UNIVERSITÀ DEGLI STUDI DI MILANO
PhD Course in Molecular and Cellular Biology

XXX Ciclo

Role of the SMYD3 methylase in embryonic stem cells differentiation and zebrafish development

Pamela Floris

PhD Thesis

Scientific tutor: Dr. Giuseppina Caretti

Academic year: 2016-2017

SSD: BIO/18; BIO/11; BIO/13

Thesis performed at Department of Biosciences, Università degli Studi di Milano

INDEX

PART I

ABSTRACT

SET and MYND domain containing protein 3 (SMYD3) is a methyltransferase that methylates lysine 5 of histone 4 (H4K5) and a group of non-histone proteins. Furthermore, in order to regulate downstream genes, SMYD3 contacts RNA polymerase II and promotes the recruitment of positive coactivator 4 (PC4) and BRD4 at the proximal promoter and coding regions of target genes.

SMYD3 is overexpressed in several cancers, such as liver, breast and colorectal carcinomas, in which it promotes proliferation, migration and invasion of tumor cells. in order to promote cancer progression, SMYD3 activates the transcription of several cell cycle regulators, Myc and Ctnnb1 oncogenes, components of the IL6-Jak-Stat3 cascade and Epithelial Mesenchymal Transition (EMT) markers. Recently, a novel role for SMYD3 in regulating homologous recombination (HR) repair and maintaining genome stability has been reported.

SMYD3 plays a cytoplasmic role by methylating non-histone proteins. Indeed, SMYD3 methylates MAP3K2, vascular endothelial growth factor receptor 1 (VEGFR1), v-Akt Murine Thymoma Viral Oncogene Homolog 1 (AKT1), and Human Epidermal Growth Factor Receptor 2 (HER2).

Few evidences highlighted a possible role of SMYD3 in regulating developmental processes. In zebrafish, SMYD3 is required for correct maturation of cardiomyocytes and for trunk morphogenesis. However, the SMYD3-mediated molecular mechanisms, its functional role in development and its interplay with other proteins involved in development are still elusive.

Aim of the present PhD project was to better understand SMYD3 role throughout development, by using mESCs and zebrafish as model systems. After knocking down SMYD3 in mESCs, we observed an upregulation of several mesendodermal markers at different time points of differentiation. In parallel, an increase in mesendodermal markers was detectable in *smyd3* knocked down zebrafish embryos. Since SMYD3 is involved in cardiomyocyte maturation in zebrafish, we knocked down SMYD3 and prompted mESCs to differentiate towards cardiomyocytes by forming embryoid bodies (EBs). As a result, we observed an upregulation of early and late cardiovascular markers in SMYD3 depleted EBs. These data were corroborated by an upregulation of the endothelial marker *kdrl* in *smyd3* morphants. Taken together, our results suggest that SMYD3 negatively modulates mesendodermal commitment during mESC and zebrafish development.

As a second aim, we wanted to investigate the molecular mechanisms underlying SMYD3-mediated regulation of EMT. We assessed this aim by providing a link between TGFβ/SMAD signaling pathway, which is an important regulator of EMT during tumor progression, and SMYD3. Indeed, we demonstrated that SMYD3 directly interacts with SMAD3. Furthermore, the ability of SMYD3

inhibitor BCI-121 to impair cancer cell migration has been proved in zebrafish as a xenograft cancer model.

1 INTRODUCTION

1.1 HISTONE METHYLATION

Histone methylation was reported for the first time in *Salmonella typhimuriumhas* flagellar proteins (Ambler & Rees, 1959) and some years later on histone proteins (Murray, 1964).

Histone methylation takes place primarily on the side chains of arginine (R) and lysine (K) residues, although it can occur on other residues, such as histidine, aspartic acid, and glutamic acid (Alam et al., 2015). Unlike other post-translational modifications (PTMs), such as acetylation and phosphorylation, histone methylation does not modify protein charge (Greer & Shi, 2012).

The ω nitrogen atoms of arginine can be mono- or di-methylated, either symmetrically or asymmetrically. Up to now, H3R2, H3R8, H3R17, H3R26, H4R3, H2AR11 and H2AR29 have been reported to be methylated (Alam et al., 2015; Wesche, Kuhn, Kessler, Salton, & Wolf, 2017).

Lysine is methylated on its ε-amine and can be mono-, di- or tri-methylated. The conventional sites for methylation include several lysine residues of histone 3 (K4, K9, K27, K36 and K79), K20 of histone 4 and K26 of histone 1. With the exception of H3K79, these modifications are all placed in

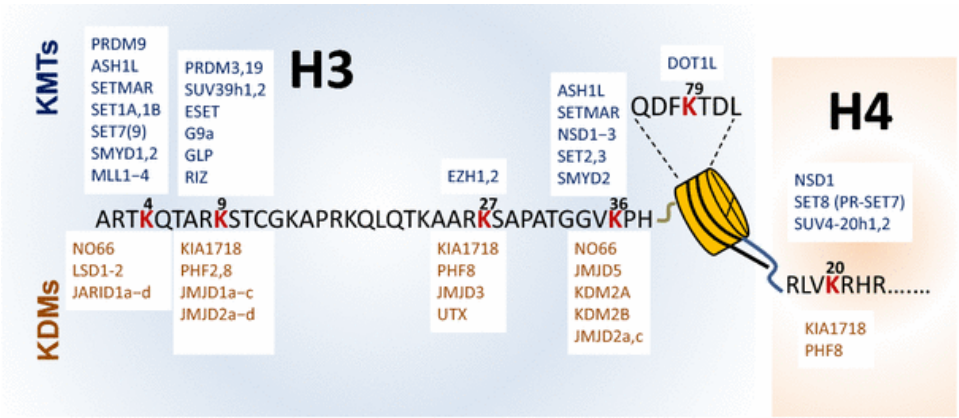


Figure 1.1. Schematic representation of main lysine methylation sites on histones H3 and H4 (Alam, Gu, & Lee, 2015)
Histone lysine methyltransferases (KMTs) and lysine demethylases (KDMs) are associated to their lysine methylation site.

the N-terminal tails of histone proteins (Figure 1) (Alam et al., 2015).

Histone lysine methylation regulates transcription and other chromatin-related processes (replication, recombination (Y. Zhang & Reinberg, 2001) and DNA-damage response (Sanders et al., 2004)). Moreover, histone methylation is involved in heterochromatin formation, X inactivation, genomic imprinting and homeotic genes silencing, demonstrating its importance in many biological processes (Martin & Zhang, 2005).

Unlike acetylation and arginine methylation, which generally correlate with transcriptional activation, histone lysine methylation can be associated with either activation or repression of the transcription process. This dual function depends on the methylation site, as well as on the number of methyl groups added. H3K4me3 is believed to be important for the initiation of transcription, given that this modified residue is highly enriched at the transcription start site (TSS) of active promoters. The sequence immediately flanking the TSS is also enriched in H3K4me1 and H3K4me2. H3K4me1 is widely distributed in enhancer regions (Calo & Wysocka, 2013). Moreover, active promoters are also associated with gene bodies enriched in H4K20me1, H2BK5me1, and H3K36me3, while inactive ones are often marked by H3K27me3 or H3K9me3. A peculiar subset of promoters, named bivalent promoters, are characterized by both H3K4me3 and H3K27me3 on different H3 tails within the same nucleosome, suggesting that intra-nucleosomal tail-to-tail crosstalk might be a pivotal regulatory mechanism (Black, Van Rechem, & Whetstine, 2012).

If compared to other PTMs such as acetylation, which possesses a fast turnover, histone lysine methylation has a slower turnover rate, similar to the one of the total histone. Furthermore, methylation displays a higher turnover rate when associated with active genes, if compared to repressive marks, which instead exhibit a slower turnover (Zee, Levin, DiMaggio, & Garcia, 2010). This particular slow kinetics of methylated histone lysine residues was suggested to confer epigenetic stability (Zee, Levin, DiMaggio, et al., 2010; Zee, Levin, Xu, et al., 2010).

Given its slow turnover rate, histone methylation was thought to be an irreversible process. The discovery of Lys-specific demethylase 1 (LSD1, also known as KDM1A), which has substrate specificity for H3K4me1/2 and H3K9me1/2, instead, revealed the real dynamic nature of histone methylation (Figure 1) (Metzger et al., 2005; Shi et al., 2004). Some methylation events might need to be stably maintained, depending on the biological context (e.g. inheritance through mitosis of a silenced heterochromatin state), whereas others may have to be responsive to change (for example, during cell differentiation in response to environmental stimuli) (Greer & Shi, 2012).

Indeed, the diverse array of methylation events provides a regulatory potential. According to the current model, histone methylation carries out its biological function primarily indirectly, through the identification of methyl marks by effector proteins (called ‘readers’). Readers recognize histone methyl marks and lead to transcriptional changes also by the recruitment of other proteins. Growing evidences also support direct actions of histone methylation, which include protein conformational changes. In fact, H3K79me2 alters the nucleosomal surface, while H4K20me3 is known to affect higher-order structure of chromatin (Lu et al., 2008). Another important feature is the interaction with other PTMs, that could represent a collaborative or an antagonistic relationship. In this context,

methylation can occur at multiple sites on the same histones. Importantly, some histone marks are thought to be mutually exclusive. For example, loss of the Polycomb repressive complex 2 (PRC2), that catalyzes H3K27me3 in embryonic stem cells, results in a global increase in H3K27 acetylation (Pasini et al., 2010). Combinatorial histone marks can also affect the recognition and binding by methyl-binding proteins. In fact, H3S10 phosphorylation prevents H3K9me3-binding protein heterochromatin protein 1 (HP1) from binding to histone tails during the M phase of the cell cycle (Fischle et al., 2005). Therefore, histones PTMs can play an important role in the recruitment of methyl-modifying enzymes to specific genomic locations and, in some instances, also alter their substrate specificity (Greer & Shi, 2012).

Recent evidences suggest that histone methylation is also affected by fluctuations in metabolism. Changes in the availability of cofactors (such as SAM, α-ketoglutarate (α-KG) and flavin adenine dinucleotide (FAD)) of methyltransferases and demethylases might induce adaptive responses affecting the methylation of proteins, especially histones (Murn & Shi, 2017). A reduced production of SAM in ES cells, induced by restricted metabolism of Thr, substantially decreases the global level of H3K4me3 leading to a slower growth and increased differentiation (Shyh-Chang et al., 2013).

Reader proteins contain methyl-lysine-binding motifs, comprising plant homeodomain (PHD), chromo, tudor, PWWP (Pro-Trp-Trp-Pro), WD40, bromo-adjacent homology BAH, ADD (ATRX-DNMT3-DNMT3L), ankyrin repeat, malignant brain tumor (MBT) and the zinc-finger CW (zn-CW) domains, through which they are capable of distinguishing methylated lysine residues based on their methylation state and surrounding amino acid sequences (Musselman, Lalonde, Cote, & Kutateladze, 2012). The recognition of methyl-lysine histone sequences is a conserved mechanism, in which the side chain of the methylated lysine places itself into an aromatic cage in the reader domain. The specific recognition of methylated histones by reader proteins recruits several components of the nuclear signaling network to chromatin, mediating fundamental processes such as gene transcription, DNA replication and recombination, DNA damage response and chromatin remodeling. Chromatin-associated complexes often contain different readers within one or more subunits that display specificities for distinct PTMs (Musselman et al., 2012).

Histone demethylases can be classified in two different functional catalytic families: monoamine oxidases and jumonji C (JmjC) domain-containing proteins. The first enzymatic family comprises LSD2, the only homologue of LSD1, and, in a flavin-dependent manner, catalyzes the demethylation of mono- and di-methyl lysine residues. On the other hand, JmjC domain-containing proteins demethylate mono-, di- and tri-methylated lysine residues. These enzymes are conserved

from yeast to human and are capable of demethylating histone and non-histone substrates (Hamamoto, Saloura, & Nakamura, 2015).

1.2 LYSINE METHYLTRANSFERASES

The addition of methyl groups at histone lysine residues is catalyzed by specific lysine methyltransferases (KMTs). These enzymes can be divided in two different classes on the basis of their catalytic domains. One class is characterized by a 130-140 amino acid long SET domain, which was originally identified in the Su(var)3–9, Enhancer-of-zeste and Trithorax proteins (Jenuwein, Laible, Dorn, & Reuter, 1998; Tschiersch et al., 1994). The SET domain is a highly evolutionarily conserved catalytic motif that catalyzes lysine methylation by adding methyl groups to lysine residues of other proteins using S-adenosylmethionine as a donor substrate. SET containing KMTs are distributed in seven families (SUV39, SET1, SET2, EZ, RIZ, SMYD AND SUV4-20) and few other members, such as SET7/9 and SET8 (Yi, Jiang, Li, & Jiang, 2015).

The second class of KMTs does not contain a SET domain but includes highly conserved proteins, such as yeast disruptor of telomeric silencing-1, DOT1 (also known as KMT4), its eukaryotic homologs (human and murine DOT1-Like, or DOT1L), and other members like methyltransferase-like 10 (METTL10) and METTL21A (Hamamoto et al., 2015). Most of the non-SET-domain enzymes belong to the seven-β-strand (7BS) methyltransferase family, which is characterized by a twisted β-sheet structure (Falnes, Jakobsson, Davydova, Ho, & Malecki, 2016).

The two classes of KMTs also differ for the molecular mechanism of substrate recognition. SET-domain containing proteins recognize their substrates through interaction with short and linear peptides, while non-SET-domain containing KMTs may recognize the tertiary structure of their targets (Falnes et al., 2016).

While SET containing KMTs generally methylate lysine residues within the histone N-terminal tails, DOT1 and DOT1L methylate lysine 79 within the globular core of histone 3 (Alam et al., 2015).

Moreover, SET domain frequently co-occurs with many other protein domains, suggesting that the SET domain doesn't act individually to determine substrate specificity (Herz, Garruss, & Shilatifard, 2013).

To date, the involvement of diverse KMTs in embryonic development has been demonstrated. In mammals, the Mixed lineage leukaemias (MLLs) are an evolutionarily conserved protein family, which bear homology with yeast Set1 and *Drosophila* Trx. Six MLL family members (MLL1, MLL2, MLL3, MLL4, Set1A, Set1B) show a H3K4-specific methyltransferase activity. MLL1, MLL2, MLL3, MLL4, Set1A and Set1B form different complexes in association with a central

scaffold consisting of Ash2, Wdr5, Rbbp5 and Dpy30, which are essential for the MLL-associated KMT activity (Dou et al., 2006). Furthermore, MLLs epigenetically modulate genes associated with cell-cycle regulation, embryogenesis and development. Importantly, among the genes regulated by MLLs, we find *Hox* genes (*HOXA-D*). *Hox* genes are finely regulated at different stages of development and are essential for normal physiological functions and growth (Ansari & Mandal, 2010). The importance of MLL function during development is highlighted by the fact that homozygous *Mll1* knockout mice die during embryogenesis (day 10.5). The mutant phenotype is characterized by defects in neural crest-derived structures of the branchial arches, cranial nerves and ganglia (B. D. Yu, Hanson, Hess, Horning, & Korsmeyer, 1998; B. D. Yu, Hess, Horning, Brown, & Korsmeyer, 1995). The expression of *Hox* genes is correctly initiated in *Mll1* knockout mice, but after day 9 the expression becomes dependent on Mll1 (B. D. Yu et al., 1998). Mll1 SET domain deletion in mice through homologous recombination in ES cells is not lethal: mice are viable and fertile and show skeletal defects and altered transcription of *Hox* genes during development. *Hox loci* are also characterized by a reduction of H3K4me1 and defects in DNA methylation (Terranova, Agherbi, Boned, Meresse, & Djabali, 2006).

1.2.1 NON-HISTONE PROTEIN METHYLTRANSFERASES

Much of what we know about KMTs derive from histone methylation, however growing evidences demonstrate that many of these enzymes can methylate also non-histone proteins (Huang & Berger, 2008; Z. Wu, Connolly, & Biggar, 2017; X. Zhang, Huang, & Shi, 2015; X. Zhang, Wen, & Shi, 2012). Similar to histones, lysine methylation of non-histone proteins can be reversed by demethylases (Huang & Berger, 2008) and it can be read by specific methyl-binding motifs in order to produce downstream outcomes (X. Zhang et al., 2015). Depending on the position of the methylated residue in the protein, the biological function can vary. Furthermore, many evidences displayed a cross-talk between protein lysine methylation and other PTMs. In fact, non-histone lysine methylation can block lysine ubiquitylation at specific sites, increasing protein half-life (Z. Wu et al., 2017). On the contrary, lysine methylation may also decrease the half-life of the protein promoting its ubiquitylation: this is the case of the DCAF1/DDB1/CUL4 E3 ubiquitin ligase complex (J. M. Lee et al., 2012). Moreover, it has been reported that a cross-talk with phosphorylation exists, which provides a diversification of the downstream signaling (Z. Wu et al., 2017). Lysine methylation in non-histone proteins has been associated with regulation of a plethora of biological processes, among which there are signal transduction, DNA damage response, protein folding, metabolism and also cell growth (Z. Wu et al., 2017).

G9a (also known as euchromatin histone methyltransferase 2, or EHMT2) forms homodimers and heterodimeric complexes with its homolog G9a-like protein (GLP) (Tachibana et al., 2005). It catalyzes H3K9me1 and H3K9me2, repressing genes at euchromatic regions (Rice et al., 2003; Tachibana et al., 2002). In addition, it has been shown that G9a methylates H3K27, H3K56 and residues of H1 variants (Weiss et al., 2010; H. Wu et al., 2011; Y. Yu et al., 2012). Beyond histones, G9a can also modify non-histone proteins, among which the first identified was itself. The autocatalytic G9a methylation recruits, *in vivo*, the epigenetic regulator heterochromatin protein 1 (HP1) and this interaction is abolished by an adjacent G9a phosphorylation (S. C. Sampath et al., 2007). Several G9a substrates are transcription factors, such as C/EBPβ, p53 and the myogenic regulatory factor MyoD (Huang et al., 2010; Ling et al., 2012; Pless et al., 2008). Additionally, G9a

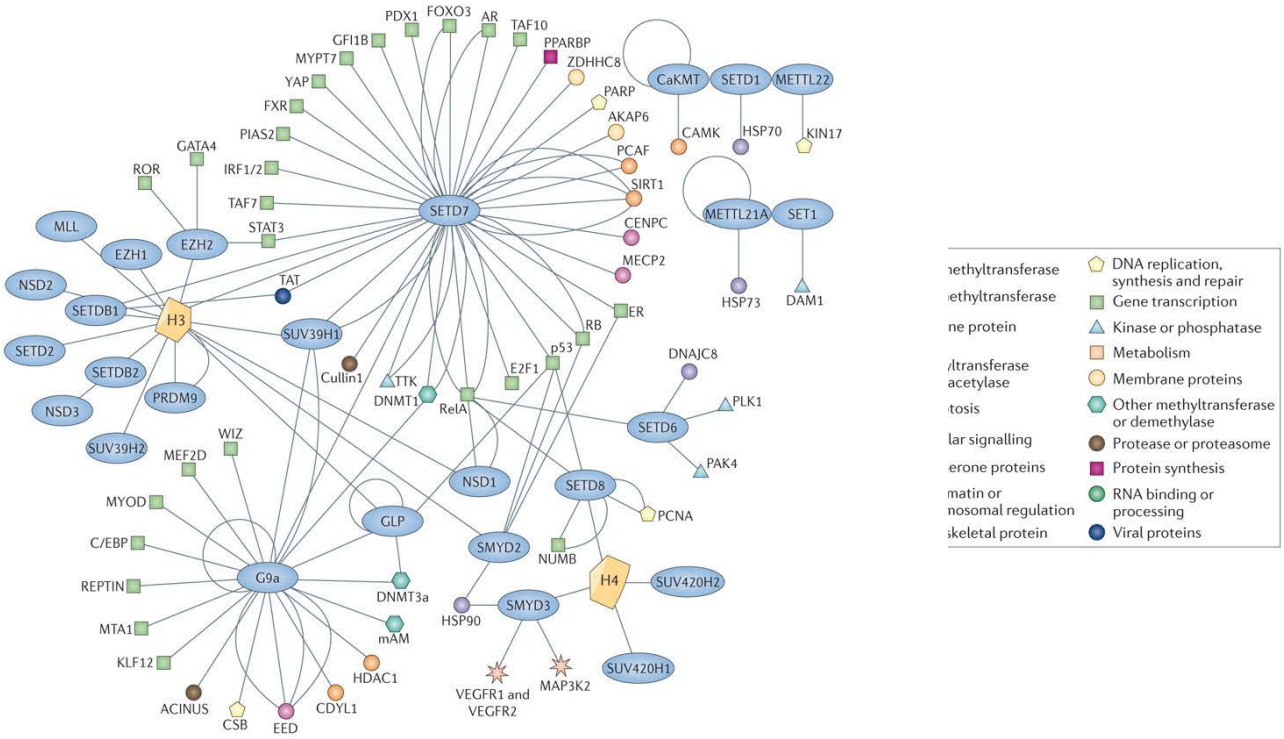


Figure 1.2. Lysine methyltransferase–substrate network (Adapted by (Biggar & Li, 2015)).
Lysine methyltransferases (KMTs) are associated with their histone and non-histone substrates as found on published literature. KMTs and their protein substrates are color-coded according to their functions.

methylates the chromatin-remodeling factor Reptin and it reversibly methylates MEF2D and metastatic tumor antigen 1 MTA1 (Figure 2) (J. Choi et al., 2014; J. S. Lee et al., 2010; Nair, Li, & Kumar, 2013).

Through a peptide array approach, Rathert and collaborators found that G9a recognizes an Arg-Lys site on its targets and the G9a catalytic activity is inhibited by methylation of the arginine residue. The peptides derived from chromodomain Y-like protein (CDYL1), widely interspaced zinc finger motifs protein (WIZ), Cockayne syndrome group B protein (CSB) showed the strongest methylation, whereas a weaker methylation was found with the peptides derived from ACINUS, histone deacetylase-1 (HDAC1), DNA methyltransferase-1 (Dnmt1) and Kruppel-like factor-12 (Kruppel) (Rathert et al., 2008).

1.2.2 SMYD FAMILY OF METHYLASES

SMYD family proteins contain evolutionary conserved SET and MYND domains. Currently, five members of SMYD family have been identified (SMYD1-5) (Figure 3). Unlike other histone methyltransferases, they are characterized by a distinctive SET domain, which is split in two by a myeloid-Nervy-DEAF-1 (MYND) domain followed by a cysteine-rich post-SET portion (J. M. Kim et al., 2015). As previously described for KMTs, the SET domain of SMYD proteins catalyzes

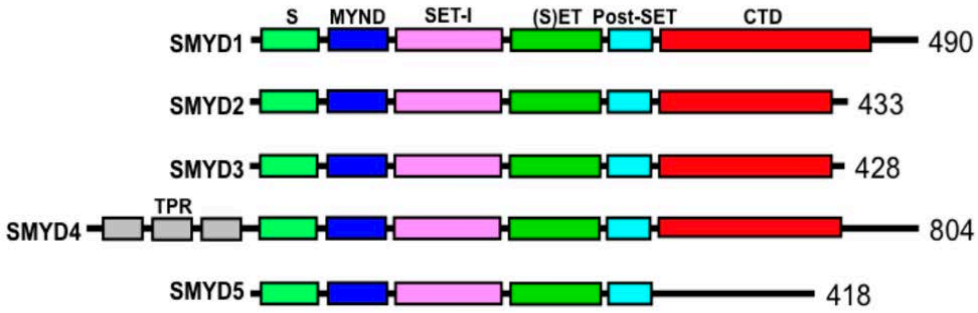


Figure 1.3. SMYD family members domains (Spellmon, Holcomb, Trescott, Sirinupong, & Yang, 2015).
S-sequence of the SET domain (S); MYND domain; insertion SET domain (SET-I); core SET domain ((S)ET; SET C-terminal flanking domain (Post-SET); C-terminal domain (CTD).

lysine methylation. Their MYND domain, instead is a cysteine-rich zinc finger motif involved in protein-protein interactions and DNA binding (Figure 3). A proper MYND domain folding is required for the catalytic activity of the SET domain through proper three-dimensional conformation (Foreman et al., 2011). With the exception of SMYD5, all family members share the presence of tetratricopeptide repeats (TPR) at the C-terminal domain (CTD) (Figure 3). Few evidences demonstrated that the CTD domain plays a regulatory role in modulating the catalytic activity of SMYD proteins (Du, Tan, & Zhang, 2014).

Several reports demonstrate that SMYD proteins play important roles in heart development and cardiomyogenesis (reviewed in (Du et al., 2014)).

SMYD1, previously known as BOP, has a histone methyltransferase activity and acts as a transcriptional repressor (Gottlieb et al., 2002; Tan, Rotllant, Li, De Deyne, & Du, 2006). SMYD1 is a direct downstream target of serum response factor (SRF) and myogenin, acting as a key regulator of myogenic differentiation (D. Li et al., 2009). Furthermore, *Smyd1* expression in the mouse developing heart depends on the direct interaction with the myocyte enhancer factor 2C (MEF2C) (Phan et al., 2005). *Smyd1* targeted deletion in mice displayed a role in maturation of ventricular cardiomyocytes and right ventricular development (Gottlieb et al., 2002). Consistent with its crucial role in mice cardiomyogenesis, *smyd1* knockdown in zebrafish (*Danio rerio*) embryos results in absence of heartbeat and disrupted myofibril organization in cardiac muscles (H. Li et al., 2013; Tan et al., 2006). Nonsense mutations in the zebrafish *smyd1* gene in the recessive mutant *flatline* (*fla*) confirmed the cardiac and skeletal muscle dysfunction (Just et al., 2011).

SMYD2 is strongly expressed in heart, brain, and skeletal muscle (Brown, Sims, Gottlieb, & Tucker, 2006). The methyltransferase catalyzes H3K36me2 and H3K4me1 less efficiently, *in vitro* (Abu-Farha et al., 2008; Brown et al., 2006). Its H3K4 catalytic activity is enhanced in presence of the heat shock protein HSP90A and promotes expression of genes involved in regulation of cell cycle and transcriptional regulation (Abu-Farha et al., 2008). Unexpectedly, SMYD2 is not required for heart development in mice. Indeed, conditional knockout mice with a cardiomyocyte-specific deletion of *Smyd2* had no effect on heart formation (Diehl et al., 2010). However, it has been reported that *smyd2* zebrafish morphants showed an impaired myofibril organization in cardiac muscles (Voelkel et al., 2013). SMYD2 methylates also non-histone targets: the tumor suppressor p53 at lysine K370, repressing its transcriptional regulation function (Huang et al., 2006); the retinoblastoma tumor suppressor RB at lysine K860 (Saddic et al., 2010); and cytoplasmic Hsp90 (Donlin et al., 2012).

SMYD4 has been identified as a potential tumor suppressor that inhibits platelet-derived growth factor receptor A polypeptide (PDGFR-A) in breast cancer (Hu, Zhu, Qi, & Zhu, 2009). During *Drosophila melanogaster* embryogenesis, the SMYD4 homologue, dSmyd4, is expressed in the mesoderm, with highest levels in the somatic musculature (Thompson & Travers, 2008). Knockdown of dSmyd4 expression using muscle-specific RNA interference gets to eclosion failure and late pupal lethality, suggesting a role for dSmyd4 in development or function of adult muscles (Thompson & Travers, 2008).

In contrast to the other members, **SMYD5** does not bind HSP90A (Abu-Farha et al., 2011). With a mainly nuclear role, SMYD5 catalyzes H4K20me3 controlling pro-inflammatory genes (Stender et

al., 2012). Recently, it has also been demonstrated that SMYD5 targets H4K20me3 in heterochromatin regions, in mouse embryonic stem cells (mESCs), leading to repression of lineage-specific genes and maintenance of the self-renewal gene expression (Kidder, Hu, Cui, & Zhao, 2017). SMYD5 also plays a role as epigenetic regulator of hematopoiesis during zebrafish embryogenesis (Fujii et al., 2016).

1.3 THE SMYD3 METHYLASE

Initially, SMYD3 methylation activity was studied through *in vitro* assays performed using a wild-type form of the protein purified from cell cultures, and it was described as a histone H3-methyltransferase (Hamamoto et al., 2004). Afterwards, it has been demonstrated that SMYD3 preferentially catalyzes H4K5me1, H4K5me2 and H4K5me3, employing bacterially expressed recombinant proteins (Van Aller et al., 2012). In some instances, SET-domain-containing histone methyltransferases require co-factors for their catalytic activity (Du et al., 2014). SMYD3 interacts with the heat shock-protein HSP90A through its C-terminal domain (Brown et al., 2015) and its activity is enhanced by the presence of the chaperone (Hamamoto et al., 2004).

The SET domain has been showed to be responsible for SMYD3 catalytic activity (Figure 4).

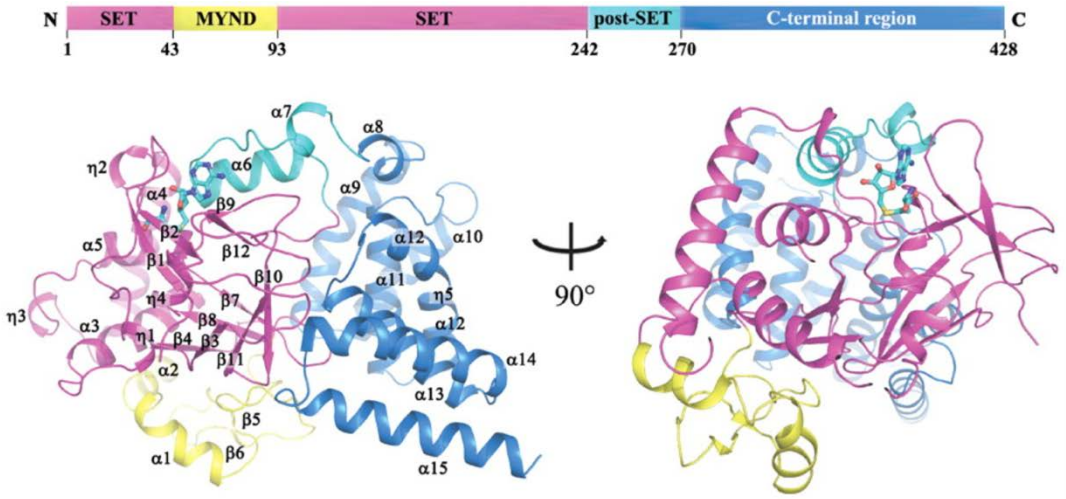


Figure 1.4. Structure of SMYD3-AdoHcy complex (Adapted by (Xu, Wu, Sun, Zhong, & Ding, 2011)). Top: a schematic representation of the full-length SMYD3. Bottom: two views of the structure of the Smyd3–AdoHcy complex are represented. The N-terminal SET domain (residues 1–43 and 94–242), the MYND domain (residues 44–93), the post-SET domain (residues 243–270), and the C-terminal region (residues 271–428) are colored in magenta, yellow, cyan and blue, respectively. The cofactor product AdoHcy is shown with a ball-and-stick model and colored in cyan.

Indeed, SMYD3 recombinant proteins carrying point mutations in the SET domain (F183A, N205A, Y239A and deletions of NHSC and EEL sequences) fail to methylate recombinant histone proteins *in vitro* (Hamamoto et al., 2004; Van Aller et al., 2012). Structural studies confirmed that SET domain forms the substrate binding pocket in collaboration with post-SET and TPR domains (Xu et al., 2011).

Several lines of evidence pointed out that SMYD3 exerts its function also in the cytoplasm through the methylation of non-histone proteins. SMYD3 methylates MAP3K2, potentiating the activation of Ras signalling in cancer cells (Mazur et al., 2014), and lysine 831 of vascular endothelial growth

factor receptor 1 (VEGFR1) *in vitro*, resulting in increased kinase activity (Kunizaki et al., 2007). Recently, other SMYD3 target proteins have been identified: v-Akt Murine Thymoma Viral Oncogene Homolog 1 (AKT1), in which lysine 14 methylation causes the kinase activation (Yoshioka et al., 2016), and Human Epidermal Growth Factor Receptor 2 (HER2), whose tri-methylation enhances the receptor homodimerization and the resulting autophosphorylation (Yoshioka et al., 2017).

Besides its methyltransferase activity towards histone and non-histone proteins, SMYD3 shows a role as a transcriptional cofactor. In order to regulate downstream genes, SMYD3 contacts RNA polymerase II (RNA Pol II) through a direct interaction with an RNA helicase (HELZ) (Hamamoto et al., 2004). Moreover, SMYD3 promotes the recruitment of positive coactivator 4 (PC4) (J. M. Kim et al., 2015), and BRD4 at the proximal promoter and coding regions of a subset of target genes, positively regulating their expression (Proserpio, Fittipaldi, Ryall, Sartorelli, & Caretti, 2013),

In contrast to other histone methyltransferases, SMYD3 can directly bind a 5′-CCCTCC-3′ *consensus* sequence, which can be found in the promoter of downstream genes (Hamamoto et al., 2004). Despite this DNA binding preference was confirmed in few regulatory regions at different loci, *de novo* motif search of SMYD3 ChIp-Seq analysis didn't confirm these findings (Sarris, Moulos, Haroniti, Giakountis, & Talianidis, 2016). The results of *in vitro* SMYD3-catalyzed methylation assays (Xu et al., 2011) and its crystal structure (Sirinupong, Brunzelle, Doko, & Yang, 2011; Xu et al., 2011) imply that the MYND domain can stimulate the SET-dependent histone methylase activity through direct DNA binding. However, SMYD3 recruitment to chromatin regulatory regions of the estrogen receptor (ER) target genes (H. Kim et al., 2009; Y. Li et al., 2011) does not require a consensus DNA sequence, indicating that a 5′-CCCTCC-3′ binding is not always a necessary step for histone methylation.

SMYD3 regulates the expression of several target genes like the homeobox *Nkx2.8* (Hamamoto et al., 2004), *WNT10B* (Hamamoto et al., 2006), the *Telomere Reverse Transcriptase* (*hTERT*) (Liu et al., 2007), *c-Met* (Proserpio et al., 2013; Zou et al., 2009) and retinoblastoma protein-interacting zinc finger gene 1 (RIZ1) (Dong et al., 2014).

Based on expression data in public datasets, SMYD3 is expressed at low levels ubiquitously in adult tissues. It is over-expressed in tumors, when compared to the matching tissue of origin (http://firebrowse.org/viewGene.html?gene=SMYD3).

1.3.1 SMYD3 IN DEVELOPMENT

While several studies focused on the SMYD3 role as a transcriptional co-activator in cancer, little is known on its possible role in regulating developmental processes. In zebrafish embryos, two transcripts of *smyd3* are ubiquitously expressed from early developmental stages (0.75 hpf) to 96 hpf. Through a morpholino microinjection approach, it has been demonstrated that SMYD3 is required for trunk morphogenesis in zebrafish (Figure 5, left panel) (Fujii, Tsunesumi, Yamaguchi, Watanabe, & Furukawa, 2011). Moreover, expression analyses of cardiac chamber markers, such as ventricular myosin heavy chain (*vmhc*), atrial myosin heavy chain (*amhc*) and cardiac myosin light

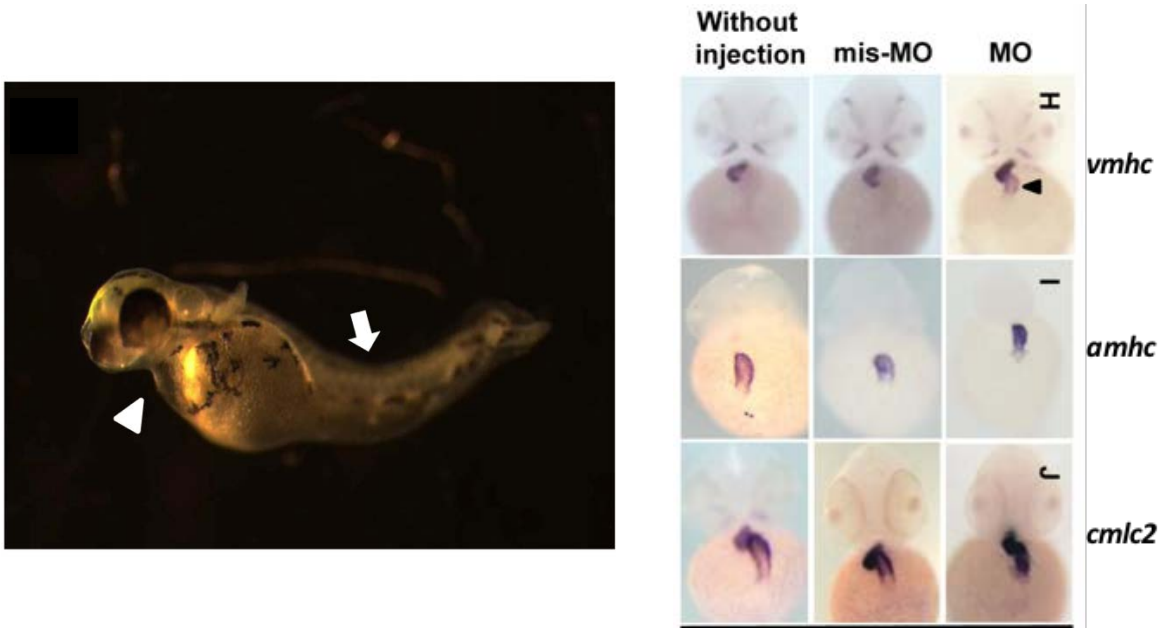


Figure 5. Effect of Smyd3 knockdown in zebrafish embryos. (Adapted by Fujii et al., 2011).
On the left, phenotype of embryos injected with smyd3-MO at 72 hpf. On the right, *in situ* hybridization of *vmhc, amhc and cmlc2* in morphants, control embryos and without injection embryos at 24 hpf.

chain2 (*cmlc2*) showed cardiac defects in SMYD3 morphants (Figure 5, right panel). Fujii and collaborators suggested that morphant phenotype may result from impaired maturation and/or delayed development of cardiomyocytes (Fujii et al., 2011).

During vertebrate development, heart and skeletal muscle arise from mesoderm (Kiecker, Bates, & Bell, 2016), through regulation mediated by the TGF-β signaling pathway and its downstream effector SMAD transcription factors (Gaarenstroom & Hill, 2014; Sakaki-Yumoto, Katsuno, & Derynck, 2013). However, the molecular mechanisms of SMYD3 role in development and its interplay with other proteins involved in development are still elusive.

