UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Istologia, Microbiologia e Biotecnologie mediche
Sezione di Istologia ed Embriologia

SCUOLA DI DOTTORATO DI RICERCA IN BIOMEDICINA
CICLO XXIV

Role of TGFβ signaling in embrionic development and as regulator of Hypoxia in metastasis progression

Direttore della Scuola: Ch.mo Prof. Giorgio Palù
Supervisore: Ch.mo Prof. Stefano Piccolo

Dottoranda: Dr.ssa Elena Enzo

INDEX

ABSTRACT

The focus of my PhD work has been centered on the biological roles of TGFβ signaling in vivo. In particular, I concentrated on the mechanisms at the basis of two of the most critical aspects of this pathway: TGFβ as master of cell fate decisions in early embryonic development and the mechanism by which TGFβ serves as prometastatic factor in tumor cells.

The definition of embryonic potency and induction of specific cell fates are intimately linked to the tight control over TGFβ signaling and, surprisingly, little is known on the intracellular factors that negatively control Smad activity in mammalian tissues. By means of genetic ablation (gene-knockout), we found that the Smad4 ubiquitin-ligase Ectodermin (Ecto, also known as Tif1γ) is required to limit TGFβ responsiveness in early mouse embryo. This analysis represents a novelty because prior of this work all the available mouse genetic models concentrated on positive element of this pathway (TGFβ ligands, receptors, Smads). Here we found that negative regulation of Smad is as important as their activation: loss of Ecto invariably drives Smad activity to the highest threshold; new phenotypes, linked to excessive TGFβ activity, emerge from such a modified landscape of Smad responsiveness.

In the second part of my thesis I focused instead on a mechanism that cancer cells adopt to induce metastasis downstream of TGFβ. We had previously identified p63 as metastasis suppressor (Adorno et al., 2009), whose function is opposed by TGFβ, and here we asked about the nature of p63 metastasis suppression in Triple Negative Breast cancer (TNBC), whose malignancy is known to be driven by

autocrine TGFβ production. By interrogating TNBC clinical datasets, we found that Hypoxia-Inducible-Factors (HIFs), key regulators of the invasive and metastatic cancer cell phenotype are unleashed by p63 inactivation. The p63 target Sharp1 is at the center of this crosstalk. Sharp1 inhibits invasive cell behaviors in vitro and acts as metastasis suppressor in vivo through inhibition of HIF-1α/HIF-2α. Mechanistically, Sharp1 promotes HIF-1α/HIF-2α proteasomal degradation by serving as HIFs presenting factor to the proteasome. Importantly, this process is independent of pVHL, hypoxia and the ubiquitination machinery. As such, Sharp1 is a required determinant for the intrinsic instability of HIFs proteins acting in parallel to and cooperating with oxygen levels. This work sheds light on the mechanisms and pathways by which TNBC acquire invasiveness and metastatic propensity.

RIASSUNTO

Durante il mio peridodo di dottorato mi sono occupata di studiare il rolo biologico della via di segnale TGFβ in vivo. In particolare mi sono concentrata su due meccanismi critici di questa via di segnale: TGFβ come principale regolatore del destino cellulare durante i processi di formazione dell'embrione e come fattore prometastatico nelle cellule tumorali.

La regolazione della pluripotenza e l'induzione di specifici percorsi di differenziamento cellulare sono intimamente connessi a uno stretto controllo della via di segnale del TGFβ. Pochissimo è noto sull'importanza di fattori che controllino in modo negativo l'attività delle Smad nelle cellule dei tessuti di mammifero. Con questo studio, utilizzando tecniche di ablazione genetica, dimostriamo che l'inibizione di Smad4 da parte di Ectodermin/Tif1γ/TRIM33 (Ecto) è richiesta per permettere a Nodal (il principale ligando TGFβ durante lo sviluppo embrionale di mammifero) di svolgere le sue funzioni morfogenetiche. La delezione di Ecto sposta le risposte Smad4 dipendenti nella finestra di massima responsività. Inoltre, la delezione di Ecto ristetta all'epiblasto incanala il differenziamento del mesoderma verso destini di nodo, mostrando come sia necessaria un'inibizione di Smad4 per permettere l'allocazione ordinata di cellule ai vari destini lungo la stria primitiva.

Nella seconda parte della mia tesi mi sono focalizzata sullo studio dei meccanismi, indotti da TGFβ, che promuovono la metastasi. In particolare, tramite analisi bioinformatica su librerie di dati provenienti da tumori alla mammella tripli negativi, abbiamo identificato quali vie di segnale guidano la crescita metastatica in

questa classe di tumori così vasta ed eterogenea. A valle di TGFβ abbiamo identificato p53-mutante\p63 e la via di HIFs (Hypoxia-inducible-Factor), già noti come principali regolatori del comportamento metastatico e invasivo delle cellule tumorali. Al centro di questo quadro si pone Sharp1, indotto da p63 e in grado di inibire l'azione di HIF inducendo la degradazione di HIF stesso tramite il proteasoma, in modo indipendente da ubiquitina e pVHL. Sharp1 è quindi un fondamentale freno contro l'attività di HIF che coopera con l'ipossia.

PUBBLICATIONS

Sharp1 opposes breast cancer metastasis by presenting Hypoxia-inducible-Factor to the proteosome; Montagner M, **Enzo E**., Parenti A, Rosato A, Bicciato S., Cordenonsi M, and Piccolo S; in preparation

USP15 is a Deubiquitinating Enzyme for Receptor-activated Smads.; Inui M, Manfrin A., Mamidi A, Martello G, Morsut L, Soligo S, **Enzo E**, Moro S, Polo S, Dupont S, Cordenonsi M and Piccolo S; Nature Cell Biology, published online 25 September 2011.

