taken together we have uncovered the previously unknown function of a transcription factor KLF10 in regulating the pro-metastatic function of TGFβ signaling by inhibiting TGFβ-induced EMT. Furthermore, we also show that KLF10 is required for recruitment of HDAC1 to SNAI2 promoter and consequently causing its repression. KLF10 expression further correlated with poor outcome in lung adenocarcinoma and breast carcinoma (luminal B) patients which implicates a potential for KLF10 as a prognostic marker. In a more global approach we have investigated the efficacy of a small

molecule inhibitor against epigenetic modifiers. Several studies have highlighted the importance of LSD1 in tumorigenesis and promoting undifferentiated phenotype in cancer cells especially in breast cancer. Combined inhibition of LSD1 and HDACs is been considered as a potential approach for targeted therapy against certain types cancer. In our study we have utilized a dual LSD1/HDAC inhibitor 4SC-202 and show that combined inhibition of LSD1 and HDACs blocks the TGFβ-induced EMT and significantly decreases the tumor growth *in vivo.* Currently available small molecule inhibitors against LSD1 show poor selectivity and *in vivo* toxicity thus limiting their use in the patients. Further in-depth investigation of 4SC-202 is required to establish it as a potential targeted therapy option.

6. Reference list

Adamo, A., Sesé, B., Boue, S., Castaño, J., Paramonov, I., Barrero, M.J., and Belmonte, J.C.I. (2011). LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. Nat. Cell Biol. *13*, 652–659.

Aghdassi, A., Sendler, M., Guenther, A., Mayerle, J., Behn, C.-O., Heidecke, C.-D., Friess, H., Büchler, M., Evert, M., Lerch, M.M., et al. (2012a). Recruitment of histone deacetylases HDAC1 and HDAC2 by the transcriptional repressor ZEB1 downregulates E-cadherin expression in pancreatic cancer. Gut *61*, 439–448.

Aghdassi, A., Sendler, M., Guenther, A., Mayerle, J., Behn, C.-O., Heidecke, C.-D., Friess, H., Büchler, M., Evert, M., Lerch, M.M., et al. (2012b). Recruitment of histone deacetylases HDAC1 and HDAC2 by the transcriptional repressor ZEB1 downregulates E-cadherin expression in pancreatic cancer. Gut *61*, 439–448.

Akhtar-Zaidi, B., Cowper-Sal·lari, R., Corradin, O., Saiakhova, A., Bartels, C.F., Balasubramanian, D., Myeroff, L., Lutterbaugh, J., Jarrar, A., Kalady, M.F., et al. (2012). Epigenomic Enhancer Profiling Defines a Signature of Colon Cancer. Science *336*, 736–739.

Akiyoshi, S., Inoue, H., Hanai, J., Kusanagi, K., Nemoto, N., Miyazono, K., and Kawabata, M. (1999). c-Ski Acts as a Transcriptional Co-repressor in Transforming Growth Factor-β Signaling through Interaction with Smads. J. Biol. Chem. *274*, 35269–35277.

Alexandrow, M.G., Kawabata, M., Aakre, M., and MosEs, H.L. (1995). Overexpression of the c-Myc oncoprotein blocks the growth-inhibitory response but is required for the mitogenic effects of transforming growth factor beta 1. Proc. Natl. Acad. Sci. *92*, 3239–3243.

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

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Antony, M.L., Nair, R., Sebastian, P., and Karunagaran, D. (2009). Changes in expression, and/or mutations in TGF-β receptors (TGF-β RI and TGF-β RII) and Smad 4 in human ovarian tumors. J. Cancer Res. Clin. Oncol. *136*, 351–361.

Armstrong, J.A., Bieker, J.J., and Emerson, B.M. (1998). A SWI/SNF–Related Chromatin Remodeling Complex, E-RC1, Is Required for Tissue-Specific Transcriptional Regulation by EKLF In Vitro. Cell *95*, 93–104.

Ashcroft, G.S. (1999). Bidirectional regulation of macrophage function by TGF-β. Microbes Infect. *1*, 1275–1282.

Ateeq, B., Unterberger, A., Szyf, M., and Rabbani, S.A. (2008). Pharmacological Inhibition of DNA Methylation Induces Proinvasive and Prometastatic Genes In Vitro and In Vivo. Neoplasia N. Y. N *10*, 266–278.

Bailey, J.M., Singh, P.K., and Hollingsworth, M.A. (2007). Cancer metastasis facilitated by developmental pathways: Sonic hedgehog, Notch, and bone morphogenic proteins. J. Cell. Biochem. *102*, 829–839.

Bannister, A.J., Schneider, R., and Kouzarides, T. (2002). Histone Methylation: Dynamic or Static? Cell *109*, 801–806.

Batlle, E., Sancho, E., Francí, C., Domínguez, D., Monfar, M., Baulida, J., and García de Herreros, A. (2000). The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nat. Cell Biol. *2*, 84–89.

Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome — biological and translational implications. Nat. Rev. Cancer *11*, 726–734.

Bedi, U., Mishra, V.K., Wasilewski, D., Scheel, C., and Johnsen, S.A. (2014). Epigenetic plasticity: A central regulator of epithelial-to-mesenchymal transition in cancer. Oncotarget *5*, 2016–2029.

Bender, H., Wang, Z., Schuster, N., and Krieglstein, K. (2004). TIEG1 facilitates transforming growth factor-β-mediated apoptosis in the oligodendroglial cell line OLI-neu. J. Neurosci. Res. *75*, 344–352.

Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A., and Weinberg, R.A. (2008). An embryonic stem cell–like gene expression signature in poorly differentiated aggressive human tumors. Nat. Genet. *40*, 499–507.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes Dev. *23*, 781–783.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell *125*, 315–326.

Bhat-Nakshatri, P., Appaiah, H., Ballas, C., Pick-Franke, P., Goulet, R., Badve, S., Srour, E.F., and Nakshatri, H. (2010). SLUG/SNAI2 and Tumor Necrosis Factor Generate Breast Cells With CD44+/CD24- Phenotype. BMC Cancer *10*, 411.

Bieker, J.J. (2001). Krüppel-like Factors: Three Fingers in Many Pies. J. Biol. Chem. *276*, 34355–34358.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Black, A.R., Black, J.D., and Azizkhan-Clifford, J. (2001). Sp1 and krüppel-like factor family of transcription factors in cell growth regulation and cancer. J. Cell. Physiol. *188*, 143–160.

Blok, L.J., Grossmann, M.E., Perry, J.E., and Tindall, D.J. (1995). Characterization of an early growth response gene, which encodes a zinc finger transcription factor, potentially involved in cell cycle regulation. Mol. Endocrinol. *9*, 1610–1620.

Bloushtain-Qimron, N., Yao, J., Snyder, E.L., Shipitsin, M., Campbell, L.L., Mani, S.A., Hu, M., Chen, H., Ustyansky, V., Antosiewicz, J.E., et al. (2008). Cell type-specific DNA methylation patterns in the human breast. Proc. Natl. Acad. Sci. *105*, 14076–14081.

Bolós, V., Peinado, H., Pérez-Moreno, M.A., Fraga, M.F., Esteller, M., and Cano, A. (2003). The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. J. Cell Sci. *116*, 499–511.

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

Bonnomet, A., Syne, L., Brysse, A., Feyereisen, E., Thompson, E.W., Noël, A., Foidart, J.-M., Birembaut, P., Polette, M., and Gilles, C. (2012). A dynamic in vivo model of epithelial-to-mesenchymal transitions in circulating tumor cells and metastases of breast cancer. Oncogene *31*, 3741–3753.

Bozic, I., Reiter, J.G., Allen, B., Antal, T., Chatterjee, K., Shah, P., Moon, Y.S., Yaqubie, A., Kelly, N., Le, D.T., et al. (2013). Evolutionary dynamics of cancer in response to targeted combination therapy. eLife *2*, e00747.

Brabletz, T. (2012). To differentiate or not — routes towards metastasis. Nat. Rev. Cancer *12*, 425–436.

Brugarolas, J., Chandrasekaran, C., Gordon, J.I., Beach, D., Jacks, T., and Hannon, G.J. (1995). Radiation-induced cell cycle arrest compromised by p21 deficiency. Nature *377*, 552–557.

von Burstin, J., Eser, S., Paul, M.C., Seidler, B., Brandl, M., Messer, M., von Werder, A., Schmidt, A., Mages, J., Pagel, P., et al. (2009). E-Cadherin Regulates Metastasis of Pancreatic Cancer In Vivo and Is Suppressed by a SNAIL/HDAC1/HDAC2 Repressor Complex. Gastroenterology *137*, 361–371.e5.

Byles, V., Zhu, L., Lovaas, J.D., Chmilewski, L.K., Wang, J., Faller, D.V., and Dai, Y. (2012). SIRT1 induces EMT by cooperating with EMT transcription factors and enhances prostate cancer cell migration and metastasis. Oncogene *31*, 4619–4629.

Caestecker, M.P. de, Piek, E., and Roberts, A.B. (2000). Role of Transforming Growth Factor-β Signaling in Cancer. J. Natl. Cancer Inst. *92*, 1388–1402.

Campos, E.I., and Reinberg, D. (2009). Histones: Annotating Chromatin. Annu. Rev. Genet. *43*, 559–599.

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

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. *10*, 295–304.

Chaffer, C.L., and Weinberg, R.A. (2011). A Perspective on Cancer Cell Metastasis. Science *331*, 1559–1564.

Chaffer, C.L., Brennan, J.P., Slavin, J.L., Blick, T., Thompson, E.W., and Williams, E.D. (2006). Mesenchymal-to-Epithelial Transition Facilitates Bladder Cancer Metastasis: Role of Fibroblast Growth Factor Receptor-2. Cancer Res. *66*, 11271–11278.