Recently, Smyd3 knockout mice have been generated. However, despite the widespread functions of SMYD3, mutant mice develop normally and are viable and fertile, suggesting compensatory roles of other SMYD family members or other methyltransferases (Mazur et al., 2014).
Therefore, we decided to better investigate SMYD3 involvement during development by using two different models, mESCs and zebrafish, which recapitulate vertebrate embryonic development.

1.4 MODEL SYSTEMS TO STUDY DEVELOPMENT

1.4.1 MOUSE EMBRYONIC STEM CELLS (mESCs)

Mouse embryonic stem cells (mESC) derive from the inner cell mass of the preimplantation embryo and are capable of self-renewal, meaning that they can be propagated *in vitro* in an undifferentiated state. When allowed to differentiate, they can go towards endo-, ecto-, and meso-dermal derivatives and give origin to all somatic cell types of the organism (Thomson et al., 1998; Zhou, Kim, Yuan, & Braun, 2011). To mimic the *in vivo* differentiation of totipotent stem cells, mESCs cultivated *in vitro* are differentiated towards embryoid aggregates, also known as embryoid bodies (EBs). The generation of ES cell lines initially require the presence of a monolayer of inactivated mouse embryonic fibroblasts and/or the addition of the cytokine leukemia inhibitory factor (LIF) to the medium, which can both promote self-renewal and suppress differentiation. Therefore, the self-

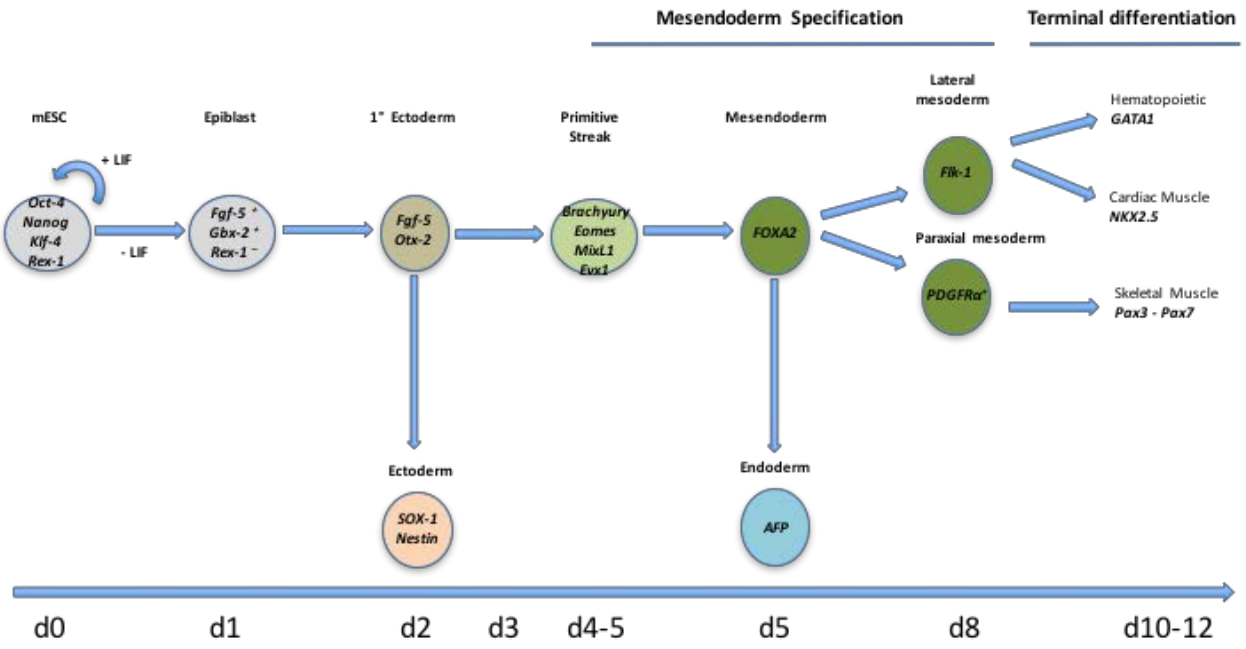


Figure 1.6. Schematic summary of factors involved in mESC differentiation in culture.
Markers that identify the progression through the major steps of mESC lineage differentiation are depicted.

renewal property of mESCs depends on the balance among diverse signaling molecules, characterized by an integrated transcriptional network regulated by the pluripotency-associated transcription factors Oct4, Nanog and Sox2. The activity of these transcription factors combined with LIF signaling allow mESCs to grow in an undifferentiated state (Figure 6) (Jaenisch & Young, 2008). LIF removal is sufficient to promote differentiation (Toyooka, Shimosato, Murakami, Takahashi, & Niwa, 2008). Moreover, growth factors that affect proliferation and survival of specific cell types (such as: basic fibroblast growth factor, bFGF; transforming growth factor, TGF; activin-A; bone morphogenic protein 4, BMP-4; hepatocyte growth factor, HGF; epidermal growth factor, EGF; nerve growth factor, NGF; retinoic acid, RA) are often added to the medium to promote differentiation towards a specific lineage (Kurosawa, 2007).

In mouse, mesoderm and endoderm specification occurs during gastrulation, and this process involves the movement of epiblast cells through a structure called primitive streak (Gadue, Huber, Paddison, & Keller, 2006): in fact, epiblast cells acquire mesenchymal features and migrate along the primitive streak, in between the endoderm and ectoderm layers, thus forming the mesoderm. Further patterning of the newly formed mesoderm leads to the formation of mesendoderm in the anterior primitive streak, and posterior extraembryonic mesoderm in the posterior primitive streak. While mesendoderm will lead to the somitic or paraxial mesoderm and the definitive endoderm, posterior mesoderm will contribute to tail and cardiac mesoderm (Willems & Leyns, 2008). Several of the molecular players that control patterning of the embryo have been characterized by complementing mouse models with ES cell models (Tam & Loebel, 2007). In fact, cultured early primitive ectoderm-like (EPL) cells, deriving from ESCs, share many properties with early primitive ectoderm and also with cells derived from the late primitive ectoderm, including morphology and increased expression of early and late primitive ectoderm markers, Fgf5 and Otx2 (Figure 6) (Vassilieva et al., 2012). In the mouse model, the ectoderm specification is then followed by the expression of late mesoderm and endoderm markers. Wnt signaling is required for the induction of mesoderm by inducing epiblast cells to adopt a paraxial mesodermal fate and to maintain the expression of the mesodermal transcription factor *T (Brachyury)* (Yamaguchi, Takada, Yoshikawa, Wu, & McMahon, 1999). The regulation of *Brachyury* expression by β-catenin activity is also confirmed by using differentiating mESCs (Turner, Rue, Mackenzie, Davies, & Martinez Arias, 2014). Brachyury is considered the earliest marker of mesodermal and also endodermal differentiation, both in ESC studies and during gastrulation (Figure 6) (Faial et al., 2015; Murry & Keller, 2008). Therefore, Nodal signaling induces differentiation towards mesendoderm in ESCs by inducing the expression of Goosecoid (Gsc), Mix-like homeodomain protein 1 (Mixl1), and other key differentiation genes (Figure 6) (X. Chen et al., 1997; Hart et al., 2002).

With the progression of embryonic development, mesoderm migrates and develops in tissues including endothelium and blood (Williamson et al., 2008). Hemangioblasts are the common progenitors for hematopoietic and endothelial cells and they can be detected at the midstreak stage of gastrulation (Huber, Kouskoff, Fehling, Palis, & Keller, 2004). Besides, ESC-derived embryoid bodies contain hemangioblasts (K. Choi, Kennedy, Kazarov, Papadimitriou, & Keller, 1998). Both in the ES-cell differentiation system and the developing embryo, the expression of *Brachyury* overlaps with the expression of the vascular endothelial growth factor 1 receptor 2, *Flk1* and the formation of the hemangioblast-like cells ($Bry^+/Flk\text{-}1^+$ cells) (Faloon et al., 2000; Williamson et al., 2008). This is followed by the development of $Bry^+/Flk\text{-}1^+$ cardiovascular progenitors which are capable of generating cardiomyocytes, smooth muscle cells, and endothelium (Kattman, Huber, & Keller, 2006). Subsequently, cardiovascular progenitors express the *Isl1* and *Nkx2.5* transcription factors, which are common markers in the three major cardiovascular lineages: endothelial, cardiac, and smooth muscle cells (Moretti et al., 2006).

mESCs have been widely used as a model to decipher the epigenetic mechanisms of cell fate determination and cell homeostasis (Zhou et al., 2011). Cardiac lineage commitment during heart development requires the intimate interaction between cardiac-specific transcription networks and epigenetic modifiers. In fact, several master cardiac transcription factors, such as Nkx2.5, Isl-1, Mef2c, GATA-4 and Tbx-5, have been shown to form complexes with epigenetic modifiers. These interactions result in epigenetic modifications of cardiac-specific genes' promoters, thus regulating gene expression during cardiac lineage specification (Zhou et al., 2011).

The molecular mechanisms underlying ES cells identity and their potential for differentiation are still not fully understood (Sese, Barrero, Fabregat, Sander, & Izpisua Belmonte, 2013). It has been shown that in undifferentiated ES cells, pluripotent genes are located in euchromatic transcriptionally active *loci*, whereas genes associated with differentiation are transcriptionally silent. Moreover, an involvement of histone PTMs has recently emerged as a pivotal epigenetic event for the regulation of the pluripotent state of ES cells and for cell fate determination (Meissner, 2010).

1.4.2 ZEBRAFISH AS A MODEL SYSTEM FOR DEVELOPMENT

Zebrafish (*Danio rerio*) represents a monophyletic species in the Cyprinidae family of teleost fish. Zebrafish are tropical fish living in small rivers, streams, paddy fields and channels in Southeast Asia. They possess a unique combination of features that makes them particularly amenable to experimental and genetic analyses of early vertebrate development (Holtzman, Iovine, Liang, & Morris, 2016).

Like other teleost, zebrafish underwent a whole-genome duplication event about 270 million years ago. As a result, 71% of human genes have at least one orthologue in zebrafish. Among orthologous human genes, 47% have a single zebrafish counterpart, whereas 24% have two (Howe et al., 2013). Moreover, the Genome Reference Consortium provided the latest zebrafish reference genome assembly, GRCz11 (https://www.ncbi.nlm.nih.gov/grc/zebrafish). Extensive expressed sequence tag (EST) sequencing and mapping projects are still on-going (https://genecollections.nci.nih.gov/ZGC/).

Since zebrafish adults are less than 3.5 cm in length, many thousands can be housed in a confined laboratory space. Zebrafish have a relatively short generation time, reaching the adulthood in nearly three months after fertilization, with a lifespan of two-three years. They are prolific breeders, generating approximately two hundred of embryos *per* week. As a result, hundreds of eggs from several breeding pairs can be gathered for genetic or experimental analyses.

In particular, zebrafish embryos are well suited for cell biology studies because they are fertilized and develop externally. These features are also a great advantage over mammalian systems for imaging during development. The fertilized embryos develop rapidly, making it possible to observe the entire course of early development in a short time. For instance, gastrulation is complete within 10 hours post fertilization (hpf) and somitogenesis begins at about 9 hpf. At 24 hpf the zebrafish embryo has already formed all the major tissues and many organ precursors, such as a beating heart, circulating blood, nervous system, eyes and ears, all of which can be readily observed under a simple dissecting microscope. Larvae hatch by about 2.5 days post fertilization (dpf) and they are swimming and feeding by 5-6 dpf (Beis & Stainier, 2006; Weinstein, 2002).

To date, numerous genetic screens have been performed in zebrafish, identifying mutants with very interesting phenotypes. These investigations contributed to the comprehension of basic vertebrate biology and vertebrate development. Screens in zebrafish not only enabled the systematic definition of a large range of early developmental phenotypes, but they contributed also more generally to the understanding of factors that control the specification of cell types, organ systems and body axes of vertebrates (Amsterdam et al., 1999; Driever et al., 1996; Haffter et al., 1996).

A variety of tools and methodologies have been developed to exploit the advantages of the zebrafish system. Zebrafish embryos and early larvae are optically clear, allowing for direct, non-invasive observation or experimental manipulation at all stages of their development such as whole mount *in situ* hybridization (WISH) analysis of gene expression patterns with extraordinarily high resolution (Vogel & Weinstein, 2000). The externally developing embryos are readily accessible to experimental manipulation by techniques such as microinjection of biologically active molecules (RNA, DNA or antisense oligonucleotides), cell and tissue transplantation/ablation, fate mapping and cell lineage tracing (Holder & Xu, 1999; Mizuno, Shinya, & Takeda, 1999; Nicoli, Ribatti, Cotelli, & Presta, 2007; Stainier, Lee, & Fishman, 1993). Different tissue-specific transgenic lines which label developing tissues or organs are available and they allow the observation of dynamic processes that are hard to visualize in other systems (Gut et al., 2013; Holtzman et al., 2016). Therefore, clear embryos are also well suited for different imaging tools, like confocal microscopy and selective plane illumination microscopy (SPIM) (Huisken & Stainier, 2009; Mickoleit et al., 2014; Reischauer, Arnaout, Ramadass, & Stainier, 2014).

Several strategies enable researchers to easily generate mutations in specific *loci*. Indeed, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR) combined with CRISPR-associated proteins 9 (Cas9), provide new genome editing approaches for establishing genetic knock-out and knock-in fish lines (M. Li, Zhao, Page-McCaw, & Chen, 2016).

ZEBRAFISH GASTRULATION

Zebrafish embryos allow many genetic and experimental advantages to study vertebrate gastrulation. For instance, several mutants showing defects in morphogenesis during gastrulation are available (Heisenberg & Tada, 2002).

Zebrafish gastrulation starts around 4 hpf, when the embryo has reached the blastula stage after meroblastic cleavage. The embryo is formed by a mass of cells, called blastoderm, residing atop the big yolk cell. The blastoderm consists of an epithelial enveloping layer (EVL) and a non-epithelial deep layer (DEL), from which all embryonic tissues arise (Heisenberg & Tada, 2002; Montero & Heisenberg, 2004). Beneath and vegetally to the blastoderm we find the yolk syncytial layer (YSL), formed by the yolk syncytial nuclei. At this time, EVL and DEL cells start to spread over the yolk cells through epibolic movements. In the DEL, epiboly is initiated by radial intercalation of blastula deep cells which move towards more superficial layers. This results in a thinning of DEL and an extension of its margin over the yolk (Figure 7). Conversely, the EVL does not go through radial intercalation, but it moves towards the vegetal pole. YSL also undergoes epiboly towards the vegetal pole (Lepage & Bruce, 2010). When the YSL and the blastoderm cover approximately half

of the yolk (50% epiboly), the germ layers start to be morphologically distinct. At this stage, mesendodermal progenitors internalize and accumulate at the blastoderm margin, forming the germ ring. The germ ring is organized in an inner hypoblast, which is formed by the internalizing mesendodermal progenitors, and a superficial epiblast layer composed by non-involuting ectodermal progenitors (Figure 7). As soon as the germ ring is formed, both hypoblast and epiblast converge towards the dorsal side of the gastrula and extend along the forming antero-posterior axis. In particular, at the dorsal side of the embryo these convergence movements establish a thickening in the germ ring, called shield, which gives rise to the embryonic organizer. Mesendodermal and ectodermal progenitors converging towards the dorsal side undergo medio–lateral cell intercalations, leading to a thinning of the forming body axis along its medio–lateral extent and a concomitant elongation along the antero–posterior axis (Figure 7) (Solnica-Krezel, 2006).

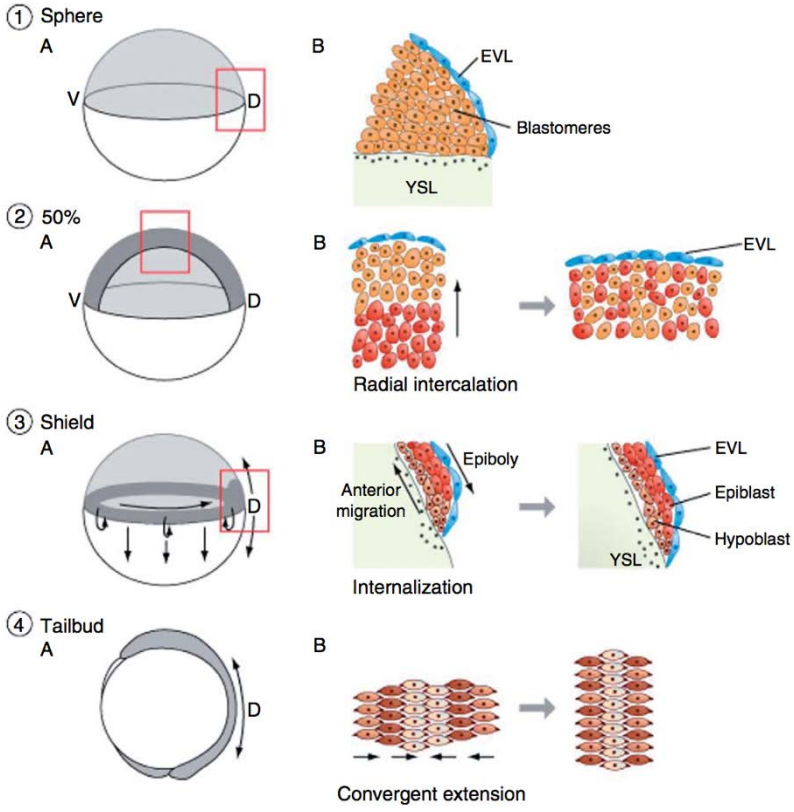


Figure 7. Zebrafish gastrulation movements (Rohde and Heisenberg, 2007).
1-4A, diagrams of gastrulation stages. 1-3B, closer views illustrate different embryonic tissue layers and gastrulation cell movements. Convergent extension movements are represented in 4B.

Like in other vertebrates, during zebrafish gastrulation, signaling molecules play pivotal roles in axis determination and germ layer induction and patterning. Indeed, the Nodal genes, *squint* (*ndr1*/*sqt*) and *cyclops* (*ndr2*/*cyc*), are expressed in both the YSL and the marginal blastoderm cells, and *cyc;sqt* double mutant embryos lack endodermal and mesodermal tissues (Erter, Solnica-Krezel, & Wright, 1998; Feldman et al., 1998; Rebagliati, Toyama, Haffter, & Dawid, 1998; K. Sampath et al., 1998). These observations suggest an essential role of Nodal ligands in zebrafish mesendoderm induction and demonstrate that the YSL is responsible for the Nodal signaling emanation, at the onset of gastrulation. It has been demonstrated that low concentrations of Nodal can induce mesodermal markers, while high Nodal concentrations induce endodermal markers (Dougan, Warga, Kane, Schier, & Talbot, 2003; Gritsman, Talbot, & Schier, 2000; Thisse, Wright, & Thisse, 2000).

Smad2 and Smad3 have been identified as specific mediators of Nodal (including Activin and Vg1) pathways, while Smad1, Smad5 (called Somitabun in zebrafish) and Smad8 mediate the BMP pathway. Phosphorylated Smads translocate into the nucleus and promote transcriptional activation of specific targets by binding to specific DNA-binding factors (Kimelman, 2006; Schier & Talbot, 2005). For instance, Smad association to the transcription factor Mixer (also known as Bonnie, Bon) results in activation of endoderm genes and endoderm formation, while the binding of Smads to FoxH1 (also known as Schmalspur, Sur) leads to mesoderm transcripts expression and mesoderm formation (Kimelman, 2006).

Endoderm induction, is regulated by Nodal signaling through activation of endoderm-specific genes. Among these genes there are the Sox gene *casanova* (*cas*), the GATA gene *faust*, and the homeobox genes *bon* and *mezzo* (Schier & Talbot, 2005). *faust* mutants show that *gata5* is required for the normal expression of *nkx2.5* and for multiple aspects of cardiac and endoderm development (Reiter et al., 1999). *cas* encodes a Sox-type transcription factor, which is expressed in endodermal progenitors in the marginal region of the late blastula embryo. *Casanova* expression is sufficient to induce endodermal differentiation. Indeed, in the absence of *cas*, endoderm progenitors differentiate towards mesodermal derivatives (Dickmeis et al., 2001; Kikuchi et al., 2001; Sakaguchi, Kuroiwa, & Takeda, 2001). It has been also demonstrated that the maternal T box transcription factor Eomesodermin (Eomes) is required for endogenous endoderm induction. Physically interacting with the Mix-type homeodomain transcription factor Bon and Gata5, Eomes promotes endoderm induction in marginal blastomeres by facilitating the assembly of a transcriptional activating complex on the *casanova* promoter (Bjornson et al., 2005).

1.5 SMYD3 IN CANCER

Cancer progression is a multistep process resulting from altered gene expression through genetic and epigenetic mechanisms (Hanahan & Weinberg, 2011). Several reports show how SMYD3 plays a role in cellular migration and invasion. This feature may be pivotal for tumor cells local invasion and ability to metastasize to distant sites. Since SMYD3 has been found to be upregulated in liver and colorectal carcinomas and to be related to the growth, proliferation and invasion of cancer cells (Hamamoto et al., 2004), most of the studies have been focused on a better characterization of SMYD3 role in tumor growth and have considered SMYD3 as an attractive candidate for therapeutic targeting in cancer.

SMYD3 is also overexpressed in breast cancer and activates the transcription of oncogenes (*WNT10B, N-myc, RIZ and hTERT*), cell cycle genes (*cyclinG1* and *CDK2*) and components of signal transduction pathways (*STAT1, MAP3K11* and *PIK3CB*) (Ren et al., 2011). Similar results were obtained also in HeLa cells, in which SMYD3 downregulation reduced the ability of cells to proliferate, form colonies and migrate (Wang et al., 2008). SMYD3 promotes up-regulation of the *MMP-9* metalloprotease, which is involved in tumor migration and invasion, both *in vitro* and in zebrafish xenograft model (Cock-Rada et al., 2012). Luo and collaborators reported that SMYD3 promotes the MRTF-A-mediated upregulation of myosin regulatory light chain 9 (MYL9), which is involved in cell movement, deformation process and cell migration (Luo et al., 2014).

SMYD3 loss in knockout mice inhibits the development of pancreatic and lung tumorigenesis, highlighting the importance of SMYD3 during tumorigenesis (Mazur et al., 2014).

Recently, Sarris and collaborators deepened the characterization of SMYD3 nuclear function, which is essential to drive cancer-related gene expression programs. SMYD3 activates the transcription of several cell cycle regulators, *Myc* and *Ctnnb1* oncogenes, components of the IL6-Jak-Stat3 cascade and EMT markers (Sarris et al., 2016). SMYD3 co-occupies already active chromatin regions occupied also by RNA Pol II and H3K4me3, and act as a “potentiator” of transcription. The proposed model suggests that SMYD3 selects its targets not necessarily with the recognition of a specific sequence, but through a combinatorial outcome of interaction with specific DNA sequence, with DNA binding factors, binding to the RNA Pol II and also interactions with H3K4me3 thanks to SMYD3 C-terminal domain. This functional specificity of SMYD3 is translated into the positive regulation of proliferation and cancer-promoting genes (Sarris et al., 2016).

A recent work reported a novel role for SMYD3 in regulating homologous recombination (HR) repair and maintaining genome stability. SMYD3 knockdown in breast cancer cell lines leads to a decreased expression of HR associated genes. This causes DNA-damage hypersensitivity,

decreased levels of repair foci, and an impairment of homologous recombination (Y. J. Chen, Tsai, Wang, & Teng, 2017).

1.6 EPITHELIAL-MESENCHYMAL TRANSITION (EMT) IN CANCER PROGRESSION

The epithelial-mesenchymal transition (EMT) is a fundamental process that occurs in both development and disease progression. During EMT, epithelial cells lose cell-cell junctions and polarity to gain a motile mesenchymal phenotype (Figure 8). EMT is a reversible process, in fact, mesenchymal cells can revert to epithelial cells via mesenchymal-epithelial transition (MET). In particular, MET is important for final developmental cellular differentiation (Skrypek, Goossens, De Smedt, Vandamme, & Berx, 2017) and for the growth of secondary metastasis (T. Chen, You, Jiang, & Wang, 2017).

Three classes of EMT are associated with different biological settings responsible for distinct functions. In primary, or type I, cells undergo several rounds of EMT and MET during implantation, embryogenesis and organogenesis. Secondary (type II) EMT is associated with wound healing, tissue regeneration, and organ development, whereas, tertiary EMT is induced in tumorigenesis and metastasis formation (Singh, Yelle, Venugopal, & Singh, 2017).

In physiological conditions, epithelial cells are characterized by cell-cell cohesion and tissue integrity, while mesenchymal cells correlate with migration and invasiveness (Figure 8). Alterations in cell morphology, cell–matrix adhesion, and acquisition of migration abilities are necessary events for the transition. Therefore, loss of the epithelial marker E-cadherin alone has been shown to initiate EMT (Singh et al., 2017). E-cadherin is responsible for maintaining intact cell–cell interactions and stabilizing the cytoskeleton, precluding tumor cell mobility, and subsequent dissemination. Thus, its downregulation destabilizes adherent junctions between cells and cell polarity, leading to a mesenchymal phenotype with migratory capabilities (Ghahhari & Babashah, 2015; Wijnhoven, Dinjens, & Pignatelli, 2000). During the transition, other key epithelial markers are lost, such as Mucin-1, Cytokeratins, Occludin, Desmoplakin and Claudins. On the contrary, mesenchymal markers, including N-cadherin, Vimentin, Smooth Muscle Actin (αSMA), Fibronectin, and Vitronectin are gained during the process (Figure 8) (Serrano-Gomez, Maziveyi, & Alahari, 2016).

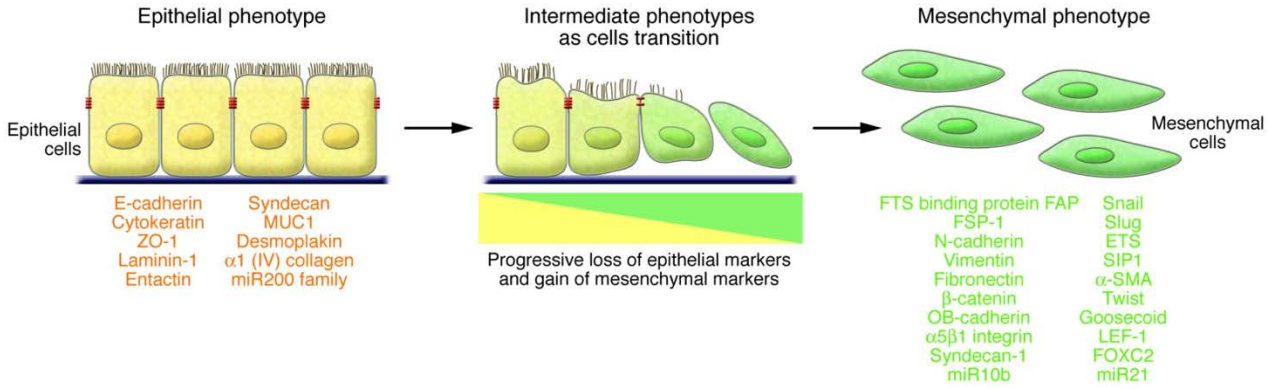


Figure 1.8. Epithelial-mesenchymal transition (Kalluri & Weinberg, 2009).
Functional transition of polarized epithelial cells into mobile and ECM component–secreting mesenchymal cells defines the EMT. The epithelial and mesenchymal cell markers commonly used are listed. Co-localization of these two sets of distinct markers defines an intermediate phenotype of EMT, indicating cells that have passed only partly through an EMT.

The interplay of different inducers, like transforming growth factor (TGF-β), fibroblast growth factor (FGF), receptor tyrosine kinases (RTKs), WNT/b-catenin, NOTCH and Hedgehog signaling pathways, is required for impairing the endothelial function (Ghahhari & Babashah, 2015; Goswami et al., 2016; Kalluri & Weinberg, 2009).

In addition, the EMT program is modulated by microRNAs and long noncoding RNAs (lncRNAs), as well as epigenetic and post-translational modifications, including methylation of cytosine residues in DNA CpG dinucleotides and histone modifications at N-terminal tails (Prieto-Garcia, Diaz-Garcia, Garcia-Ruiz, & Agullo-Ortuno, 2017; Serrano-Gomez et al., 2016).

More importantly, many transcription factors (TFs) have been identified as key regulators of EMT. They cooperate repressing epithelial genes and activating transcription of mesenchymal genes (Prieto-Garcia et al., 2017; Singh et al., 2017; Skrypek et al., 2017). For instance, SNAIL1, SNAIL2 (also known as Slug), and zinc finger E-box binding proteins (ZEBs) directly repress E-cadherin transcription, by binding to E-boxes present on the CDH1 promoter (Bolos et al., 2003). Additionally, SNAIL1 up-regulates the expression of several matrix metalloproteinase genes, including MMP9, which are essential for the basement membrane degradation and subsequent tumor cell invasion (Jorda et al., 2005; Kessenbrock, Plaks, & Werb, 2010). Other EMT-inducing TFs, are Kruppel-like factor (KLF) proteins, Forkhead box (FOX) proteins, Grainyhead-like (GRHL) proteins, the T-box protein Brachyury and Beta helix-loop-helix (βHLH) proteins, like Twist1 and Twist2 (Prieto-Garcia et al., 2017; Singh et al., 2017).

In tumorigenesis, EMT leads to the acquisition of several mesenchymal characteristics, allowing cancer cells to leave the primary tumor, to survive in the circulation and to invade a secondary site.

Once arrived at the metastatic site, cells can re-acquire an epithelial phenotype (MET) promoting colonization and developing metastasis (Figure 9) (Prieto-Garcia et al., 2017; Singh et al., 2017).

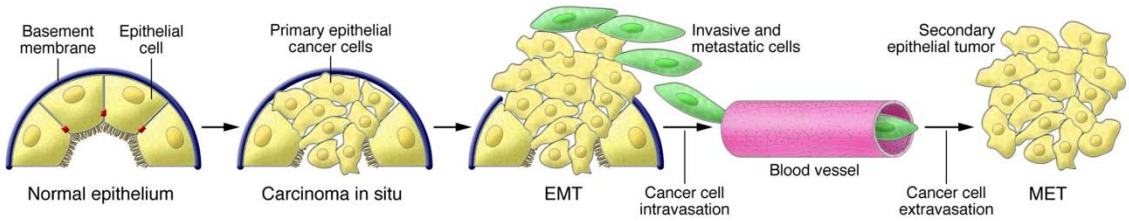


Figure 1.9. EMT in cancer progression (Kalluri and Weinberg, 2009).
Progression from normal epithelium to invasive carcinoma undergoes several stages. Epithelial cells lose their polarity and detach from the basement membrane. The composition of the basement membrane also changes, altering cell-ECM interactions and signaling networks. Subsequently, cancer cells are able to enter the circulation and exit the blood stream at a remote site, where they may form metastases, which may involve METs and thus a reversion to an epithelial phenotype.

Several studies demonstrated that TGF-β signaling induces EMT in diverse cancer types (Song, 2007). In particular, TGF-β induces EMT through activation of SMAD2/3 signaling or other non-canonical signaling pathways, such as PI3K/AKT or MAPK/ERK pathways (Massague, 2012). In the canonical TGF-β signaling pathway, TGF-β binds to its cell-surface receptors, which phosphorylate SMAD2 and SMAD3 (Massague, 2012). Once activated, SMAD2 and SMAD3 translocate into the nucleus where they downregulate epithelial genes and upregulate mesenchymal genes by associating with SMAD4 and other DNA-binding transcription factors and cofactors (Cantelli, Crosas-Molist, Georgouli, & Sanz-Moreno, 2017; Massague, 2012).

As mentioned above, recent evidences have demonstrated that SMYD3 acts as a transcriptional regulator of EMT genes in lung and colon cancer (Sarris et al., 2016). Therefore, we decided to better investigate SMYD3 transcriptional role during EMT in breast cancer cells (Figure 10).

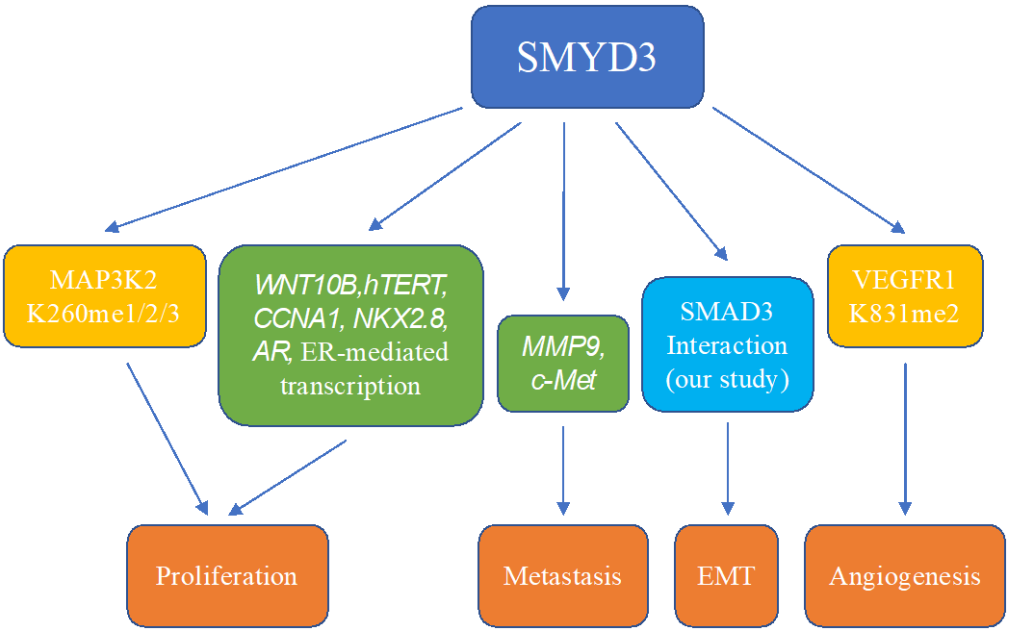


Figure 1.10. Schematic representation of SMYD3 role in cancer progression.
SMYD3 promotes cancer progression through different mechanisms.

1.7 ZEBRAFISH IN XENOGRAFT CANCER MODELS

In addition to developmental studies, zebrafish has emerged as a suitable model for cancer research. Indeed, the feasibility of xenografting human cells to zebrafish has been documented by several reports until now. A high interspecies conservation of molecular pathways exists between zebrafish and mammals. Additionally, a plethora of human oncogenes and tumor suppressor genes have zebrafish homologues (B. Zhang, Xuan, Ji, Zhang, & Wang, 2015). Compared to immunopermissive mice, xenotransplantation of human malignant cells of different origins into zebrafish may represent a helpful approach to study different aspects of tumor growth, including self-renewal of cancer stem cells (CSCs), tumor angiogenesis, invasiveness and metastatic behaviour (Konantz et al., 2012; B. Zhang et al., 2015). Moreover, a successful tumor engraftment in zebrafish can provide a readout within a few days (B. Zhang et al., 2015).

Recently, zebrafish has become a cost-effective alternative to mammals to study drug efficacy and toxicity. In fact, small molecular weight compounds can be dissolved in fish water and their impact can be investigated real-time (Konantz et al., 2012; Tobia, De Sena, & Presta, 2011; B. Zhang et al., 2015). Therefore, zebrafish xenotransplantation may also contribute to better assess drug response and high throughput screening of novel anti-cancer agents (Tobia et al., 2011; B. Zhang et al., 2015; Zhao, Huang, & Ye, 2015).

Tumor cell xenotransplantation can be performed both into adult and juvenile zebrafish, and thanks to transparent or transgenic fluorescent lines tumor engraftment, growth, and metastasis are easily observed. However, juvenile and adult zebrafish need immune suppression to avoid rejection, because immature T cells arise in the thymus by 3–4 dpf (Konantz et al., 2012; B. Zhang et al., 2015).

On the contrary, injection of human cancer cells in embryos at 48 hpf does not require immune suppression. Therefore, zebrafish embryos at 2 dpf are widely considered the ideal developmental stage for xenografting human cells. At this stage, different injection sites, such as blastodisk, yolk sac, hindbrain ventricle, bloodstream and perivitelline space, have been reported (B. Zhang et al., 2015).

Since zebrafish represents a powerful model to study tumor progression, we decided to employ zebrafish xenografts to evaluate the ability of the SMYD3 inhibitor BCI-121 (Peserico et al., 2015) to impair breast cancer cell migration. BCI-121 is a small molecule that has been identified, among other several candidates, through a virtual screening to identify new SMYD3 inhibitors. It has been reported that BCI-121 can induce a significant reduction in SMYD3 activity both *in vitro* and in colon rectal cancer cells (Peserico et al., 2015).

2 AIMS OF THE PROJECT

1. The current knowledge about SMYD3 is mainly related to its role in tumor progression. Nevertheless, there are also few evidences on its involvement in muscular and cardiac development in zebrafish. Therefore, the focus of my PhD project was to better characterize SMYD3 function in development. In order to assess this aim, we decided to use a widely used model to study development, mESCs. In addition, in order to confirm the results obtained in mESCs, we decided to use zebrafish as a second developmental model. Both of these models represent well-suited systems to investigate developmental processes.

2. SMYD3 promotes cancer cell proliferation and invasion, and it's overexpressed in several cancer types. Recently, it has been reported by Sarris and collaborators that SMYD3 acts as a "potentiator" of transcription for proliferation and EMT genes, in cancer progression. However, the molecular mechanisms through which SMYD3 promotes tumor cells migration and invasion are still elusive. Therefore, we aimed at elucidating the underlying dynamics of SMYD3-mediated regulation of EMT. The SMAD-mediated TGF-β signaling pathway is responsible for EMT in diverse types of cancer, especially breast cancer. For this reason, we employed breast cancer cells as a model system and dissected SMYD3 link with the TGF-β/SMAD signaling pathway. In addition, we also employed xenografts in zebrafish as a model system to investigate the metastatic process.

3 MAIN RESULTS AND CONCLUSIONS

3.1 SMYD3 IN DEVELOPMENT

A considerable amount of data about SMYD3 is related to its function in different cancer lines. Since a previous study suggested that SMYD3 is necessary for cardiac and skeletal muscle formation in zebrafish embryos (Fujii et al., 2011), the main focus of our studies was to clarify the role played by SMYD3 throughout embryonic development. In order to study SMYD3, we used mESCs as a model that recapitulates embryonic development. In particular, we focused our attention on mesodermal lineage differentiation. In parallel, we wanted to investigate SMYD3 role in an *in vivo* model. Therefore, we chose zebrafish to perform our analyses, because it is widely considered a well-suited developmental model.

First, we evaluated by qRT-PCR SMYD3 expression throughout mESCs differentiation. After a decrease at early stages, SMYD3 is stably expressed in mESCs. Additionally, immunofluorescence assays suggested that its expression is not restricted to nuclei, but is also detectable in cytoplasm.

Therefore, we decided to investigate whether SMYD3 has a role in mESC differentiation. We performed SMYD3 knock down through short hairpin RNA-interference and we evaluated pluripotency genes expression levels by qRT-PCR. *Nanog*, *KLF-4* and *Oct4* expression levels were comparable in both SMYD3 depleted mESCs and Sh-Scramble ones, suggesting that SMYD3 depletion does not affect mESC stemness.

We also wanted to assess if SMYD3 is involved in the regulation of mESC differentiation potential. Following leukemia inhibitory factor (LIF) removal, we prompted mESCs to form embryoid bodies (EBs). Subsequently, we analyzed mRNA expression pattern of several lineage markers at different time points of differentiation. We observed that the expression levels of the mesendodermal markers *Eomes*, *Mesp1*, *Sox17* and *FoxA2* were significantly increased in SMYD3 depleted EBs in comparison with Sh-Scramble EBs. On the contrary, the expression pattern of the primitive ectodermal markers *Fgf5* and *Otx2* was similar in SMYD3 depleted EBs and control ones.

In parallel, to investigate the role of *smyd3* in zebrafish development, we performed *smyd3* knock-down by injecting in 1-2 cell stage embryos a translation blocking morpholino already published and validated (Fujii et al., 2011). We decided to use only the translation blocking morpholino because, according to Fujii and colleagues, the injection of the ATG morpholino has a more pronounced and clearer effect (in terms of embryos' head and trunk defects) than that of the splicing morpholino. Therefore, we decided that double or co-injection of the two morpholinos in our experiments was not necessary. *smyd3* knocked-down embryos exhibited heart and trunk defects, confirming the phenotype observed by Fujii and collaborators. In particular, we observed

different grades of pericardial edema, impaired blood circulation and curved trunk in *smyd3* morphants compared to embryos injected with standard control morpholino. Therefore, we divided *smyd3* knocked-down embryos in four phenotypic classes following an increased severity.

Afterwards, we decided to verify the expression of genes involved in zebrafish early developmental stages. In particular, we were interested in analyzing genes whose expression is associated with the three germ layers during zebrafish gastrulation. Through *in situ* hybridization (ISH) analyses performed at 50% epiboly stage, we observed an increase of the mesendodermal markers *gata6*, *gata5*, *mespaa* and *snai1a* in knocked-down embryos if compared to standard control ones. Through ISH, we also observed a signal increase of *sox17* expression in *smyd3* morphants at 70% epiboly stage.

Subsequently, we focused our attention on some dorso-ventral pattern genes because we first observed by ISH an increased expression of the dorsal marker *goosecoid* (*gsc*) in knocked-down embryos at 50% epiboly stage. This effect was accompanied also by an expression enhancement of another dorsal marker, *chordin* (*chd*). As a consequence, a following reduction in the expression of the ventral marker *gata2* was detected by ISH at the same developmental stage in *smyd3* morphants. All these results suggest that SMYD3 plays a role in mesendodermal cell fate induction in zebrafish embryos, corroborating the data obtained using mESCs as a model.

As mentioned above, we previously demonstrated that SMYD3 depleted EBs are likely to differentiate towards mesendodermal lineage. The next step of our study was to investigate whether the increase of mesendodermal precursors in SMYD3 depleted EBs resulted also in an enhancement of differentiated cells deriving from mesoderm. Since Fujii and collaborators demonstrated that SMYD3 is necessary for a correct cardiomyocyte formation in zebrafish embryos, we prompted Sh-SMYD3 and control cells to differentiate towards cardiomyocytes. Therefore, we evaluated, through qRT-PCR, the expression levels of early markers of multipotent cardiovascular progenitors. Indeed, *Nkx2.5* and *Isl-1* showed an upregulation in SMYD3 depleted EBs in comparison with Sh-Scramble EBs. Regarding the second mesodermal marker *Flk1,* immunoblot analysis displayed an anticipation of Flk1 protein expression from day 8 to day 6 in SMYD3 depleted EBs. This result was also confirmed by immunofluorescence assays in Sh-SMYD3 and Sh-Scramble EBs at day 5 of differentiation.

In order to assess whether SMYD3 depletion also increases the expression of late cardiovascular lineage markers in EBs, we analyzed transcript levels of late cardiac markers, such as *TnnT2*, *Myh7*, and *Myl7*, which are important component of the contractile machinery. By qRT-PCR, we detected an increased expression of *TnnT2*, *Myh7*, and *Myl7*. This finding was also accompanied by a higher percentage of beating SMYD3 depleted EBs if compared to Sh-Scramble ones.

Since Flk1 positive cells give rise to cardiomyocytes, smooth muscle and endothelial cells populations, we decided to query the expression levels of smooth muscle actin (αSMA) and endothelial marker CD31. Both protein and mRNA levels of αSMA were increased in SMYD3 depleted EBs at day 10 of differentiation. Similarly, immunofluorescence experiments suggested an increase of CD31 at day 10.

In order to better investigate if SMYD3 knockdown affected mesodermal lineages *in vivo*, we decided to verify the expression of the endothelial marker *kinase insert domain receptor like*, *kdrl*, the *Flk1* orthologue in zebrafish. Through *in situ* hybridization experiments performed on 48 hpf embryos, we found an upregulation of *kdrl* in *smyd3* knocked-down ones. In particular, a *kdrl* upregulation was detectable in cardinal vein, posterior caudal vein and intersegmental vessels.

3.2 SMYD3 IN CANCER

As a side project, I also contributed to the elucidation of SMYD3 role in Epithelial Mesenchymal Transition (EMT) in breast cancer cell lines (Part III).

3.2.1 MAPPING SMYD3/SMAD3 INTERACTIONS

Recently, Sarris and collaborators demonstrated that SMYD3 potentiates transcription of many cancer-promoting genes during liver and colon cancer development, including EMT genes. In particular, SMYD3 is required for transcription of epithelial mesenchymal transition (EMT) key regulators (Sarris et al., 2016). EMT is initiated and maintained by several signaling pathways, among which TGFβ signaling plays a pivotal role. TGFβ pathway activation results in SMAD2 and SMAD3 phosphorylation and translocation into the nucleus, where they promote transcriptional activation of EMT-inducing transcription factors.

The aim of this project was to better describe the underlying dynamics of SMYD3-mediated regulation of EMT and investigate SMYD3 link with the TGFβ/SMADs signaling pathway.

During my Ph.D., I contributed to better characterize the interaction between SMYD3 and SMAD3. First, I overexpressed in HEK293T the human wild type forms of the two proteins and subsequently performed a co-Immunoprecipitation (co-IP). Wild type SMYD3 and SMAD3 co-immunoprecipitated, suggesting an interaction between them in cell extracts. Therefore, we wanted to identify SMYD3 regions responsible for the interaction with SMAD3. I performed co-IP with four SMYD3 mutants that lack different regions of the protein and full-length SMAD3. SMYD3 mutants were previously generated in our lab. I mapped the region of interaction, since only SMYD3 mutants that contained the C-terminal domain co-immunoprecipitated with SMAD3, indicating that C-termini is necessary for the interaction.

In addition, we wanted to understand whether the interaction between SMYD3 and SMAD3 was direct, or potentially mediated by additional partners. First, I produced in *E. coli* and purified a catalytically active recombinant form of SMYD3 through affinity chromatography. After GST tag cleavage, I performed the co-IP of recombinant SMYD3 and commercial available recombinant SMAD3. The result of this experiment pointed out that there is a direct interaction between the two proteins.

Since SMYD3 methylates and regulates several non-histone targets, regulating their function, we hypothesized that SMYD3 could also methylate SMAD3. Therefore, I performed an *in vitro* methylation assay and as a control I used a purified recombinant form of MAP3K2 1-350 aa, since SMYD3 efficiently methylates MAP3K2 (Mazur et al., 2016). However, I found that SMYD3 does not methylate SMAD3 *in vitro*, when compared to MAP3K2 1-350 methylation (Figure 1).

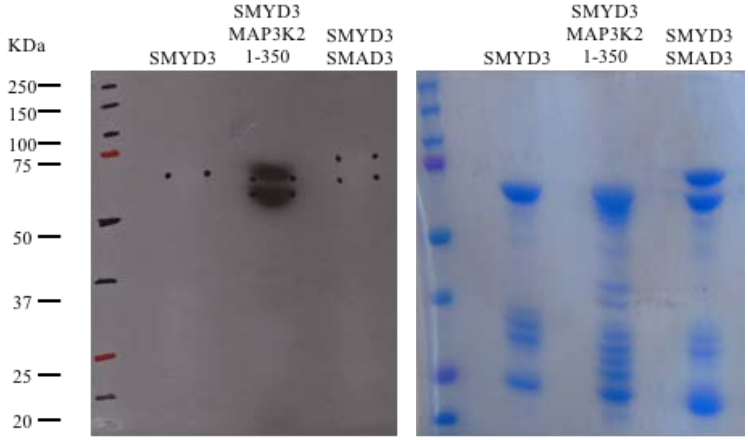


Figure 3.1. SMYD3 doesn't methylate SMAD3.
Left panel, autoradiography showing methylation signal of MAP3K2 1-350 by SMYD3. Right panel, SDS-PAGE of each sample after ON incubation.

3.2.2 MDA-MB-231 ZEBRAFISH XENOGRAFT

Besides, we wanted to assess SMYD3 inhibitor BCI-121 to impair breast cancer cell migration employing zebrafish xenografts. Together with a collaborator in the Zebrafish laboratory, we injected MDA-MB-231 cells in subperidermal space of 48 hpf zebrafish embryos. I treated the two sperimental groups with DMSO or with 50 µM BCI-121. At 48 hours post injection, I evaluated the presence of circulating grafted cells into the head and trunk/tail region. These experiments showed a significant decrease of circulating grafted cells in embryos treated with SMYD3 inhibitor BCI-121 compared to control embryos.

FINAL CONCLUSIONS

Overall, our results suggest that SMYD3 plays a negative role in modulating differentiation towards mesendoderm in mESCs. These data are also confirmed in zebrafish. We hypothesize that SMYD3 may modulate Nodal/Activin signaling pathway, which is necessary for mesendodermal commitment during embryonic development. The physical interaction between SMAD3 and SMYD3 may suggest a negative regulation in the specific context of development. However, further experiments have to be done to understand the mechanism through which SMYD3 negatively regulates SMAD3 in development.

We can speculate that *smyd3* morphants would suffer of a depletion of muscle/cardiac cells due to an anticipated expression of this lineage determinants and the heart and muscle phenotype could be associated with this depletion, however additional experiments are needed to confirm this hypothesis.

Additionally, we started to characterize SMAD2/3 phosphorylation levels during mESCs differentiation and found that p-SMAD2/3 (Ser 423/425) levels increase in Sh-SMYD3 EBs, suggesting that the Nodal signaling pathway is more active in absence of SMYD3 in mESCs. This data is consistent with increased upregulation of mesendodermal markers Brachyury, Eomes,

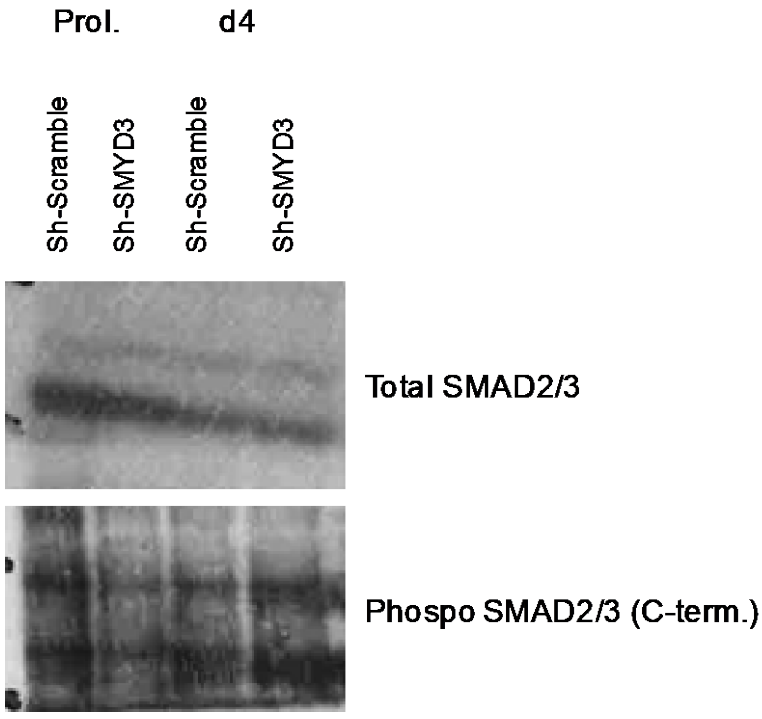


Figure 3.2. SMAD2/3 phosphorylation increases after SMYD3 knock down.
Immunoblot of phospho and total SMAD2/3 in proliferating mESCs and day 4 EBs, in control and SMYD3 depleted cells.

Mesp1.

Our findings in cancer cells show that SMYD3 interacts with SMAD3 and that SMYD3 promotes SMAD3 transcriptional targets in EMT. However, in mESCs depleted of SMYD3 a parallel mechanism should lead to downregulation of SMAD2/3 targets, which include mesendodermal genes. Therefore, SMYD3 seems to have an opposite mechanism in mESCs and also in zebrafish if compared to its role observed in EMT. Consequently, different molecular mechanisms have to be investigated to reconcile these findings.

4 FUTURE PERSPECTIVES

4.1 SMYD3 IN DEVELOPMENT

According to our data obtained in mESCs and zebrafish, *smyd3* overexpression should lead to a decreased expression of mesendodermal markers. Therefore, gain of function experiment can be performed to confirm our results.

On the contrary, we hypothesize that an overexpression of *smyd3* and a *smad3* constitutive active form together could lead to a smad3 blockade by smyd3. Mesendodermal marker expression will be analyzed through *in situ* hybridization analyses.

Furthermore, changes in the expression of phosphorylated-smad3 may be assessed, after *smyd3* morpholino microinjection, by using a transgenic zebrafish reporter line responsive to smad3 activity.

In mESCs, we would like to better characterize Nodal/Activin, BMP, and WNT pathways alterations during the early stages of differentiation, using inducers and inhibitors of the different pathways. Additionally, we are interested in better studying late cardiovascular marker expression after *smyd3* knock down in zebrafish.

In order to address if muscle/cardiac cells are depleted in *smyd3* morphants, we could inject *smyd3* morpholino in transgenic fluorescent lines. For instance, *cmlc2*:dsRed and *nkx2.5*:EGFP transgenic lines can be used. Subsequently, morphants will be analyzed in somitogenesis in order to assay if cardiac precursors are affected. In parallel, muscle determinants could be analyzed in transgenic lines, which express EGFP under *her1* or *actc1a* promoter.

4.2 SMYD3 IN CANCER

We are planning to perform additional xenograft experiment by injecting MDA-MB-231 silenced for SMYD3 in zebrafish embryos.

REFERENCES

Abu-Farha, M., Lambert, J. P., Al-Madhoun, A. S., Elisma, F., Skerjanc, I. S., & Figeys, D. (2008). The tale of two domains: proteomics and genomics analysis of SMYD2, a new histone methyltransferase. *Mol Cell Proteomics, 7*(3), 560-572. doi:10.1074/mcp.M700271-MCP200

Abu-Farha, M., Lanouette, S., Elisma, F., Tremblay, V., Butson, J., Figeys, D., & Couture, J. F. (2011). Proteomic analyses of the SMYD family interactomes identify HSP90 as a novel target for SMYD2. *J Mol Cell Biol, 3*(5), 301-308. doi:10.1093/jmcb/mjr025

Alam, H., Gu, B., & Lee, M. G. (2015). Histone methylation modifiers in cellular signaling pathways. *Cell Mol Life Sci, 72*(23), 4577-4592. doi:10.1007/s00018-015-2023-y

Ambler, R. P., & Rees, M. W. (1959). Epsilon-N-Methyl-lysine in bacterial flagellar protein. *Nature, 184*, 56-57.

Amsterdam, A., Burgess, S., Golling, G., Chen, W., Sun, Z., Townsend, K., . . . Hopkins, N. (1999). A large-scale insertional mutagenesis screen in zebrafish. *Genes Dev, 13*(20), 2713-2724.

Ansari, K. I., & Mandal, S. S. (2010). Mixed lineage leukemia: roles in gene expression, hormone signaling and mRNA processing. *FEBS J, 277*(8), 1790-1804. doi:10.1111/j.1742-4658.2010.07606.x

Beis, D., & Stainier, D. Y. (2006). In vivo cell biology: following the zebrafish trend. *Trends Cell Biol, 16*(2), 105-112. doi:10.1016/j.tcb.2005.12.001

Biggar, K. K., & Li, S. S. (2015). Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol, 16*(1), 5-17. doi:10.1038/nrm3915

Bjornson, C. R., Griffin, K. J., Farr, G. H., 3rd, Terashima, A., Himeda, C., Kikuchi, Y., & Kimelman, D. (2005). Eomesodermin is a localized maternal determinant required for endoderm induction in zebrafish. *Dev Cell, 9*(4), 523-533. doi:10.1016/j.devcel.2005.08.010

Black, J. C., Van Rechem, C., & Whetstine, J. R. (2012). Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell, 48*(4), 491-507. doi:10.1016/j.molcel.2012.11.006

Bolos, V., Peinado, H., Perez-Moreno, M. A., Fraga, M. F., Esteller, M., & 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*(Pt 3), 499-511.

Brown, M. A., Foreman, K., Harriss, J., Das, C., Zhu, L., Edwards, M., . . . Tucker, H. (2015). C-terminal domain of SMYD3 serves as a unique HSP90-regulated motif in oncogenesis. *Oncotarget, 6*(6), 4005-4019. doi:10.18632/oncotarget.2970

Brown, M. A., Sims, R. J., 3rd, Gottlieb, P. D., & Tucker, P. W. (2006). Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Mol Cancer, 5*, 26. doi:10.1186/1476-4598-5-26

Calo, E., & Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? *Mol Cell, 49*(5), 825-837. doi:10.1016/j.molcel.2013.01.038

Cantelli, G., Crosas-Molist, E., Georgouli, M., & Sanz-Moreno, V. (2017). TGFBeta-induced transcription in cancer. *Semin Cancer Biol, 42*, 60-69. doi:10.1016/j.semcancer.2016.08.009

Chen, T., You, Y., Jiang, H., & Wang, Z. Z. (2017). Epithelial-mesenchymal transition (EMT): A biological process in the development, stem cell differentiation, and tumorigenesis. *J Cell Physiol, 232*(12), 3261-3272. doi:10.1002/jcp.25797

Chen, X., Weisberg, E., Fridmacher, V., Watanabe, M., Naco, G., & Whitman, M. (1997). Smad4 and FAST-1 in the assembly of activin-responsive factor. *Nature, 389*(6646), 85-89. doi:10.1038/38008

Chen, Y. J., Tsai, C. H., Wang, P. Y., & Teng, S. C. (2017). SMYD3 Promotes Homologous Recombination via Regulation of H3K4-mediated Gene Expression. *Sci Rep, 7*(1), 3842. doi:10.1038/s41598-017-03385-6

Choi, J., Jang, H., Kim, H., Lee, J. H., Kim, S. T., Cho, E. J., & Youn, H. D. (2014). Modulation of lysine methylation in myocyte enhancer factor 2 during skeletal muscle cell differentiation. *Nucleic Acids Res, 42*(1), 224-234. doi:10.1093/nar/gkt873

Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J. C., & Keller, G. (1998). A common precursor for hematopoietic and endothelial cells. *Development, 125*(4), 725-732.

Cock-Rada, A. M., Medjkane, S., Janski, N., Yousfi, N., Perichon, M., Chaussepied, M., . . . Weitzman, J. B. (2012). SMYD3 promotes cancer invasion by epigenetic upregulation of the metalloproteinase MMP-9. *Cancer Res, 72*(3), 810-820. doi:10.1158/0008-5472.CAN-11-1052

Dickmeis, T., Mourrain, P., Saint-Etienne, L., Fischer, N., Aanstad, P., Clark, M., . . . Rosa, F. (2001). A crucial component of the endoderm formation pathway, CASANOVA, is encoded by a novel sox-related gene. *Genes Dev, 15*(12), 1487-1492. doi:10.1101/gad.196901

Diehl, F., Brown, M. A., van Amerongen, M. J., Novoyatleva, T., Wietelmann, A., Harriss, J., . . . Engel, F. B. (2010). Cardiac deletion of Smyd2 is dispensable for mouse heart development. *PLoS One, 5*(3), e9748. doi:10.1371/journal.pone.0009748

Dong, S. W., Zhang, H., Wang, B. L., Sun, P., Wang, Y. G., & Zhang, P. (2014). Effect of the downregulation of SMYD3 expression by RNAi on RIZ1 expression and proliferation of esophageal squamous cell carcinoma. *Oncol Rep, 32*(3), 1064-1070. doi:10.3892/or.2014.3307

Donlin, L. T., Andresen, C., Just, S., Rudensky, E., Pappas, C. T., Kruger, M., . . . Linke, W. A. (2012). Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization. *Genes Dev, 26*(2), 114-119. doi:10.1101/gad.177758.111

Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., . . . Roeder, R. G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol, 13*(8), 713-719. doi:10.1038/nsmb1128

Dougan, S. T., Warga, R. M., Kane, D. A., Schier, A. F., & Talbot, W. S. (2003). The role of the zebrafish nodal-related genes squint and cyclops in patterning of mesendoderm. *Development, 130*(9), 1837-1851.

Driever, W., Solnica-Krezel, L., Schier, A. F., Neuhauss, S. C., Malicki, J., Stemple, D. L., . . . Boggs, C. (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development, 123*, 37-46.

Du, S. J., Tan, X., & Zhang, J. (2014). SMYD proteins: key regulators in skeletal and cardiac muscle development and function. *Anat Rec (Hoboken), 297*(9), 1650-1662. doi:10.1002/ar.22972

Erter, C. E., Solnica-Krezel, L., & Wright, C. V. (1998). Zebrafish nodal-related 2 encodes an early mesendodermal inducer signaling from the extraembryonic yolk syncytial layer. *Dev Biol, 204*(2), 361-372. doi:10.1006/dbio.1998.9097

Faial, T., Bernardo, A. S., Mendjan, S., Diamanti, E., Ortmann, D., Gentsch, G. E., . . . Pedersen, R. A. (2015). Brachyury and SMAD signalling collaboratively orchestrate distinct mesoderm and endoderm gene regulatory networks in differentiating human embryonic stem cells. *Development, 142*(12), 2121-2135. doi:10.1242/dev.117838

Falnes, P. O., Jakobsson, M. E., Davydova, E., Ho, A., & Malecki, J. (2016). Protein lysine methylation by seven-beta-strand methyltransferases. *Biochem J, 473*(14), 1995-2009. doi:10.1042/BCJ20160117

Faloon, P., Arentson, E., Kazarov, A., Deng, C. X., Porcher, C., Orkin, S., & Choi, K. (2000). Basic fibroblast growth factor positively regulates hematopoietic development. *Development, 127*(9), 1931-1941.

Feldman, B., Gates, M. A., Egan, E. S., Dougan, S. T., Rennebeck, G., Sirotkin, H. I., . . . Talbot, W. S. (1998). Zebrafish organizer development and germ-layer formation require nodal-related signals. *Nature, 395*(6698), 181-185. doi:10.1038/26013

Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., . . . Allis, C. D. (2005). Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature, 438*(7071), 1116-1122. doi:10.1038/nature04219

Foreman, K. W., Brown, M., Park, F., Emtage, S., Harriss, J., Das, C., . . . Tucker, P. (2011). Structural and functional profiling of the human histone methyltransferase SMYD3. *PLoS One, 6*(7), e22290. doi:10.1371/journal.pone.0022290

Fujii, T., Tsunesumi, S., Sagara, H., Munakata, M., Hisaki, Y., Sekiya, T., . . . Watanabe, S. (2016). Smyd5 plays pivotal roles in both primitive and definitive hematopoiesis during zebrafish embryogenesis. *Sci Rep, 6*, 29157. doi:10.1038/srep29157

Fujii, T., Tsunesumi, S., Yamaguchi, K., Watanabe, S., & Furukawa, Y. (2011). Smyd3 is required for the development of cardiac and skeletal muscle in zebrafish. *PLoS One, 6*(8), e23491. doi:10.1371/journal.pone.0023491

Gaarenstroom, T., & Hill, C. S. (2014). TGF-beta signaling to chromatin: how Smads regulate transcription during self-renewal and differentiation. *Semin Cell Dev Biol, 32*, 107-118. doi:10.1016/j.semcdb.2014.01.009

Gadue, P., Huber, T. L., Paddison, P. J., & Keller, G. M. (2006). Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. *Proc Natl Acad Sci U S A, 103*(45), 16806-16811. doi:10.1073/pnas.0603916103

Ghahhari, N. M., & Babashah, S. (2015). Interplay between microRNAs and WNT/beta-catenin signalling pathway regulates epithelial-mesenchymal transition in cancer. *Eur J Cancer, 51*(12), 1638-1649. doi:10.1016/j.ejca.2015.04.021

Goswami, M. T., Reka, A. K., Kurapati, H., Kaza, V., Chen, J., Standiford, T. J., & Keshamouni, V. G. (2016). Regulation of complement-dependent cytotoxicity by TGF-beta-induced epithelial-mesenchymal transition. *Oncogene, 35*(15), 1888-1898. doi:10.1038/onc.2015.258

Gottlieb, P. D., Pierce, S. A., Sims, R. J., Yamagishi, H., Weihe, E. K., Harriss, J. V., . . . Srivastava, D. (2002). Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nat Genet, 31*(1), 25-32. doi:10.1038/ng866

Greer, E. L., & Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet, 13*(5), 343-357. doi:10.1038/nrg3173

Gritsman, K., Talbot, W. S., & Schier, A. F. (2000). Nodal signaling patterns the organizer. *Development, 127*(5), 921-932.

Gut, P., Baeza-Raja, B., Andersson, O., Hasenkamp, L., Hsiao, J., Hesselson, D., . . . Stainier, D. Y. (2013). Whole-organism screening for gluconeogenesis identifies activators of fasting metabolism. *Nat Chem Biol, 9*(2), 97-104. doi:10.1038/nchembio.1136

Haffter, P., Granato, M., Brand, M., Mullins, M. C., Hammerschmidt, M., Kane, D. A., . . . Nusslein-Volhard, C. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development, 123*, 1-36.

Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F. P., Li, M., . . . Nakamura, Y. (2004). SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol, 6*(8), 731-740. doi:10.1038/ncb1151

Hamamoto, R., Saloura, V., & Nakamura, Y. (2015). Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat Rev Cancer, 15*(2), 110-124. doi:10.1038/nrc3884

Hamamoto, R., Silva, F. P., Tsuge, M., Nishidate, T., Katagiri, T., Nakamura, Y., & Furukawa, Y. (2006). Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci, 97*(2), 113-118. doi:10.1111/j.1349-7006.2006.00146.x

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell, 144*(5), 646-674. doi:10.1016/j.cell.2011.02.013

Hart, A. H., Hartley, L., Sourris, K., Stadler, E. S., Li, R., Stanley, E. G., . . . Robb, L. (2002). Mixl1 is required for axial mesendoderm morphogenesis and patterning in the murine embryo. *Development, 129*(15), 3597-3608.

Heisenberg, C. P., & Tada, M. (2002). Zebrafish gastrulation movements: bridging cell and developmental biology. *Semin Cell Dev Biol, 13*(6), 471-479.

Herz, H. M., Garruss, A., & Shilatifard, A. (2013). SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends Biochem Sci, 38*(12), 621-639. doi:10.1016/j.tibs.2013.09.004

Holder, N., & Xu, Q. (1999). Microinjection of DNA, RNA, and protein into the fertilized zebrafish egg for analysis of gene function. *Methods Mol Biol, 97*, 487-490. doi:10.1385/1-59259-270-8:487

Holtzman, N. G., Iovine, M. K., Liang, J. O., & Morris, J. (2016). Learning to Fish with Genetics: A Primer on the Vertebrate Model Danio rerio. *Genetics, 203*(3), 1069-1089. doi:10.1534/genetics.116.190843

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., . . . Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature, 496*(7446), 498-503. doi:10.1038/nature12111

Hu, L., Zhu, Y. T., Qi, C., & Zhu, Y. J. (2009). Identification of Smyd4 as a potential tumor suppressor gene involved in breast cancer development. *Cancer Res, 69*(9), 4067-4072. doi:10.1158/0008-5472.CAN-08-4097

Huang, J., & Berger, S. L. (2008). The emerging field of dynamic lysine methylation of non-histone proteins. *Curr Opin Genet Dev, 18*(2), 152-158. doi:10.1016/j.gde.2008.01.012

Huang, J., Dorsey, J., Chuikov, S., Perez-Burgos, L., Zhang, X., Jenuwein, T., . . . Berger, S. L. (2010). G9a and Glp methylate lysine 373 in the tumor suppressor p53. *J Biol Chem, 285*(13), 9636-9641. doi:10.1074/jbc.M109.062588

Huang, J., Perez-Burgos, L., Placek, B. J., Sengupta, R., Richter, M., Dorsey, J. A., . . . Berger, S. L. (2006). Repression of p53 activity by Smyd2-mediated methylation. *Nature, 444*(7119), 629-632. doi:10.1038/nature05287

Huber, T. L., Kouskoff, V., Fehling, H. J., Palis, J., & Keller, G. (2004). Haemangioblast commitment is initiated in the primitive streak of the mouse embryo. *Nature, 432*(7017), 625-630. doi:10.1038/nature03122

Huisken, J., & Stainier, D. Y. (2009). Selective plane illumination microscopy techniques in developmental biology. *Development, 136*(12), 1963-1975. doi:10.1242/dev.022426

Jaenisch, R., & Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell, 132*(4), 567-582. doi:10.1016/j.cell.2008.01.015

Jenuwein, T., Laible, G., Dorn, R., & Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci, 54*(1), 80-93.

Jorda, M., Olmeda, D., Vinyals, A., Valero, E., Cubillo, E., Llorens, A., . . . Fabra, A. (2005). Upregulation of MMP-9 in MDCK epithelial cell line in response to expression of the Snail transcription factor. *J Cell Sci, 118*(Pt 15), 3371-3385. doi:10.1242/jcs.02465

Just, S., Meder, B., Berger, I. M., Etard, C., Trano, N., Patzel, E., . . . Rottbauer, W. (2011). The myosin-interacting protein SMYD1 is essential for sarcomere organization. *J Cell Sci, 124*(Pt 18), 3127-3136. doi:10.1242/jcs.084772

Kalluri, R., & Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *J Clin Invest, 119*(6), 1420-1428. doi:10.1172/JCI39104

Kattman, S. J., Huber, T. L., & Keller, G. M. (2006). Multipotent flk-1+ cardiovascular progenitor cells give rise to the cardiomyocyte, endothelial, and vascular smooth muscle lineages. *Dev Cell, 11*(5), 723-732. doi:10.1016/j.devcel.2006.10.002

Kessenbrock, K., Plaks, V., & Werb, Z. (2010). Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell, 141*(1), 52-67. doi:10.1016/j.cell.2010.03.015

Kidder, B. L., Hu, G., Cui, K., & Zhao, K. (2017). SMYD5 regulates H4K20me3-marked heterochromatin to safeguard ES cell self-renewal and prevent spurious differentiation. *Epigenetics Chromatin, 10*, 8. doi:10.1186/s13072-017-0115-7

Kiecker, C., Bates, T., & Bell, E. (2016). Molecular specification of germ layers in vertebrate embryos. *Cell Mol Life Sci, 73*(5), 923-947. doi:10.1007/s00018-015-2092-y

Kikuchi, Y., Agathon, A., Alexander, J., Thisse, C., Waldron, S., Yelon, D., . . . Stainier, D. Y. (2001). casanova encodes a novel Sox-related protein necessary and sufficient for early endoderm formation in zebrafish. *Genes Dev, 15*(12), 1493-1505. doi:10.1101/gad.892301

Kim, H., Heo, K., Kim, J. H., Kim, K., Choi, J., & An, W. (2009). Requirement of histone methyltransferase SMYD3 for estrogen receptor-mediated transcription. *J Biol Chem, 284*(30), 19867-19877. doi:10.1074/jbc.M109.021485

Kim, J. M., Kim, K., Schmidt, T., Punj, V., Tucker, H., Rice, J. C., . . . An, W. (2015). Cooperation between SMYD3 and PC4 drives a distinct transcriptional program in cancer cells. *Nucleic Acids Res, 43*(18), 8868-8883. doi:10.1093/nar/gkv874

Kimelman, D. (2006). Mesoderm induction: from caps to chips. *Nat Rev Genet, 7*(5), 360-372. doi:10.1038/nrg1837

Konantz, M., Balci, T. B., Hartwig, U. F., Dellaire, G., Andre, M. C., Berman, J. N., & Lengerke, C. (2012). Zebrafish xenografts as a tool for in vivo studies on human cancer. *Ann N Y Acad Sci, 1266*, 124-137. doi:10.1111/j.1749-6632.2012.06575.x

Kunizaki, M., Hamamoto, R., Silva, F. P., Yamaguchi, K., Nagayasu, T., Shibuya, M., . . . Furukawa, Y. (2007). The lysine 831 of vascular endothelial growth factor receptor 1 is a novel target of methylation by SMYD3. *Cancer Res, 67*(22), 10759-10765. doi:10.1158/0008-5472.CAN-07-1132

Kurosawa, H. (2007). Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *J Biosci Bioeng, 103*(5), 389-398. doi:10.1263/jbb.103.389

Lee, J. M., Lee, J. S., Kim, H., Kim, K., Park, H., Kim, J. Y., . . . Baek, S. H. (2012). EZH2 generates a methyl degron that is recognized by the DCAF1/DDB1/CUL4 E3 ubiquitin ligase complex. *Mol Cell, 48*(4), 572-586. doi:10.1016/j.molcel.2012.09.004

Lee, J. S., Kim, Y., Kim, I. S., Kim, B., Choi, H. J., Lee, J. M., . . . Baek, S. H. (2010). Negative regulation of hypoxic responses via induced Reptin methylation. *Mol Cell, 39*(1), 71-85. doi:10.1016/j.molcel.2010.06.008

Lepage, S. E., & Bruce, A. E. (2010). Zebrafish epiboly: mechanics and mechanisms. *Int J Dev Biol, 54*(8-9), 1213-1228. doi:10.1387/ijdb.093028sl

Li, D., Niu, Z., Yu, W., Qian, Y., Wang, Q., Li, Q., . . . Liu, M. (2009). SMYD1, the myogenic activator, is a direct target of serum response factor and myogenin. *Nucleic Acids Res, 37*(21), 7059-7071. doi:10.1093/nar/gkp773

Li, H., Zhong, Y., Wang, Z., Gao, J., Xu, J., Chu, W., . . . Du, S. J. (2013). Smyd1b is required for skeletal and cardiac muscle function in zebrafish. *Mol Biol Cell, 24*(22), 3511-3521. doi:10.1091/mbc.E13-06-0352

Li, M., Zhao, L., Page-McCaw, P. S., & Chen, W. (2016). Zebrafish Genome Engineering Using the CRISPR-Cas9 System. *Trends Genet, 32*(12), 815-827. doi:10.1016/j.tig.2016.10.005

Li, Y., Sun, L., Zhang, Y., Wang, D., Wang, F., Liang, J., . . . Shang, Y. (2011). The histone modifications governing TFF1 transcription mediated by estrogen receptor. *J Biol Chem, 286*(16), 13925-13936. doi:10.1074/jbc.M111.223198

Ling, B. M., Bharathy, N., Chung, T. K., Kok, W. K., Li, S., Tan, Y. H., . . . Taneja, R. (2012). Lysine methyltransferase G9a methylates the transcription factor MyoD and regulates skeletal muscle differentiation. *Proc Natl Acad Sci U S A, 109*(3), 841-846. doi:10.1073/pnas.1111628109

Liu, C., Fang, X., Ge, Z., Jalink, M., Kyo, S., Bjorkholm, M., . . . Xu, D. (2007). The telomerase reverse transcriptase (hTERT) gene is a direct target of the histone methyltransferase SMYD3. *Cancer Res, 67*(6), 2626-2631. doi:10.1158/0008-5472.CAN-06-4126

Lu, X., Simon, M. D., Chodaparambil, J. V., Hansen, J. C., Shokat, K. M., & Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat Struct Mol Biol, 15*(10), 1122-1124. doi:10.1038/nsmb.1489

Luo, X. G., Zhang, C. L., Zhao, W. W., Liu, Z. P., Liu, L., Mu, A., . . . Zhang, T. C. (2014). Histone methyltransferase SMYD3 promotes MRTF-A-mediated transactivation of MYL9 and migration of MCF-7 breast cancer cells. *Cancer Lett, 344*(1), 129-137. doi:10.1016/j.canlet.2013.10.026

Martin, C., & Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol, 6*(11), 838-849. doi:10.1038/nrm1761

Massague, J. (2012). TGFbeta signalling in context. *Nat Rev Mol Cell Biol, 13*(10), 616-630. doi:10.1038/nrm3434

Mazur, P. K., Reynoird, N., Khatri, P., Jansen, P. W., Wilkinson, A. W., Liu, S., . . . Gozani, O. (2014). SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature, 510*(7504), 283-287. doi:10.1038/nature13320

Meissner, A. (2010). Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol, 28*(10), 1079-1088. doi:10.1038/nbt.1684

Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., . . . Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. *Nature, 437*(7057), 436-439. doi:10.1038/nature04020

Mickoleit, M., Schmid, B., Weber, M., Fahrbach, F. O., Hombach, S., Reischauer, S., & Huisken, J. (2014). High-resolution reconstruction of the beating zebrafish heart. *Nat Methods, 11*(9), 919-922. doi:10.1038/nmeth.3037

Mizuno, T., Shinya, M., & Takeda, H. (1999). Cell and tissue transplantation in zebrafish embryos. *Methods Mol Biol, 127*, 15-28. doi:10.1385/1-59259-678-9:15

Montero, J. A., & Heisenberg, C. P. (2004). Gastrulation dynamics: cells move into focus. *Trends Cell Biol, 14*(11), 620-627. doi:10.1016/j.tcb.2004.09.008

Moretti, A., Caron, L., Nakano, A., Lam, J. T., Bernshausen, A., Chen, Y., . . . Chien, K. R. (2006). Multipotent embryonic isl1+ progenitor cells lead to cardiac, smooth muscle, and endothelial cell diversification. *Cell, 127*(6), 1151-1165. doi:10.1016/j.cell.2006.10.029

Murn, J., & Shi, Y. (2017). The winding path of protein methylation research: milestones and new frontiers. *Nat Rev Mol Cell Biol, 18*(8), 517-527. doi:10.1038/nrm.2017.35

Murray, K. (1964). The Occurrence of Epsilon-N-Methyl Lysine in Histones. *Biochemistry, 3*, 10-15.