Chaffer, C.L., Thompson, E.W., and Williams, E.D. (2007). Mesenchymal to Epithelial Transition in Development and Disease. Cells Tissues Organs *185*, 7–19.

Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised Chromatin at the ZEB1 Promoter Enables Breast Cancer Cell Plasticity and Enhances Tumorigenicity. Cell *154*, 61–74.

Chalaux, E., López-Rovira, T., Rosa, J.L., Pons, G., Boxer, L.M., Bartrons, R., and Ventura, F. (1999). A zinc-finger transcription factor induced by TGF-β promotes apoptotic cell death in epithelial Mv1Lu cells. FEBS Lett. *457*, 478–482.

Chambers, A.F., Groom, A.C., and MacDonald, I.C. (2002). Metastasis: Dissemination and growth of cancer cells in metastatic sites. Nat. Rev. Cancer *2*, 563–572.

Chang, H.W., Chow, V., Lam, K.Y., Wei, W.I., and Wing Yuen, A.P. (2002). Loss of E-cadherin expression resulting from promoter hypermethylation in oral tongue carcinoma and its prognostic significance. Cancer *94*, 386–392.

Chao, Y.L., Shepard, C.R., and Wells, A. (2010). Breast carcinoma cells re-express E-cadherin during mesenchymal to epithelial reverting transition. Mol. Cancer *9*, 179.

Chen, C.-L., Liu, S.S., Ip, S.-M., Wong, L.C., Ngan, H.Y.S., and Ng, T.Y. (2003). E-cadherin expression is silenced by DNA methylation in cervical cancer cell lines and tumours. Eur. J. Cancer *39*, 517–523.

Chen, C.-R., Kang, Y., and Massagué, J. (2001). Defective repression of c-myc in breast cancer cells: A loss at the core of the transforming growth factor β growth arrest program. Proc. Natl. Acad. Sci. *98*, 992–999.

Chen, J., Chan, A.W.H., To, K.-F., Chen, W., Zhang, Z., Ren, J., Song, C., Cheung, Y.-S., Lai, P.B.S., Cheng, S.-H., et al. (2013a). SIRT2 overexpression in hepatocellular carcinoma mediates epithelial to mesenchymal transition by protein kinase B/glycogen synthase kinase-3β/β-catenin signaling. Hepatology *57*, 2287–2298.

Chen, X., Xiao, W., Chen, W., Luo, L., Ye, S., and Liu, Y. (2013b). The epigenetic modifier trichostatin A, a histone deacetylase inhibitor, suppresses proliferation and epithelial–mesenchymal transition of lens epithelial cells. Cell Death Dis. *4*, e884.

Chen, Y., Yang, Y., Wang, F., Wan, K., Yamane, K., Zhang, Y., and Lei, M. (2006). Crystal structure of human histone lysine-specific demethylase 1 (LSD1). Proc. Natl. Acad. Sci. *103*, 13956–13961.

Cheung, P., and Lau, P. (2005). Epigenetic Regulation by Histone Methylation and Histone Variants. Mol. Endocrinol. *19*, 563–573.

Chung, Y.-J., Song, J.-M., Lee, J.-Y., Jung, Y.-T., Seo, E.-J., Choi, S.-W., and Rhyu, M.-G. (1996). Microsatellite Instability-associated Mutations Associate Preferentially with the Intestinal Type of Primary Gastric Carcinomas in a High-Risk Population. Cancer Res. *56*, 4662–4665.

Clarke, M.F., Dick, J.E., Dirks, P.B., Eaves, C.J., Jamieson, C.H.M., Jones, D.L., Visvader, J., Weissman, I.L., and Wahl, G.M. (2006). Cancer Stem Cells—Perspectives on Current Status and Future Directions: AACR Workshop on Cancer Stem Cells. Cancer Res. *66*, 9339–9344.

Coffey, R.J., Bascom, C.C., Sipes, N.J., Graves-Deal, R., Weissman, B.E., and Moses, H.L. (1988). Selective inhibition of growth-related gene expression in murine keratinocytes by transforming growth factor beta. Mol. Cell. Biol. *8*, 3088–3093.

Collins, A.T., Berry, P.A., Hyde, C., Stower, M.J., and Maitland, N.J. (2005). Prospective Identification of Tumorigenic Prostate Cancer Stem Cells. Cancer Res. *65*, 10946–10951.

Côme, C., Magnino, F., Bibeau, F., Barbara, P.D.S., Becker, K.F., Theillet, C., and Savagner, P. (2006). Snail and Slug Play Distinct Roles during Breast Carcinoma Progression. Clin. Cancer Res. *12*, 5395–5402.

Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., Mareel, M., Huylebroeck, D., and van Roy, F. (2001). The Two-Handed E Box Binding Zinc Finger Protein SIP1 Downregulates E-Cadherin and Induces Invasion. Mol. Cell *7*, 1267–1278.

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

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. *107*, 21931–21936.

Daigle, S.R., Olhava, E.J., Therkelsen, C.A., Majer, C.R., Sneeringer, C.J., Song, J., Johnston, L.D., Scott, M.P., Smith, J.J., Xiao, Y., et al. (2011). Selective Killing of Mixed Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. Cancer Cell *20*, 53–65.

Dalal, B.I., Keown, P.A., and Greenberg, A.H. (1993). 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. *143*, 381–389.

Dang, D.T., Pevsner, J., and Yang, V.W. (2000). The biology of the mammalian Krüppel-like family of transcription factors. Int. J. Biochem. Cell Biol. *32*, 1103–1121.

Datto, M.B., Li, Y., Panus, J.F., Howe, D.J., Xiong, Y., and Wang, X.F. (1995). Transforming growth factor beta induces the cyclin-dependent kinase inhibitor p21 through a p53-independent mechanism. Proc. Natl. Acad. Sci. U. S. A. *92*, 5545–5549.

Dawson, M.A., and Kouzarides, T. (2012). Cancer Epigenetics: From Mechanism to Therapy. Cell *150*, 12–27.

Dean, M., Fojo, T., and Bates, S. (2005). Tumour stem cells and drug resistance. Nat. Rev. Cancer *5*, 275–284.

Deng, C., Zhang, P., Wade Harper, J., Elledge, S.J., and Leder, P. (1995). Mice Lacking p21CIP1/WAF1 undergo normal development, but are defective in G1 checkpoint control. Cell *82*, 675–684.

Derynck, R., and Akhurst, R.J. (2007). Differentiation plasticity regulated by TGF-β family proteins in development and disease. Nat. Cell Biol. *9*, 1000–1004.

Derynck, R., and Zhang, Y.E. (2003). Smad-dependent and Smad-independent pathways in TGF-β family signalling. Nature *425*, 577–584.

Derynck, R., Akhurst, R.J., and Balmain, A. (2001). TGF-beta signaling in tumor suppression and cancer progression. Nat Genet *29*, 117–129.

Diaz Jr, L.A., Williams, R.T., Wu, J., Kinde, I., Hecht, J.R., Berlin, J., Allen, B., Bozic, I., Reiter, J.G., Nowak, M.A., et al. (2012). The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature *486*, 537–540.

Dong, C., Wu, Y., Yao, J., Wang, Y., Yu, Y., Rychahou, P.G., Evers, B.M., and Zhou, B.P. (2012). G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. J. Clin. Invest. *122*, 1469–1486.

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

Duro de Oliveira, K., Vannucci Tedardi, M., Cogliati, B., Zaidan Dagli, M.L., cia, Duro de Oliveira, K., Vannucci Tedardi, M., Cogliati, B., Zaidan Dagli, M.L., and cia (2013). Higher Incidence of Lung Adenocarcinomas Induced by DMBA in Connexin 43 Heterozygous Knockout Mice, Higher Incidence of Lung Adenocarcinomas Induced by DMBA in Connexin 43 Heterozygous Knockout Mice. BioMed Res. Int. BioMed Res. Int. *2013*, *2013*, e618475.

Ebisawa, T., Fukuchi, M., Murakami, G., Chiba, T., Tanaka, K., Imamura, T., and Miyazono, K. (2001). Smurf1 Interacts with Transforming Growth Factor-β Type I Receptor through Smad7 and Induces Receptor Degradation. J. Biol. Chem. *276*, 12477–12480.

Egeblad, M., and Werb, Z. (2002). New functions for the matrix metalloproteinases in cancer progression. Nat. Rev. Cancer *2*, 161–174.

Eger, A., Aigner, K., Sonderegger, S., Dampier, B., Oehler, S., Schreiber, M., Berx, G., Cano, A., Beug, H., and Foisner, R. (2005). DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene *24*, 2375–2385.

Eramo, A., Lotti, F., Sette, G., Pilozzi, E., Biffoni, M., Di Virgilio, A., Conticello, C., Ruco, L., Peschle, C., and De Maria, R. (2007). Identification and expansion of the tumorigenic lung cancer stem cell population. Cell Death Differ. *15*, 504–514.

Espada, J., Peinado, H., Lopez-Serra, L., Setién, F., Lopez-Serra, P., Portela, A., Renart, J., Carrasco, E., Calvo, M., Juarranz, A., et al. (2011). Regulation of SNAIL1 and E-cadherin function by DNMT1 in a DNA methylation-independent context. Nucleic Acids Res. *39*, 9194–9205.

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. Nat. Rev. Genet. *8*, 286–298.

Evans, P.M., Chen, X., Zhang, W., and Liu, C. (2010). KLF4 Interacts with β-Catenin/TCF4 and Blocks p300/CBP Recruitment by β-Catenin. Mol. Cell. Biol. *30*, 372–381.

Ewen, M.E., Oliver, C.J., Sluss, H.K., Miller, S.J., and Peeper, D.S. (1995). p53-dependent repression of CDK4 translation in TGF-beta-induced G1 cell-cycle arrest. Genes Dev. *9*, 204–217.

Fang, J., Shing, Y., Wiederschain, D., Yan, L., Butterfield, C., Jackson, G., Harper, J., Tamvakopoulos, G., and Moses, M.A. (2000). Matrix metalloproteinase-2 is required for the switch to the angiogenic phenotype in a tumor model. Proc. Natl. Acad. Sci. U. S. A. *97*, 3884–3889.

Fautsch, M.P., Vrabel, A., Subramaniam, M., Hefferen, T.E., Spelsberg, T.C., and Wieben, E.D. (1998). TGFβ-Inducible Early Gene (TIEG) Also Codes for Early Growth Response α (EGRα): Evidence of Multiple Transcripts from Alternate Promoters. Genomics *51*, 408–416.

Feinberg, A.P., and Tycko, B. (2004). The history of cancer epigenetics. Nat. Rev. Cancer *4*, 143–153.

Ferrari-Amorotti, G., Fragliasso, V., Esteki, R., Prudente, Z., Soliera, A.R., Cattelani, S., Manzotti, G., Grisendi, G., Dominici, M., Pieraccioli, M., et al. (2013). Inhibiting interactions of lysine demethylase LSD1 with Snail/Slug blocks cancer cell invasion. Cancer Res. *73*, 235–245.

Fidler, I.J. (2003a). The pathogenesis of cancer metastasis: the "seed and soil" hypothesis revisited. Nat. Rev. Cancer *3*, 453–458.

Fidler, I.J. (2003b). The pathogenesis of cancer metastasis: the "seed and soil" hypothesis revisited. Nat. Rev. Cancer *3*, 453–458.

Fiskus, W., Sharma, S., Shah, B., Portier, B.P., Devaraj, S.G.T., Liu, K., Iyer, S.P., Bearss, D., and Bhalla, K.N. (2014). Highly effective combination of LSD1 (KDM1A) antagonist and pan-histone deacetylase inhibitor against human AML cells. Leukemia *28*, 2155–2164.

Forneris, F., Binda, C., Vanoni, M.A., Battaglioli, E., and Mattevi, A. (2005). Human Histone Demethylase LSD1 Reads the Histone Code. J. Biol. Chem. *280*, 41360–41365.

Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., et al. (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat. Genet. *37*, 391–400.

Friedman, E., Gold, L.I., Klimstra, D., Zeng, Z.S., Winawer, S., and Cohen, A. (1995). High levels of transforming growth factor beta 1 correlate with disease progression in human colon cancer. Cancer Epidemiol. Biomarkers Prev. *4*, 549–554.

Ginestier, C., Hur, M.H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., Jacquemier, J., Viens, P., Kleer, C.G., Liu, S., et al. (2007). ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome. Cell Stem Cell *1*, 555–567.

Glozak, M.A., and Seto, E. (2007). Histone deacetylases and cancer. Oncogene *26*, 5420–5432.

Gorsch, S.M., Memoli, V.A., Stukel, T.A., Gold, L.I., and Arrick, B.A. (1992). Immunohistochemical Staining for Transforming Growth Factor β1 Associates with Disease Progression in Human Breast Cancer. Cancer Res. *52*, 6949–6952.

Graff, J.R., Gabrielson, E., Fujii, H., Baylin, S.B., and Herman, J.G. (2000). Methylation Patterns of the E-cadherin 5' CpG Island Are Unstable and Reflect the Dynamic, Heterogeneous Loss of E-cadherin Expression during Metastatic Progression. J. Biol. Chem. *275*, 2727–2732.

Griffon, A., Barbier, Q., Dalino, J., van Helden, J., Spicuglia, S., and Ballester, B. (2014). Integrative analysis of public ChIP-seq experiments reveals a complex multi-cell regulatory landscape. Nucleic Acids Res. gku1280.

Gröschel, S., Sanders, M.A., Hoogenboezem, R., de Wit, E., Bouwman, B.A.M., Erpelinck, C., van der Velden, V.H.J., Havermans, M., Avellino, R., van Lom, K., et al. (2014). A Single Oncogenic Enhancer Rearrangement Causes Concomitant EVI1 and GATA2 Deregulation in Leukemia. Cell *157*, 369–381.

Guo, P., Dong, X.-Y., Zhao, K.-W., Sun, X., Li, Q., and Dong, J.-T. (2010). Estrogen-induced interaction between KLF5 and estrogen receptor (ER) suppresses the function of ER in ER-positive breast cancer cells. Int. J. Cancer J. Int. Cancer *126*, 81–89.

Gupta, G.P., and Massagué, J. (2006). Cancer Metastasis: Building a Framework. Cell *127*, 679–695.

Hajra, K.M., Chen, D.Y., and Fearon, E.R. (2002). The SLUG zinc-finger protein represses E-cadherin in breast cancer. Cancer Res. *62*, 1613–1618.

Hakimi, M.-A., Bochar, D.A., Chenoweth, J., Lane, W.S., Mandel, G., and Shiekhattar, R. (2002). A core–BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. Proc. Natl. Acad. Sci. *99*, 7420–7425.

Hakimi, M.-A., Dong, Y., Lane, W.S., Speicher, D.W., and Shiekhattar, R. (2003a). A Candidate X-linked Mental Retardation Gene Is a Component of a New Family of Histone Deacetylase-containing Complexes. J. Biol. Chem. *278*, 7234–7239.

Hakimi, M.-A., Dong, Y., Lane, W.S., Speicher, D.W., and Shiekhattar, R. (2003b). A Candidate X-linked Mental Retardation Gene Is a Component of a New Family of Histone Deacetylase-containing Complexes. J. Biol. Chem. *278*, 7234–7239.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. Cell *144*, 646–674.

Hannon, G.J., and Beach, D. (1994). pl5INK4B is a potentia| effector of TGF-β-induced cell cycle arrest. Nature *371*, 257–261.

Harms, K.L., and Chen, X. (2007). Histone Deacetylase 2 Modulates p53 Transcriptional Activities through Regulation of p53-DNA Binding Activity. Cancer Res. *67*, 3145–3152.

Harris, W.J., Huang, X., Lynch, J.T., Spencer, G.J., Hitchin, J.R., Li, Y., Ciceri, F., Blaser, J.G., Greystoke, B.F., Jordan, A.M., et al. (2012). The Histone Demethylase KDM1A Sustains the Oncogenic Potential of MLL-AF9 Leukemia Stem Cells. Cancer Cell *21*, 473–487.

Hayes, J.J. (2002). Changing chromatin from the inside. Nat. Struct. Mol. Biol. *9*, 161–163.

He, H.H., Meyer, C.A., Shin, H., Bailey, S.T., Wei, G., Wang, Q., Zhang, Y., Xu, K., Ni, M., Lupien, M., et al. (2010). Nucleosome dynamics define transcriptional enhancers. Nat. Genet. *42*, 343–347.

Hefferan, T.E., Reinholz, G.G., Rickard, D.J., Johnsen, S.A., Waters, K.M., Subramaniam, M., and Spelsberg, T.C. (2000). Overexpression of a Nuclear Protein, TIEG, Mimics Transforming Growth Factor- Action in Human Osteoblast Cells. J. Biol. Chem. *275*, 20255–20259.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature *459*, 108–112.

Hermann, P.C., Huber, S.L., Herrler, T., Aicher, A., Ellwart, J.W., Guba, M., Bruns, C.J., and Heeschen, C. (2007). Distinct Populations of Cancer Stem Cells Determine Tumor Growth and Metastatic Activity in Human Pancreatic Cancer. Cell Stem Cell *1*, 313–323.

Herranz, N., Pasini, D., Díaz, V.M., Francí, C., Gutierrez, A., Dave, N., Escrivà, M., Hernandez-Muñoz, I., Croce, L.D., Helin, K., et al. (2008). Polycomb Complex 2 Is Required for E-cadherin Repression by the Snail1 Transcription Factor. Mol. Cell. Biol. *28*, 4772–4781.

Hitchins, M.P., Rapkins, R.W., Kwok, C.-T., Srivastava, S., Wong, J.J.L., Khachigian, L.M., Polly, P., Goldblatt, J., and Ward, R.L. (2011). Dominantly Inherited Constitutional Epigenetic Silencing of MLH1 in a Cancer-Affected Family Is Linked to a Single Nucleotide Variant within the 5′UTR. Cancer Cell *20*, 200–213.

Horn, P.J., and Peterson, C.L. (2002). Chromatin Higher Order Folding--Wrapping up Transcription. Science *297*, 1824–1827.

Hrzenjak, A., Moinfar, F., Kremser, M.-L., Strohmeier, B., Staber, P.B., Zatloukal, K., and Denk, H. (2006). Valproate inhibition of histone deacetylase 2 affects differentiation and decreases proliferation of endometrial stromal sarcoma cells. Mol. Cancer Ther. *5*, 2203–2210.

Hu, W., Hofstetter, W.L., Li, H., Zhou, Y., He, Y., Pataer, A., Wang, L., Xie, K., Swisher, S.G., and Fang, B. (2009). Putative Tumor-Suppressor Function of Krüppel-Like Factor 4 in Primary Lung Carcinoma. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. *15*, 5688–5695.

Huang, X., and Guo, B. (2006). Adenomatous Polyposis Coli Determines Sensitivity to Histone Deacetylase Inhibitor–Induced Apoptosis in Colon Cancer Cells. Cancer Res. *66*, 9245–9251.