Murry, C. E., & Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell, 132*(4), 661-680. doi:10.1016/j.cell.2008.02.008

Musselman, C. A., Lalonde, M. E., Cote, J., & Kutateladze, T. G. (2012). Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol, 19*(12), 1218-1227. doi:10.1038/nsmb.2436

Nair, S. S., Li, D. Q., & Kumar, R. (2013). A core chromatin remodeling factor instructs global chromatin signaling through multivalent reading of nucleosome codes. *Mol Cell, 49*(4), 704-718. doi:10.1016/j.molcel.2012.12.016

Nicoli, S., Ribatti, D., Cotelli, F., & Presta, M. (2007). Mammalian tumor xenografts induce neovascularization in zebrafish embryos. *Cancer Res, 67*(7), 2927-2931. doi:10.1158/0008-5472.CAN-06-4268

Pasini, D., Malatesta, M., Jung, H. R., Walfridsson, J., Willer, A., Olsson, L., . . . Helin, K. (2010). Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res, 38*(15), 4958-4969. doi:10.1093/nar/gkq244

Peserico, A., Germani, A., Sanese, P., Barbosa, A. J., Di Virgilio, V., Fittipaldi, R., . . . Simone, C. (2015). A SMYD3 Small-Molecule Inhibitor Impairing Cancer Cell Growth. *J Cell Physiol, 230*(10), 2447-2460. doi:10.1002/jcp.24975

Phan, D., Rasmussen, T. L., Nakagawa, O., McAnally, J., Gottlieb, P. D., Tucker, P. W., . . . Olson, E. N. (2005). BOP, a regulator of right ventricular heart development, is a direct transcriptional target of MEF2C in the developing heart. *Development, 132*(11), 2669-2678. doi:10.1242/dev.01849

Pless, O., Kowenz-Leutz, E., Knoblich, M., Lausen, J., Beyermann, M., Walsh, M. J., & Leutz, A. (2008). G9a-mediated lysine methylation alters the function of CCAAT/enhancer-binding protein-beta. *J Biol Chem, 283*(39), 26357-26363. doi:10.1074/jbc.M802132200

Prieto-Garcia, E., Diaz-Garcia, C. V., Garcia-Ruiz, I., & Agullo-Ortuno, M. T. (2017). Epithelial-to-mesenchymal transition in tumor progression. *Med Oncol, 34*(7), 122. doi:10.1007/s12032-017-0980-8

Proserpio, V., Fittipaldi, R., Ryall, J. G., Sartorelli, V., & Caretti, G. (2013). The methyltransferase SMYD3 mediates the recruitment of transcriptional cofactors at the myostatin and c-Met genes and regulates skeletal muscle atrophy. *Genes Dev, 27*(11), 1299-1312. doi:10.1101/gad.217240.113

Rathert, P., Dhayalan, A., Murakami, M., Zhang, X., Tamas, R., Jurkowska, R., . . . Jeltsch, A. (2008). Protein lysine methyltransferase G9a acts on non-histone targets. *Nat Chem Biol, 4*(6), 344-346. doi:10.1038/nchembio.88

Rebagliati, M. R., Toyama, R., Haffter, P., & Dawid, I. B. (1998). cyclops encodes a nodal-related factor involved in midline signaling. *Proc Natl Acad Sci U S A, 95*(17), 9932-9937.

Reischauer, S., Arnaout, R., Ramadass, R., & Stainier, D. Y. (2014). Actin binding GFP allows 4D in vivo imaging of myofilament dynamics in the zebrafish heart and the identification of Erbb2 signaling as a remodeling factor of myofibril architecture. *Circ Res, 115*(10), 845-856. doi:10.1161/CIRCRESAHA.115.304356

Reiter, J. F., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N., & Stainier, D. Y. (1999). Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev, 13*(22), 2983-2995.

Ren, T. N., Wang, J. S., He, Y. M., Xu, C. L., Wang, S. Z., & Xi, T. (2011). Effects of SMYD3 over-expression on cell cycle acceleration and cell proliferation in MDA-MB-231 human breast cancer cells. *Med Oncol, 28 Suppl 1*, S91-98. doi:10.1007/s12032-010-9718-6

Rice, J. C., Briggs, S. D., Ueberheide, B., Barber, C. M., Shabanowitz, J., Hunt, D. F., . . . Allis, C. D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell, 12*(6), 1591-1598.

Saddic, L. A., West, L. E., Aslanian, A., Yates, J. R., 3rd, Rubin, S. M., Gozani, O., & Sage, J. (2010). Methylation of the retinoblastoma tumor suppressor by SMYD2. *J Biol Chem, 285*(48), 37733-37740. doi:10.1074/jbc.M110.137612

Sakaguchi, T., Kuroiwa, A., & Takeda, H. (2001). A novel sox gene, 226D7, acts downstream of Nodal signaling to specify endoderm precursors in zebrafish. *Mech Dev, 107*(1-2), 25-38.

Sakaki-Yumoto, M., Katsuno, Y., & Derynck, R. (2013). TGF-beta family signaling in stem cells. *Biochim Biophys Acta, 1830*(2), 2280-2296. doi:10.1016/j.bbagen.2012.08.008

Sampath, K., Rubinstein, A. L., Cheng, A. M., Liang, J. O., Fekany, K., Solnica-Krezel, L., . . . Wright, C. V. (1998). Induction of the zebrafish ventral brain and floorplate requires cyclops/nodal signalling. *Nature, 395*(6698), 185-189. doi:10.1038/26020

Sampath, S. C., Marazzi, I., Yap, K. L., Sampath, S. C., Krutchinsky, A. N., Mecklenbrauker, I., . . . Tarakhovsky, A. (2007). Methylation of a histone mimic within the histone methyltransferase G9a regulates protein complex assembly. *Mol Cell, 27*(4), 596-608. doi:10.1016/j.molcel.2007.06.026

Sanders, S. L., Portoso, M., Mata, J., Bahler, J., Allshire, R. C., & Kouzarides, T. (2004). Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell, 119*(5), 603-614. doi:10.1016/j.cell.2004.11.009

Sarris, M. E., Moulos, P., Haroniti, A., Giakountis, A., & Talianidis, I. (2016). Smyd3 Is a Transcriptional Potentiator of Multiple Cancer-Promoting Genes and Required for Liver and Colon Cancer Development. *Cancer Cell, 29*(3), 354-366. doi:10.1016/j.ccell.2016.01.013

Schier, A. F., & Talbot, W. S. (2005). Molecular genetics of axis formation in zebrafish. *Annu Rev Genet, 39*, 561-613. doi:10.1146/annurev.genet.37.110801.143752

Serrano-Gomez, S. J., Maziveyi, M., & Alahari, S. K. (2016). Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. *Mol Cancer, 15*, 18. doi:10.1186/s12943-016-0502-x

Sese, B., Barrero, M. J., Fabregat, M. C., Sander, V., & Izpisua Belmonte, J. C. (2013). SMYD2 is induced during cell differentiation and participates in early development. *Int J Dev Biol, 57*(5), 357-364. doi:10.1387/ijdb.130051ji

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., . . . Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell, 119*(7), 941-953. doi:10.1016/j.cell.2004.12.012

Shyh-Chang, N., Locasale, J. W., Lyssiotis, C. A., Zheng, Y., Teo, R. Y., Ratanasirintrawoot, S., . . . Cantley, L. C. (2013). Influence of threonine metabolism on S-adenosylmethionine and histone methylation. *Science, 339*(6116), 222-226. doi:10.1126/science.1226603

Singh, M., Yelle, N., Venugopal, C., & Singh, S. K. (2017). EMT: Mechanisms and therapeutic implications. *Pharmacol Ther*. doi:10.1016/j.pharmthera.2017.08.009

Sirinupong, N., Brunzelle, J., Doko, E., & Yang, Z. (2011). Structural insights into the autoinhibition and posttranslational activation of histone methyltransferase SmyD3. *J Mol Biol, 406*(1), 149-159. doi:10.1016/j.jmb.2010.12.014

Skrypek, N., Goossens, S., De Smedt, E., Vandamme, N., & Berx, G. (2017). Epithelial-to-Mesenchymal Transition: Epigenetic Reprogramming Driving Cellular Plasticity. *Trends Genet*. doi:10.1016/j.tig.2017.08.004

Solnica-Krezel, L. (2006). Gastrulation in zebrafish -- all just about adhesion? *Curr Opin Genet Dev, 16*(4), 433-441. doi:10.1016/j.gde.2006.06.009

Song, J. (2007). EMT or apoptosis: a decision for TGF-beta. *Cell Res, 17*(4), 289-290. doi:10.1038/cr.2007.25

Spellmon, N., Holcomb, J., Trescott, L., Sirinupong, N., & Yang, Z. (2015). Structure and function of SET and MYND domain-containing proteins. *Int J Mol Sci, 16*(1), 1406-1428. doi:10.3390/ijms16011406

Stainier, D. Y., Lee, R. K., & Fishman, M. C. (1993). Cardiovascular development in the zebrafish. I. Myocardial fate map and heart tube formation. *Development, 119*(1), 31-40.

Stender, J. D., Pascual, G., Liu, W., Kaikkonen, M. U., Do, K., Spann, N. J., . . . Glass, C. K. (2012). Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. *Mol Cell, 48*(1), 28-38. doi:10.1016/j.molcel.2012.07.020

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., . . . Shinkai, Y. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev, 16*(14), 1779-1791. doi:10.1101/gad.989402

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., . . . Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev, 19*(7), 815-826. doi:10.1101/gad.1284005

Tam, P. P., & Loebel, D. A. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet, 8*(5), 368-381. doi:10.1038/nrg2084

Tan, X., Rotllant, J., Li, H., De Deyne, P., & Du, S. J. (2006). SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proc Natl Acad Sci U S A, 103*(8), 2713-2718. doi:10.1073/pnas.0509503103

Terranova, R., Agherbi, H., Boned, A., Meresse, S., & Djabali, M. (2006). Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. *Proc Natl Acad Sci U S A, 103*(17), 6629-6634. doi:10.1073/pnas.0507425103

Thisse, B., Wright, C. V., & Thisse, C. (2000). Activin- and Nodal-related factors control antero-posterior patterning of the zebrafish embryo. *Nature, 403*(6768), 425-428. doi:10.1038/35000200

Thompson, E. C., & Travers, A. A. (2008). A Drosophila Smyd4 homologue is a muscle-specific transcriptional modulator involved in development. *PLoS One, 3*(8), e3008. doi:10.1371/journal.pone.0003008

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science, 282*(5391), 1145-1147.

Tobia, C., De Sena, G., & Presta, M. (2011). Zebrafish embryo, a tool to study tumor angiogenesis. *Int J Dev Biol, 55*(4-5), 505-509. doi:10.1387/ijdb.103238ct

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., & Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development, 135*(5), 909-918. doi:10.1242/dev.017400

Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., & Reuter, G. (1994). The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J, 13*(16), 3822-3831.

Turner, D. A., Rue, P., Mackenzie, J. P., Davies, E., & Martinez Arias, A. (2014). Brachyury cooperates with Wnt/beta-catenin signalling to elicit primitive-streak-like behaviour in differentiating mouse embryonic stem cells. *BMC Biol, 12*, 63. doi:10.1186/s12915-014-0063-7

Van Aller, G. S., Reynoird, N., Barbash, O., Huddleston, M., Liu, S., Zmoos, A. F., . . . Kruger, R. G. (2012). Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics, 7*(4), 340-343. doi:10.4161/epi.19506

Vassilieva, S., Goh, H. N., Lau, K. X., Hughes, J. N., Familari, M., Rathjen, P. D., & Rathjen, J. (2012). A system to enrich for primitive streak-derivatives, definitive endoderm and mesoderm, from pluripotent cells in culture. *PLoS One, 7*(6), e38645. doi:10.1371/journal.pone.0038645

Voelkel, T., Andresen, C., Unger, A., Just, S., Rottbauer, W., & Linke, W. A. (2013). Lysine methyltransferase Smyd2 regulates Hsp90-mediated protection of the sarcomeric titin springs and cardiac function. *Biochim Biophys Acta, 1833*(4), 812-822. doi:10.1016/j.bbamcr.2012.09.012

Vogel, A. M., & Weinstein, B. M. (2000). Studying vascular development in the zebrafish. *Trends Cardiovasc Med, 10*(8), 352-360.

Wang, S. Z., Luo, X. G., Shen, J., Zou, J. N., Lu, Y. H., & Xi, T. (2008). Knockdown of SMYD3 by RNA interference inhibits cervical carcinoma cell growth and invasion in vitro. *BMB Rep, 41*(4), 294-299.

Weinstein, B. (2002). Vascular cell biology in vivo: a new piscine paradigm? *Trends Cell Biol, 12*(9), 439-445.

Weiss, T., Hergeth, S., Zeissler, U., Izzo, A., Tropberger, P., Zee, B. M., . . . Schneider, R. (2010). Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. *Epigenetics Chromatin, 3*(1), 7. doi:10.1186/1756-8935-3-7

Wesche, J., Kuhn, S., Kessler, B. M., Salton, M., & Wolf, A. (2017). Protein arginine methylation: a prominent modification and its demethylation. *Cell Mol Life Sci, 74*(18), 3305-3315. doi:10.1007/s00018-017-2515-z

Wijnhoven, B. P., Dinjens, W. N., & Pignatelli, M. (2000). E-cadherin-catenin cell-cell adhesion complex and human cancer. *Br J Surg, 87*(8), 992-1005. doi:10.1046/j.1365-2168.2000.01513.x

Willems, E., & Leyns, L. (2008). Patterning of mouse embryonic stem cell-derived pan-mesoderm by Activin A/Nodal and Bmp4 signaling requires Fibroblast Growth Factor activity. *Differentiation, 76*(7), 745-759. doi:10.1111/j.1432-0436.2007.00257.x

Williamson, A. J., Smith, D. L., Blinco, D., Unwin, R. D., Pearson, S., Wilson, C., . . . Whetton, A. D. (2008). Quantitative proteomics analysis demonstrates post-transcriptional regulation of embryonic stem cell differentiation to hematopoiesis. *Mol Cell Proteomics, 7*(3), 459-472. doi:10.1074/mcp.M700370-MCP200

Wu, H., Chen, X., Xiong, J., Li, Y., Li, H., Ding, X., . . . Zhu, B. (2011). Histone methyltransferase G9a contributes to H3K27 methylation in vivo. *Cell Res, 21*(2), 365-367. doi:10.1038/cr.2010.157

Wu, Z., Connolly, J., & Biggar, K. K. (2017). Beyond histones - the expanding roles of protein lysine methylation. *FEBS J, 284*(17), 2732-2744. doi:10.1111/febs.14056

Xu, S., Wu, J., Sun, B., Zhong, C., & Ding, J. (2011). Structural and biochemical studies of human lysine methyltransferase Smyd3 reveal the important functional roles of its post-SET and TPR domains and the regulation of its activity by DNA binding. *Nucleic Acids Res, 39*(10), 4438-4449. doi:10.1093/nar/gkr019

Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N., & McMahon, A. P. (1999). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev, 13*(24), 3185-3190.

Yi, X., Jiang, X. J., Li, X. Y., & Jiang, D. S. (2015). Histone methyltransferases: novel targets for tumor and developmental defects. *Am J Transl Res, 7*(11), 2159-2175.

Yoshioka, Y., Suzuki, T., Matsuo, Y., Nakakido, M., Tsurita, G., Simone, C., . . . Hamamoto, R. (2016). SMYD3-mediated lysine methylation in the PH domain is critical for activation of AKT1. *Oncotarget, 7*(46), 75023-75037. doi:10.18632/oncotarget.11898

Yoshioka, Y., Suzuki, T., Matsuo, Y., Tsurita, G., Watanabe, T., Dohmae, N., . . . Hamamoto, R. (2017). Protein lysine methyltransferase SMYD3 is involved in tumorigenesis through regulation of HER2 homodimerization. *Cancer Med, 6*(7), 1665-1672. doi:10.1002/cam4.1099

Yu, B. D., Hanson, R. D., Hess, J. L., Horning, S. E., & Korsmeyer, S. J. (1998). MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proc Natl Acad Sci U S A, 95*(18), 10632-10636.

Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A., & Korsmeyer, S. J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. *Nature, 378*(6556), 505-508. doi:10.1038/378505a0

Yu, Y., Song, C., Zhang, Q., DiMaggio, P. A., Garcia, B. A., York, A., . . . Grunstein, M. (2012). Histone H3 lysine 56 methylation regulates DNA replication through its interaction with PCNA. *Mol Cell, 46*(1), 7-17. doi:10.1016/j.molcel.2012.01.019

Zee, B. M., Levin, R. S., DiMaggio, P. A., & Garcia, B. A. (2010). Global turnover of histone post-translational modifications and variants in human cells. *Epigenetics Chromatin, 3*(1), 22. doi:10.1186/1756-8935-3-22

Zee, B. M., Levin, R. S., Xu, B., LeRoy, G., Wingreen, N. S., & Garcia, B. A. (2010). In vivo residue-specific histone methylation dynamics. *J Biol Chem, 285*(5), 3341-3350. doi:10.1074/jbc.M109.063784

Zhang, B., Xuan, C., Ji, Y., Zhang, W., & Wang, D. (2015). Zebrafish xenotransplantation as a tool for in vivo cancer study. *Fam Cancer, 14*(3), 487-493. doi:10.1007/s10689-015-9802-3

Zhang, X., Huang, Y., & Shi, X. (2015). Emerging roles of lysine methylation on non-histone proteins. *Cell Mol Life Sci, 72*(22), 4257-4272. doi:10.1007/s00018-015-2001-4

Zhang, X., Wen, H., & Shi, X. (2012). Lysine methylation: beyond histones. *Acta Biochim Biophys Sin (Shanghai), 44*(1), 14-27. doi:10.1093/abbs/gmr100

Zhang, Y., & Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev, 15*(18), 2343-2360. doi:10.1101/gad.927301

Zhao, S., Huang, J., & Ye, J. (2015). A fresh look at zebrafish from the perspective of cancer research. *J Exp Clin Cancer Res, 34*, 80. doi:10.1186/s13046-015-0196-8

Zhou, Y., Kim, J., Yuan, X., & Braun, T. (2011). Epigenetic modifications of stem cells: a paradigm for the control of cardiac progenitor cells. *Circ Res, 109*(9), 1067-1081. doi:10.1161/CIRCRESAHA.111.243709

Zou, J. N., Wang, S. Z., Yang, J. S., Luo, X. G., Xie, J. H., & Xi, T. (2009). Knockdown of SMYD3 by RNA interference down-regulates c-Met expression and inhibits cells migration and invasion induced by HGF. *Cancer Lett, 280*(1), 78-85. doi:10.1016/j.canlet.2009.02.015

PART II: SMYD3 IN DEVELOPMENT

SMYD3 methylase contributes to the regulation of early developmental stages, in embryonic stem cells and zebrafish (Manuscript in preparation)

Raffaella Fittipaldi[1,*], Pamela Floris[1,*], Monica Beltrame[1], Franco Cotelli[1], Giuseppina Caretti[1, 2]

[1]Department of Biosciences, University of Milan, Via Celoria 26, 20133, Milan, Italy

* These authors contributed equally

[2] To whom correspondence should be addressed:
Dept. of Biosciences, University of Milano, Via Celoria 26, 20133, Milan, Italy
Tel:+390250315002
Fax:+390250315044
giuseppina.caretti@unimi.it

Key words: embryonic stem cells, SMYD3, zebrafish, development

Abstract

SMYD3 (SET and MYND domain containing protein 3), is a methylase over-expressed in cancer cells and involved in oncogenesis. SMYD3 is expressed in normal tissue and in the developing embryo. Despite this, SMYD3 role in physiological conditions has not been fully investigated yet. Importantly, it has been recently reported that SMYD3 is be involved in cardiac development and myogenesis in zebrafish.

Here, we report that SMYD3 depletion promotes the induction of mesodermal pattern during *in vitro* differentiation of mouse embryonic stem cells, and causes an upregulation of cardiovascular lineage markers at later stages. In vivo, *smyd3* knockdown in zebrafish favors the upregulation of mesendodermal markers during zebrafish gastrulation.

Overall, our study reveals that SMYD3 may modulate genes involved in the early stages of development and of embryonic stem cells differentiation.

Introduction

Epigenetic regulation of gene expression plays a pivotal role in the establishment of developmental programs and the maintenance of the differentiated state (Fisher and Fisher, 2011). During development, transcription factors and chromatin modifiers are either activated or repressed to commit specific cells to defined lineages in response to external cues (Ong and Corces, 2011).

Pluripotent embryonic stem cells (ESCs) possess the unique ability to self-renew and to commit towards differentiation of several lineages (Young, 2011). Because of this feature, they represent a powerful tool that became an amenable *in vitro* system to study molecular mechanisms underlying lineage commitment and study embryonic development. The maintenance of pluripotency is endowed with the interconnected circuitry of the pluripotency transcription factors Oct4, Nanog, Klf4 and Sox2. These core transcription factors concurrently recruit several chromatin regulators to the genomic regions of stemness genes to maintain transcriptional activation and sustain the repression of differentiation genes through repressive epigenetic complexes (Delmore et al., 2011; Orkin and Hochedlinger, 2011). Several chromatin modulators such as Polycomb and MLL complexes, Mof and Jmj proteins, have been described as pivotal regulators of stem cells biology (Li et al., 2012; Ohtani et al.). Some of these factors are under the direct transcriptional control of the pluripotency network, controlled by Oct4, Nanog, Sox2, and Klf4 (Boyer et al., 2005). Two reports have recently uncovered SMYD3 as an Oct4/Nanog/Sox2 target in mouse embryonic stem cells (Campbell et al., 2007; Chen et al., 2008), hinting for a possible role for SMYD3 in ESCs biology.

SMYD3 belongs to the SMYD family of methylases, which is composed of five members. SMYD proteins share a unique architecture of their SET domain, which is split into two parts by a MYND (myeloid translocation protein 8, nervy, DEAF-1) domain and it's followed by a cysteine-rich post-SET domain. Several reports demonstrate that SMYD proteins play important roles in heart development and cardiomyogenesis (Du et al., 2014). *Smyd1* deletion in mice displayed a role in maturation of ventricular cardiomyocytes and right ventricular development (Gottlieb et al., 2002) and *smyd1* knockdown in zebrafish (*Danio rerio*) embryos results in absence of heartbeat and disrupted myofibril organization in cardiac muscles (Tan et al., 2006). *smyd2* zebrafish morphants show an impaired myofibril organization in cardiac muscles (Voelkel *et al.*, 2013) and knockdown of drosophila dSmyd4 expression suggests a role for dSmyd4 in development or function of adult muscles (Thompson and Travers, 2008). SMYD5 negatively regulates lineage-specific genes and plays a role in the maintenance of self-renewal in mESCs (Kidder et al., 2017) and in hematopoiesis during zebrafish embryogenesis (Fujii et al., 2016).

SMYD3 is expressed in the mouse embryo (Hamamoto et al., 2004a) and it becomes over-

expressed in several tumors where it promotes invasiveness, proliferation and metastasis (Hamamoto et al., 2004b; Zou et al., 2009) and multiple reports have focused on SMYD3 role in promoting tumorigenesis (Cock-Rada et al., 2012; Luo et al., 2014).

Despite its expression in normal tissues, very little is known about SMYD3 function in the physiological state. Insights in SMYD3 role in the normal embryo were provided by from a report by Fujii et al., showing that SMYD3 depletion correlates with heart and skeletal muscle abnormalities in adult zebrafish (Fujii et al., 2011). SMYD3 expression in the mouse embryo and adult normal tissues prompted us to investigate SMYD3 role during development, employing mESCs and zebrafish as amenable models.

We show that SMYD3 depletion results in upregulation of mesendodermal markers and in increased expression of cardiovascular markers, in differentiating ESCs. Additionally, we tested the impact of SMYD3 depletion in an *in vivo* model, and found that SMYD3 knockdown promotes mesendodermal markers over-expression, during zebrafish gastrulation.

Collectively, our results show that SMYD3 plays a role during the early steps of differentiation, during the commitment towards distinct germ layers fates.

Material and methods

Cell cultures

All cell lines were obtained from ATCC and were cultured at 37°C with 5% CO_2 in a humid incubator.

Undifferentiated R1 mouse embryonic stem cells were plated on mitomycin C-treated mouse embryonic fibroblast (MEF) feeders for 2 passages and then plated on tissue culture dishes coated with 0.1% gelatin (Millipore), in DMEM (LONZA) supplemented with 15% fetal bovine serum (HyClone), 1% non-essential amino acids (NEAA) (GIBCO), 1% penicillin/streptomycin (P/S) (GIBCO), 1% Glutamax (GIBCO), 0.1 mM β-mercaptoethanol (Sigma) and 1000 U/ml leukemia inhibitory factor LIF (Millipore).

HEK-293T cells used for virus production and MEF cells used for ES feeder were cultured in DMEM (LONZA), containing 10% FBS (GIBCO), 1% P/S (Invitrogen) and 100uM β-mercaptoethanol (Sigma) (MEF only).

In vitro differentiation of ES cells

To induce EB formation, undifferentiated ES cells were cultured in hanging drops for 3days (d) at a density of 800 cells/20µl of differentiation medium (DM), which consisted of DMEM supplemented with penicillin/streptomycin, 0.1 mM nonessential amino acids, 0.1 mM 2-β-mercaptonethanol and 10% FBS (LONZA). EBs were transferred to suspension cultures for an additional 3d (d3+3) in non-adherent Petri dishes. For cardiac differentiation of ESC, EBs were plated in differentiation medium in 0.1% gelatin (Millipore) coated plates. The medium was changed every two days.

Zebrafish lines and maintenance

Zebrafish (*Danio rerio*) embryos were raised and maintained under standard conditions and national guidelines (Italian decree 4th March 2014, n.26). All experimental procedures were approved by IACUC (Institutional Animal Care and Use Committee). Zebrafish AB wild-type strains were obtained from the Wilson lab, University College London, London, United Kingdom. For *in vivo* observations and imaging, embryos beyond 24 hpf were washed, dechorionated and anaesthetized with 0.016% tricaine (ethyl 3-aminobenzoate methanesulfonate salt; Sigma-Aldrich) before observations and picture acquisition. Embryos were staged according to morphological criteria (Kimmel et al., 1995).

Whole mount *in situ* hybridization, embryo sectioning and imaging

Whole mount *in situ* hybridizations were carried out as described (Thisse and Thisse, 2008). The riboprobes were synthesized using the Ambion® MAXIscript® SP6/T7 in Vitro Transcription Kit (Thermo Fisher Scientific). Images of stained embryos were taken on a Leica MZFLIII epifluorescence stereomicroscope equipped with a DFC 480 digital camera and LAS Leica imaging software (Leica, Wetzlar, Germany).

Morpholino microinjection

Antisense morpholino was purchased from Gene Tools (LLC, Philomath, OR). Morpholinos were diluted in Danieau solution (Nasevicius and Ekker, 2000) and injected into 1-to-2 cell-stage embryos using Eppendorf FemtoJet Micromanipulator 5171. Rhodamine dextran (Molecular Probes, Life technology) was co-injected as dye tracer. *smyd3* knockdown was performed injecting 0,3 pmol/embryo of a translation blocking MO (*smyd3*-MO 5'-CCTCTCCATAATCACAGCCTCCATC-3') previously described (Fujii et al, 2011). A standard control oligo (ctrl-MO, 5'-CCTCTTACCTCAGTTACAATTTATA-3', against human β-globin gene) with no target in zebrafish embryos, was also used, to check for non-specific effects due to the injection procedure.

Virus production and plasmids

mES cells were infected with viral supernatant produced as following: HEK293T were transiently transfected with pCL-ECO (Addgene) and retroviral vector pSh-Scramble or pSh-SMYD3 (Proserpio et al., 2013) by using the calcium phosphate technique. After 10 hours, the medium was changed, harvested 36h after transfection, and used to infect mES cells in the presence of 4μg/mL polybrene (Sigma). We performed spinoculation at 4000rpm for 45 minutes. Infected cells were selected with 1μg/mL puromycin (Sigma) for 3 days.

HEK293T were transfected with psPAX2, pMD2.G and pLKO-puro-sh_Scramble or lentivirus brachyury-eGFP (Addgene plasmid 2122) (Kita-Matsuo et al., 2009).

Reverse transcriptase PCR and Real-time PCR

Total RNA was isolated by the Tryzol method (SIGMA), reverse transcribed to cDNA with High Capacity cDNA Reverse Transcription Kit (Applied Bio-System) and subjected to qPCR analysis. Primers were designed to amplify regions of 80-120bp in size and include two exons. Oligonucleotides sequence used in qRT-PCR is reported below. Quantitative PCR was performed using SYBR green IQ reagent (GeneSpin) in the iCycler IQ detection system (Biorad). qPCR was

performed in triplicate for each of the biological replicates.

For embryonic cDNAs, pregnant mice were sacrificed between at day 10.5 and 12.5 of pregnancy, embryos were collected by caesarean section, dissected and processed as above.

GENE		Sequence
NANOG	F	GCGGACTGTGTTCTCTCAGG
	R	CCACCGCTTGCACTTCATCC
KLF 4	F	CGTCCCAGTCACAGTGGTAA
	R	AAAAGAACAGCCACCCACAC
EOMES	F	CCCACGTCTACCTGTGCAAC
	R	GGTGGGGTTGAGTCCGTTTA
GAPDH	F	CACCATCTTCCAGGAGCGAG
	R	CCTTCTCCATGGTGGTGAAGAC
Oct 4	F	AAGCAACTCAGAGGGAACCT
	R	GGTGATCCTCTTCTGCTTCA
EOMES	F	CCCACGTCTACCTGTGCAAC
	R	GGTGGGGTTGAGTCCGTTTA
Mesp 1	F	GTTCCTGTACGCAGAAACAGCAT
	R	GTTTCTAGAAGAGCCAGCATGTC
SOX-17	F	GCCGATGAACGCCTTTATGGTG
	R	TCTCTGCCAAGGTCAACGCCTT
Brachyury	F	TGCCTACCAGAATGAGGAGATTA
	R	CCATTACATCTTTGTGGTCGTTT
TnnT	F	AGCGCGTGGAGAAGGACCTGA
	R	CCGCTCTGCCCGACGCTTT
MYH7	F	CCAAGGGCCTGAATGAGGAG
	R	GCAAAGGCTCCAGGTCTGAG
MYL7	F	AGGAAGCCATCCTGAGTGCCTT
	R	CATGGGTGTCAGCGCAAACAGT
ACTA2	F	AGGCACCACTGAACCCTAAG
	R	ACAGCACAGCCTGAATAGCC
NKX2.5	F	GTCCAGCTCCACTGCCTTCT
	R	CAAGTGCTCTCCTGCTTTCC
Isl 1	F	GGCTACACAGCGGAAACACT
	R	ACGTGCTTTGTTAGGGATGG
Snail 1	F	ACTGGTGAGAAGCCATTCTCCT
	R	CTGGTATCTCTTCACATCCGAGT
E-cadherin	F	GTCTACCAAAGTGACGCTGAAGT
	R	TCTCGTTTCTGTCTTCTGAGACC
OTX2	F	CTTCATGAGGGAAGAGGTGGCAC
	R	TGGCGGCACTTAGCTCTTCGATTC
FGF5	F	GCTGTGTCTCAGGGGATTGT
	R	CACTCTCGGCCTGTCTTTTC
Smyd3	F	AGAGGTGTGCAAGTGATGAAAGT
	R	ATCAAATCTTCAATCAGGCTGTG
GAPDH	F	CACCATCTTCCAGGAGCGAG
	R	CCTTCTCCATGGTGGTGAAGAC

Protein extracts

Total cells extracts were prepared lysing the cells pellet in RIPA buffer as described (Proserpio et al., 2013). The extracts were centrifuged at maximum speed for 15 min and the supernatant was collected.

Protein concentration was determined by Bradford assay. Total cell extracts were separated by SDS-PAGE, and immunoblotted with specific antibodies (SMYD3: GeneTex, GTX121945, GAPDH: Santa Cruz, sc-32233, Nanog: Cosmobio REC-RCAB002P-F, Oct4: ab19857, αSMA:ab15694, Flk1: Novus NB100-627SS).

To isolate chromatin, sample were prepared following Mendez's protocol described in (Méndez and Stillman, 2000). Briefly, cells were resuspended in buffer A (10 mM HEPES pH 7.9, 10 mM KCl,

1.5 mM MgCl2, 0.34 M sucrose, 10% glycerol, 1 mM DTT and Protease inhibitor cocktail). Triton X-100 (0.1%) was added, and the cells were incubated for 5 min on ice. Nuclei were collected in pellet 1 (P1) by low-speed centrifugation. The supernatant (S1) was further clarified by high-speed centrifugation to remove cell debris and insoluble aggregates. Nuclei were washed once in buffer A, and then lysed in buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, protease inhibitors). Insoluble chromatin was collected by centrifugation, washed once in buffer B, and centrifuged again under the same conditions. The final chromatin pellet (P3) was resuspended in Laemmli buffer and sonicated.

Immunofluorescence analysis

Immunostaining of cultured cells were carried out as described previously (Proserpio et al., 2013). Briefly, the cells were fixed for 10 min in 4% paraformaldehyde (PFA) in PBS and permeabilized with 0.2% Triton X-100 in PBS for 10 min. After incubation in 3% BSA for at least 20 min at room temperature, cells were incubated for 2 h at room temperature with antibodies. After washing the samples with 0.1% Triton X-100 in PBS for 10 min, cells were incubated with secondary antibodies, which are fluorescent labeled. DAPI was used for nuclear staining. The samples were examined with a fluorescence microscope (Carl Zeiss, Italy). Pictures of staining were obtained using an AxioCam (Carl Zeiss Vision, Italy).

Statistical analysis

Data are represented as mean ± standard error of the mean (SEM). Data are analyzed by using a two-way ANOVA followed by post-hoc Bonferroni’s test comparison. Probability values of less than 0.05 were considered statistically significant.

Results

SMYD3 is expressed in the developing mouse embryo and in embryonic stem cells

SMYD3 was mainly studied in different cancer cell lines, where it promotes proliferation and invasion (Giakountis et al., 2017; Hamamoto et al., 2004a). To investigate if SMYD3 may play a role in development, we first analyzed SMYD3 expression levels in non-tumor derived cells. SMYD3 transcripts were high in mESCs and C2C12 myoblasts, when compared to the fibroblast cells NIH3T3. SMYD3 mRNA was also expressed in the mouse embryo at d10.5 (Fig.1 A).

To investigate SMYD3 role during development, we decided to use mESCs as a model system that recapitulates early stages of embryo formation.

We first analyzed SMYD3 protein levels by Western blot throughout differentiation of embryoid bodies (EBs) and we observed that SMYD3 levels decrease at early stages of differentiation and it is maintained to stable levels till day 12 (Fig. 1B).

We also checked SMYD3 distribution in different intracellular compartments by chromatin fractionation and immunofluorescence assays. These experiments revealed that SMYD3 is expressed in mESCs and it is detectable in cytoplasm, nuclei and chromatin fraction (Fig. 1C, 1D).

Together, these data indicate that SMYD3 is expressed in mouse embryonic stem cells and EBs, which can be used to further investigate SMYD3 role in stemness and lineage commitment.

SMYD3 depletion does not affect stemness

To test if SMYD3 plays a role in the differentiation of mESCs, we performed loss-of-function experiments by reducing SMYD3 levels with a short hairpin RNA (shRNA) targeting SMYD3 (Fig. 2A). Pluripotency markers *Oct4*, *Nanog* and *KLF4* maintained the same expression profile, SMYD3-depleted cells compared to the control, in the early stages of differentiation. These results were confirmed by immunoblot analysis, with antibodies raised against Nanog and Oct4 (Fig. 2B). Overall, we concluded that SMYD3 does not play a role in mESCs stemness maintenance.

SMYD3 depletion accelerates mesendodermal markers expression, in mESCs

Next, we sought to elucidate whether SMYD3 regulates the differentiation potential of mESCs. We examined and compared mRNA expression patterns for various lineage markers at different time points following LIF (leukemia inhibitory factor) withdrawal and embryoid bodies formation. The mesendodermal markers *Eomes*, *Mesp1*, *Sox17*, *FoxA2* were upregulated in control cells and peaked through day 4-5 of differentiation. Strikingly, in SMYD3 depleted cells the expression of *Eomes*, *Mesp1* and *FoxA2* was significantly higher compare to the control ($p < 0.05$) (Fig. 3A). To further characterize SMYD3 role in the early stages of lineage commitment, we analyzed the pan-

mesodermal marker Brachyury/T expression levels, which was upregulated in SMYD3-depleted cells (Fig. 3A). Moreover, we co-infected undifferentiated Sh-SMYD3 or Sh-Scramble mESCs with a virus carrying eGFP cDNA under the control of the Brachyury/T promoter. Consistently, we observed eGFP fluorescence peaking at days 6 of differentiation, and Sh-SMYD3 cells displayed more foci of GFP positive cells (Fig. 3B).

Furthermore, markers of the primitive ectoderm *Fgf5* and *Otx2* as well as the *Sox1* ectoderm marker behaved in a similar manner in SMYD3 knocked-down and control cells (Fig. 3C). Overall, these data show that SMYD3 depletion favors mesendodermal commitment in differentiating EBs.

smyd3 knockdown promotes mesendodermal fate during zebrafish gastrulation

To test whether our findings were confirmed in an in vivo model, we injected *smyd3* morpholino (Smyd3_MO) in zebrafish embryos, at the 1-2 cell stage. Morphants displays curved trunk, pericardial edema, impaired circulation compared to the standard control embryos and were divided in four phenotypic classes, with increasing severity (Supplemental Fig. 1A and B). These defects reflected the previously reported phenotype for *smyd3* morphants (Fujii et al., 2011; Sarris et al., 2016).