Huang, A., Gilmour, J.W., Imami, N., Amjadi, P., Henderson, D.C., and Allen-Mersh, T.G. (2003). Increased serum transforming growth factor-beta1 in human colorectal cancer correlates with reduced circulating dendritic cells and increased colonic Langerhans cell infiltration. Clin. Exp. Immunol. *134*, 270–278.

Huang, B.H., Laban, M., Leung, C.H.-W., Lee, L., Lee, C.K., Salto-Tellez, M., Raju, G.C., and Hooi, S.C. (2005). Inhibition of histone deacetylase 2 increases apoptosis and p21Cip1/WAF1 expression, independent of histone deacetylase 1. Cell Death Differ. *12*, 395–404.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Huang, J., Sengupta, R., Espejo, A.B., Lee, M.G., Dorsey, J.A., Richter, M., Opravil, S., Shiekhattar, R., Bedford, M.T., Jenuwein, T., et al. (2007). p53 is regulated by the lysine demethylase LSD1. Nature *449*, 105–108.

Hugo, H., Ackland, M.L., Blick, T., Lawrence, M.G., Clements, J.A., Williams, E.D., and Thompson, E.W. (2007). Epithelial—mesenchymal and mesenchymal—epithelial transitions in carcinoma progression. J. Cell. Physiol. *213*, 374–383.

Huh, S.J., Chen, Y.-L., Friedman, S.L., Liao, J., Huang, H.-J.S., Cavenee, W.K., and Robertson, G.P. (2010). KLF6 Gene and Early Melanoma Development in a Collagen I-Rich Extracellular Environment. JNCI J. Natl. Cancer Inst. *102*, 1131–1147.

Humphrey, G.W., Wang, Y., Russanova, V.R., Hirai, T., Qin, J., Nakatani, Y., and Howard, B.H. (2001). Stable Histone Deacetylase Complexes Distinguished by the Presence of SANT Domain Proteins CoREST/kiaa0071 and Mta-L1. J. Biol. Chem. *276*, 6817–6824.

Hüsemann, Y., Geigl, J.B., Schubert, F., Musiani, P., Meyer, M., Burghart, E., Forni, G., Eils, R., Fehm, T., Riethmüller, G., et al. (2008). Systemic Spread Is an Early Step in Breast Cancer. Cancer Cell *13*, 58–68.

Itoh, T., Tanioka, M., Yoshida, H., Yoshioka, T., Nishimoto, H., and Itohara, S. (1998). Reduced Angiogenesis and Tumor Progression in Gelatinase A-deficient Mice. Cancer Res. *58*, 1048–1051.

Javaid, S., Zhang, J., Anderssen, E., Black, J.C., Wittner, B.S., Tajima, K., Ting, D.T., Smolen, G.A., Zubrowski, M., Desai, R., et al. (2013). Dynamic Chromatin Modification Sustains Epithelial-Mesenchymal Transition following Inducible Expression of Snail-1. Cell Rep. *5*, 1679–1689.

Jenuwein, T., and Allis, C.D. (2001). Translating the Histone Code. Science *293*, 1074–1080.

Jethwa, P., Naqvi, M., Hardy, R.G., Hotchin, N.A., Roberts, S., Spychal, R., and Tselepis, C. (2008). Overexpression of Slug is associated with malignant progression of esophageal adenocarcinoma. World J. Gastroenterol. WJG *14*, 1044–1052.

Ji, M., Lee, E., Kim, K., Kim, Y., Sung, R., Lee, S.-J., Kim, D., and Park, S. (2015). HDAC inhibitors induce epithelial-mesenchymal transition in colon carcinoma cells. Oncol. Rep.

Jiang, W., Sui, X., Zhang, D., Liu, M., Ding, M., Shi, Y., and Deng, H. (2011). CD24: A Novel Surface Marker for PDX1-Positive Pancreatic Progenitors Derived from Human Embryonic Stem Cells. STEM CELLS *29*, 609–617.

Jie, D., Zhongmin, Z., Guoqing, L., Sheng, L., Yi, Z., Jing, W., and Liang, Z. (2013). Positive Expression of LSD1 and Negative Expression of E-cadherin Correlate with Metastasis and Poor Prognosis of Colon Cancer. Dig. Dis. Sci. *58*, 1581–1589.

Jin, W., Chen, B., Li, J., Zhu, H., Huang, M., Gu, S., Wang, Q., Chen, J., Yu, S., Wu, J., et al. (2012). TIEG1 Inhibits Breast Cancer Invasion and Metastasis by Inhibition of Epidermal Growth Factor Receptor (EGFR) Transcription and the EGFR Signaling Pathway. Mol. Cell. Biol. *32*, 50–63.

Johnsen, S.A., Subramaniam, M., Janknecht, R., and Spelsberg, T.C. (2002a). TGFβ inducible early gene enhances TGFβ/Smad-dependent transcriptional responses. Oncogene *21*, 5783–5790.

Johnsen, S.A., Subramaniam, M., Katagiri, T., Janknecht, R., and Spelsberg, T.C. (2002b). Transcriptional regulation of Smad2 is required for enhancement of TGF?/Smad signaling by TGF? inducible early gene. J. Cell. Biochem. *87*, 233–241.

Johnsen, S.A., Subramaniam, M., Effenberger, K.E., and Spelsberg, T.C. (2004). The TGFβ inducible early gene plays a central role in the anti-proliferative response to TGFβ. Signal Transduct. *4*, 29–35.

Kaczynski, J., Cook, T., and Urrutia, R. (2003). Sp1- and Krüppel-like transcription factors. Genome Biol. *4*, 206.

Kahl, P., Gullotti, L., Heukamp, L.C., Wolf, S., Friedrichs, N., Vorreuther, R., Solleder, G., Bastian, P.J., Ellinger, J., Metzger, E., et al. (2006). Androgen Receptor Coactivators Lysine-Specific Histone Demethylase 1 and Four and a Half LIM Domain Protein 2 Predict Risk of Prostate Cancer Recurrence. Cancer Res. *66*, 11341–11347.

Kaimori, A., Potter, J.J., Choti, M., Ding, Z., Mezey, E., and Koteish, A.A. (2010). Histone deacetylase inhibition suppresses the transforming growth factor β1–induced epithelial-to-mesenchymal transition in hepatocytes. Hepatology *52*, 1033–1045.

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

Kang, Y., and Massagué, J. (2004). Epithelial-Mesenchymal Transitions: Twist in Development and Metastasis. Cell *118*, 277–279.

Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. Nucleic Acids Res. *32*, D493–D496.

Kavsak, P., Rasmussen, R.K., Causing, C.G., Bonni, S., Zhu, H., Thomsen, G.H., and Wrana, J.L. (2000). Smad7 Binds to Smurf2 to Form an E3 Ubiquitin Ligase that Targets the TGFβ Receptor for Degradation. Mol. Cell *6*, 1365–1375.

Kerenyi, M.A., Shao, Z., Hsu, Y.-J., Guo, G., Luc, S., O'Brien, K., Fujiwara, Y., Peng, C., Nguyen, M., and Orkin, S.H. (2013). Histone demethylase Lsd1 represses hematopoietic stem and progenitor cell signatures during blood cell maturation. eLife *2*, e00633.

Kessenbrock, K., Plaks, V., and Werb, Z. (2010). Matrix Metalloproteinases: Regulators of the Tumor Microenvironment. Cell *141*, 52–67.

KIESSLICH, T., PICHLER, M., and NEUREITER, D. (2013). Epigenetic control of epithelial-mesenchymal-transition in human cancer. Mol. Clin. Oncol. *1*, 3–11.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kim, J., Shin, S., Subramaniam, M., Bruinsma, E., Kim, T.-D., Hawse, J.R., Spelsberg, T.C., and Janknecht, R. (2010). Histone demethylase JARID1B/KDM5B is a corepressor of TIEG1/KLF10. Biochem. Biophys. Res. Commun. *401*, 412–416.

Kingsley, L.A., Fournier, P.G.J., Chirgwin, J.M., and Guise, T.A. (2007). Molecular Biology of Bone Metastasis. Mol. Cancer Ther. *6*, 2609–2617.

Klose, R.J., Kallin, E.M., and Zhang, Y. (2006). JmjC-domain-containing proteins and histone demethylation. Nat. Rev. Genet. *7*, 715–727.

Ko, H., So, Y., Jeon, H., Jeong, M.-H., Choi, H.-K., Ryu, S.-H., Lee, S.-W., Yoon, H.-G., and Choi, K.-C. (2013). TGF-β1-induced epithelial–mesenchymal transition and acetylation of Smad2 and Smad3 are negatively regulated by EGCG in Human A549 lung cancer cells. Cancer Lett. *335*, 205–213.

Koizume, S., Tachibana, K., Sekiya, T., Hirohashi, S., and Shiraishi, M. (2002). Heterogeneity in the modification and involvement of chromatin components of the CpG island of the silenced human CDH1 gene in cancer cells. Nucleic Acids Res. *30*, 4770–4780.

Komarova, N.L., and Boland, C.R. (2013). Cancer: Calculated treatment. Nature *499*, 291–292.

Komarova, N.L., and Wodarz, D. (2005). Drug resistance in cancer: Principles of emergence and prevention. Proc. Natl. Acad. Sci. U. S. A. *102*, 9714–9719.

Kong, D., Ahmad, A., Bao, B., Li, Y., Banerjee, S., and Sarkar, F.H. (2012). Histone Deacetylase Inhibitors Induce Epithelial-to-Mesenchymal Transition in Prostate Cancer Cells. PLOS ONE *7*, e45045.

Kornberg, R.D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA. Science *184*, 868–871.

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

Kruidenier, L., Chung, C., Cheng, Z., Liddle, J., Che, K., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H., et al. (2012). A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. Nature *488*, 404–408.

Kummar, S., Gutierrez, M., Gardner, E.R., Donovan, E., Hwang, K., Chung, E.J., Lee, M.-J., Maynard, K., Kalnitskiy, M., Chen, A., et al. (2007). Phase I Trial of MS-275, a Histone Deacetylase Inhibitor, Administered Weekly in Refractory Solid Tumors and Lymphoid Malignancies. Clin. Cancer Res. *13*, 5411–5417.

Lagger, G., Doetzlhofer, A., Schuettengruber, B., Haidweger, E., Simboeck, E., Tischler, J., Chiocca, S., Suske, G., Rotheneder, H., Wintersberger, E., et al. (2003). The Tumor Suppressor p53 and Histone Deacetylase 1 Are Antagonistic Regulators of the Cyclin-Dependent Kinase Inhibitor p21/WAF1/CIP1 Gene. Mol. Cell. Biol. *23*, 2669–2679.

Laiho, M., DeCaprio, J.A., Ludlow, J.W., Livingston, D.M., and Massague, J. (1990). Growth inhibition by TGF-β linked to suppression of retinoblastoma protein phosphorylation. Cell *62*, 175–185.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

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

Laybourn, P.J., and Kadonaga, J.T. (1991). Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science *254*, 238–245.

Lee, J.-S., Smith, E., and Shilatifard, A. (2010). The Language of Histone Crosstalk. Cell *142*, 682–685.

Lee, M.G., Wynder, C., Bochar, D.A., Hakimi, M.-A., Cooch, N., and Shiekhattar, R. (2006). Functional Interplay between Histone Demethylase and Deacetylase Enzymes. Mol. Cell. Biol. *26*, 6395–6402.

Lei, W., Zhang, K., Pan, X., Hu, Y., Wang, D., Yuan, X., Shu, G., and Song, J. (2010). Histone deacetylase 1 is required for transforming growth factor-β1-induced epithelial–mesenchymal transition. Int. J. Biochem. Cell Biol. *42*, 1489–1497.

Letterio, J.J., and Roberts, and A.B. (1998). REGULATION OF IMMUNE RESPONSES BY TGF-β. Annu. Rev. Immunol. *16*, 137–161.

Levy, L., and Hill, C.S. (2006). Alterations in components of the TGF-β superfamily signaling pathways in human cancer. Cytokine Growth Factor Rev. *17*, 41–58.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. Nat. Rev. Genet. *3*, 662–673.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, 1000 Genome Project Data Processing (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, Q., Gao, Y., Jia, Z., Mishra, L., Guo, K., Li, Z., Le, X., Wei, D., Huang, S., and Xie, K. (2012). Dysregulated Krüppel-Like Factor 4 and Vitamin D Receptor Signaling Contribute to Progression of Hepatocellular Carcinoma. Gastroenterology *143*, 799–810.e2.

Li, Z., Zhao, J., Li, Q., Yang, W., Song, Q., Li, W., and Liu, J. (2010). KLF4 promotes hydrogen-peroxide-induced apoptosis of chronic myeloid leukemia cells involving the bcl-2/bax pathway. Cell Stress Chaperones *15*, 905–912.

Lim, S., Janzer, A., Becker, A., Zimmer, A., Schüle, R., Buettner, R., and Kirfel, J. (2010). Lysine-specific demethylase 1 (LSD1) is highly expressed in ER-negative breast cancers and a biomarker predicting aggressive biology. Carcinogenesis *31*, 512–520.

Limame, R., Op de Beeck, K., Lardon, F., De Wever, O., and Pauwels, P. (2010). Krüppel-like factors in cancer progression: three fingers on the steering wheel. Oncotarget *1*, 29–48.

Lin, T., Ponn, A., Hu, X., Law, B.K., and Lu, J. (2010). Requirement of the Histone Demethylase LSD1 in Snai1-mediated Transcriptional Repression during Epithelial-Mesenchymal Transition. Oncogene *29*, 4896–4904.

Lombaerts, M., van Wezel, T., Philippo, K., Dierssen, J.W.F., Zimmerman, R.M.E., Oosting, J., van Eijk, R., Eilers, P.H., van de Water, B., Cornelisse, C.J., et al. (2006). E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. Br. J. Cancer *94*, 661–671.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lu, S.-L., Zhang, W.-C., Akiyama, Y., Nomizu, T., and Yuasa, Y. (1996). Genomic Structure of the Transforming Growth Factor β Type II Receptor Gene and Its Mutations in Hereditary Nonpolyposis Colorectal Cancers. Cancer Res. *56*, 4595–4598.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260.

Luo, K., Stroschein, S.L., Wang, W., Chen, D., Martens, E., Zhou, S., and Zhou, Q. (1999). The Ski oncoprotein interacts with the Smad proteins to repress TGFβ signaling. Genes Dev. *13*, 2196–2206.

Lv, T., Yuan, D., Miao, X., Lv, Y., Zhan, P., Shen, X., and Song, Y. (2012). Over-Expression of LSD1 Promotes Proliferation, Migration and Invasion in Non-Small Cell Lung Cancer. PLOS ONE *7*, e35065.

Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. Cell *133*, 704–715.

Marks, P.A., Rifkind, R.A., Richon, V.M., Breslow, R., Miller, T., and Kelly, W.K. (2001). Histone deacetylases and cancer: causes and therapies. Nat. Rev. Cancer *1*, 194–202.

Massague, J. (2000). NEW EMBO MEMBERS REVIEW: Transcriptional control by the TGF-beta/Smad signaling system. EMBO J. *19*, 1745–1754.

Massagué, J. (2000). How cells read TGF-β signals. Nat. Rev. Mol. Cell Biol. *1*, 169–178.

Massagué, J. (2008). TGFβ in Cancer. Cell *134*, 215–230.

Massagué, J. (2012). TGFβ signalling in context. Nat. Rev. Mol. Cell Biol. *13*, 616–630.

Massagué, J., Blain, S.W., and Lo, R.S. (2000). TGFβ Signaling in Growth Control, Cancer, and Heritable Disorders. Cell *103*, 295–309.

McCabe, M.T., Ott, H.M., Ganji, G., Korenchuk, S., Thompson, C., Van Aller, G.S., Liu, Y., Graves, A.P., Iii, A.D.P., Diaz, E., et al. (2012). EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature *492*, 108–112.

McDonald, O.G., Wu, H., Timp, W., Doi, A., and Feinberg, A.P. (2011). Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. Nat. Struct. Mol. Biol. *18*, 867–874.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Metzger, E., Wissmann, M., Yin, N., Müller, J.M., Schneider, R., Peters, A.H.F.M., Günther, T., Buettner, R., and Schüle, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436–439.

Miller, I.J., and Bieker, J.J. (1993). A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins. Mol. Cell. Biol. *13*, 2776–2786.

Mishra, V.K., and Johnsen, S.A. (2014). Targeted therapy of epigenomic regulatory mechanisms controlling the epithelial to mesenchymal transition during tumor progression. Cell Tissue Res. *356*, 617–630.

Moreno-Bueno, G., Portillo, F., and Cano, A. (2008). Transcriptional regulation of cell polarity in EMT and cancer. Oncogene *27*, 6958–6969.

Moses, H.L., Yang, E.Y., and Pietenpol, J.A. (1990). TGF-β stimulation and inhibition of cell proliferation: New mechanistic insights. Cell *63*, 245–247.

Moustakas, A., and Heldin, C.-H. (2007). Signaling networks guiding epithelial–mesenchymal transitions during embryogenesis and cancer progression. Cancer Sci. *98*, 1512–1520.

Müller, B.M., Jana, L., Kasajima, A., Lehmann, A., Prinzler, J., Budczies, J., Winzer, K.-J., Dietel, M., Weichert, W., and Denkert, C. (2013). Differential expression of histone deacetylases HDAC1, 2 and 3 in human breast cancer - overexpression of HDAC2 and HDAC3 is associated with clinicopathological indicators of disease progression. BMC Cancer *13*, 215.

Munshi, A., Tanaka, T., Hobbs, M.L., Tucker, S.L., Richon, V.M., and Meyn, R.E. (2006). Vorinostat, a histone deacetylase inhibitor, enhances the response of human tumor cells to ionizing radiation through prolongation of γ-H2AX foci. Mol. Cancer Ther. *5*, 1967–1974.

Myeroff, L.L., Parsons, R., Kim, S.-J., Hedrick, L., Cho, K.R., Orth, K., Mathis, M., Kinzler, K.W., Lutterbaugh, J., Park, K., et al. (1995). A Transforming Growth Factor β Receptor Type II Gene Mutation Common in Colon and Gastric but Rare in Endometrial Cancers with Microsatellite Instability. Cancer Res. *55*, 5545–5547.

Naber, H.P.H., Drabsch, Y., Snaar-Jagalska, B.E., ten Dijke, P., and van Laar, T. (2013). Snail and Slug, key regulators of TGF-β-induced EMT, are sufficient for the induction of single-cell invasion. Biochem. Biophys. Res. Commun. *435*, 58–63.

Nagasawa, S., Sedukhina, A.S., Nakagawa, Y., Maeda, I., Kubota, M., Ohnuma, S., Tsugawa, K., Ohta, T., Roche-Molina, M., Bernal, J.A., et al. (2015). LSD1 Overexpression Is Associated with Poor Prognosis in Basal-Like Breast Cancer, and Sensitivity to PARP Inhibition. PLoS ONE *10*.