To analyze zebrafish mesendodermal markers in *smyd3* morphants, we performed in situ hybridization (ISH) at the onset of gastrulation (50% epiboly). *Gata6, gata5, mespaa* and *snai1a* signal was enhanced in the blastodermal margin, in *smyd3* morphants (Fig. 4A). Conversely, the ectodermal marker *otx3* expression was comparable in morphant and standard control embryos (Fig. 4B). In midgastrulation (80% epiboly), *sox17* levels were enhanced in the endodermal precursors cells (Fig. 4C). Additionally, the dorsal markers *goosecoid* (*gsc*) and *chordin* (*chd*) showed an expansion in the blastodermal margin of *smyd3* knocked down embryos (Fig. 4D), while the signal for the ventral marker *gata2* was reduced. Overall, these data suggest that *smyd3* knockdown favors the induction of mesendodermal fate in the zebrafish embryo, supporting the in vitro results in mESCs.

SMYD3 knockdown influences later stages of differentiation

SMYD3-depleted ESC appeared to be more prone to differentiate towards the mesendodermal lineage. To test whether upregulation in mesendodermal lineage persisted also at later stages of differentiation, we differentiated Sh-SMYD3 and Sh-Scramble EBs toward cardiovascular fate. This lineage choice was particularly relevant, because of the possible role of SMYD3 in cardiomyocyte formation in zebrafish (Fujii et al., 2011).

Early markers of multipotent cardiovascular progenitors such as *Nkx2.5* and *Isl-1* were more

robustly expressed in SMYD3-depleted EBs when compared to Sh-Scramble control EBs (Fig. 5A). In parallel, we also investigated the expression of a second mesodermal marker, *Flk-1*. Flk-1 protein expression in SMYD3 depleted EBs was anticipated from day 8 to day 6, when compared to Sh-Scramble cells (Fig. 5B). This finding was confirmed by immunofluorescence experiments in Sh-SMYD3 and control EBs at day 5 of differentiation. At this stage, Flk1-positive cells were more abundant in SMYD3-depleted EBs, when compared to control Sh-Scramble cells, in immunofluorescence assays (Fig. 5C).

In order to investigate whether SMYD3 depletion enhanced the expression of cardiac lineage markers in EBs, we analyzed the transcript levels of late cardiac markers, e.g. as *TnnT2*, *Myh7*, and *Myl7*, which are essential components of contractile cardiac machinery. We detected elevated mRNA levels of *TnnT2, Myh7* and *Myl7* in SMYD3 depleted EBs (Fig. 6A). This increase was paralleled by a higher percentage of beating EBs in SMYD3 depleted EBs compare to the control ones (Fig. 6B).

Since Flk1 positive cells give rise to cardiomyocytes, smooth muscle and endothelial populations, we also explored the expression levels of smooth muscle actin (αSMA) and endothelial marker CD31. αSMA mRNA and protein levels were increased at d10 of differentiation (Fig. 6C and D). The endothelial marker CD31 was also increased in Sh-SMYD3 EBs compared to Sh-Scramble ones, at day 10 (Fig. 6E).

Our results show that SMYD3 depletion promotes cardiac, smooth muscle and endothelial marker expression in differentiating ESCs.

In addition, we analyzed the expression of the endothelial marker *kinase insert domain receptor like* (*kdrl*), which is the Flk1 orthologous in zebrafish. Through *in situ* hybridization analyses we observed a slight increase in *kdrl* expression in *smyd3* morphants at 24 hpf. *Kdrl* increased expression persists and intensifies at 48 hpf (Fig. 6F). Enhanced expression was particularly evident in the cardinal vein, dorsal aorta, posterior caudal vein, intersegmental vessels and head vessels of *smyd3* knocked down embryos at 48 hpf, suggesting that the kinase receptor is upregulated in the endothelial cells lacking *smyd3*.

Taken together, these findings hint for a tendency of SMYD3-depleted cells to differentiate into mesodermal lineages.

Discussion

Fujii et al. (2011) recently showed that SMYD3 may be involved in cardiomyocyte differentiation and myogenesis in zebrafish (Fujii et al., 2011), but its potential contribution in developmental processes has not been fully investigated yet.

We detected robust SMYD3 expression in the developing embryo as well as in mouse embryonic stem cells, which were used as suitable *in vitro* model to study lineages differentiation.

Our data show that SMYD3 knockdown promotes mesendodermal markers expression, both in differentiating mESCs and in gastrulating zebrafish embryos.

While zebrafish *smyd3* morphants have phenotypic defects, Smyd3 knockout mice don't show an apparent phenotype (Mazur et al., 2014; Sarris et al., 2016). SMYD2 knockdown in ESCs leads to the upregulation of endodermal markers, but not of mesodermal ones. We speculate that SMYD3 functions can be partially compensated by other SMYD family members or unrelated proteins, during mouse development, where SMYD3 loss does not result in any obvious defects in the mouse. Alternatively, SMYD3 functions might be not conserved among species.

SMYD3 has both nuclear and cytoplasmic functions, which may depend or be not on its methylation activity (Mazur et al., 2014; Sarris et al., 2016). We speculate that SMYD3 might participate in the induction of repressors of the mesendodermal fate.

Several signaling pathways modulate mesendoderm formation: FGF and Wnt pathways are required for the proper formation of germ layers (Willems and Leyns, 2008) and Activin/Nodal and BMP4 pathways further influence mesoderm and endoderm specification (Lindsley et al., 2006). Wnt and Nodal/Activin pathways modulate mesodermal markers transcription (Gadue et al., 2006). It appears that in mESCs SMYD3 may affect one of these pathways to diminish mesendodermal genes transcription.

Recently, SMYD3 was shown to methylate the MAP3K2 kinase and affect Ras-activated pathway in certain types of cancer (Mazur et al., 2014). It is tempting to speculate that SMYD3 modulates other cytoplasmic kinases along the Wnt or Nodal/Activin pathway to modulate the outcome of their signaling.

Interestingly, SMYD3 localization in ESCs is not restricted to the cytoplasm, but it is also present in the nucleus. Therefore, SMYD3 nuclear functions may play a direct role on the chromatin of pivotal regulatory genes, which negatively modulate mesodermal gene transcription.

Bibliography

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R., Young, R.A., 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947-956.
Campbell, P.A., Perez-Iratxeta, C., Andrade-Navarro, M.A., Rudnicki, M.A., 2007. Oct4 targets regulatory nodes to modulate stem cell function. PLoS One 2, e553.
Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., Loh, Y.H., Yeo, H.C., Yeo, Z.X., Narang, V., Govindarajan, K.R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W.K., Clarke, N.D., Wei, C.L., Ng, H.H., 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133, 1106-1117.
Cock-Rada, A.M., Medjkane, S., Janski, N., Yousfi, N., Perichon, M., Chaussepied, M., Chluba, J., Langsley, G., Weitzman, J.B., 2012. SMYD3 promotes cancer invasion by epigenetic upregulation of the metalloproteinase MMP-9. Cancer Res 72, 810-820.
Delmore, J.E., Issa, G.C., Lemieux, M.E., Rahl, P.B., Shi, J., Jacobs, H.M., Kastritis, E., Gilpatrick, T., Paranal, R.M., Qi, J., Chesi, M., Schinzel, A.C., McKeown, M.R., Heffernan, T.P., Vakoc, C.R., Bergsagel, P.L., Ghobrial, I.M., Richardson, P.G., Young, R.A., Hahn, W.C., Anderson, K.C., Kung, A.L., Bradner, J.E., Mitsiades, C.S., 2011. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. Cell 146, 904-917.
Du, S.J., Tan, X., Zhang, J., 2014. SMYD proteins: key regulators in skeletal and cardiac muscle development and function. Anat Rec (Hoboken) 297, 1650-1662.
Fisher, C.L., Fisher, A.G., 2011. Chromatin states in pluripotent, differentiated, and reprogrammed cells. Curr Opin Genet Dev 21, 140-146.
Fujii, T., Tsunesumi, S., Sagara, H., Munakata, M., Hisaki, Y., Sekiya, T., Furukawa, Y., Sakamoto, K., Watanabe, S., 2016. Smyd5 plays pivotal roles in both primitive and definitive hematopoiesis during zebrafish embryogenesis. Sci Rep 6, 29157.
Fujii, T., Tsunesumi, S., Yamaguchi, K., Watanabe, S., Furukawa, Y., 2011. Smyd3 is required for the development of cardiac and skeletal muscle in zebrafish. PLoS One 6, e23491.
Gadue, P., Huber, T.L., Paddison, P.J., Keller, G.M., 2006. Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. Proc Natl Acad Sci U S A 103, 16806-16811.
Giakountis, A., Moulos, P., Sarris, M.E., Hatzis, P., Talianidis, I., 2017. Smyd3-associated regulatory pathways in cancer. Semin Cancer Biol 42, 70-80.
Gottlieb, P.D., Pierce, S.A., Sims, R.J., Yamagishi, H., Weihe, E.K., Harriss, J.V., Maika, S.D., Kuziel, W.A., King, H.L., Olson, E.N., Nakagawa, O., Srivastava, D., 2002. Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. Nature genetics 31, 25-32.
Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R., Nakamura, Y., 2004a. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat Cell Biol 6, 731-740.
Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R., Nakamura, Y., 2004b. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat Cell Biol 6, 731-740.
Kidder, B.L., He, R., Wangsa, D., Padilla-Nash, H.M., Bernardo, M.M., Sheng, S., Ried, T., Zhao, K., 2017. SMYD5 controls heterochromatin and chromosome integrity during embryonic stem cell differentiation. Cancer Res.
Kita-Matsuo, H., Barcova, M., Prigozhina, N., Salomonis, N., Wei, K., Jacot, J.G., Nelson, B., Spiering, S., Haverslag, R., Kim, C., Talantova, M., Bajpai, R., Calzolari, D., Terskikh, A., McCulloch, A.D., Price, J.H., Conklin, B.R., Chen, H.S., Mercola, M., 2009. Lentiviral vectors and

protocols for creation of stable hESC lines for fluorescent tracking and drug resistance selection of cardiomyocytes. PLoS One 4, e5046.
Li, X., Li, L., Pandey, R., Byun, J.S., Gardner, K., Qin, Z., Dou, Y., 2012. The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. Cell Stem Cell 11, 163-178.
Lindsley, R.C., Gill, J.G., Kyba, M., Murphy, T.L., Murphy, K.M., 2006. Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. Development 133, 3787-3796.
Luo, X.G., Zhang, C.L., Zhao, W.W., Liu, Z.P., Liu, L., Mu, A., Guo, S., Wang, N., Zhou, H., Zhang, T.C., 2014. Histone methyltransferase SMYD3 promotes MRTF-A-mediated transactivation of MYL9 and migration of MCF-7 breast cancer cells. Cancer Lett 344, 129-137.
Mazur, P.K., Reynoird, N., Khatri, P., Jansen, P.W., Wilkinson, A.W., Liu, S., Barbash, O., Van Aller, G.S., Huddleston, M., Dhanak, D., Tummino, P.J., Kruger, R.G., Garcia, B.A., Butte, A.J., Vermeulen, M., Sage, J., Gozani, O., 2014. SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. Nature.
Méndez, J., Stillman, B., 2000. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. Mol Cell Biol 20, 8602-8612.
Ohtani, K., Zhao, C., Dobreva, G., Manavski, Y., Kluge, B., Braun, T., Rieger, M.A., Zeiher, A.M., Dimmeler, S., 2013. Jmjd3 controls mesodermal and cardiovascular differentiation of embryonic stem cells. Circ Res 113, 856-862.
Ong, C.T., Corces, V.G., 2011. Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat Rev Genet 12, 283-293.
Orkin, S.H., Hochedlinger, K., 2011. Chromatin connections to pluripotency and cellular reprogramming. Cell 145, 835-850.
Proserpio, V., Fittipaldi, R., Ryall, J.G., Sartorelli, V., Caretti, G., 2013. The methyltransferase SMYD3 mediates the recruitment of transcriptional cofactors at the myostatin and c-Met genes and regulates skeletal muscle atrophy. Genes Dev 27, 1299-1312.
Sarris, M.E., Moulos, P., Haroniti, A., Giakountis, A., Talianidis, I., 2016. Smyd3 Is a Transcriptional Potentiator of Multiple Cancer-Promoting Genes and Required for Liver and Colon Cancer Development. Cancer Cell 29, 354-366.
Tan, X., Rotllant, J., Li, H., De Deyne, P., DeDeyne, P., Du, S.J., 2006. SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. Proc Natl Acad Sci U S A 103, 2713-2718.
Thompson, E.C., Travers, A.A., 2008. A Drosophila Smyd4 homologue is a muscle-specific transcriptional modulator involved in development. PloS one 3, e3008.
Willems, E., Leyns, L., 2008. Patterning of mouse embryonic stem cell-derived pan-mesoderm by Activin A/Nodal and Bmp4 signaling requires Fibroblast Growth Factor activity. Differentiation 76, 745-759.
Young, R.A., 2011. Control of the embryonic stem cell state. Cell 144, 940-954.
Zou, J.N., Wang, S.Z., Yang, J.S., Luo, X.G., Xie, J.H., Xi, T., 2009. Knockdown of SMYD3 by RNA interference down-regulates c-Met expression and inhibits cells migration and invasion induced by HGF. Cancer Lett 280, 78-85.

Figure legends

Figure 1. SMYD3 is expressed in mouse embryo and in mouse embryonic stem cells (mESCs). **A**, mRNA was extracted from NIH3T3, C2C12 myoblasts and mESC in parallel with d10.5 embryo and SMYD3 transcript levels were measured by qRT-PCR. Data were normalized to GAPDH. Means from three independent experiments are shown. **B**, Immunoblot analysis of SMYD3 protein levels at different time points of ESCs differentiation. GAPDH served as a loading control. **C**, immunofluorescence was performed on undifferentiated mESCs with an antibody raised against SMYD3. Nuclei were visualized by DAPI staining (blue). **D**, ESCs were analyzed by biochemical fractionation, followed by immunoblot, to characterize SMYD3 distribution in different fractions: cytoplasmic or soluble fraction (S2), solubilized nuclear proteins fraction (S3), and chromatin-nuclear matrix-bound fraction (Chr). GAPDH served as S2 control, while H3 served as Chr control.

Figure 2. SMYD3 depletion does not interfere with stemness. A, mRNAs were extracted at different stages of ESCs differentiation, and SMYD3 or pluripotency-related genes (*Nanog*, *Klf4*, *Oct4*) transcript levels were evaluated throughout differentiation, by qRT-PCR. **B**, Nanog, Oct4 and SMYD3 protein levels were analyzed by immunoblot. Data were normalized to GAPDH. Means and SEM from three independent experiments are shown.

Figure 3. SMYD3 knockdown increase mesendodermal marker expression, in differentiating mESCs. A, mRNA levels of mesendodermal markers were measured by qRT-PCR. Data were normalized to GAPDH. Means and SEM from three independent experiments are shown. Significance was calculated by 2 way Anova, followed by Bonferroni post-hoc test. *$p<0.05$ value for the significance is shown in the plot.
B, eGFP reporter fluorescence was analyzed in EBs infected with Sh-Scramble/brachyury-eGFP and Sh-SMYD3/brachyury-eGFP were checked for eGFP expression, at day 6 of differentiation.
C, primitive ectoderm genes (*Otx2*, *FGF5*, *Sox1*) expression was measured by qRT-PCR at different time points. Means and SEM from three independent experiments are shown.

Figure 4. *smyd3* knockdown promotes mesendodermal fate in zebrafish
A, expression of the mesendodermal markers *gata6*, *gata5*, *mespaa* and *snai1a* (side view), was evaluated by ISH in *smyd3* morphants at 50% of epiboly, in comparison to standard control embryos. Embryos are shown in dorsal view. **B**, ISH of the ectodermal marker *otx3* was performed in morphants and standard control embryos at 50% epiboly stage. **C**, the endodermal marker *sox17* was examined at 80% epiboly and shown in side view. D, ISH of dorsal markers *gsc*, *chd* and of ventral marker *gata2* was performed at 50% epiboly and shown in dorsal view.

Figure 5. SMYD3 depletion enhances early cardiac differentiation in mESCs. A, Gene expression of early cardiac markers *Isl1* and *Nkx2.5* was analyzed by qRT-PCR. Data were normalized to GAPDH. **B**, immunofluorescence analysis was performed with Flk1 antibodies at d5 of EBs differentiation. DAPI stains nuclei. **C**, Western Blot analysis was performed on mESCs protein extracts at different time of differentiation, with antibodies raised against Flk1. GAPDH serves as a loading control.

Figure 6. SMYD3 modulates late cardiovascular markers. A, Cardiac markers *TnnT2*, *Myh7* and *Myl7* were measured by qRT-PCR at d0, d6, d8, d10 and d12 of mESC differentiation. Data were normalized to GAPDH **B,** Beating EBs were counted at d10, d11, d12 of EBs formation in Sh-Scramble and Sh-SMYD3. **C**, Western Blot analysis of αSMA protein extracts of Sh-Scramble and Sh-SMYD3 EBs at d10, d12, d14. GAPDH serves as a loading control**. D,** αSMA mRNA levels were measured by qRT-PCR as in A. **E**, Immunofluorescence microscopy was performed on d10 EBs with antibodies against the endothelial marker CD31. DAPI stains nuclei. **F,** *kdrl* expression

was analyzed at 24 and 48 hpf, in standard control and SMYD3 morphants. Arrows show caudal artery (CA), dorsal aorta (DA), posterior caudal vein (PCV), intersegmental vessels (ISVs) and cardinal vein (CV).

Supplementary Fig. S1. SMYD3 morphant phenotypic classes.
A, Schematic representation of morpholino-mediated targeting of *smyd3* transcript.
B, Representative images of the four observed phenotypic classes at 48hpf are shown.
C, Morphants distribution in the four phenotypic classes. Morphants were analyzed at two morpholino does (0.2 and 0.4pmol/embryo).

Fig.1

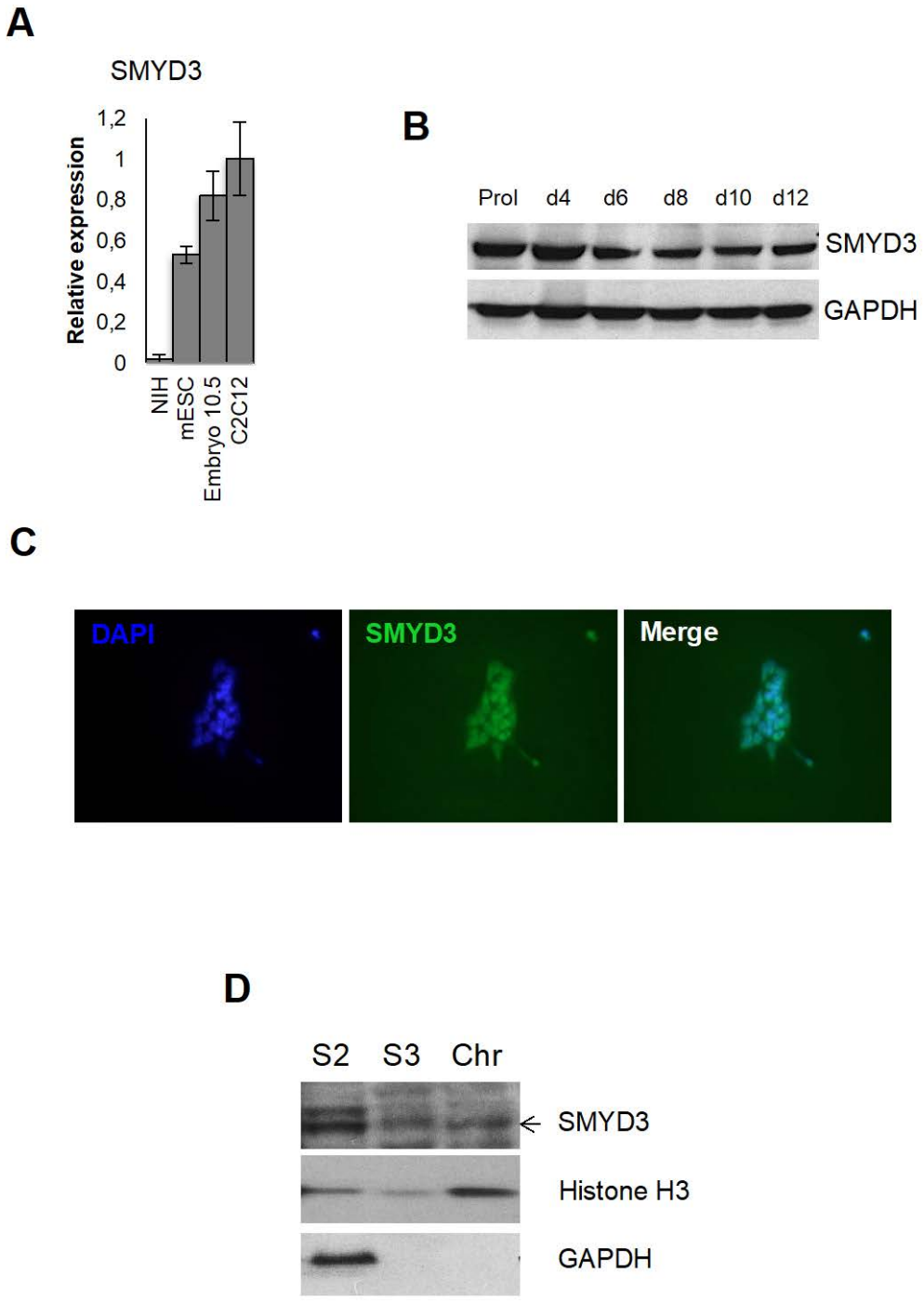
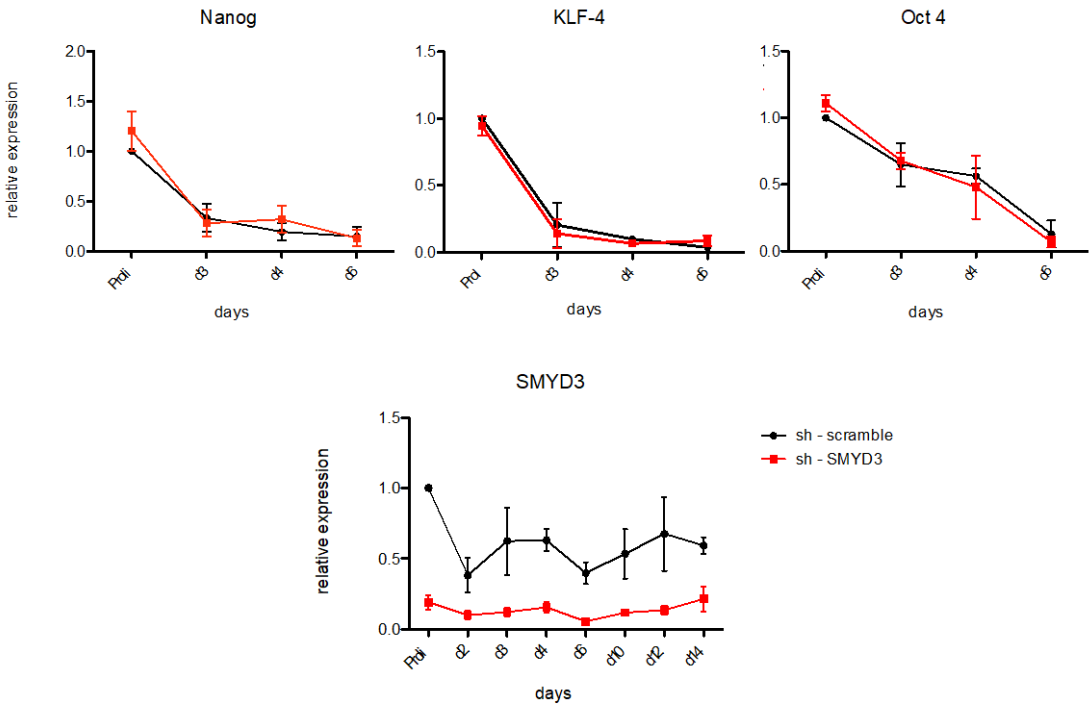

A
SMYD3
Relative expression
1,2
1
0,8
0,6
0,4
0,2
0
NIH
mESC
Embryo 10.5
C2C12
B
Prol
d4
d6
d8
d10
d12
SMYD3
GAPDH
C
DAPI
SMYD3
Merge
D
S2
S3
Chr
SMYD3
Histone H3
GAPDH

Fig.2

A


Nanog
KLF-4
Oct 4
relative expression
days
Prol
d3
d4
d6
SMYD3
sh - scramble
sh - SMYD3
d2
d10
d12
d14

B

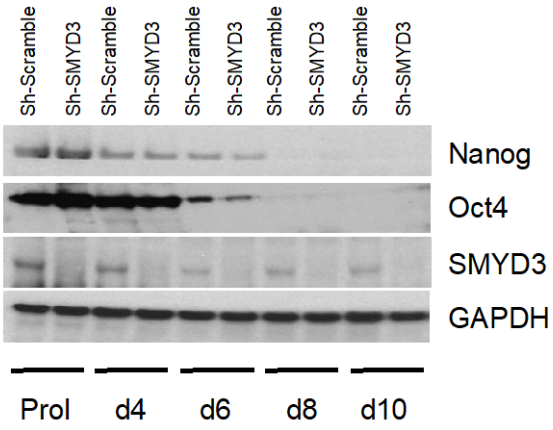

Sh-Scramble
Sh-SMYD3
Nanog
Oct4
SMYD3
GAPDH
Prol
d4
d6
d8
d10

Fig.3

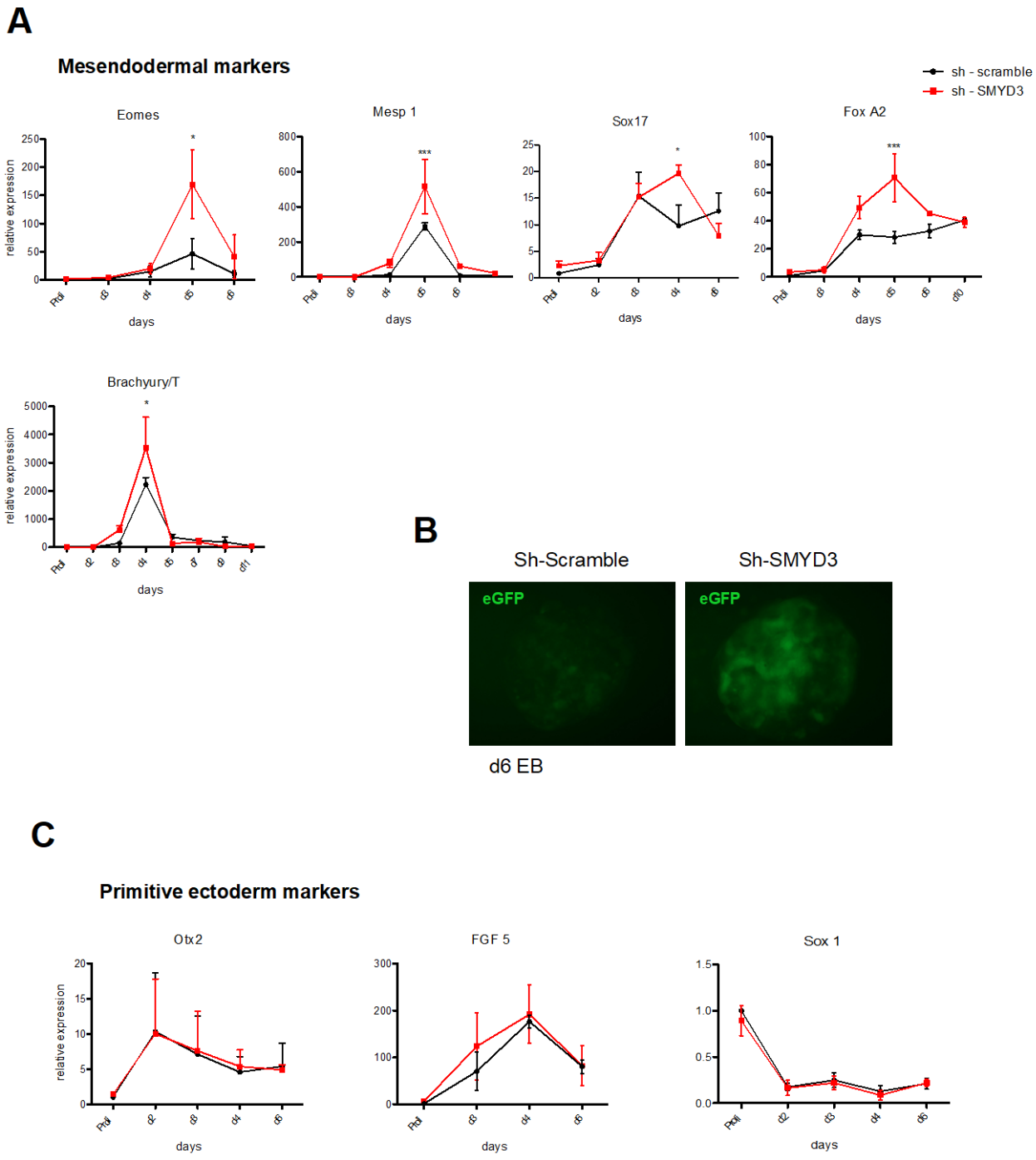

A
Mesendodermal markers
sh - scramble
sh - SMYD3
Eomes
Mesp 1
Sox17
Fox A2
Brachyury/T
relative expression
days
B
Sh-Scramble
Sh-SMYD3
eGFP
eGFP
d6 EB
C
Primitive ectoderm markers
Otx2
FGF 5
Sox 1
relative expression
days

Fig.4

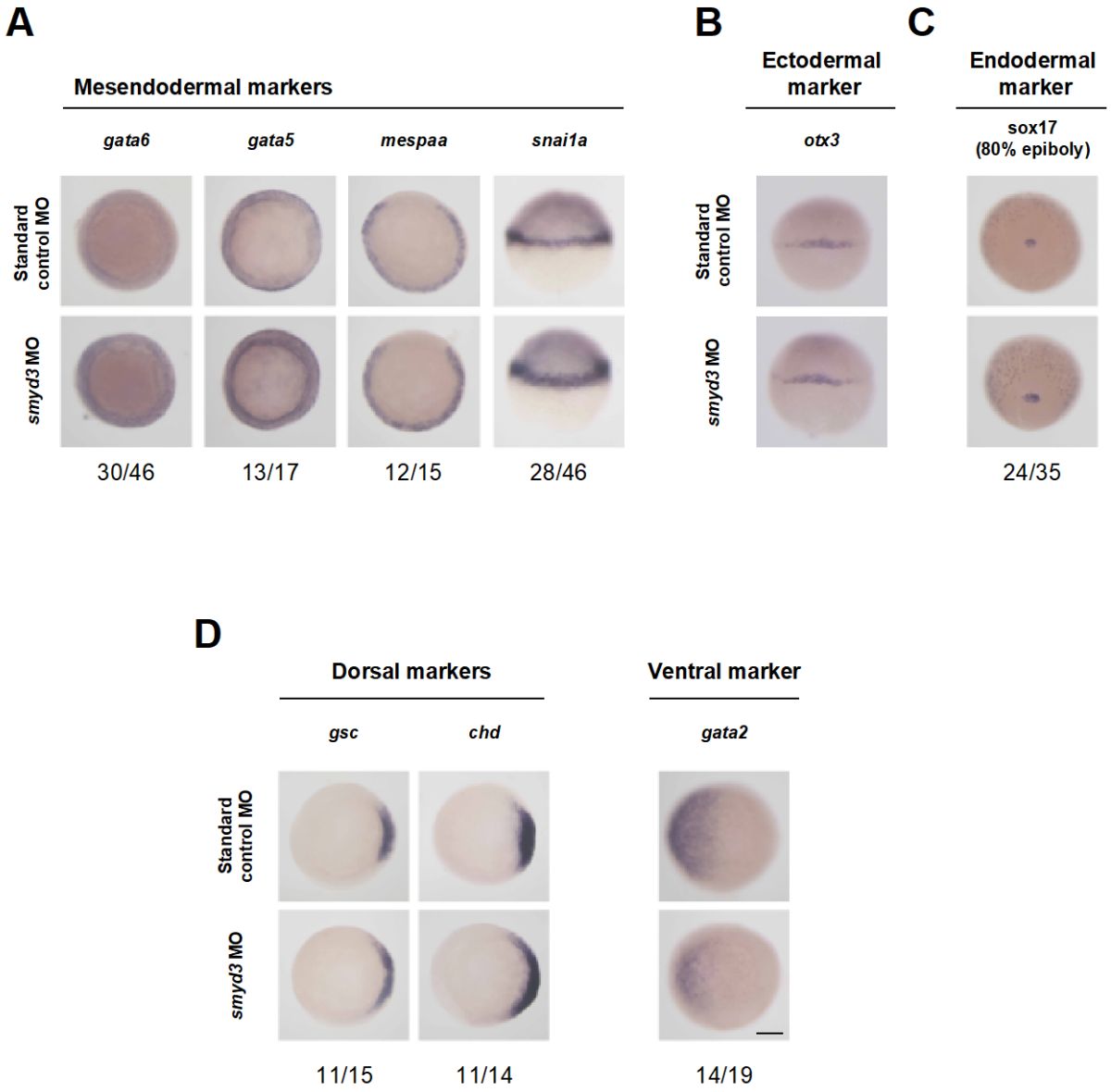
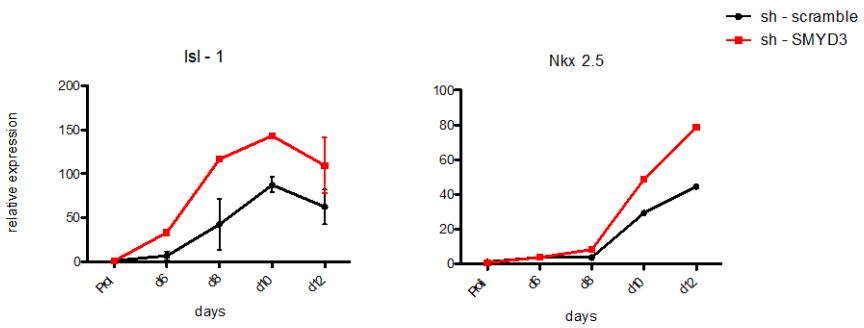
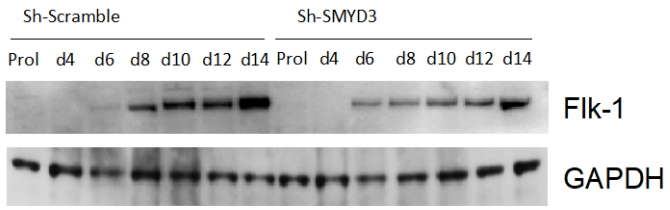

A
Mesendodermal markers
gata6
gata5
mespaa
snai1a
Standard control MO
smyd3 MO
30/46
13/17
12/15
28/46
B
Ectodermal marker
otx3
Standard control MO
smyd3 MO
C
Endodermal marker
sox17 (80% epiboly)
24/35
D
Dorsal markers
gsc
chd
Ventral marker
gata2
Standard control MO
smyd3 MO
11/15
11/14
14/19

Fig.5

A


Early markers of Multipotent Cardiovascular Progenitors
sh - scramble
sh - SMYD3
Isl - 1
Nkx 2.5
relative expression
days
days

B


Sh-Scramble
Sh-SMYD3
Prol d4 d6 d8 d10 d12 d14 Prol d4 d6 d8 d10 d12 d14
Flk-1
GAPDH

C

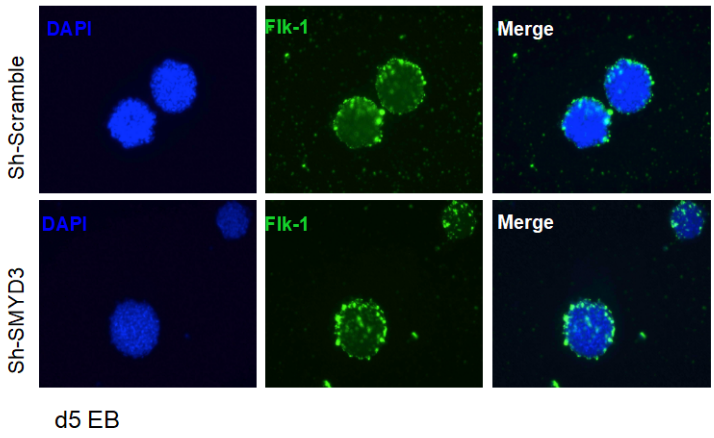
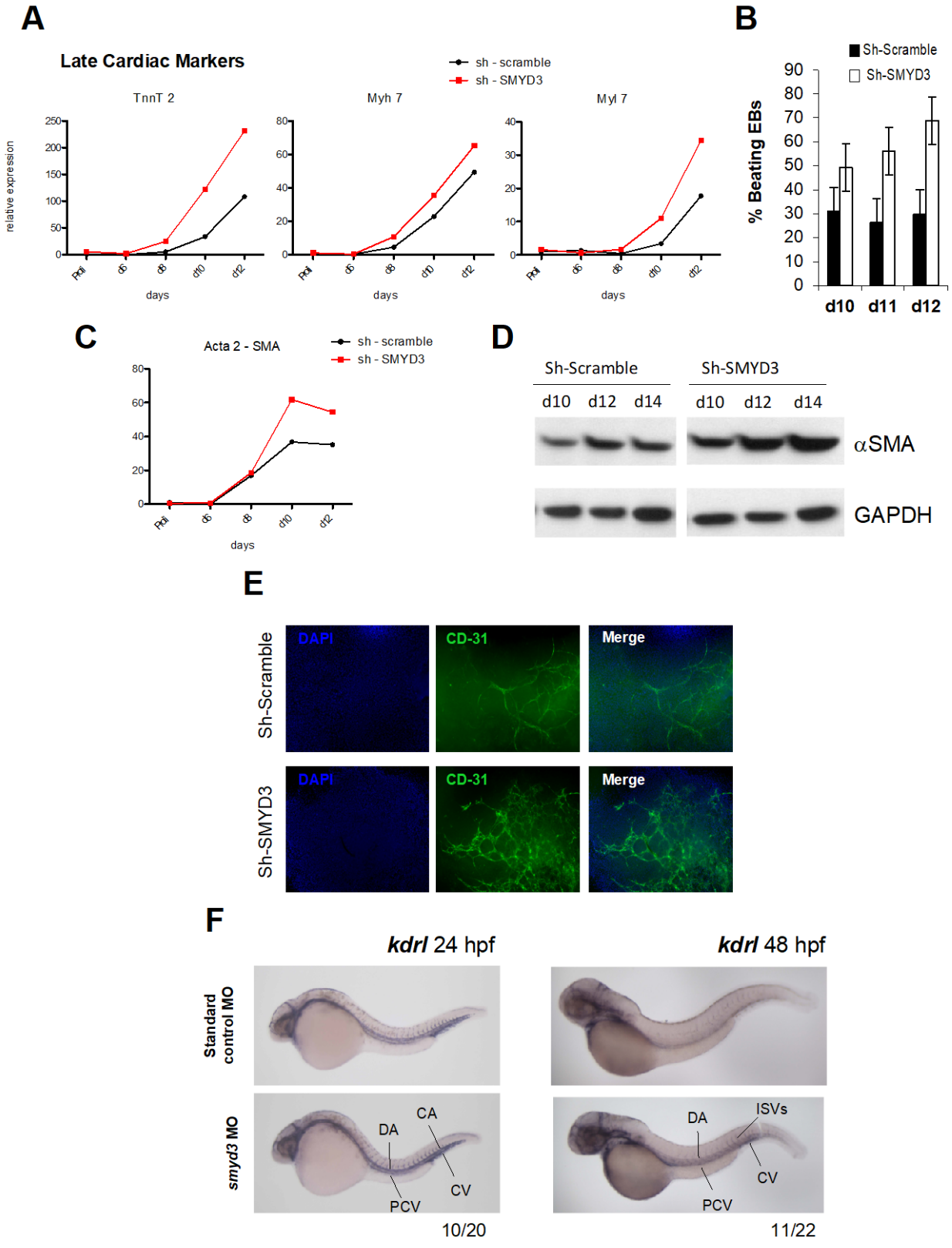
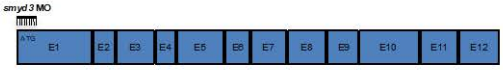

Sh-Scramble
Sh-SMYD3
DAPI
Flk-1
Merge
DAPI
Flk-1
Merge
d5 EB

Fig.6


A
Late Cardiac Markers
sh - scramble
sh - SMYD3
TnnT 2
Myh 7
Myl 7
relative expression
days
B
Sh-Scramble
Sh-SMYD3
% Beating EBs
d10 d11 d12
C
Acta 2 - SMA
sh - scramble
sh - SMYD3
days
D
Sh-Scramble
Sh-SMYD3
d10 d12 d14
d10 d12 d14
αSMA
GAPDH
E
Sh-Scramble
Sh-SMYD3
DAPI
CD-31
Merge
F
kdrl 24 hpf
kdrl 48 hpf
Standard control MO
smyd3 MO
DA
CA
CV
PCV
ISVs
10/20
11/22

Supplementary Fig.S1

A



smyd3 transcript

B

Control

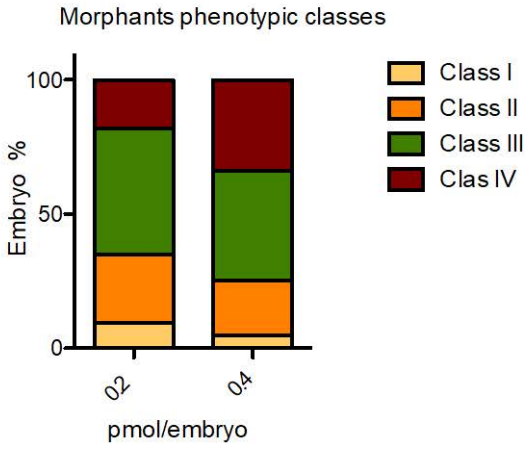
Class I

Class II
Mild cardiac edema, slightly curved trunk, impaired circulation;

Class III
Cardiac edema, lack of circulation, curved trunk;

Class IV
Severe

C



PART III: SMYD3 IN CANCER

SMYD3 promotes the Epithelial-Mesenchymal-Transition in Breast Cancer (Paper in revision)

Claudio Fenizia[1,*], Cinzia Bottino[1,*], Raffaella Fittipaldi[1], Pamela Floris[1], Silvia Corbetta[1], Germano Gaudenzi[2], Silvia Carra[1], Franco Cotelli[1], Giovanni Vitale[2,3], Alberto Del Rio[4], Giuseppina Caretti[1, #]

[1]Department of Biosciences, Universita' degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy
[2]Department of Clinical Sciences and Community Health (DISCCO), University of Milan, Milan, Italy.
[3]Laboratory of Endocrine and Metabolic Research, Istituto Auxologico Italiano IRCCS, Milan, Italy.
[4]Institute of Organic Synthesis and Photoreactivity (ISOF), National Research Council (CNR), Bologna, Italy.