Nakao, A., Afrakhte, M., Morn, A., Nakayama, T., Christian, J.L., Heuchel, R., Itoh, S., Kawabata, M., Heldin, N.-E., Heldin, C.-H., et al. (1997). Identification of Smad7, a TGFβ-inducible antagonist of TGF-β signalling. Nature *389*, 631–635.

Narla, G., Heath, K.E., Reeves, H.L., Li, D., Giono, L.E., Kimmelman, A.C., Glucksman, M.J., Narla, J., Eng, F.J., Chan, A.M., et al. (2001). KLF6, a Candidate Tumor Suppressor Gene Mutated in Prostate Cancer. Science *294*, 2563–2566.

Narla, G., DiFeo, A., Reeves, H.L., Schaid, D.J., Hirshfeld, J., Hod, E., Katz, A., Isaacs, W.B., Hebbring, S., Komiya, A., et al. (2005). A Germline DNA Polymorphism Enhances Alternative Splicing of the KLF6 Tumor Suppressor Gene and Is Associated with Increased Prostate Cancer Risk. Cancer Res. *65*, 1213–1222.

Narlikar, G.J., Fan, H.-Y., and Kingston, R.E. (2002). Cooperation between Complexes that Regulate Chromatin Structure and Transcription. Cell *108*, 475–487.

Nass, S.J., Herman, J.G., Gabrielson, E., Iversen, P.W., Parl, F.F., Davidson, N.E., and Graff, J.R. (2000). Aberrant methylation of the estrogen receptor and E-cadherin 5' CpG islands increases with malignant progression in human breast cancer. Cancer Res. *60*, 4346–4348.

Nieto, M.A. (2013). Epithelial Plasticity: A Common Theme in Embryonic and Cancer Cells. Science *342*, 1234850.

Noboru Ueki, Masaru Nakazato, Toshihisa Ohkawa, Tatsuhiko Ikeda, Yoshiki Amuro, Toshikazu Hada, and Kazuya Higashino (1992). Excessive production of transforming growth-factor β1 can play an important role in the development of tumorigenesis by its action for angiogenesis: validity of neutralizing antibodies to block tumor growth. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1137*, 189–196.

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

Ocaña, O.H., Córcoles, R., Fabra, Á., Moreno-Bueno, G., Acloque, H., Vega, S., Barrallo-Gimeno, A., Cano, A., and Nieto, M.A. (2012). Metastatic Colonization Requires the Repression of the Epithelial-Mesenchymal Transition Inducer Prrx1. Cancer Cell *22*, 709–724.

O'Connor, O.A., Heaney, M.L., Schwartz, L., Richardson, S., Willim, R., MacGregor-Cortelli, B., Curly, T., Moskowitz, C., Portlock, C., Horwitz, S., et al. (2006). Clinical Experience With Intravenous and Oral Formulations of the Novel Histone Deacetylase Inhibitor Suberoylanilide Hydroxamic Acid in Patients With Advanced Hematologic Malignancies. J. Clin. Oncol. *24*, 166–173.

Oft, M., Heider, K.-H., and Beug, H. (1998). TGFβ signaling is necessary for carcinoma cell invasiveness and metastasis. Curr. Biol. *8*, 1243–1252.

Padua, D., Zhang, X.H.-F., Wang, Q., Nadal, C., Gerald, W.L., Gomis, R.R., and Massagué, J. (2008). TGFβ Primes Breast Tumors for Lung Metastasis Seeding through Angiopoietin-like 4. Cell *133*, 66–77.

Parant, J.M., and Lozano, G. (2003). Disrupting TP53 in mouse models of human cancers. Hum. Mutat. *21*, 321–326.

Pece, S., Tosoni, D., Confalonieri, S., Mazzarol, G., Vecchi, M., Ronzoni, S., Bernard, L., Viale, G., Pelicci, P.G., and Di Fiore, P.P. (2010). Biological and Molecular Heterogeneity of Breast Cancers Correlates with Their Cancer Stem Cell Content. Cell *140*, 62–73.

Peinado, H., Ballestar, E., Esteller, M., and Cano, A. (2004a). Snail Mediates E-Cadherin Repression by the Recruitment of the Sin3A/Histone Deacetylase 1 (HDAC1)/HDAC2 Complex. Mol. Cell. Biol. *24*, 306–319.

Peinado, H., Ballestar, E., Esteller, M., and Cano, A. (2004b). Snail Mediates E-Cadherin Repression by the Recruitment of the Sin3A/Histone Deacetylase 1 (HDAC1)/HDAC2 Complex. Mol. Cell. Biol. *24*, 306–319.

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

Peinado, H., Lavotshkin, S., and Lyden, D. (2011). The secreted factors responsible for pre-metastatic niche formation: Old sayings and new thoughts. Semin. Cancer Biol. *21*, 139–146.

Pérez-Mancera, P.A., González-Herrero, I., Pérez-Caro, M., Gutiérrez-Cianca, N., Flores, T., Gutiérrez-Adán, A., Pintado, B., Sánchez-Martín, M., and Sánchez-García, I. (2005). SLUG in cancer development. Oncogene *24*, 3073–3082.

Pérez-Moreno, M.A., Locascio, A., Rodrigo, I., Dhondt, G., Portillo, F., Nieto, M.A., and Cano, A. (2001). A New Role for E12/E47 in the Repression ofE-cadherin Expression and Epithelial-Mesenchymal Transitions. J. Biol. Chem. *276*, 27424–27431.

Perillo, B., Ombra, M.N., Bertoni, A., Cuozzo, C., Sacchetti, S., Sasso, A., Chiariotti, L., Malorni, A., Abbondanza, C., and Avvedimento, E.V. (2008). DNA Oxidation as Triggered by H3K9me2 Demethylation Drives Estrogen-Induced Gene Expression. Science *319*, 202–206.

Perlman, R., Schiemann, W.P., Brooks, M.W., Lodish, H.F., and Weinberg, R.A. (2001). TGF-β-induced apoptosis is mediated by the adapter protein Daxx that facilitates JNK activation. Nat. Cell Biol. *3*, 708–714.

Petkova, N., Hennenlotter, J., Sobiesiak, M., Todenhöfer, T., Scharpf, M., Stenzl, A., Bühring, H.-J., and Schwentner, C. (2013). Surface CD24 distinguishes between low differentiated and transit-amplifying cells in the basal layer of human prostate. The Prostate *73*, 1576–1590.

Polyak, K., and Weinberg, R.A. (2009). Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat. Rev. Cancer *9*, 265–273.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Ramadoss, S., Chen, X., and Wang, C.-Y. (2012). Histone Demethylase KDM6B Promotes Epithelial-Mesenchymal Transition. J. Biol. Chem. *287*, 44508–44517.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191.

Reinholz, M.M., An, M.-W., Johnsen, S.A., Subramaniam, M., Suman, V.J., Ingle, J.N., Roche, P.C., and Spelsberg, T.C. (2004). Differential Gene Expression of TGFβ Inducible Early Gene (TIEG), Smad7, Smad2 and Bard1 in Normal and Malignant Breast Tissue. Breast Cancer Res. Treat. *86*, 75–88.

Reynisdóttir, I., Polyak, K., Iavarone, A., and Massagué, J. (1995). Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta. Genes Dev. *9*, 1831–1845.

Rhim, A.D., Mirek, E.T., Aiello, N.M., Maitra, A., Bailey, J.M., McAllister, F., Reichert, M., Beatty, G.L., Rustgi, A.K., Vonderheide, R.H., et al. (2012). EMT and Dissemination Precede Pancreatic Tumor Formation. Cell *148*, 349–361.

Ribeiro, A., Bronk, S.F., Roberts, P.J., Urrutia, R., and Gores, G.J. (1999). The transforming growth factor β1–inducible transcription factor, TIEG1, mediates apoptosis through oxidative stress. Hepatology *30*, 1490–1497.

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

Rice, J.C., and Allis, C.D. (2001). Histone methylation versus histone acetylation: new insights into epigenetic regulation. Curr. Opin. Cell Biol. *13*, 263–273.

Richon, V.M. (2006). Cancer biology: mechanism of antitumour action of vorinostat (suberoylanilide hydroxamic acid), a novel histone deacetylase inhibitor. Br. J. Cancer *95*, S2–S6.

Robertson, K.D. (2005). DNA methylation and human disease. Nat. Rev. Genet. *6*, 597–610.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., Zabierowski, S.E., Brafford, P.A., Vultur, A., Basu, D., Gimotty, P., Vogt, T., and Herlyn, M. (2010). A Temporarily Distinct Subpopulation of Slow-Cycling Melanoma Cells Is Required for Continuous Tumor Growth. Cell *141*, 583–594.

Ropero, S., and Esteller, M. (2007). The role of histone deacetylases (HDACs) in human cancer. Mol. Oncol. *1*, 19–25.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R., et al. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389–393.

Rotello, R.J., Lieberman, R.C., Purchio, A.F., and Gerschenson, L.E. (1991). Coordinated regulation of apoptosis and cell proliferation by transforming growth factor beta 1 in cultured uterine epithelial cells. Proc. Natl. Acad. Sci. *88*, 3412–3415.

Rowland, B.D., and Peeper, D.S. (2006). KLF4, p21 and context-dependent opposing forces in cancer. Nat. Rev. Cancer *6*, 11–23.

Saijo, K., Katoh, T., Shimodaira, H., Oda, A., Takahashi, O., and Ishioka, C. (2012). Romidepsin (FK228) and its analogs directly inhibit phosphatidylinositol 3-kinase activity and potently induce apoptosis as histone deacetylase/phosphatidylinositol 3-kinase dual inhibitors. Cancer Sci. *103*, 1994–2001.

Saltzman, A., Munro, R., Searfoss, G., Franks, C., Jaye, M., and Ivashchenko, Y. (1998). Transforming Growth Factor-β-Mediated Apoptosis in the Ramos B-Lymphoma Cell Line Is Accompanied by Caspase Activation and Bcl-XLDownregulation. Exp. Cell Res. *242*, 244–254.

Sambucetti, L.C., Fischer, D.D., Zabludoff, S., Kwon, P.O., Chamberlin, H., Trogani, N., Xu, H., and Cohen, D. (1999). Histone Deacetylase Inhibition Selectively Alters the Activity and Expression of Cell

Cycle Proteins Leading to Specific Chromatin Acetylation and Antiproliferative Effects. J. Biol. Chem. *274*, 34940–34947.

Sandhu, C., Garbe, J., Bhattacharya, N., Daksis, J., Pan, C.H., Yaswen, P., Koh, J., Slingerland, J.M., and Stampfer, M.R. (1997). Transforming growth factor beta stabilizes p15INK4B protein, increases p15INK4B-cdk4 complexes, and inhibits cyclin D1-cdk4 association in human mammary epithelial cells. Mol. Cell. Biol. *17*, 2458–2467.

Sandoval, J., and Esteller, M. (2012). Cancer epigenomics: beyond genomics. Curr. Opin. Genet. Dev. *22*, 50–55.

Scheel, C., and Weinberg, R.A. (2012). Cancer stem cells and epithelial–mesenchymal transition: Concepts and molecular links. Semin. Cancer Biol. *22*, 396–403.

Schenk, T., Chen, W.C., Göllner, S., Howell, L., Jin, L., Hebestreit, K., Klein, H.-U., Popescu, A.C., Burnett, A., Mills, K., et al. (2012). Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-trans-retinoic acid differentiation pathway in acute myeloid leukemia. Nat. Med. *18*, 605–611.

Schepers, A., and Clevers, H. (2012). Wnt Signaling, Stem Cells, and Cancer of the Gastrointestinal Tract. Cold Spring Harb. Perspect. Biol. *4*, a007989.

Schulte, J.H., Lim, S., Schramm, A., Friedrichs, N., Koster, J., Versteeg, R., Ora, I., Pajtler, K., Klein-Hitpass, L., Kuhfittig-Kulle, S., et al. (2009). Lysine-Specific Demethylase 1 Is Strongly Expressed in Poorly Differentiated Neuroblastoma: Implications for Therapy. Cancer Res. *69*, 2065–2071.

Schwarte-Waldhoff, I., Volpert, O.V., Bouck, N.P., Sipos, B., Hahn, S.A., Klein-Scory, S., Lüttges, J., Klöppel, G., Graeven, U., Eilert-Micus, C., et al. (2000). Smad4/DPC4-mediated tumor suppression through suppression of angiogenesis. Proc. Natl. Acad. Sci. *97*, 9624–9629.

Seligson, D.B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M., and Kurdistani, S.K. (2005). Global histone modification patterns predict risk of prostate cancer recurrence. Nature *435*, 1262–1266.

Selvakumaran, M., Lin, H.K., Sjin, R.T., Reed, J.C., Liebermann, D.A., and Hoffman, B. (1994). The novel primary response gene MyD118 and the proto-oncogenes myb, myc, and bcl-2 modulate transforming growth factor beta 1-induced apoptosis of myeloid leukemia cells. Mol. Cell. Biol. *14*, 2352–2360.

Senisterra, G., Wu, H., Allali-Hassani, A., Wasney, G.A., Barsyte-Lovejoy, D., Dombrovski, L., Dong, A., Nguyen, K.T., Smil, D., Bolshan, Y., et al. (2013). Small-molecule inhibition of MLL activity by disruption of its interaction with WDR5. Biochem. J. *449*, 151–159.

Serce, N., Gnatzy, A., Steiner, S., Lorenzen, H., Kirfel, J., and Buettner, R. (2012). Elevated expression of LSD1 (Lysine-specific demethylase 1) during tumour progression from pre-invasive to invasive ductal carcinoma of the breast. BMC Clin. Pathol. *12*, 13.

Serrano-Gomez, S.J., Maziveyi, M., and Alahari, S.K. (2016). Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. Mol. Cancer *15*, 18.

Sharma, S., Kelly, T.K., and Jones, P.A. (2010). Epigenetics in cancer. Carcinogenesis *31*, 27–36.

Shi, Y., and Massagué, J. (2003). Mechanisms of TGF-β Signaling from Cell Membrane to the Nucleus. Cell *113*, 685–700.

Shi, Y., and Whetstine, J.R. (2007). Dynamic Regulation of Histone Lysine Methylation by Demethylases. Mol. Cell *25*, 1–14.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone Demethylation Mediated by the Nuclear Amine Oxidase Homolog LSD1. Cell *119*, 941–953.

Shi, Y.-J., Matson, C., Lan, F., Iwase, S., Baba, T., and Shi, Y. (2005). Regulation of LSD1 Histone Demethylase Activity by Its Associated Factors. Mol. Cell *19*, 857–864.

Shih, J.-Y. (2005). Transcription Repressor Slug Promotes Carcinoma Invasion and Predicts Outcome of Patients with Lung Adenocarcinoma. Clin. Cancer Res. *11*, 8070–8078.

Shin, H., Liu, T., Manrai, A.K., and Liu, X.S. (2009). CEAS: cis-regulatory element annotation system. Bioinformatics *25*, 2605–2606.

Shipitsin, M., Campbell, L.L., Argani, P., Weremowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., et al. (2007). Molecular Definition of Breast Tumor Heterogeneity. Cancer Cell *11*, 259–273.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. *15*, 272–286.

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

Sleeman, J.P., Christofori, G., Fodde, R., Collard, J.G., Berx, G., Decraene, C., and Rüegg, C. (2012). Concepts of metastasis in flux: The stromal progression model. Semin. Cancer Biol. *22*, 174–186.

Song, J., Noh, J.H., Lee, J.H., Eun, J.W., Ahn, Y.M., Kim, S.Y., Lee, S.H., Park, W.S., Yoo, N.J., Lee, J.Y., et al. (2005). Increased expression of histone deacetylase 2 is found in human gastric cancer. APMIS *113*, 264–268.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. Nat. Rev. Cancer *6*, 846–856.

Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Möröy, T., Bartek, J., Massagué, J., Hänel, F., et al. (2001). Repression of p15INK4b expression by Myc through association with Miz-1. Nat. Cell Biol. *3*, 392–399.

Stavropoulos, P., Blobel, G., and Hoelz, A. (2006). Crystal structure and mechanism of human lysine-specific demethylase-1. Nat. Struct. Mol. Biol. *13*, 626–632.

Stearns, M.E., Garcia, F.U., Fudge, K., Rhim, J., and Wang, M. (1999). Role of Interleukin 10 and Transforming Growth Factor β1 in the Angiogenesis and Metastasis of Human Prostate Primary Tumor Lines from Orthotopic Implants in Severe Combined Immunodeficiency Mice. Clin. Cancer Res. *5*, 711–720.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Subramaniam, M., Harris, S.A., Oursler, M.J., Rasmussen, K., Riggs, B.L., and Spelsberg, T.C. (1995). Identification of a novel TGF-β-regulated gene encoding a putative zinc finger protein in human osteoblasts. Nucleic Acids Res. *23*, 4907–4912.

Subramaniam, M., Hefferan, T. e., Tau, K., Peus, D., Pittelkow, M., Jalal, S., Riggs, B. l., Roche, P., and Spelsberg, T. c. (1998). Tissue, cell type, and breast cancer stage-specific expression of a TGF-β inducible early transcription factor gene. J. Cell. Biochem. *68*, 226–236.

Subramaniam, M., Hawse, J.R., Johnsen, S.A., and Spelsberg, T.C. (2007). Role of TIEG1 in biological processes and disease states. J. Cell. Biochem. *102*, 539–548.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. *102*, 15545–15550.

Suske, G., Bruford, E., and Philipsen, S. (2005). Mammalian SP/KLF transcription factors: Bring in the family. Genomics *85*, 551–556.

Suzuki, C., Murakami, G., Fukuchi, M., Shimanuki, T., Shikauchi, Y., Imamura, T., and Miyazono, K. (2002). Smurf1 Regulates the Inhibitory Activity of Smad7 by Targeting Smad7 to the Plasma Membrane. J. Biol. Chem. *277*, 39919–39925.

Symonds, H., Krall, L., Remington, L., Saenz-Robles, M., Lowe, S., Jacks, T., and Van Dyke, T. (1994). p53-Dependent apoptosis suppresses tumor growth and progression in vivo. Cell *78*, 703–711.

Tachibana, I., Imoto, M., Adjei, P.N., Gores, G.J., Subramaniam, M., Spelsberg, T.C., and Urrutia, R. (1997). Overexpression of the TGFbeta-regulated zinc finger encoding gene, TIEG, induces apoptosis in pancreatic epithelial cells. J. Clin. Invest. *99*, 2365–2374.

Tam, W.L., and Weinberg, R.A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat. Med. *19*, 1438–1449.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat. Struct. Mol. Biol. *14*, 1025–1040.

Tetreault, M.-P., Yang, Y., and Katz, J.P. (2013). Kruppel-like factors in cancer. Nat. Rev. Cancer *13*, 701–713.

Thiery, J.P. (2002). Epithelial–mesenchymal transitions in tumour progression. Nat. Rev. Cancer *2*, 442–454.

Thiery, J.P., and Chopin, D. (1999). Epithelial Cell Plasticity in Development and Tumor Progression. Cancer Metastasis Rev. *18*, 31–42.