* These authors contributed equally

[#] To whom correspondence should be addressed:
Dept. of Biosciences, Università degli Studi Milano, Via Celoria 26, 20133, Milan, Italy
Tel:+390250315002
Fax:+390250315044
giuseppina.caretti@unimi.it

Abstract

SMYD3 is a methylase previously linked to cancer cell invasion and migration. Here we show that SMYD3 favors TGFβ-induced epithelial-mesenchymal transition (EMT) in mammary epithelial cells, promoting mesenchymal and EMT-TFs expression.

SMYD3 directly interacts with SMAD3 but it is unnecessary for SMAD2/3 phosphorylation and nuclear translocation. Conversely, SMYD3 is indispensable for SMAD3 direct association to EMT genes regulatory regions. Accordingly, SMYD3 knockdown or its pharmacological blockade with the BCI121 inhibitor dramatically reduce TGFβ-induced SMAD3 association to the chromatin.

Remarkably, BCI121 treatment attenuates mesenchymal genes transcription in MDA-MB-231 cells and reduces their invasive ability *in vivo*, in a zebrafish xenograft model.

In addition, clinical dataset analysis revealed that higher SMYD3 levels are linked to a less favorable prognosis in claudin-low breast cancers and to a reduced metastasis free survival in breast cancer patients.

Overall, our data point at SMYD3 as a pivotal SMAD3 cofactor that promotes TGFβ-dependent mesenchymal gene expression and cell migration, in breast cancer.

Introduction

A crucial event occurring in the metastatic process is the epithelial-mesenchymal transition (EMT), which endows tumor cells with the ability to trans-differentiate from epithelial to mesenchymal-like cells and to acquire the capability to leave the primary tumor mass and disseminate at distant sites (1,2). EMT has been extensively linked to both metastatic progression of cancer (1,3) and acquisition of stem-cell properties (4,5), strengthening the hypothesis that reprogramming from epithelial to mesenchymal phenotype leads to acquired migration and self-renewal abilities, which foster the formation of secondary tumors at distant sites.

Specific extracellular signals activate the EMT and promote epithelial cells reprogramming towards a mesenchymal phenotype (6). During this process, epithelial features such as cell-cell adhesion, polarity and lack of motility are lost, and cells acquire mesenchymal characteristics, including motility and ability to invade. In breast cancer progression, EMT is triggered and maintained by various extracellular signals promoting either autocrine or paracrine stimuli. Among these signals, the transforming growth beta (TGFβ) signaling pathway plays a crucial role (1,3,7). Evolution to invasive and metastatic forms of breast cancer correlates with the activation of EMT (6,8,9) and with increased levels of TGF-β1 in plasma of breast cancer patients and at invasive front in human breast cancer tissue sections (10,11).

TGFβ signaling pathway activation results in SMAD2 and SMAD3 phosphorylation, their association with SMAD4, and translocation to the nucleus, where SMAD2/3 promote transcriptional activation of EMT-inducing transcription factors (EMT-TFs): Snail1, Slug, ZEB1/2, Twist1/2 and Sox4 (12,13). SMAD2/3/4 transcription factors orchestrate the epithelial cell reprogramming in concert with epigenetic factors that regulate epigenome plasticity and coordinate the expression of molecular signatures associated with the EMT program (14,15). Different DNA-binding partners confer SMADs target gene selectivity and influence the recruitment of transcriptional coactivators or co-repressors. SMADs can bind to acetyltransferases as p300, P/CAF, but also HDACs and chromatin repressors, which contribute to target genes regulation (8,(16).

Expression profiles analysis of human breast cancers allowed the identification at least five different molecular subtypes (luminal A, luminal B, Her2^{+}, basal and claudin-low) (17,18). Among these subtypes, claudin-low tumors express characteristic mesenchymal genes, and the EMT signature is particularly enriched in this set of tumors, which associate with poorer prognosis (19,20).

The methylase SMYD3 was reported to promote cancer cells proliferation and invasion, and to be over-expressed in several cancers, e.g. colorectal, pancreatic, liver and breast cancer cells, and was

also identified in gene signatures of metastatic pancreatic cancer cells (21-25). SMYD3 has been reported to directly regulate the metalloproteinase MMP-9 and cancer invasion (26-29). Recently, Sarris *et al.* (2016) described that SMYD3 is recruited at chromatin regulatory regions of proliferation and EMT genes, in mouse models of liver and colon cancers and they depict SMYD3 as a transcriptional "potentiator" for proliferation and EMT genes, during cancer progression (23).
Nevertheless, although SMYD3 oncogenic function is well established, the molecular mechanism through which it promotes tumor cells migration and invasion hasn't been fully described yet. To elucidate the underlying dynamics of SMYD3-mediated regulation of EMT, we employed breast cancer cells as a model system and investigated SMYD3 link with the TGFβ/SMADs signaling pathway. Here we report that SMYD3 is indispensable for SMAD3 mediated regulation of target genes, in TGFβ treated breast cancer cells. SMYD3 blockade with the BCI121 inhibitor reduced cell motility, both in cell cultures and in an *in vivo* model of zebrafish xenograft. Our study provides novel insight in TGFβ-induced transcriptional activation and it supports SMYD3 as a promising therapeutic target for cells that undergo EMT.

Results

SMYD3 is required for efficient TGFβ-induced EMT

Our analysis of the Human Cancer Genome Atlas (TCGA) database (30,31) revealed that SMYD3 transcripts are over-expressed in a variety of cancers, and amplification frequency measured by RNA transcript levels is especially apparent in different breast cancers cohorts (Supplementary Fig. S1A).
In addition to its role in cancer cells proliferation, SMYD3 was shown to regulate EMT genes (23). Since EMT is causatively linked to distant metastases for epithelial cancers including breast cancer, we tested SMYD3 function in TGFβ-induced EMT, in NMuMG mouse mammary epithelial cells. We knocked down SMYD3 levels by Sh-interference. As expected, following TGFβ treatment ShScramble NMuMG cells morphology acquired a mesenchymal phenotype (Fig. 1A, left panel), accompanied by the down-regulation of epithelial transcripts E-cadherin and Claudin 6 and upregulation of mesenchymal transcripts N-Cadherin and Fibronectin (Fig. 1B). Remarkably, ShSMYD3 cells maintained a more epithelial morphology (Fig. 1A, right panel) and the epithelial transcripts E-Cadherin and Claudin6 failed to decrease (Fig. 1C). Immunoblot analysis revealed that E-Cadherin and Occludin protein levels were maintained at higher levels following TGFβ stimulation, when compared to ShScramble cells (Fig. 1C). Moreover, ShSMYD3 cells failed to to increase Fibronectin1 and N-cadherin transcript and protein levels (Fig. 1B and C). Snail1, which plays a critical role in the transcriptional repression of epithelial genes and transition to a

mesenchymal phenotype during EMT (32), failed to be upregulated both at the transcriptional and protein level in ShSMYD3 cells (Fig. 1D). Similar results were obtained when we employed SMYD3-depleted NMuMG cells, generated using an unrelated retrovirus targeting SMYD3 (Supplementary Fig. S2A). Conversely, AKT and ERK phosphorylation levels, which were shown to be affected by SMYD3 depletion in previous reports (25,33,34), remained unchanged in our model system (Supplementary Fig. S2B).

F-actin and Paxillin immunostaining revealed that SMYD3 depletion did not affect the epithelial morphology of NMuMG cells in the absence of TGFβ. Following TGFβ treatment, instead, expression of the epithelial marker ZO-1 was retained at the cell membrane of SMYD3-depleted cells, and the signal for the mesenchymal marker N-cadherin was barely detectable in ShSMYD3 cells (Fig. 1E). TGFβ induced stress fibers and focal adhesion formation, revealed by F-actin and paxillin staining respectively, were almost absent in ShSMYD3 cells and revealed reduced remodeling of cytoskeleton, when compared to ShScramble TGFβ-treated cells (Fig. 1E).

Since cells undergoing TGFβ-induced EMT acquire the ability to migrate, we evaluated SMYD3 impact on cell migration through a wound-healing assay. We observed that ShSMYD3 cells migrated significantly less than control shScramble cells (Fig. 1F).

Together, these results demonstrate that SMYD3 is required for efficient TGFβ-induced EMT in murine mammary epithelial cells.

Similar results were obtained in immortalized benign MCF10A human mammary epithelial cells, in which the cuboidal cobblestone epithelial cells switch to a fibroblastoid morphology following TGFβ□treatment (Fig. 2A). SMYD3 depletion by RNA interference led to a persistency of the epithelial morphology, upon TGFβ treatment (Fig. 2A, right panel); this lack of morphological change was accompanied by a delay in mesenchymal markers expression, such as Snail1, N-cadherin and Vimentin (Fig. 2B). Concurrently, protein levels of the epithelial marker Occludin were maintained at higher level in SMYD3-depleted cells, when compared to control scramble interfered cells. In TGFβ-treated SMYD3-knockdown cells, N-cadherin, Vimentin, Fibronectin1 and MMP9 transcripts were lower than in control cells, suggesting that SMYD3 positively regulates transcriptional regulation of a set of mesenchymal genes (Fig. 2C). Snail1 and other transcription factors (e.g. Sox4, Snail2, Zeb1,2, Twist1,2) induce EMT through repression of epithelial genes and are defined as EMT transcription factors (EMT-TFs) (15). When SMYD3-deficient and control cells were exposed to TGFβ for 6 hours we observed that mRNA levels of a subset of EMT-TFs (Snail1, Snail2, Sox4, Twist1) failed to be upregulated in SMYD3-knockdown cells (Fig. 2D). Overall, these data suggest that SMYD3 plays a role in regulating EMT-TFs and mesenchymal genes transcription and modulates TGF-β induced EMT in human and mouse mammary cells.

SMYD3 pharmacological blockade prevents EMT, independently of SMYD3 methylation activity

SMYD3 function was recently challenged by the small inhibitor BCI121 (34). In a parallel set of experiments, SMYD3 function was hindered by pharmacological blockade, by treating both NMuMG and MCF10A cells with increasing concentration of the SMYD3 inhibitor BCI121(34).

Both cell lines were initially treated with different doses of BCI121 for 3 hours and subsequently co-treated with TGFβ for additional 48 hours. At the 10μM BCI121 dose, TGFβ–treated NMuMG cells failed to upregulate mesenchymal transcripts such as Fibronectin and to down-regulate mRNA levels of the epithelial genes E-cadherin and Claudin6 (Fig. 3A). Likewise, SMYD3 pharmacological blockade in MCF10A cells resulted in a modest increase in the mesenchymal transcripts N-cadherin and MMP9, following TGFβ treatment, starting from the 5μM BCI121 dose (Fig. 3B).

As previously reported for other cell lines, 72 hour BCI121 treatment led to decreased cell viability(34), both in NMuMG and MCF10A cells. With the exception of the 50μM dose in NMuMG cells, BCI121 did not significantly affect cell growth at 48 hour treatment, suggesting that BCI121 effects on EMT were largely independent on its impact on cell proliferation at earlier stages (Supplemental Fig. S3A and S3B).

By interacting with the lysine-binding pocket, BCI121 exerts a dual function on SMYD3, both inhibiting its catalytic activity and preventing its recruitment to the chromatin (34). We therefore tested whether SMYD3 methylase activity was indispensable for EMT regulation. We evaluated the ability of SMYD3 catalytically inactive mutants to promote EMT, when compared to wild type SMYD3 (SMYD3-WT). Both SMYD3-WT and a catalytically defective mutant (SMYD3-ΔEEL) promoted upregulation of the mesenchymal genes N-cadherin and Vimentin, and equally reduced E-cadherin and Occludin protein levels, as well as membrane-associated ZO-1 (Supplementary Fig. S4A and Fig. S4B). These results are in agreement with previously reported data (Sarris *et al.*, 2016) and they suggest that SMYD3-mediated EMT genes regulation is independent on SMYD3 methylation activity.

SMYD3 directly interacts with SMAD3

Transcriptional regulation of TGFβ–modulated genes is largely controlled by SMAD2/3 transcription factors during EMT (32). Therefore, we first checked whether SMYD3 interacted with SMAD3 in HEK293T cells, transfected with Flag-SMYD3 and HA-SMAD3 plasmids.

Immunoprecipitation experiments revealed that over-expressed SMYD3 was able to associate to

SMAD3, in the absence of TGFβ stimulation (Fig. 4A). Next, we determined whether SMYD3 was able to interact with endogenous SMAD3 and performed immunoprecipitation experiments using NMuMG extracts. The experiments with endogenous proteins confirmed the interaction, and indicated that the association is independent of TGFβ treatment in NMuMG cells and it is not affected by the presence of the TGFβ inhibitor SB43152 in growth medium (Fig. 4B). To define the protein domains involved in the interaction, we performed immunoprecipitation experiments in HEK293T cells over-expressing SMAD3 deletion mutants with full-length SMYD3 and found that the MH2 domain is required for SMYD3 association (Fig. 4C and Supplementary Fig. S5A). Co-immunoprecipitation experiments with full-length HA-SMAD3 and Flag-SMYD3 deletion mutants revealed that mutant SMYD3 encompassing aa 219-428 and aa 111-428 associate with SMAD3 as the full-length protein. Conversely SMYD3 proteins lacking the C-terminal region (SMYD3_1-219 and SMYD3_1-380) weakly interacted with SMAD3 (Fig. 4D and Supplementary Fig. S5B). To investigate whether SMYD3 c-terminal region was relevant in the modulation of TGFβ induced genes, we transfected MCF10A cells either with full-length SMYD3 or SMYD3_1-380 plasmids. Following TGFβ treatment, Occludin levels were reduced and mesenchymal proteins Snail1, Snail2 and N-cadherin were increased in SMYD3 full-length over-expressing cells, compared to the empty vector transfected cells. SMYD3_1-380 over-expression instead didn't promote mesenchymal genes expression or Occludin down regulation (Fig.4E). Additionally, we performed immnuoprecipitation experiments with recombinant GST-SMAD3 and SMYD3 proteins and we were able to confirm a direct association between the two proteins *in vitro* (Fig. 4F). Overall, these data suggest that SMYD3 directly interact with SMAD3 through its C-terminal domain, and this region is crucial for EMT genes modulation.

SMYD3 is indispensable for SMAD3 association to the chromatin

To deeper delve in the molecular mechanism underlying SMYD3 regulation at mesenchymal genes, we tested SMYD3 ability to associate to a subset of genes, in TGFβ treated MCF10A cells. We found that SMYD3 is recruited to Snai1, Sox4, MMP9 and Vimentin regulatory regions in MCF10A cells. The BCI121 inhibitor competes with histones for SMYD3 binding (34) and as expected BCI121 treatment prevented SMYD3 occupancy at these regulatory regions. To provide further mechanistic insight, we assayed SMAD2/3 recruitment at mesenchymal target genes, in BCI121-treated cells. Chromatin immunoprecipitation (ChIP) assays revealed that SMAD2/3 engagement at mesenchymal genes was significantly reduced by SMYD3 pharmacological blockade (Fig. 5A). However, BCI121 treatment did not prevent SMAD2/3 and SMYD3 co-immunoprecipitation in MCF10A TGFβ-treated cells (Supplementary Fig. S6). Comparable

results were obtained when SMAD2/3 and SMYD3 recruitment was evaluated at regulatory regions of MMP9, Snail1, Sox4, Vimentin in ShSMYD3 MCF10A cells treated with TGFβ, and compared to ShScramble control cells (Fig. 5B).

Remarkably, both immunoblot and immunofluorescence experiments revealed that TGFβ dependent SMAD2/3 phosphorylation was unaffected by SMYD3 knockdown (Fig. 5C and 5D). Likewise, SMAD2/3 nuclear translocation was not influenced by SMYD3 reduction (Fig. 5E), suggesting that SMYD3 depletion didn't impair SMAD2/3 nuclear localization.

ChIP assays in NMuMG cells, confirmed a reduction in SMAD2/3 and SMYD3 occupancy at the Snail1 promoter, in ShSMYD3 TGFβ-treated cells compared to ShScramble cells. Additionally, histone marks analysis at Snail1 regulatory regions, revealed that SMYD3 depletion was accompanied by a local decrease in histone H3K9ac and H3K4me3. Conversely, repressive histone marks H3K27me3 and H3K9me3 occupancy was comparable in ShSMYD3 and ShScramble cells. Concurrently, RNA Polymerase II recruitment was reduced at mesenchymal genes in ShSMYD3 NMuMG cells (Fig. 5F) further supporting Snail1 reduced transcription. Taken together, these findings suggest that SMYD3 directly takes part in the transcriptional regulation of a subset of mesenchymal genes, and that its engagement correlates with SMAD2/3 and RNAPolII recruitment and regulatory regions. Overall, SMYD3 appears to promote EMT-TFs and mesenchymal genes transcription, independently of its methylase activity and by hindering SMAD2/3 chromatin association.

SMYD3 inhibition decreases MDA-MB-231 cells migration ability and mesenchymal phenotype

We next asked whether SMYD3 depletion could attenuate the mesenchymal phenotype in a metastatic mesenchymal-like breast cancer cell line. We therefore treated human MDA-MB-231 cells with BCI121 and asked whether SMYD3 blockade could reduce the MDA-MB-231 mesenchymal gene transcription and promote epithelial markers expression. We treated MDA-MB-231 cells with two doses of BCI121 (10 and 50 μM) for 48 hours and found that SMYD3 inhibition increased the expression of epithelial markers Occludin and Claudin6 and reduced the expression of mesenchymal markers Snail2, Fibronectin1 and N-cadherin (Fig. 6A). In agreement with reduced mRNA, Occludin protein levels increased, whereas Snail1, Snail2, Vimentin and Fibronectin levels decreased following BCI121 administration (Fig. 6B). Functionally, BCI121 treatment decreased MDA-MB-231 ability to migrate *in vitro,* as shown by in wound-healing assays (Fig. 6C).

MDA-MB-231 cells have been employed to evaluate *in vivo* migration, in zebrafish xenografts. MDA-MB-231 cells were injected in the sub-peridermal space, close to the sub-intestinal vessels

(SIV) plexus, in zebrafish embryos at 48 hours post fertilization (hpf). After cells injection, embryos were treated either with 50μM BCI121 or DMSO for 48 hours. We then evaluated the number of single cells that invaded the zebrafish body, either in the tail, heart or head, at 2 days post-injection (dpi). We found that the number of invading MDA-MB-231 cells was significantly reduced in zebrafish embryos treated with BCI121, when compared to DMSO treated embryos (Fig. 6D and E).

Overall, these results indicate that BCI121 treatment decreases MDA-MB-231 cells mesenchymal signature and it reduces their invasion ability, both *in vitro* and *in vivo*.

SMYD3 is highly expressed in human breast cancers and correlates with increased metastasis

To assess whether our findings could be related to the pathogenesis of human breast cancer and metastasis formation, we studied SMYD3 expression in one of the publicly available data sets in the TCGA repository (Supplementary Fig. S1). In the Metabric cohort (35,36) SMYD3 was altered in 24.1% of the 2501 cases (Supplementary Fig. S1).

Analysis of RNA-seq expression profiles of breast invasive carcinomas obtained from the TCGA database and analyzed with the Firebrowse portal showed that SMYD3 expression was significantly enhanced in breast cancer tumors (1093 tumor samples) when compared to healthy tissue (112 normal samples) (Fig. 7A). Remarkably, we couldn't point out any specific SMYD3 enrichment in relation to any molecular subtype (Supplementary Fig. S7A). We next focused on EMT signature genes and interrogated expression data from 1394 tumors obtained from the Metabric dataset to investigate whether SMYD3 transcript levels was associated with EMT genes expression. We stratified patients for SMYD3 levels (Supplementary Fig. S7B), and found that Snail2, Zeb1, Zeb2, Twist1, Twist2, N-cadherin, Fibronectin and Vimentin were significantly upregulated in High SMYD3 tumors compared to the low SMYD3 group (Supplementary Fig. S7C). We observed a positive correlation between SMYD3 and mesenchymal genes expression (Snail2, Zeb2, Fibronectin, N-Cadherin and Vimentin) in the Metabric dataset (Supplementary Fig. S7D). To confirm our findings, we extended our analysis to the NKI295 breast cancer dataset (37), and confirmed a positive correlation between SMYD3 and Snail2, N-cadherin and Fibronectin expression (Supplementary Fig. S7E).

Core-EMT signature genes are highly expressed in claudin-low breast cancers, when compared to other molecular breast cancers subtypes (20). We therefore queried for a correlation between SMYD3 expression and a subset of core-EMT genes in claudin-low tumors, within the Metabric dataset. Our analysis confirmed a positive correlation between SMYD3 expression and EMT-TFs (Snail2, ZEB2, ZEB1, Twist1), as well as mesenchymal gene transcripts (Fibronectin, N-Cadherin)

and the degree of correlation was more robust in claudin-low subtype than in the whole Metabric cohort (Fig. 7B). Transcript levels for Snail2, ZEB1, ZEB2, Twist1, Twist2, Vimentin and Fibronectin were significantly higher in claudin-low tumors over-expressing SMYD3 compared to the *SMYD3 low* tumors (Fig. 7C), and the differences between the two groups were more prominent than in the analysis on the entire dataset (Supplementary Fig. S7C).

SMYD3 higher expression levels did not correlate with a poorer prognosis in the Metabric cohort as a whole; however, *SMYD3 high* patients had a significantly lower survival probability in the claudin-low tumor subgroup, when compared to the claudin-low *SMYD3 low* patients' group (Fig. 7C).

To further investigate SMYD3 prognostic value, we performed a parallel analysis in different clinical microarray dataset from 1354 breast tumors (38). We first focused our analysis to an advanced tumor stage, and found that distant metastasis free survival was significantly higher for patients with *low SMYD3* grade 3 tumors, independently of lymph node status and systemic treatment they underwent (n=458, p value 0.022) (Fig. 7D). Likewise, *SMYD3 high* levels in grade 1, lymph node negative tumors, in patients that didn't undergo chemotherapy or endocrine therapy, displayed a significantly lower probability of distant metastasis free survival (n=94, p=0.031) (Supplementary Fig. S7F).

Overall, clinical dataset analysis suggests that SMYD3 levels positively correlate with EMT-TFs and mesenchymal gene expression in breast tumors. Higher SMYD3 levels in breast tumors associate with reduced metastasis free survival, and with a worst overall survival probability in claudin-low tumors.

Discussion

Results presented in this paper highlight SMYD3 nuclear functions in regulating EMT-transcription factors and mesenchymal gene expression, in the context of TGFβ stimulation. We describe a direct TGFβ-independent interaction between SMYD3 and SMAD3 and a functional interplay between SMYD3 and SMAD3 in EMT genes transcriptional regulation. We propose that SMYD3 is indispensable for TGFβ–induced SMAD3 recruitment at chromatin regulatory regions of mesenchymal targets and EMT-TFs.

Several lines of evidence previously linked SMYD3 to migration and invasion, in different cell lines (12-16). Our findings confirm this indication in TGFβ-induced EMT and provide a molecular mechanism that supports recently reported *in vivo* findings (23), highlighting a crucial role for SMYD3 in promoting EMT genes activation. Sarris *et al.* (2016) showed by ChIP-Seq assay that

SMYD3 occupies several regulatory regions in liver tumors and revealed that SMYD3 association to a subset of proliferation and EMT loci promotes their transcriptional activation (23). We propose that SMYD3 cooperates with SMAD3 at EMT genes driving mesenchymal and EMT-TFs transcription.

Following TGFβ-mediated stimulation, SMAD2 and SMAD3 activate, translocate to the nucleus and bind to different transcription factors, co-activators or co-repressors thus achieving higher affinity and selectivity for target genes regulatory regions (16). In different cell types SMAD3 occupancy is directed by cell-type-specific master transcription factors, which are responsible for orchestrating the cell-type-specific effects of TGFβ signaling (39). Therefore, cell-type-specific effects of TGFβ signaling are in large part determined by SMAD2/3 interactions with transcription factors and cofactors and they can vary in a cell/tissue-specific manner. In this scenario, we propose that SMYD3 takes part in the orchestrated SMAD3 recruitment at a subset of mesenchymal and EMT-TFs genes, following TGFβ stimulation. SMAD3 interacts with SMYD3 through the MH2 domain, which is responsible for association to various transcription factors and cofactors, as well as for the interaction between SMAD proteins and their transcriptional activation (40). SMYD3 interaction domain, instead, encompasses SMYD3 c-terminal region, which was previously reported to be involved in HSP90 (41) and H3K4me3 interaction *in vitro* (23).

Our ChIP data show that SMYD3 depletion doesn't affect distribution of repressive histone marks H3K9me3 and H3K27me3 at the Snail1 promoter. These findings are in agreement with the concept that SMYD3 is recruited at open-chromatin regions (23) and suggest that SMYD3/SMAD2/3 independent mechanisms are involved in H3K9me3 and H3K27me3 removal (42) and. Conversely, SMYD3 promotes H3K4me3 and H3K9Ac marks deposition at the SNAI1 promoter, as well as SMAD3 and RNAPolII recruitment and transcriptional activation. These data suggest that SMYD3 association to the SNAI1 gene is subsequent to repressive histone marks removal and that SMYD3 association potentiates transcription through SMAD2/3 and RNAPolII recruitment, at EMT loci. SMAD3 was previously reported to mediate COMPASS recruitment and precede H3K4 trimethylation at the SNAI1 promoter, in TGFβ-treated DU145 cells (43). Additionally, SMAD3 interacts with p300/CBP and TGFβ treatment promotes acetylation of SMAD2/3 target genes (44,45), further suggesting that H3 SMAD2/3 association to the chromatin is required to promote K4me3 and H3K9Ac. Additionally, SMYD3 associates with the RNAPolII complex (46,47), and this interaction may favor the transcriptional machinery recruitment at EMT genes.

Remarkably, SMYD3-mediated regulation of EMT is independent of its catalytic activity, suggesting that SMYD3 co-activator functions are not mediated by SMYD3 methylation activity at the mesenchymal/EMT-TFs genes. Likewise, SMYD3 knockdown didn't lead to ERK and AKT

phosphorylation modulation (25,33,34), indicating that SMYD3 cytosolic and methylation-dependent mechanisms are finely tuned and operate differently in different context.

In the patho-physiological context, patients' data on survival and metastasis free survival support the experimentally described SMYD3 role in promoting EMT. SMYD3 mRNA levels are upregulated in breast cancers compared to normal breast tissues, but SMYD3 transcripts appear to be significantly upregulated in every breast cancer subtypes. SMYD3 also promotes transcriptional activation of proliferation genes and its augmented levels may confer an advantage for tumor growth in the different breast cancer subtypes. Additionally, SMYD3 was shown regulate HER2 dimerization through direct methylation (48), suggesting that in different breast cancer subtypes SMYD3 may provide selective growth advantages.

Claudin-$_{low}$ breast cancer cells were shown to be more likely than other breast cancer subtypes to promote the breast tumor initiating cancer cell (BTICs) population in response to TGFβ(49). We propose that in TGFβ sensitive cells, such as claudin-low cancer cell, SMYD3 may promote EMT genes activation through its SMAD3 interplay.

Recent evidences show that EMT-associated markers are highly expressed in cancer stem cells and promote their self-renewal (4). Remarkably, BTICs are enriched within the claudin-low molecular subtype of breast cancer, where TGFβ/SMAD signaling promote BTICs expansion (49). Further investigations will reveal whether SMYD3 likewise promotes BTICs stemness in claudin-low cells.

Overall, our data further support SMYD3 as a promising pharmacological target for anti-cancer therapy.

Materials and Methods

Cell cultures and reagents

NMuMG, MCF10A and MDAMB231 cell lines were from were purchased from the American Type Culture Collection, grown in DMEM supplemented with 10% FBS, 100 U/ml penicillin and 100 mg/ml streptomycin. MCF10A cell line was grown in DMEM F-12 supplemented with 5% HS, 20 ng/ml EGF, 0.5 mg/ml hydrocortisone, 100 ng/ml cholera toxin, 10 μg/ml insulin, 100 IU/ml penicillin and 100 mg/ml streptomycin. Cells were grown in a humidified incubator with 5% CO_2 at 37°C. Cellules were starved for 12 hrs in serum free growth medium, fresh medium containing FBS was changed and treated with 5ng/ml TGFβ.

TGFβ was reconstituted in 10 mM citric acid (pH 3.0) to a final concentration of 0.1 mg/ml, then further diluted in PBS containing 0.1% BSA to a final concentration of 0.01 mg/ml and stored at -

20°C. BCI-121 (Innovamol, Italy) was dissolved in dimethyl sulfoxide (DMSO) and stored at -20°C. Unless differently described, BCI121 was used at a final concentration of 10 μM.

All cell lines were periodically tested for mycoplasma with MycoAlert Mycoplasma detection kit (Euroclone, Italy). All cell lines were fed every 48/72 hours, for a maximum number of 30 passages.

Cell proliferation and wound healing assays

Cells growth was determined with a Bürker chamber, counting cells after 48 h or 72 h of BCI-121 or DMSO exposure.

Wound healing assays was performed using Dish Culture–Inserts (Ibidi). 50000 cells per well were plated and dish were incubated at 37°C and 5% CO_2. After 24 hours the Culture-Insert was removed and medium was changed. Pictures were taken at time 0, 24 or 48 h, to evaluate migration ability. The "wounded" area was manually selected (blue lines) and quantified wih ImageJ.

RNA Isolation and Real Time PCR (qRT-PCR)

Total RNA was extracted using TRI reagent (Sigma) according to the manufacturer's instruction. cDNA was synthesized from RNA (1μg) using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystem). qRT-PCR was performed in triplicate using SYBR Green PCR Master Mix (Bio-Rad) or 2X Xtra Master Mix (GeneSpin) on a CFX Connect Real-Time PCR Detection System (Bio-Rad). The qRT-PCR reactions were normalized using GAPDH as housekeeping gene and relative quantification was done using the ddCT method.

The oligonucleotides used in qRT-PCR on NMuMG cells were:

CDH1	Forward: 5'-GTCTACCAAAGTGACGCTGAAGT-3' Reverse: 5'-TCTCGTTTCTGTCTTCTGAGACC-3'
CDH2	Forward: 5'-AATAGACCCCGTGAATGGGCAGATC-3' Reverse: 5'-AGGCGGGATTCCATTGTCAGAAG-3'
CLD6	Forward: 5'-CTCATCTCTGGCATCATCTTTG-3' Reverse: 5'-AGGGGTTGTAGAAGTCCTGGAT-3'
FN1	Forward: 5'-GCTGGATGATGGTGGACTGT-3' Reverse: 5'-CTCGGTTGTCCTTCTTGCTC-3'
SNAI1	Forward: 5'-ACTGGTGAGAAGCCATTCTCCT-3' Reverse: 5'-CTGGTATCTCTTCACATCCGAGT-3'

The oligonucleotides used in qRT-PCR on MCF10 and MDAMB-231 cells were:

CDH2	Forward: 5'-AGGGGACCTTTTCCTCAAGA-3' Reverse: 5'-CTACTGCATGTGCCCTCAAA-3'
CLD6	Forward: 5'-GATGCAGTGCAAGGTGTACG-3' Reverse: 5'-GCCTTGGAATCCTTCTCCTC-3'
FN1	Forward: 5'-TGAAAGACCAGCAGAGGCATAAG-3' Reverse: 5'-CTCATCTCCAACGGCATAATGG-3'
MMP9	Forward: 5'-TGACAGCGACAAGAAGTGGG-3' Reverse: 5'-TTCAGGGCGAGGACCATAGA-3'

OCLN	Forward: 5'-GAAGCCAAACCTCTGTGAGC-3' Reverse: 5'-GAAGACATCGTCTGGGGTGT-3'
SMYD3	Forward: 5'-TGAATGTGACTGTTTCCGTTGC-3' Reverse: 5'-ATTGCTGCTTATGATCGCCTGG-3'
SNAI1	Forward: 5'-ATCGGAAGCCTAACTACAGCGAGC-3' Reverse: 5'-CAGAGTCCCAGATGAGCATTGG-3'
SNAI2	Forward: 5'-CTTTTTCTTGCCCTCACTGC-3' Reverse: 5'-GCTTCGGAGTGAAGAAATGC-3'
SOX4	Forward: 5'-CCAGCAAGAAGGCGAGTTAG-3' Reverse: 5'-CGGAGCCTTCTGTCTTCATC-3'
VIM	Forward: 5'-CCCTCACCTGTGAAGTGGAT-3' Reverse: 5'-GCTTCAACGGCAAAGTTCTC-3'

The oligonucleotides used in ChIP qRT-PCR were:

hMMP9	Forward: 5'-GGTGGTGTAAGCCCTTTCTCAT-3' Reverse: 5'-ATGGTGAGGGCAGAGGTGTCT-3'
hSNAI1	Forward: 5'-GAGTGGTTCTTCTGCGCTACTG-3' Reverse: 5'-GCTGTAGTTAGGCTTCCGATTGG-3'
SOX4	Forward: 5'-ATTGTTTTGTGGCTTTCTCTTCCC-3' Reverse: 5'-TCAGCATTGGAATAAAGAATCAGCC-3'
hVIM	Forward: 5'-CTCTTCTCCGGGAGCCAGTC-3' Reverse: 5'-CGGTAGGAGGACGAGGACAC-3'
mSNAI1	Forward: 5'- CGGAGTTGACTACCGACCTT -3' Reverse: 5'- GACCTAGGTAGTCGGGGTCAC-3'

RNA interference

siRNAs targeting human SMYD3 (5'-GAUUGAAGAUUUGAUUCUA-3') (Eurofins Genomics). siRNA (150nM) were transfected with Lipofectamine 2000 according to the manufacturer's instructions. Scrambled siRNA (5'-GCGUUGCUCGGAUCAGAAA-3') was used as negative control. Lentiviral infections in MCF10A cells were performed with shRNA targeting SMYD3 (Mission ShRNA, Sigma) and shRNA used for retroviral infections in NMuMG cells was previously described (47). Retroviral and lentiviral infections were performed as in (50).