Thiery, J.P., and Sleeman, J.P. (2006). Complex networks orchestrate epithelial–mesenchymal transitions. Nat. Rev. Mol. Cell Biol. *7*, 131–142.

Tiwari, N., Gheldof, A., Tatari, M., and Christofori, G. (2012). EMT as the ultimate survival mechanism of cancer cells. Semin. Cancer Biol. *22*, 194–207.

Trojer, P., and Reinberg, D. (2006). Histone Lysine Demethylases and Their Impact on Epigenetics. Cell *125*, 213–217.

Tsai, J.H., Donaher, J.L., Murphy, D.A., Chau, S., and Yang, J. (2012). Spatiotemporal Regulation of Epithelial-Mesenchymal Transition Is Essential for Squamous Cell Carcinoma Metastasis. Cancer Cell *22*, 725–736.

Turner, J., and Crossley, M. (1999). Mammalian Krüppel-like transcription factors: more than just a pretty finger. Trends Biochem. Sci. *24*, 236–240.

Turner, N.C., Lord, C.J., Iorns, E., Brough, R., Swift, S., Elliott, R., Rayter, S., Tutt, A.N., and Ashworth, A. (2008). A synthetic lethal siRNA screen identifying genes mediating sensitivity to a PARP inhibitor. EMBO J. *27*, 1368–1377.

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

Vega, S., Morales, A.V., Ocaña, O.H., Valdés, F., Fabregat, I., and Nieto, M.A. (2004). Snail blocks the cell cycle and confers resistance to cell death. Genes Dev. *18*, 1131–1143.

Vesuna, F., van Diest, P., Chen, J.H., and Raman, V. (2008). Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer. Biochem. Biophys. Res. Commun. *367*, 235–241.

Vincan, E., and Barker, N. (2008). The upstream components of the Wnt signalling pathway in the dynamic EMT and MET associated with colorectal cancer progression. Clin. Exp. Metastasis *25*, 657–663.

Vogelstein, B., and Kinzler, K.W. (2004). Cancer genes and the pathways they control. Nat. Med. *10*, 789–799.

Wade Harper, J., Adami, G.R., Wei, N., Keyomarsi, K., and Elledge, S.J. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell *75*, 805–816.

Wang, Y., and Shang, Y. (2013). Epigenetic control of epithelial-to-mesenchymal transition and cancer metastasis. Exp. Cell Res. *319*, 160–169.

Wang, Z., Banerjee, S., Li, Y., Rahman, K.M.W., Zhang, Y., and Sarkar, F.H. (2006). Down-regulation of Notch-1 Inhibits Invasion by Inactivation of Nuclear Factor-κB, Vascular Endothelial Growth Factor, and Matrix Metalloproteinase-9 in Pancreatic Cancer Cells. Cancer Res. *66*, 2778–2784.

Warner, B.J., Blain, S.W., Seoane, J., and Massagué, J. (1999). Myc Downregulation by Transforming Growth Factor β Required for Activation of the p15Ink4b G1 Arrest Pathway. Mol. Cell. Biol. *19*, 5913–5922.

Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. Science *193*, 848–856.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell *153*, 307–319.

Wikström, P., Stattin, P., Franck-Lissbrant, I., Damber, J.-E., and Bergh, A. (1998). Transforming growth factor β1 is associated with angiogenesis, metastasis, and poor clinical outcome in prostate cancer. The Prostate *37*, 19–29.

Wilson, B.G., and Roberts, C.W.M. (2011). SWI/SNF nucleosome remodellers and cancer. Nat. Rev. Cancer *11*, 481–492.

Witta, S.E., Gemmill, R.M., Hirsch, F.R., Coldren, C.D., Hedman, K., Ravdel, L., Helfrich, B., Dziadziuszko, R., Chan, D.C., Sugita, M., et al. (2006). Restoring E-Cadherin Expression Increases Sensitivity to Epidermal Growth Factor Receptor Inhibitors in Lung Cancer Cell Lines. Cancer Res. *66*, 944–950.

Wu, C.-Y., Tsai, Y.-P., Wu, M.-Z., Teng, S.-C., and Wu, K.-J. (2012a). Epigenetic reprogramming and post-transcriptional regulation during the epithelial–mesenchymal transition. Trends Genet. *28*, 454–463.

Wu, Z.-Q., Li, X.-Y., Hu, C.Y., Ford, M., Kleer, C.G., and Weiss, S.J. (2012b). Canonical Wnt signaling regulates Slug activity and links epithelial–mesenchymal transition with epigenetic Breast Cancer 1, Early Onset (BRCA1) repression. Proc. Natl. Acad. Sci. *109*, 16654–16659.

Xu, J., Lamouille, S., and Derynck, R. (2009). TGF-β-induced epithelial to mesenchymal transition. Cell Res. *19*, 156–172.

Yamazaki, H., Suzuki, M., Otsuki, A., Shimizu, R., Bresnick, E.H., Engel, J.D., and Yamamoto, M. (2014). A Remote GATA2 Hematopoietic Enhancer Drives Leukemogenesis in inv(3)(q21;q26) by Activating EVI1 Expression. Cancer Cell *25*, 415–427.

Yang, J., and Weinberg, R.A. (2008a). Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis. Dev. Cell *14*, 818–829.

Yang, J., and Weinberg, R.A. (2008b). Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis. Dev. Cell *14*, 818–829.

Yang, F., Sun, L., Li, Q., Han, X., Lei, L., Zhang, H., and Shang, Y. (2012). SET8 promotes epithelial-mesenchymal transition and confers TWIST dual transcriptional activities: SET8 promotes TWIST-induced EMT. EMBO J. *31*, 110–123.

Yang, M.-H., Hsu, D.S.-S., Wang, H.-W., Wang, H.-J., Lan, H.-Y., Yang, W.-H., Huang, C.-H., Kao, S.-Y., Tzeng, C.-H., Tai, S.-K., et al. (2010). Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition. Nat. Cell Biol. *12*, 982–992.

Yao, D., Dai, C., and Peng, S. (2011). Mechanism of the Mesenchymal–Epithelial Transition and Its Relationship with Metastatic Tumor Formation. Mol. Cancer Res. *9*, 1608–1620.

Yoo, C.B., and Jones, P.A. (2006). Epigenetic therapy of cancer: past, present and future. Nat. Rev. Drug Discov. *5*, 37–50.

Yoshikawa, M., Hishikawa, K., Marumo, T., and Fujita, T. (2007). Inhibition of Histone Deacetylase Activity Suppresses Epithelial-to-Mesenchymal Transition Induced by TGF-β1 in Human Renal Epithelial Cells. J. Am. Soc. Nephrol. *18*, 58–65.

Yoshiura, K., Kanai, Y., Ochiai, A., Shimoyama, Y., Sugimura, T., and Hirohashi, S. (1995). Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. Proc. Natl. Acad. Sci. *92*, 7416–7419.

You, J.S., and Jones, P.A. (2012). Cancer Genetics and Epigenetics: Two Sides of the Same Coin? Cancer Cell *22*, 9–20.

Yu, Y., Wang, B., Zhang, K., Lei, Z., Guo, Y., Xiao, H., Wang, J., Fan, L., Lan, C., Wei, Y., et al. (2013). High expression of lysine-specific demethylase 1 correlates with poor prognosis of patients with esophageal squamous cell carcinoma. Biochem. Biophys. Res. Commun. *437*, 192–198.

Yuan, Y., Wang, Q., Paulk, J., Kubicek, S., Kemp, M.M., Adams, D.J., Shamji, A.F., Wagner, B.K., and Schreiber, S.L. (2012). A Small-Molecule Probe of the Histone Methyltransferase G9a Induces Cellular Senescence in Pancreatic Adenocarcinoma. ACS Chem. Biol. *7*, 1152–1157.

Zammarchi, F., Morelli, M., Menicagli, M., Di Cristofano, C., Zavaglia, K., Paolucci, A., Campani, D., Aretini, P., Boggi, U., Mosca, F., et al. (2011). KLF4 is a Novel Candidate Tumor Suppressor Gene in Pancreatic Ductal Carcinoma. Am. J. Pathol. *178*, 361–372.

Zhang, W., and Bieker, J.J. (1998). Acetylation and modulation of erythroid Krüppel-like factor (EKLF) activity by interaction with histone acetyltransferases. Proc. Natl. Acad. Sci. *95*, 9855–9860.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev. *15*, 2343–2360.

Zhang, J.-S., Moncrieffe, M.C., Kaczynski, J., Ellenrieder, V., Prendergast, F.G., and Urrutia, R. (2001). A Conserved α-Helical Motif Mediates the Interaction of Sp1-Like Transcriptional Repressors with the Corepressor mSin3A. Mol. Cell. Biol. *21*, 5041–5049.

Zhang, X., Choi, P.S., Francis, J.M., Imielinski, M., Watanabe, H., Cherniack, A.D., and Meyerson, M. (2016). Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. Nat. Genet. *48*, 176–182.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhao, L., Li, W., Zang, W., Liu, Z., Xu, X., Yu, H., Yang, Q., and Jia, J. (2013). JMJD2B Promotes Epithelial–Mesenchymal Transition by Cooperating with β-Catenin and Enhances Gastric Cancer Metastasis. Clin. Cancer Res. *19*, 6419–6429.

Zheng, H., and Kang, Y. (2014). Multilayer control of the EMT master regulators. Oncogene *33*, 1755–1763.

Zhou, S., Schuetz, J.D., Bunting, K.D., Colapietro, A.-M., Sampath, J., Morris, J.J., Lagutina, I., Grosveld, G.C., Osawa, M., Nakauchi, H., et al. (2001). The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat. Med. *7*, 1028–1034.