Cell extracts and immunoblot analysis

Cells collected and homogenized in RIPA lysis buffer (50 mM Tris-HCl pH 7.4, 0,5% sodium deoxycholate, 0,1% SDS, 250 mM NaCl, and 1% NP40) supplemented with protease and phosphatase inhibitors (Sigma). Homogenates were solubilised in Laemmli Sample buffer and 30 μg proteins were separated on 8%, 10% or 12% SDS-PAGE, and transferred to nitrocellulose membranes using Trans-Blot Turbo Transfer System (BioRad). Membranes were blocked with Phosphate-buffered saline (PBS) containing 0.1% Tween and 5% nonfat dry milk and incubated with primary antibody, overnight. E-caderin (Cell Signaling, 24E10), Occludin (Santa Cruz, sc-5562), Fibronectin (Santa Cruz, sc-9068), SMYD3 (GeneTex, GTX121945), SMYD3 (Cell

Signaling, D2Q4V), Snail1 (Cell Signaling, C15D3), Actin (Santa Cruz, sc-8432), p-SMAD2/3 (Santa Cruz, sc-11763), Vimentin (Santa Cruz, sc-6260), histone H3 (Santa Cruz, sc-10809), Flag (Sigma, F3165), HA (Santa Cruz, sc-805), Snail2 (Santa Cruz, sc-166476), Lamin A/C (Cell Signaling, 4C11), c-myc (Santa Cruz, sc-789), AKT1/2/3 (Santa Cruz, sc-8312), p-AKT1/2/3 (Santa Cruz, sc-7985-R), ERK (Santa Cruz, sc-94), p-ERK (Santa Cruz, sc-sc-7383), GAPDH (Santa Cruz, sc-32233). Nuclear and cytoplasmic extracts were performed with Nucleat Extract Kit (Active Motif).

Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation was performed as previously described (Proserpio *et al*., 2013). MCF10A and NMuMG cells were treated with 5 ng/ml TFGβ for 6 hrs before formaldehyde crosslinking. 2μg of SMYD3 (NovusBio, NBP1-79393) and SMAD2/3 (Cell Signaling, D7G7) Cell Signaling) antibodies were used in MCF10A cells, SMYD3 (GeneTex, GTX121945), SMAD3 (Abcam, ab28379) RNAPolII (GeneTex), H3K4me3 (Millipore, 05-745), H3K27me3 (Active Motif, 39536), H3K9me3 (Active Motif, 39161), H3K9Ac (Millipore, 06-942).

Recombinant proteins

GST-SMYD3 (kind gift form dr. N. Reynoird, University of Grenoble, France) was expressed in *E. coli* TUNER cells, previously transfected with the Chaperone pKJE7 plasmid (Takara), by induction with 0.1 mM IPTG O/N at 4°C. Recombinant GST-SMYD3 was purified using Pierce Glutathione Superflow Agarose resin (Thermo Scientific) and eluted with 125 mM Tris/HCl, 150 mM NaCl, 10 mM reduced glutathione, 0.2% Triton X100, pH 8.0. GST tag was removed with PreScission Protease (GE Healthcare), according to the manufacturer's instructions.

Immunoprecipitation

Whole cell extracts were obtained from HEK293T transfected cells or NMuMG cells collected and homogenized in IP lysis buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM $MgCl_2$, 2 mM EDTA, 10% glycerol, 0.1% NP40) supplemented with protease and phosphatase inhibitors (Sigma, Italy). Protein contents of all the samples were determined by the Bradford's method. 500 μg of whole extract were incubated with Myc (Millipore, 05-419), SMYD3 (Santa Cruz, sc-49519X), normal IgG (Santa Cruz, sc-2027, sc-2028) at 4°C for 2 hours and Protein G Agarose beads (Pierce, Italy) were added for 45 additional minutes. IP for Flag were performed with Flag-M2 affinity gel (Sigma, A2220). For IP with recombinant proteins, 200ng of GST-SMAD3 (Sigma, Italy) and SMYD3 were incubated with 2μg of anti-SMYD3 antibody (Santa Cruz, sc-49519X) or with goat IgG (Santa Cruz, sc-20208) in IP Buffer (50 mM Tris/HCl, 1 mM EDTA, 150 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP40, pH 8.0) and incubated at 4°C for 2 hours and Protein G Agarose beads (Pierce, Italy) were added for 45 additional minutes. Immuno-complexes were washed 3

times and immune complexes were resolved by SDS-PAGE and analyzed by Western blot with SMAD2/3 (Cell Signaling, D7G7), HA (Santa Cruz, sc-805), Flag (Santa Cruz, sc-807), SMYD3 (GeneTex, GTX121945).

Immunofluorescence

Cells seeded on coverslips were washed with PBS, fixed in methanol and blocked with 3% BSA in 1XPBS 0,02% Triton for 1 hour at room temperature. Samples were then probed overnight with Paxillin (Santa Cruz, 5574), ZO1 (Santa Cruz 10804), N-cadherin (Santa Cruz, 7939), F-actin Phalloidin (Santa Cruz, sc-363795), p-SMAD3 (Cell Signaling, C25A9). After three 5 minute washes with PBS, cells were incubated for 1 hour at room temperature with rabbit or mouse Alexa Fluor secondary antibodies. Cells were then washed three time for 5 minute with PBS, coverslips were mounted on slides with Fluoroshield with DAPI (Sigma) and examined with The samples wereexamined with a fluorescence microscope (Carl Zeiss). Pictures of staining were obtained using an AxioCam (Carl Zeiss Vision).

Zebrafish care and tumor xenograft

Tg(fli1a:EGFP)^{y1} transgenic embryos, collected by natural spawning, were staged and raised at 28°C in fish water (Instant Ocean, 0.1% methylene blue), according to National (Italian D.lgs 26/2014) and European laws (2010/63/EU and 86/609/EEC). Dechorionated embryos at 48 h postfertilization (hpf) were anesthetized with 0.04 mg/mL of tricaine (Sigma-Aldrich). Tumor cells, labeled with a red fluorescent dye for cell viability (CellTracker™ CM-DiI, Invitrogen) and resuspended in PBS, were implanted in the sub-peridermal space, close to the sub-intestinal vessels (SIV) plexus, of 48 hpf zebrafish embryos (100 cells for each embryo). After the cell implantation, embryos showing cells into the yolk sac and/or in the vasculature embryos were excluded from further analyses. Correctly grafted embryos were split in two experimental groups that were treated with 50 µM BCI121 and the vehicle alone (DMSO) respectively. BCI121 and DMSO were diluted directly into the fish water. Thus, treated and control embryos were incubated at 32°C. At 48 hours post injection (hpi), the presence of circulating grafted cells into the trunk/tail region was evaluated through a fluorescence stereomicroscope (Leica DM6000B equipped with LAS Leica imaging software). The spread of injected tumor cells throughout the embryo was quantified in both 50 µM BCI121treated and control embryos by the "Analyze Particle" plugin of Fiji software. Data from two independent experiments were pooled. Statistical significance was calculated with unpaired Student's t-test. * $p<0.05$.

Clinical datasets analysis

The website FireBrowse (http://firebrowse.org) was used to interrogate the TGCA database for the average SMYD3 expression in breast invasive carcinomas versus normal tissue.

The website cBioPortal (http://www.cbioportal.org) was used for meta-analysis (30,31). The Metabric dataset contains 2509 patients with breast cancer with Illumina v3 RNA microarray, clinical data and a follow-up period up to 350 months (Pereira 2016 nature; Curtis nature 2012). Only 1866 patients were included in the present study, because they displayed all the parameters considered. Patients were stratified based on SMYD3 z-score and quartiles were identified as SMYD3 High (q1: SMYD3 z-score ≥ 0.998. N=468) and SMYD3 Low (q3 SMYD3 z-score ≤ -0.327. N=467). SMYD3 expression levels were correlated with EMT genes transcript levels by Pearson correlation analysis. Estimated probability of overall survival and distant metastasis free survival was calculated with the Kaplan-Meier method.

NKI295 dataset was downloaded from http://ccb.nki.nl/data/ (51). The subgroup of patients with BRCA1 mutations was excluded from the analysis.

Statistical analysis

The statistical significance of the results was analyzed using the unpaired Student's t-tail test, 1 way Anova followed by Tukey post-test, 2 way Anova, followed by Bonferroni post-test using GraphPad (Prism6). Survival curves were analysed by Log-rank (Mantel-Cox) Test and Gehan-Breslow-Wilcoxon Test. Correlations were analyzed by Pearson r test. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$ were considered statistically significant.

Acknowledgments:

We thanks Dr. C. Simone (University of Bari, Italy) for helpful discussions throughout the development of the project and Dr. G. Pavesi in our Department for support with the bioinformatic analysis.

We thank Dr. Nicolas Reynoird (University of Grenoble, France) for GST-SMYD3 plasmid and Dr. M. Sandri (University of Padua, Italy) for SMAD3 expression vectors.

Authors' Contributions:

Conception and design: C. Fenizia, C. Bottino, G. Caretti

Development of methodology: A. Del Rio, F. Cotelli, G. Vitale, C. Fenizia, C. Bottino, G. Caretti

Acquisition of data: C. Fenizia, C. Bottino, R.Fittipaldi, P.Floris, G. Gaudenzi, S. Corbetta, S. Carra, G. Caretti

Analysis and interpretation of data: C. Fenizia, C.Bottino, R. Fittipaldi, P. Floris, G. Gaundenzi, G.Caretti

Writing, review and/or revision of the manuscript: C. Fenizia, C. Bottino, G. Caretti

Administrative, technical, or material support: F. Cotelli, A. Del Rio, G.Caretti

Study supervision: G.Caretti

Conflict of Interest: Dr. Del Rio is the Managing Director of Innovamol (Italy). All other authors declare no potential conflicts of interest.

Grant Support: This work was supported by grant from Italian Association for Cancer Research (AIRC_Trideo) to GC and in part by grants from Telethon and Worldwide Cancer Research to GC. R. Fittipaldi was supported by a FIRC fellowship.

References

1. Nieto MA, Huang RY, Jackson RA, Thiery JP. EMT: 2016. Cell **2016**;166:21-45
2. Imamura T, Hikita A, Inoue Y. The roles of TGF-β signaling in carcinogenesis and breast cancer metastasis. Breast Cancer **2012**;19:118-24
3. Thiery JP. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer **2002**;2:442-54
4. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell **2008**;133:704-15
5. Scheel C, Weinberg RA. Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. Semin Cancer Biol **2012**;22:396-403
6. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest **2009**;119:1420-8
7. Barcellos-Hoff MH, Akhurst RJ. Transforming growth factor-beta in breast cancer: too much, too late. Breast Cancer Res **2009**;11:202
8. López-Novoa JM, Nieto MA. Inflammation and EMT: an alliance towards organ fibrosis and cancer progression. EMBO Mol Med **2009**;1:303-14
9. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. Cell **2009**;139:871-90
10. Dalal BI, Keown PA, Greenberg AH. Immunocytochemical localization of secreted transforming growth factor-beta 1 to the advancing edges of primary tumors and to lymph node metastases of human mammary carcinoma. Am J Pathol **1993**;143:381-9
11. Chod J, Zavadova E, Halaska MJ, Strnad P, Fucikova T, Rob L. Preoperative transforming growth factor-beta 1 (TGF-beta 1) plasma levels in operable breast cancer patients. Eur J Gynaecol Oncol **2008**;29:613-6
12. Budi EH, Duan D, Derynck R. Transforming Growth Factor-β Receptors and Smads: Regulatory Complexity and Functional Versatility. Trends Cell Biol **2017**;27:658-72
13. Macias MJ, Martin-Malpartida P, Massagué J. Structural determinants of Smad function in TGF-β signaling. Trends Biochem Sci **2015**;40:296-308
14. Nieto MA. Epithelial plasticity: a common theme in embryonic and cancer cells. Science **2013**;342:1234850
15. Tam WL, Weinberg RA. The epigenetics of epithelial-mesenchymal plasticity in cancer. Nat Med **2013**;19:1438-49
16. Massagué J, Seoane J, Wotton D. Smad transcription factors. Genes Dev **2005**;19:2783-810
17. Network CGA. Comprehensive molecular portraits of human breast tumours. Nature **2012**;490:61-70
18. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A **2001**;98:10869-74
19. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res **2010**;12:R68
20. Sabatier R, Finetti P, Guille A, Adelaide J, Chaffanet M, Viens P, *et al.* Claudin-low breast cancers: clinical, pathological, molecular and prognostic characterization. Mol Cancer **2014**;13:228

21. Nakamura T, Fidler IJ, Coombes KR. Gene expression profile of metastatic human pancreatic cancer cells depends on the organ microenvironment. Cancer Res **2007**;67:139-48
22. Hamamoto R, Silva FP, Tsuge M, Nishidate T, Katagiri T, Nakamura Y, *et al.* Enhanced SMYD3 expression is essential for the growth of breast cancer cells. Cancer science **2006**;97:113-8
23. Sarris ME, Moulos P, Haroniti A, Giakountis A, Talianidis I. Smyd3 Is a Transcriptional Potentiator of Multiple Cancer-Promoting Genes and Required for Liver and Colon Cancer Development. Cancer Cell **2016**;29:354-66
24. Giakountis A, Moulos P, Sarris ME, Hatzis P, Talianidis I. Smyd3-associated regulatory pathways in cancer. Semin Cancer Biol **2017**;42:70-80
25. Mazur PK, Reynoird N, Khatri P, Jansen PW, Wilkinson AW, Liu S, *et al.* SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. Nature **2014**;510:283-7
26. Cock-Rada AM, Medjkane S, Janski N, Yousfi N, Perichon M, Chaussepied M, *et al.* SMYD3 promotes cancer invasion by epigenetic upregulation of the metalloproteinase MMP-9. Cancer Res **2012**;72:810-20
27. Wang SZ, Luo XG, Shen J, Zou JN, Lu YH, Xi T. Knockdown of SMYD3 by RNA interference inhibits cervical carcinoma cell growth and invasion in vitro. BMB reports **2008**;41:294-9
28. Zou JN, Wang SZ, Yang JS, Luo XG, Xie JH, Xi T. Knockdown of SMYD3 by RNA interference down-regulates c-Met expression and inhibits cells migration and invasion induced by HGF. Cancer letters **2009**;280:78-85
29. Zeng B, Li Z, Chen R, Guo N, Zhou J, Zhou Q, *et al.* Epigenetic regulation of miR-124 by hepatitis C virus core protein promotes migration and invasion of intrahepatic cholangiocarcinoma cells by targeting SMYD3. FEBS Lett **2012**;586:3271-8
30. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov **2012**;2:401-4
31. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal **2013**;6:pl1
32. Vincent T, Neve EP, Johnson JR, Kukalev A, Rojo F, Albanell J, *et al.* A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF-beta mediated epithelial-mesenchymal transition. Nat Cell Biol **2009**;11:943-50
33. Yoshioka Y, Suzuki T, Matsuo Y, Nakakido M, Tsurita G, Simone C, *et al.* SMYD3-mediated lysine methylation in the PH domain is critical for activation of AKT1. Oncotarget **2016**;7:75023-37
34. Peserico A, Germani A, Sanese P, Barbosa AJ, di Virgilio V, Fittipaldi R, *et al.* A SMYD3 Small-Molecule Inhibitor Impairing Cancer Cell Growth. J Cell Physiol **2015**;230:2447-60
35. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature **2012**;486:346-52
36. Pereira B, Chin SF, Rueda OM, Vollan HK, Provenzano E, Bardwell HA, *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nat Commun **2016**;7:11479
37. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, *et al.* Gene expression profiling predicts clinical outcome of breast cancer. Nature **2002**;415:530-6
38. Győrffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat **2010**;123:725-31

39. Mullen AC, Orlando DA, Newman JJ, Lovén J, Kumar RM, Bilodeau S, *et al.* Master transcription factors determine cell-type-specific responses to TGF-β signaling. Cell **2011**;147:565-76
40. Massagué J, Blain SW, Lo RS. TGFbeta signaling in growth control, cancer, and heritable disorders. Cell **2000**;103:295-309
41. Brown MA, Foreman K, Harriss J, Das C, Zhu L, Edwards M, *et al.* C-terminal domain of SMYD3 serves as a unique HSP90-regulated motif in oncogenesis. Oncotarget **2015**;6:4005-19
42. Ramadoss S, Chen X, Wang CY. Histone demethylase KDM6B promotes epithelial-mesenchymal transition. J Biol Chem **2012**;287:44508-17
43. Li D, Sun H, Sun WJ, Bao HB, Si SH, Fan JL, *et al.* Role of RbBP5 and H3K4me3 in the vicinity of Snail transcription start site during epithelial-mesenchymal transition in prostate cancer cell. Oncotarget **2016**;7:65553-67
44. Yuan H, Reddy MA, Sun G, Lanting L, Wang M, Kato M, *et al.* Involvement of p300/CBP and epigenetic histone acetylation in TGF-β1-mediated gene transcription in mesangial cells. Am J Physiol Renal Physiol **2013**;304:F601-13
45. Mishra VK, Subramaniam M, Kari V, Pitel KS, Baumgart SJ, Naylor RM, *et al.* Krüppel-like Transcription Factor KLF10 Suppresses TGFβ-Induced Epithelial-to-Mesenchymal Transition via a Negative Feedback Mechanism. Cancer Res **2017**;77:2387-400
46. Hamamoto R, Furukawa Y, Morita M, Iimura Y, Silva FP, Li M, *et al.* SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. Nat Cell Biol **2004**;6:731-40
47. Proserpio V, Fittipaldi R, Ryall JG, Sartorelli V, Caretti G. The methyltransferase SMYD3 mediates the recruitment of transcriptional cofactors at the myostatin and c-Met genes and regulates skeletal muscle atrophy. Genes Dev **2013**;27:1299-312
48. Yoshioka Y, Suzuki T, Matsuo Y, Tsurita G, Watanabe T, Dohmae N, *et al.* Protein lysine methyltransferase SMYD3 is involved in tumorigenesis through regulation of HER2 homodimerization. Cancer Med **2017**;6:1665-72
49. Bruna A, Greenwood W, Le Quesne J, Teschendorff A, Miranda-Saavedra D, Rueda OM, *et al.* TGFβ induces the formation of tumour-initiating cells in claudinlow breast cancer. Nat Commun **2012**;3:1055
50. Caretti G, Di Padova M, Micales B, Lyons GE, Sartorelli V. The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. Genes Dev **2004**;18:2627-38
51. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, *et al.* A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med **2002**;347:1999-2009

Figure legends

Figure 1. SMYD3 depletion impairs TGFβ-induced EMT in murine NMuMG cells. **A,** Brightfield images of shSMYD3 or shScramble NMuMG cells treated with 10 ng/ml TGFβ or vehicle for 24 hours. Scale bar=10μm. **B,** Quantitative real-time PCR analysis of epithelial (E-Cadherin, Claudin6) or mesenchymal (N-Cadherin and Fibronectin) genes. shSMYD3 or shScramble NMuMG cells were treated with 10 ng/ml TGFβ for 0, 24, 48 or 96 h. GAPDH was used as housekeeping gene.. Statistical analysis was performed with 1 way ANOVA, followed by post-hoc Tukey test. $n \geq 3$, * $p \leq 0.05$, ** $p \leq 0.01$. **C,** Western blot analysis of epithelial (E-Cadherin and Occludin) or mesenchymal (N-Cadherin and Fibronectin) genes in shSMYD3 or shScramble NMuMG cells treated with 10 ng/ml TGFβ for 0, 48 or 72 h. GAPDH serves as a loading control. **D,** Upper panel: Snail1 transcript levels were quantified by qRT-PCR, after 0, 6, 12 h of TGFβ treatment. Lower panel: Western blot to evaluate Snail1 protein levels at 0, 6 12, 24 h following TGFβ treatment. β-actin was used as loading control. **E,** immunofluorescence microscopy of mesenchymal and epithelial markers in shSMYD3 or shScramble NMuMG cells. Left panels show untreated cells and the right panel shows cells treated with 10 ng/ml TGFβ for 24hrs. Epithelial (ZO-1) and mesenchymal (N-Cadherin) were analyzed. Paxillin was used to detect focal adhesion plaques, and phalloidin to visualize the actin cytoskeleton. Scale bar=20μm. **F,** Wound healing assay on shSMYD3 or shScramble NMuMG cells treated with 10 ng/ml TGFβ for 0 and 24 hours. Migration percentage is reported in the right panel. Scale bar=10μm.

Figure 2. SMYD3 knockdown hinders TGFβ-induced EMT in human MCF10A cells. **A,** Brightfield images of siScramble or SiSMYD3 MCF10A cells treated with 5 ng/ml TGFβ or vehicle, for 24 hours. Scale bar=10μm. **B,** siSMYD3 or siScramble MCF10A cells were treated with 5 ng/ml TGFβ for 72 hours, and protein extracts were analyzed by Western blot analysis for epithelial (Occludin) or mesenchymal markers (N-Cadherin, Vimentin and Fibronectin). GAPDH serves as a loading control. **C,** MCF10A cells were treated with 10 ng/ml TGFβ for 72 hours and mRNA levels of epithelial (Occludin) and mesenchymal genes (MMP9, Vimentin, N-Cadherin and Fibronectin) were determined by quantitative real-time PCR analysis, relative to GAPDH. **D**, MCF10A cells were treated with 5 ng/ml TGFβ for 6 hours and mRNA levels of EMT-TFs was determined by quantitative real-time PCR analysis, relative to GAPDH. Statistical analysis was performed with 1 way ANOVA, followed by post-hoc Tukey test. $n \geq 3$, * $p \leq 0.05$, ** $p \leq 0.01$.

Figure 3. SMYD3 blockade with the BCI121 inhibitor hampers TGFβ-induced EMT. **A,** Epithelial (E-Cadherin, Claudin-6) or mesenchymal (N-Cadherin and Fibronectin) transcripts were quantified by quantitative real-time PCR analysis, relative to GAPDH. NMuMG cells were treated with 10 ng/ml TGFβ in the presence of 0, 1, 5 10 and 50 μM BCI121, for 48 hours. Statistical analysis was performed with 1 way ANOVA, followed by post-hoc Tukey test. $n \geq 3$, * $p \leq 0.05$, ** $p \leq 0.0$. **B,** N-Cadherin and MMP9 transcript levels were quantified by quantitative real-time PCR normalized to GAPDH in MCF10A cells treated with 5 ng/ml TGFβ in the presence of 0, 1, 5 10 and 50 μM BCI121 for 48 hours. Statistical analysis was performed with 1 way ANOVA, followed by post-hoc Tukey test. $n \geq 3$, * $p \leq 0.05$, ** $p \leq 0.01$. **C, D,** Wound healing assay on NMuMG and MCF10A cells treated with 10 ng/ml TGFβ and 10μM BCI121 for 0 and 24 hours. Migration percentage is reported in the right panel. Scale bar=10μm.

Figure 4. SMYD3 directly interacts with SMAD3. **A,** Western-blot analysis of over-expressed Flag-SMYD3 and HA-SMAD3 following co-immunoprecipitation with anti-Flag antibody. **B,** Co-immunoprecipitation analysis of endogenous SMYD3 and SMAD3 in NMuMG cells. Whole cell

extract were prepared from NMuMG cells after starvation, from cells treated with the TGFβ-signaling inhibitor SB431542, or starved and treated with 10 ng/ml TGFβ for 30 and 60 minutes. Ip was performed with antibodies raised against SMYD3. **C,** Co-immunoprecipitation with anti-Myc antibody in whole cell extracts of HEK293T cells over-expressing Myc-SMYD3 and Flag-SMAD3 WT or Flag-tagged SMAD3 domains MH1, L-MH2 and MH2. Lower panel represent a schematic representation of SMAD3 mutants. **D,** co-immunoprecipitation with anti-Flag antibody of HEK293T cells over-expressing HA-SMAD3 and Flag-SMYD3 WT or Flag-tagged SMYD3 mutants 1-219, 219-428, 111-428, 1-380. Lower panel represent a schematic representation of SMYD3 mutant. **E,** Full-length SMYD3 and SMYD3_1-380 were over-expressed in MCF10A cells and treated with 5ng/ml TGFβ for 48 h. Whole cell extracts were analyzed by immunoblot for epithelial and mesenchymal markers. **F**, *in vitro* co-immunoprecipitation of GST-SMAD3 and SMYD3 was performed with SMYD3 antibodies. Immunoblot was performed with anti-SMAD2/3 and anti-SMYD3 antibodies.

Figure 5. SMYD3 assists SMAD3 recruitment to the chromatin. **A,** SMAD2/3 and SMYD3 association to regulatory regions of EMT markers was analyzed by ChIP qPCR, in ShScramble and ShSMYD3 MCF10A cells treated with 5ng/ml TGFβ for 24 hours. Data are expressed as fold enrichment over the IgG control. Statistical analysis was performed with Student's t test. n=4, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **B**, SMAD2/3 and SMYD3 association to EMT markers was analyzed by ChIP qPCR, in MCF10A cells co-treated with 5ng/ml TGFβ and 10 μM BCI121 for 24 hours. Data are expressed as fold enrichment over the IgG control. Statistical analysis was performed with unpaired Student's t test. n=4, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. **C**, immunofluorescence microscopy of shSMYD3 or shScramble NMuMG cells treated with 10 ng/ml TGFβ for 2 h. Depicted in the left two panels the staining for the phosphorylated SMAD3 (P-SMAD3), in the middle panels DAPI and in the right panels the merge of the two stainings. Scale bar=10μm. **D,** Western-blot analysis of P-SMAD3 and SMAD2/3 levels in shSMYD3 or shScramble NMuMG cells, treated with 10 ng/ml TGFβ for for 30 min. **E,** Western-blot analysis of cytoplasmic (C) and nuclear (N) P-SMAD3 in shSMYD3 or shScramble NMuMG cells treated with 10 ng/ml TGFβ for 30 min. GAPDH was used as cytoplasmic marker, histone H3 as nuclear marker. **F,** ChIP assay was performed to evaluate SMYD3, SMAD2/3, RNAPolII, H3K9me3, H3K9Ac, H3K27me3 occupancy at the Snail1 promoter, in NMuMG cells treated with 10ng/ml TGFβ for 6 h. Data are normalized on IgG. Statistical analysis was performed with unpaired Student's t test. n=3, * $p \leq 0.05$, ** $p \leq 0.01$.

Figure 6. SMYD3 blockade by BCI121 tempers MDA-MB-231 mesenchymal phenotype and migration ability. **A,** MDA-MB-231 cells were treated with increasing doses of BCI121 (0, 10 or 50 μM) for 48 h. Epithelial (Occludin, Claudin6) or mesenchymal (Snail1, Snail2, Vimentin and Fibronectin) marker mRNA levels were measured by qRT-PCR. GAPDH was used as housekeeping gene. Statistical analysis was performed with 1 way ANOVA, followed by post-hoc Tukey test $n \geq 3$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. **B,** MDA-MB-231 cells were treated with increasing doses of BCI-121 (0, 10 or 50 μM) for 72 h and whole cell extracts were assayed for epithelial (Occludin) or mesenchymal (Snail1, Snail2, Vimentin and Fibronectin) protein levels by immunoblot. GAPDH was used as a loading control. **C,** Wound healing assay on MDA-MB-231 cells treated with 10μM BCI121 for 0 and 24 hours. Migration percentage is reported in the right panel. Scale bar=10μm. **D,** *in vivo* migration assay in zebrafish xenografts. $1x10^2$ MDA-MB-231 cells were injected in the subepidermal pocket of zebrafish embryos at 48 hpf. Embryos were treated with 50 μM BCI-121 or DMSO, for 48 h. **E,** Cells migrating in the tail, heart and head were counted in the two groups, 48 hpi. Scale bar=200μm. Statistical analysis was performed with unpaired Student's t-test. * $p \leq 0.05$

Figure 7. Prognostic value of SMYD3 in breast cancer. **A,** Box plots showing SMYD3 expression levels in normal and tumor breast tissue**.** Analysis was performed on 1098 breast cancer tissues and 112 normal tissues on TCGA data, downloaded from through Firebrowse. *** $p<0.0001$, unpaired Student's t test. **B**, Correlation between SMYD3 and the mesenchymal markers SNAIL2, ZEB1, ZEB2, TWIST1, Fibronectin and N-Cadherin expression, calculated on claudin-low tumors (Metabric), by Pearson correlation analysis. Pearsons' coefficient tests were performed to assess statistical significance. **C**, Claudin-low tumors (Metabric) were stratified for SMYD3 expression (high vs low) and Snail2, ZEB1, ZEB2, TWIST1, TWIST2, Vimentin, Fibronectin mRNA levels were compared in SMYD3-low and SMYD3-high tumors. Statistical significance was calculated with unpaired Student's t test. *$p \leq 0.05$, **$p \leq 0.001$, ***$p \leq 0.0001$. **C, D,** Kaplan-Meier plots of survival probability of claudin-low patients stratified by SMYD3 expression in the Metabric dataset. *p* value calculated by log rank test. **E,** Kaplan-Meier plots of distant metastasis-free survival of patients with grade 3 tumors, stratified by SMYD3 expression levels. *p* value was calculated by log rank test.

Supplementary Figures

Figure S1. SMYD3 over-expression is present in multiple cancers.
TCGA database was queried for SMYD3 under- or over-expressing SMYD3 tumors. TCGA data were analyzed using the open-source software cBioPortal for Cancer Genomics (http://www.cbioportal.org/).

Figure S2. SMYD3 depletion doesn't impact ERK and AKT activation.
A, Immunoblot of Epithelial (E-Cadherin, Occludin) and mesenchymal (N-Cadherin) protein levels in shSMYD3 or shScramble NMuMG cells, treated with 10 ng/ml TGFβ for 72 h. GAPDH was used as loading control.
B, Immunoblot of p-ERK, ERK, p-AKT and AKT in NMuMG cells treated with 10 ng/ml TGFβ for 72 h.

Figure S3. BCI121 impact on NMuMG and MCF10A cells growth.
A, B, NMuMG and MCF10A cells growth was evaluated in the presence of TGFβ (right panel) or vehicle (left panel), in cells treated with increasing doses of SMYD3 inhibitor BCI121.

Figure S4. SMYD3-mediated regulation of EMT genes is independent from its methylation activity.
A, NMuMG cells were transduced with pBabe-SMYD3, pBabe-SMYD3_ΔEEL and pBabe-EMPTY retroviruses and clones were treated with 10ng/ml TGFβ for 48 h. Epithelial (E-Cadherin and Occludin) and mesenchymal (N-Cadherin and Vimentin) protein levels were analyzed by immunoblot. **B**, Immunofluorescence microscopy of pBabe-SMYD3, pBabe-SMYD3_ΔEEL and pBabe-EMPTY NMuMG cells treated with TGFβ for 48 h, fixed with paraformaldehyde and stained with antibodies against the epithelial marker ZO1.

Figure S5. Inputs of Cells extracts used for co-IP experiments.
A, Input HEK293T cell extracts used for co-IP in fig 4C are shown. **B,** Input HEK293T cell extracts used for co-IP in fig 4D are shown.

Figure S6. BCI121 treatment doesn't impact SMAD3/SMYD3 interaction.
A, MCF10A cells were treated with BCI121 or DMSO for 24 h and whole cell extracts were used in immunoprecipitation experiments with antibodies raised against SMYD3. Immunoblot was performed with SMAD2/3 and SMYD3 antibodies.

Figure S7. Breast cancer clinical dataset analysis.

A, box plots represent SMYD3 mean z score across breast tumor subtypes. Data were obtained from the Metabric dataset. **B**, Patients were stratified in two group with "High" (q1) and "Low" (q3) SMYD3 expression levels. **C**, Tumors from the whole Metabric dataset were stratified for SMYD3 expression (high vs low) and Snail2, ZEB1, ZEB2, TWIST1, TWIST2, Vimentin, N-Cadherin, Fibronectin mRNA levels were compared in SMYD3-low and SMYD3-high tumors. Statistical significance was calculated with unpaired Student's t test, $*p \leq 0.05$, $**p \leq 0.001$, $***p \leq 0.0001$. **D**, Correlation between SMYD3 and the mesenchymal markers SNAIL2, ZEB2, ZEB1, TWIST1, Fibronectin and N-Cadherin expression, calculated on breast cancer tumors from the Metabric dataset, by Pearson correlation analysis. **E**, Correlation between SMYD3 expression and the mesenchymal markers SNAIL2, Fibronectin and N-Cadherin, calculated in the NKI295 breast cancers dataset by Pearson correlation analysis. **F**, Kaplan-Meier plots of distant metastasis-free survival of patients with grade 1 breast tumors, stratified by SMYD3 expression levels. *p* value was calculated by log rank test.

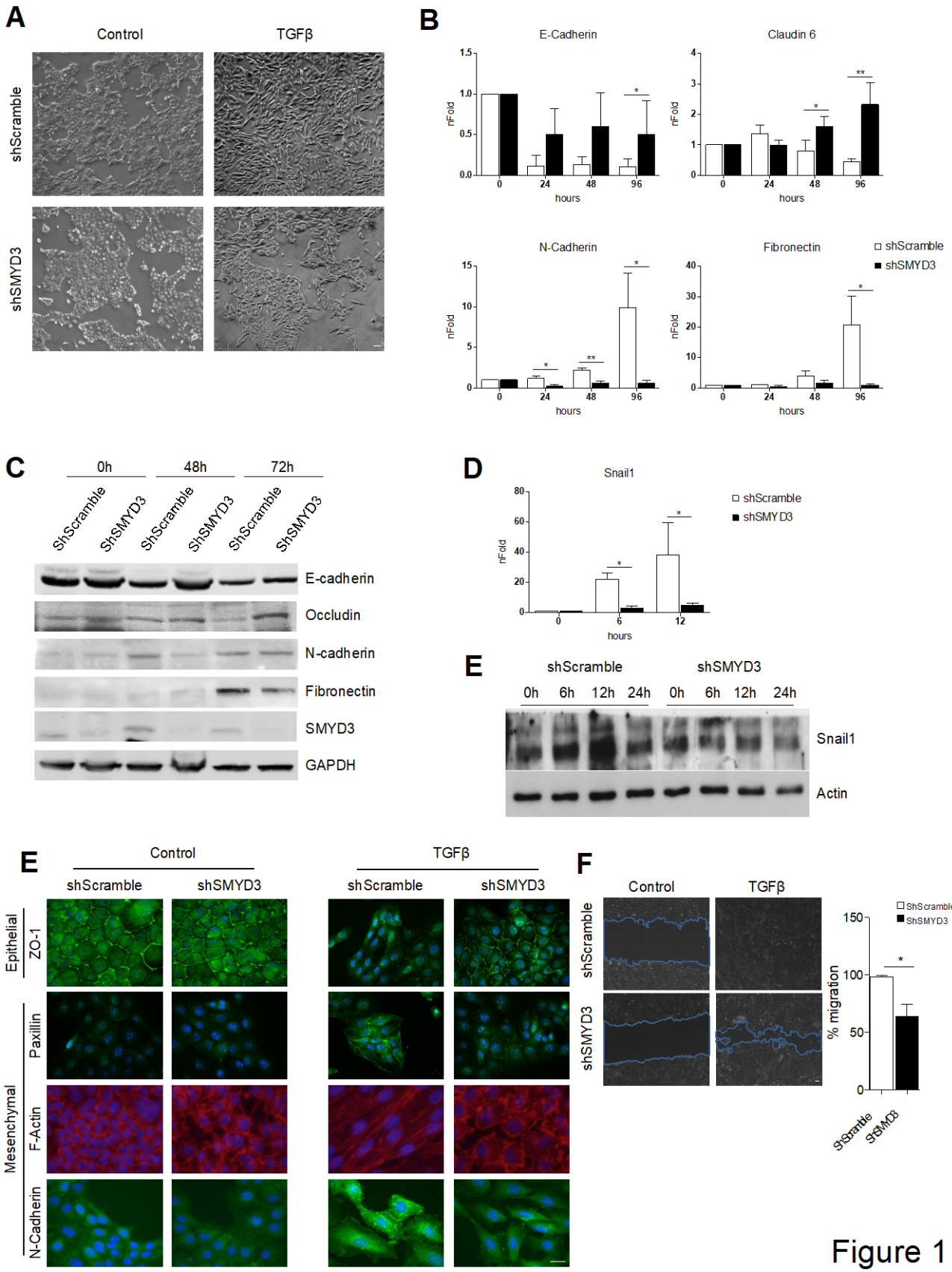
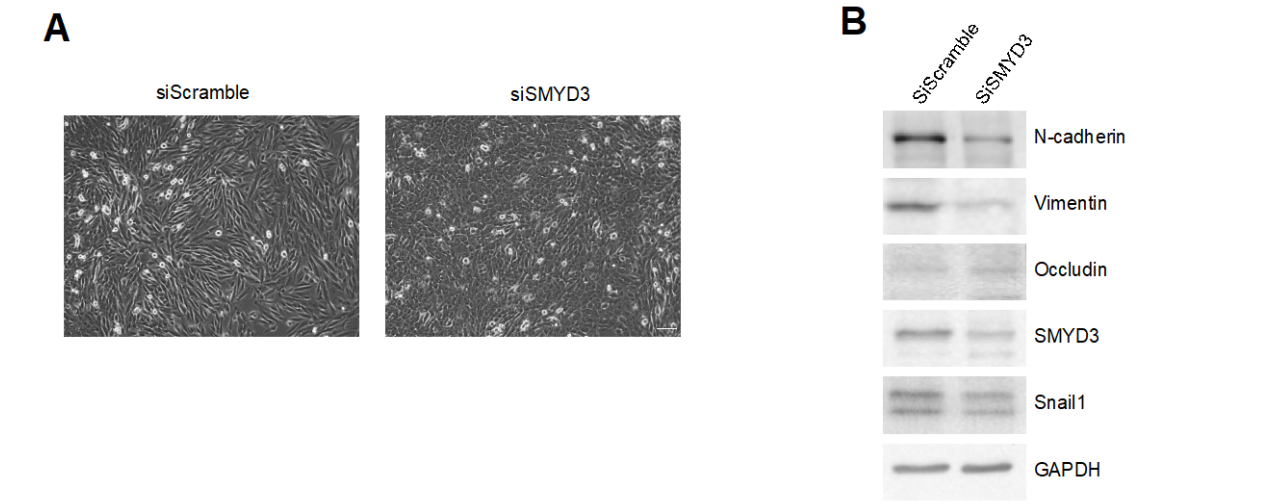
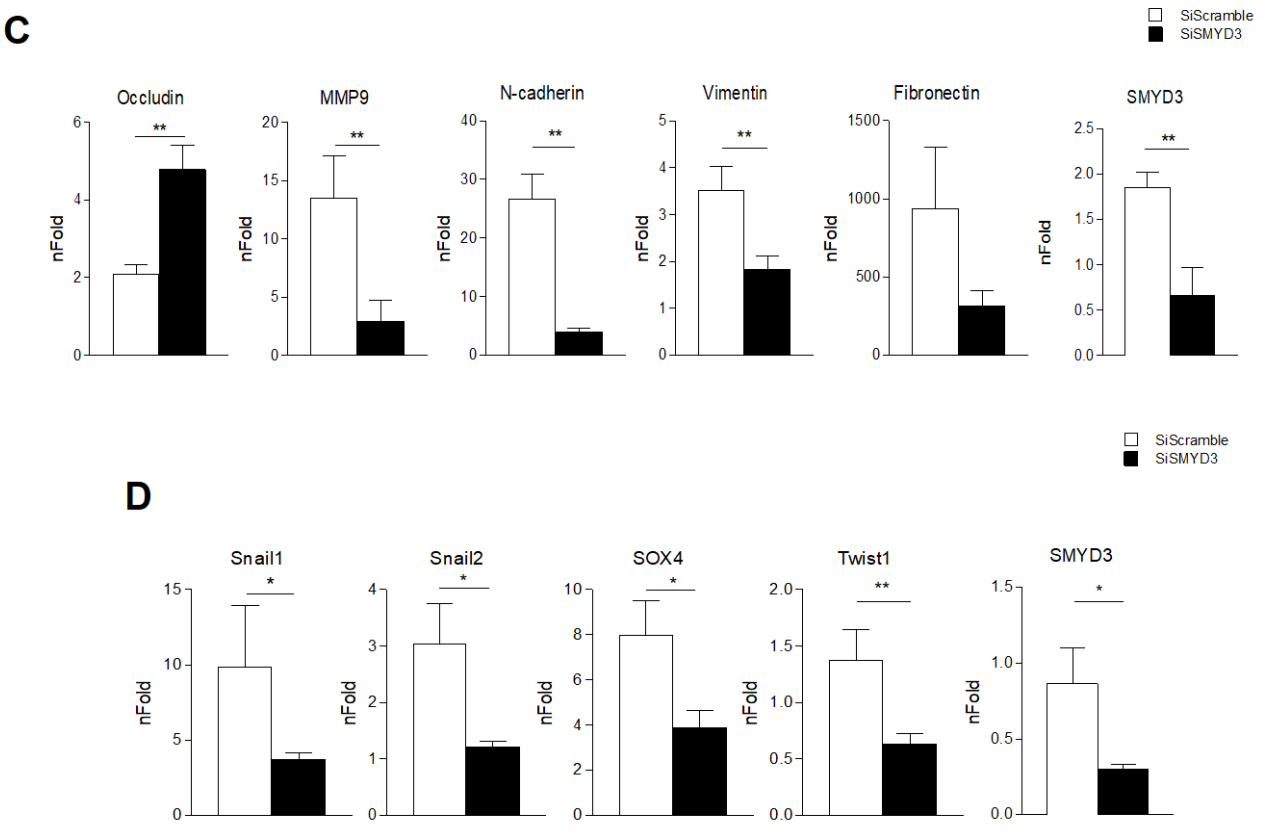

A
Control
TGFβ
shScramble
shSMYD3
B
E-Cadherin
Claudin 6
N-Cadherin
Fibronectin
nFold
hours
shScramble
shSMYD3
C
0h
48h
72h
ShScramble
ShSMYD3
E-cadherin
Occludin
N-cadherin
Fibronectin
SMYD3
GAPDH
D
Snail1
shScramble
shSMYD3
E
shScramble
shSMYD3
0h 6h 12h 24h 0h 6h 12h 24h
Snail1
Actin
E
Control
TGFβ
shScramble
shSMYD3
Epithelial
ZO-1
Mesenchymal
Paxillin
F-Actin
N-Cadherin
F
Control
TGFβ
shScramble
shSMYD3
% migration
ShScramble
ShSMD3

Figure 1


A
siScramble
siSMYD3
B
SiScramble
SiSMYD3
N-cadherin
Vimentin
Occludin
SMYD3
Snail1
GAPDH


C
SiScramble
SiSMYD3
Occludin
MMP9
N-cadherin
Vimentin
Fibronectin
SMYD3
nFold
**

D
SiScramble
SiSMYD3
Snail1
Snail2
SOX4
Twist1
SMYD3
nFold
*
**

Figure 2

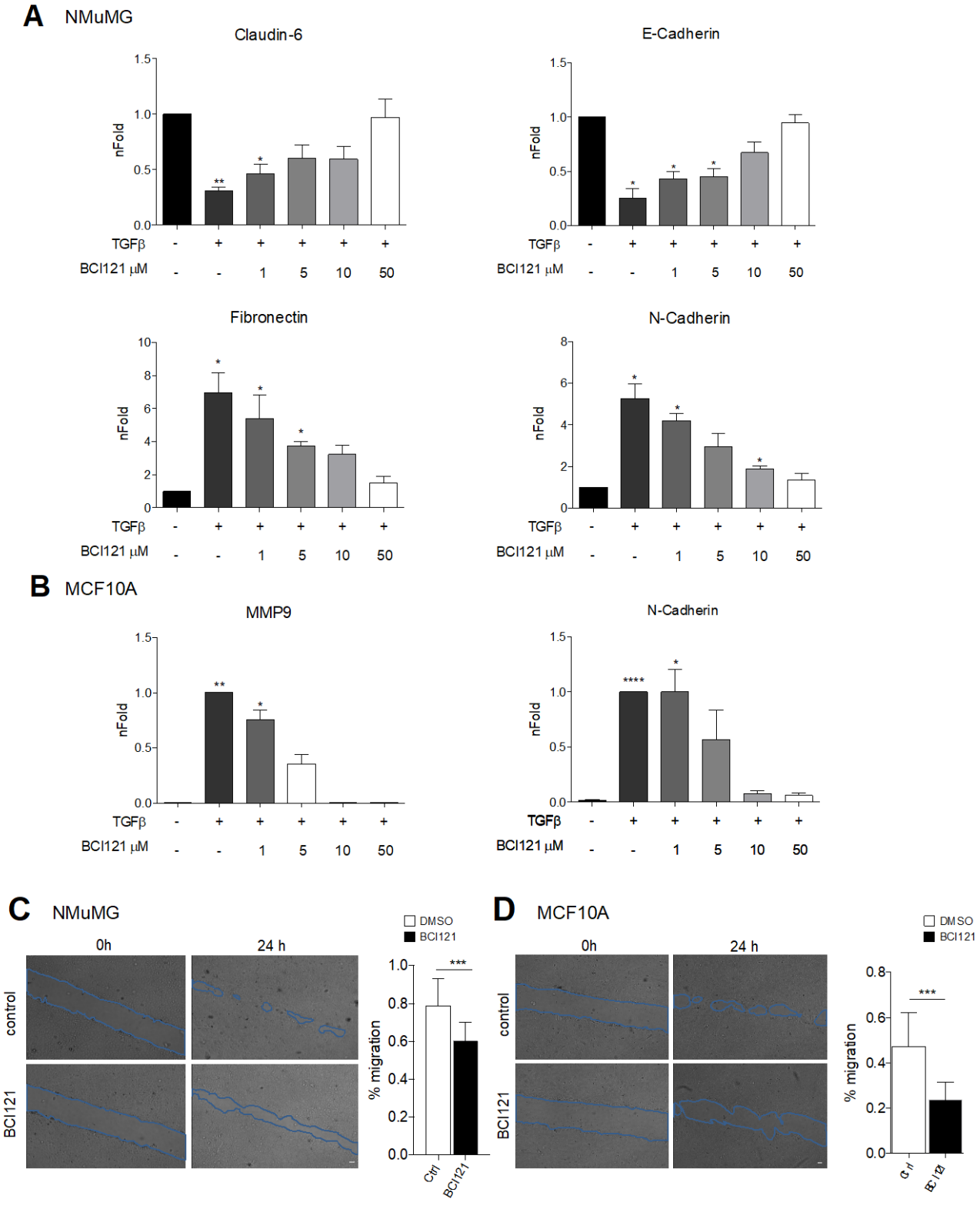

A
NMuMG
Claudin-6
E-Cadherin
Fibronectin
N-Cadherin
nFold
TGFβ
BCI121 μM
B
MCF10A
MMP9
N-Cadherin
C
NMuMG
0h
24 h
control
BCI121
DMSO
BCI121
% migration
Ctrl
BCI121
D
MCF10A
0h
24 h
control
BCI121
DMSO
BCI121
% migration

Figure 3

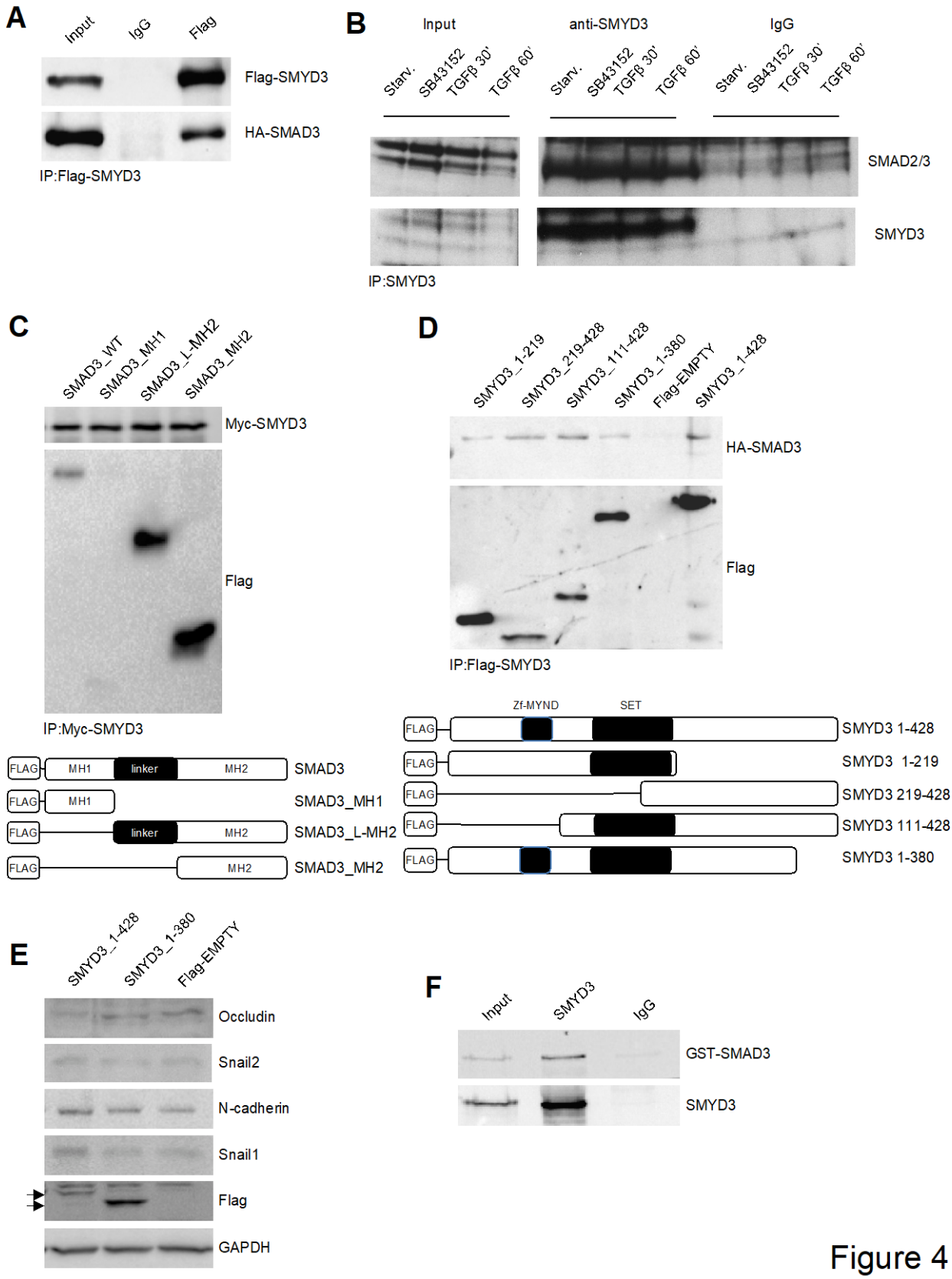

A
Input
IgG
Flag
Flag-SMYD3
HA-SMAD3
IP:Flag-SMYD3
B
Input
anti-SMYD3
IgG
Starv.
SB43152
TGFβ 30'
TGFβ 60'
SMAD2/3
SMYD3
IP:SMYD3
C
SMAD3_WT
SMAD3_MH1
SMAD3_L-MH2
SMAD3_MH2
Myc-SMYD3
Flag
IP:Myc-SMYD3
FLAG
MH1
linker
MH2
SMAD3
SMAD3_MH1
SMAD3_L-MH2
SMAD3_MH2
D
SMYD3_1-219
SMYD3_219-428
SMYD3_111-428
SMYD3_1-380
Flag-EMPTY
SMYD3_1-428
HA-SMAD3
Flag
IP:Flag-SMYD3
Zf-MYND
SET
SMYD3 1-428
SMYD3 1-219
SMYD3 219-428
SMYD3 111-428
SMYD3 1-380
E
SMYD3_1-428
SMYD3_1-380
Flag-EMPTY
Occludin
Snail2
N-cadherin
Snail1
Flag
GAPDH
F
Input
SMYD3
IgG
GST-SMAD3
SMYD3

Figure 4

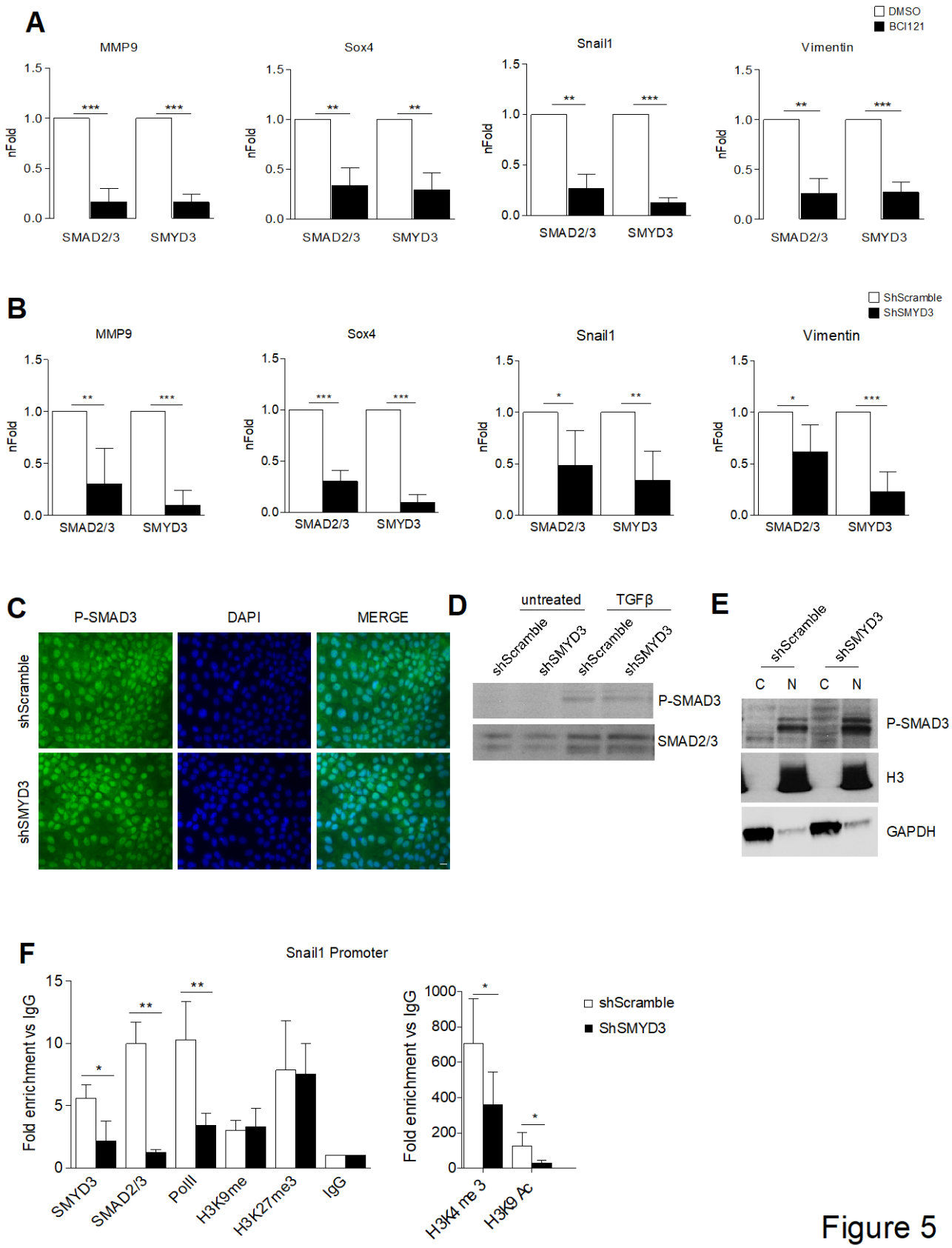

A
DMSO
BCI121
MMP9
Sox4
Snail1
Vimentin
nFold
SMAD2/3
SMYD3
B
ShScramble
ShSMYD3
C
P-SMAD3
DAPI
MERGE
shScramble
shSMYD3
D
untreated
TGFβ
shScramble
shSMYD3
P-SMAD3
SMAD2/3
E
C
N
H3
GAPDH
F
Snail1 Promoter
Fold enrichment vs IgG
SMYD3
SMAD2/3
PolII
H3K9me
H3K27me3
IgG
H3K4 me 3
H3K9 Ac

Figure 5

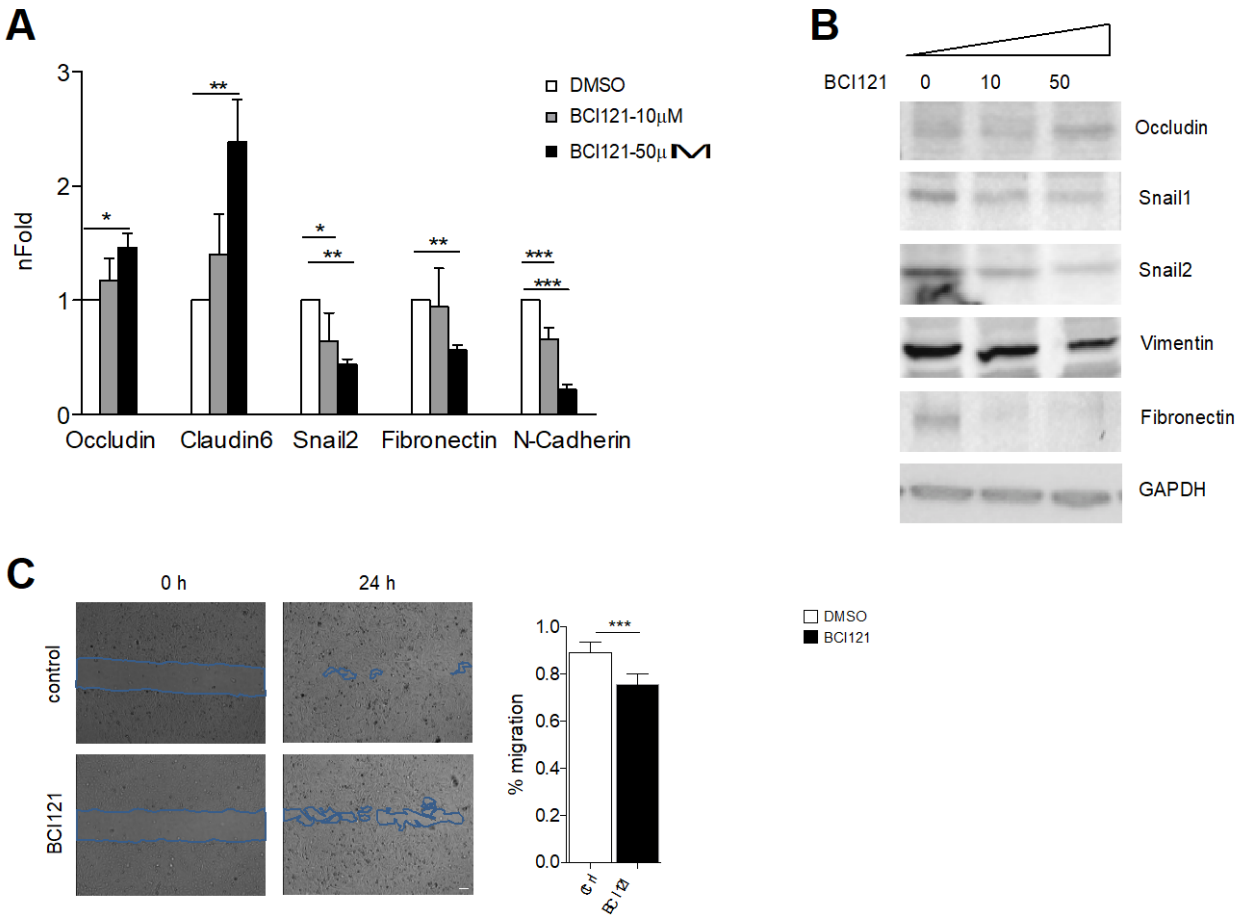
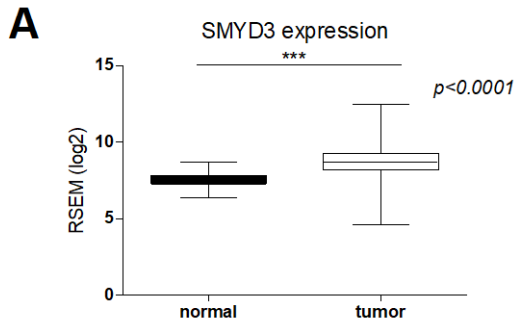
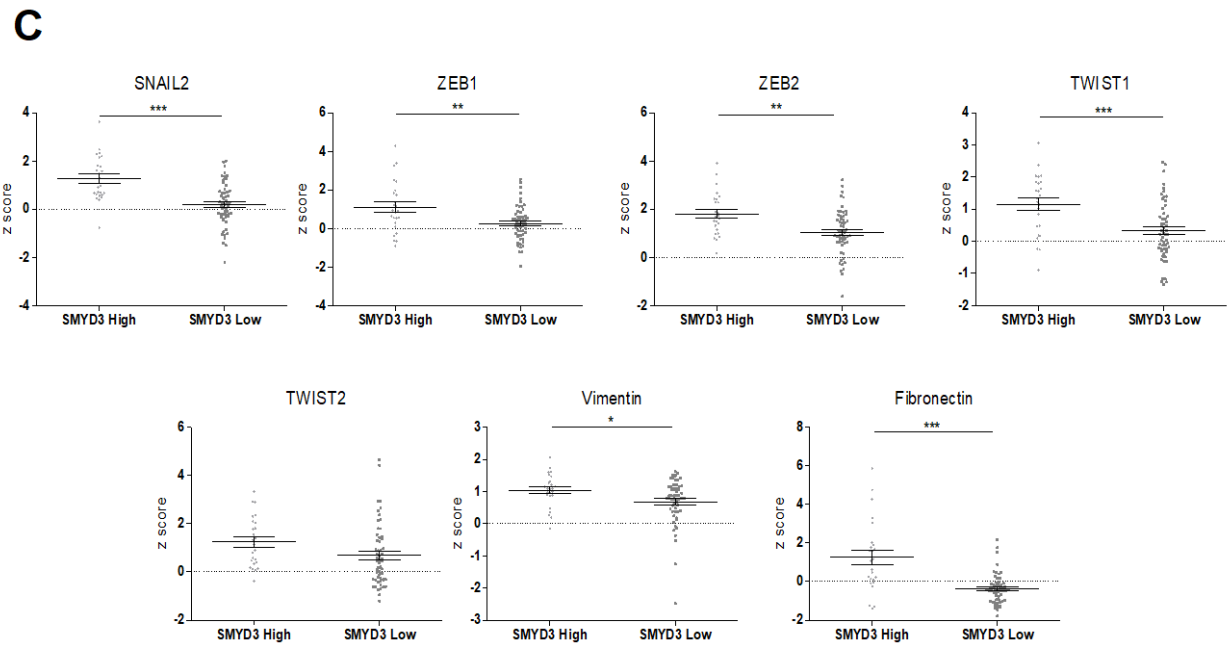

A
nFold
DMSO
BCI121-10μM
BCI121-50μM
Occludin
Claudin6
Snail2
Fibronectin
N-Cadherin
B
BCI121
0
10
50
Occludin
Snail1
Snail2
Vimentin
Fibronectin
GAPDH
C
0 h
24 h
control
BCI121
% migration
DMSO
BCI121
D
DMSO
BCI121
DMSO
BCI121
E
Number of invasive cell
DMSO
BCI121

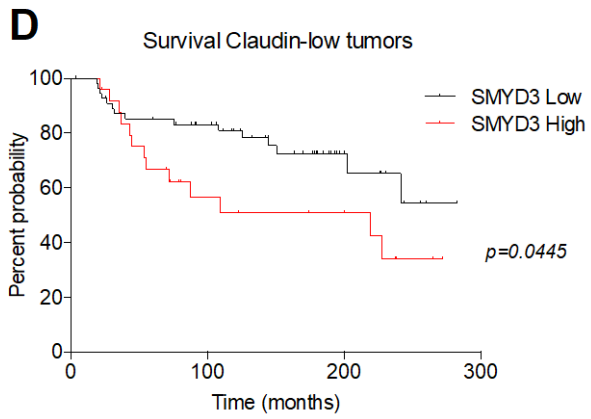
Figure 6



B

Transcript	R2	Pearson r
SNAIL2	0.1067	0.3267
ZEB1	0.0403	0.2008
ZEB2	0.0757	0.2751
TWIST1	0.0634	0.2518
N-cadherin	0.0581	0.2410
Fibronectin	0.1726	0.4155





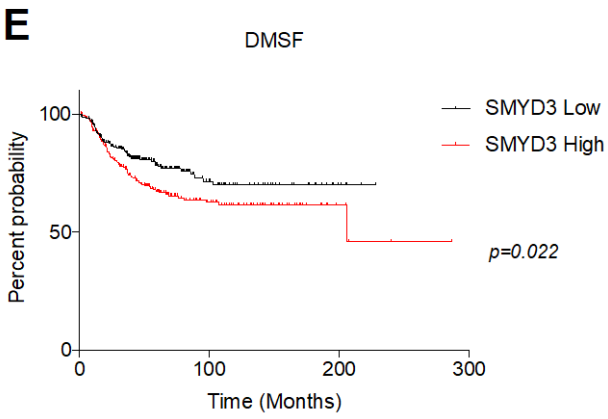
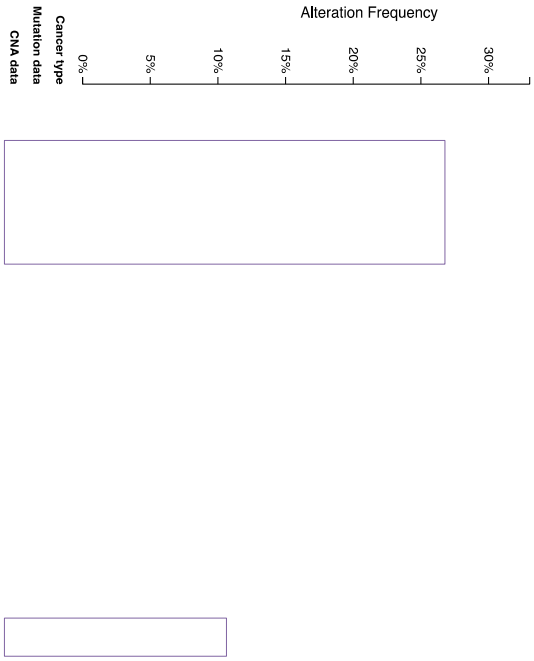
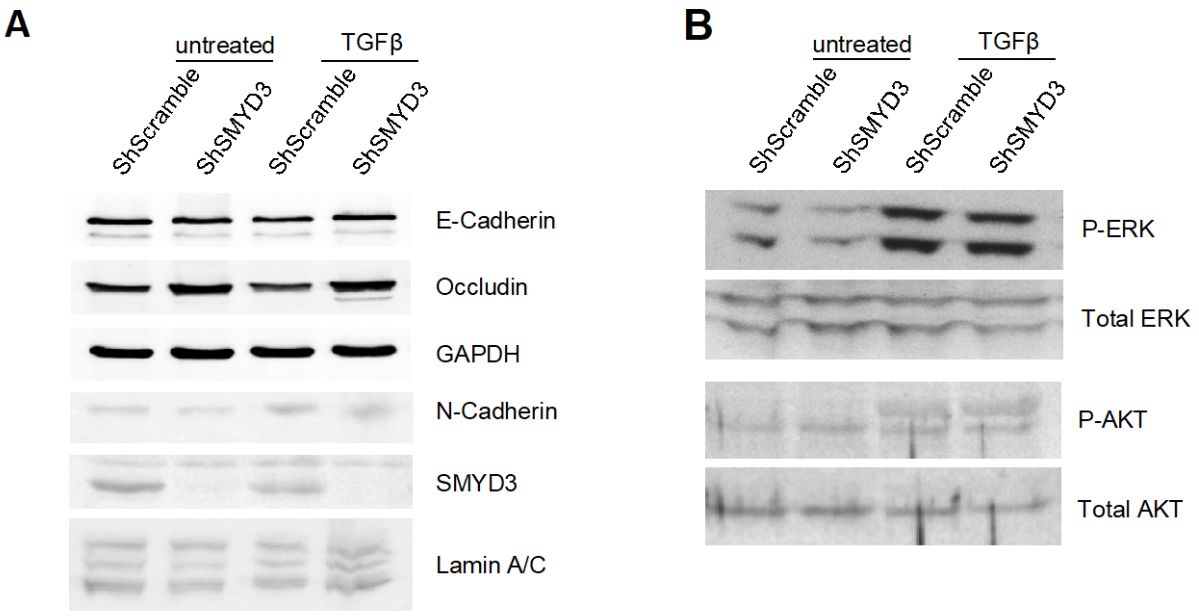


Figure 7

A

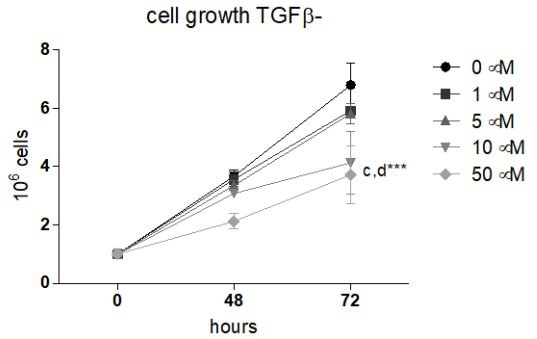


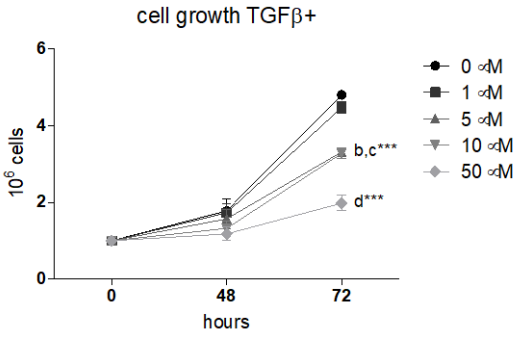
Supplemental Fig. S1



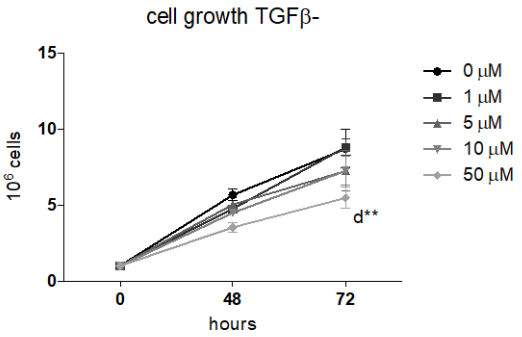
Supplemental Fig. S2

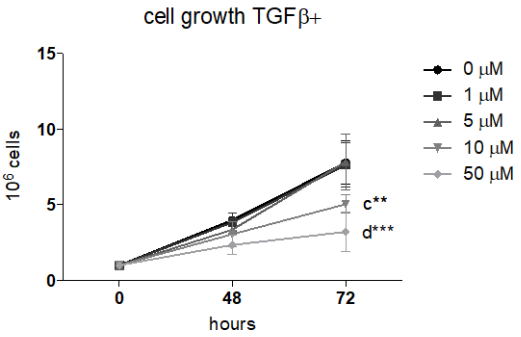
A NMuMG



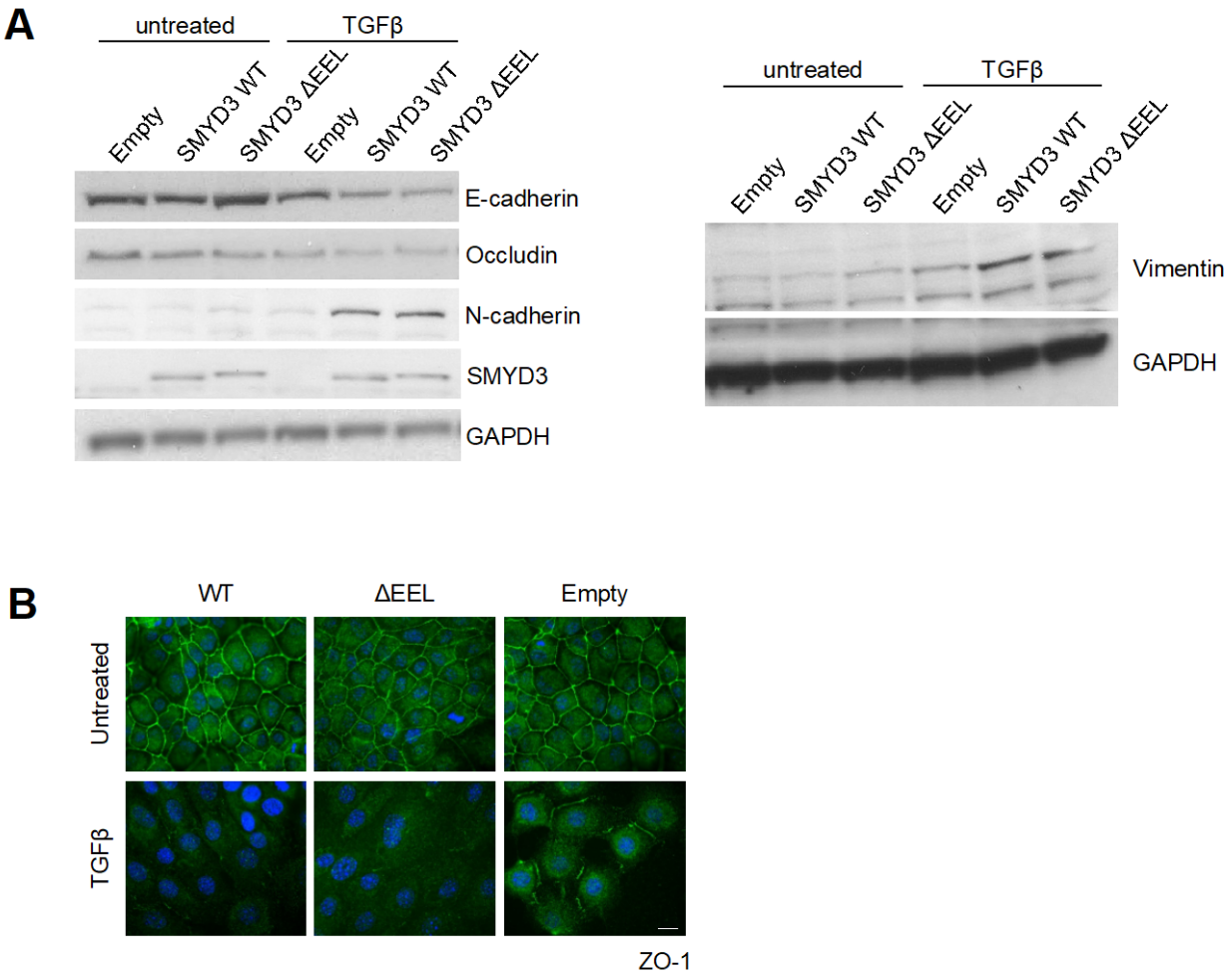


B MCF10A

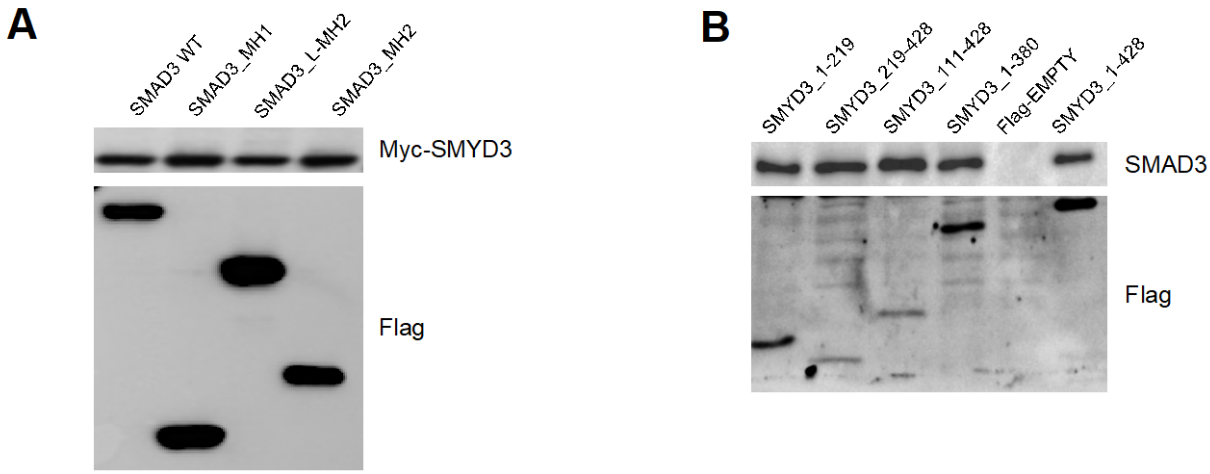




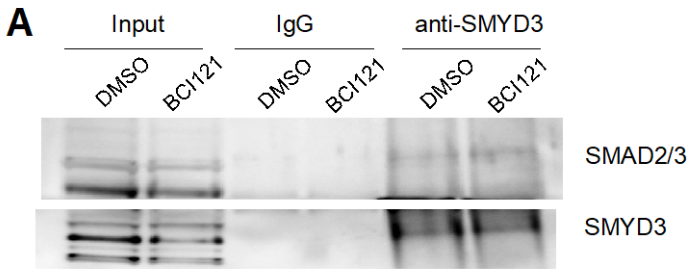
Supplemental Fig. S3



Supplemental Fig. S4



Supplemental Fig. S5



Supplemental Fig. S6

APPENDIX: zebrafish developmental stages