Role of YAP/TAZ in mechanotransduction; Dupont S, Morsut L, Aragona M, **Enzo E**, Giulitti S, Cordenonsi M, Zanconato F, Le Digabel J, Forcato M, Bicciato S, Elvassore N, Piccolo S; Nature. 2011 Jun 8;474(7350):179-83.

Negative control of Smad activity by ectodermin/Tif1gamma patterns the mammalian embryo. Morsut L, Yan KP, **Enzo E**, Aragona M, Soligo SM, Wendling O, Mark M, Khetchoumian K, Bressan G, Chambon P, Dupont S, Losson R, Piccolo S. Development. 2010 Aug 1;137(15):2571-8.

A MicroRNA targeting dicer for metastasis control; Martello G, Rosato A, Ferrari F, Manfrin A, Cordenonsi M, Dupont S, **Enzo E**, Guzzardo V, Rondina M, Spruce T, Parenti AR, Daidone MG, Bicciato S, Piccolo S. Cell. 2010 Jun 25;141(7):1195-207.

The first part was realized with the contributions of Dr. Leonardo Morsut and Dr. Sirio Dupont.

The second part of my thesis was done with the contribution of Dr. Marco Montagner. Dr. Antonio Rosato carried out the in vivo studies in nude mice, and Prof. Anna Parenti helped with the histological characterization of tumor samples. Dr. Silvio Bicciato support was crucial for the microarray set-up and corresponding bioinformatic analysis.

INTRODUCTION

The TGFβ pathway

The basics of TGFβ signaling.

The Transforming Growth Factor-β family of cytokines (TGFβ) controls a variety of cellular responses from body axis formation, left-right patterning and organogenesis in the embryo to tumor suppression, immune surveillance, cellular homeostasis and differentiation. Loss of this control contributes to cancer, metastasis, fibrosis and several other diseases (Massague, 2000; Wakefield and Roberts, 2002).

The TGFβ family can be divided into two main branches: the TGFβs, Activins, and Nodal branch and the BMP (Bone morphogenetic protein) branch. TGFβ ligands binds to specific Type I/Type II receptors, causing activation of the Receptor-Smads proteins through phosphorylation in their C-terminal domain. Activated R-Smads form a complex with Smad4, common to all R-Smads, and shuttles to the nucleus where they can activate or repress target gene expression (Fig. 1a).

TGFβ as morphogen: a key function for negative regulation of Smad activity

TGFβ are prototypic morphogens in the embryo: different cell fates are induced by different ligand concentrations and/or the time of exposure. As such, Smad activity must remain proportional to changes in extracellular ligand availability. This also implies that clearance systems must be in place to avoid Smad nuclear accumulation.

In our lab, we recently described on of such clearance system, that is, the Smad4 ubiquitin ligase, Ectodermin/Tif1γ (Ecto) (Dupont et al., 2005) (Fig. 1b).

Ectodermin is a RING-type ubiquitin ligase containing an N-terminal TRIM domain composed of a RING finger, two B-boxes, and a coiled coil, followed by a central linker region of unique sequence and a C-terminal PHD/bromodomain. Ecto can ubiquitinates Smad4, when it is in the trascriptional active complex with the RSmad. The monoubiquitination disrupts the complex that is no more able to activate the target genes. In this way, TGFβ signalling is shut down. TGFβ stimulation itsefl, through the presence of Smad4\Rsmads complex on the DNA, induces epigenetics modifications on the chromatin that bring Ecto to the DNA. The PHD and Bromo domains interact with the N-terminal region of the hystone 3, causing a structural modification in the TRIM domain, that activate the E3 ligases (Agricola et al, 2011) (Fig. 1c). Activated Ecto would then ubiquitinate chromatin-bound Smad4 at lysines 519 and 507, disrupting the complex (Dupont et al., 2009; Dupont et al., 2005).

In this thesis we have characterized the in role of Ecto *in vivo*, by using germ-line and conditional mouse knockouts. In the following chapter of the introduction I will outline some key background information on what is known about the role of TGFβ/Smad signaling in establishing germ layers and polarity in the early mouse embryo.

TGFβ signaling during mouse development

The mouse blastocyst is composed by an outside layer, the polar trophectoderm and the inner cell mass (ICM), that will originate the visceral endoderm and the epiblast (Figs. 2a). Gastrulation, the hallmark of metazoan development, in mouse starts at E6.5 and is marked by mesoderm induction that originates from a structure called primitive streak, which starts to elongate from the proximal to the distal end of the epiblast. Mesoderm cells acquire mesenchimal features and invaginate between epiblast and visceral endoderm migrating extensively, giving rise to mesoderm (which will surround all the epiblast) and definitive endoderm. The most distal part of the primitive streak originates the Node, the mammalian organizer tissue. From this structure emanates directly a column of cells called prechordal plate that will profoundly influence the patterning of neural tissue along the dorso-ventral axis. At the posterior pole gastrulation proceeds with the deposition of new endoderm,

mesoderm and neuro-ectoderm derivatives, that will aggregate in the formation of body segments (Tam and Loebel, 2007). (Fig. 2a)

During early mouse development, TGFβ ligands are able to coordinate growth and differentiation of the different cell lineages. The major ligand expressed in the early period is Nodal, expressed from blastocyst stage. Nodal signal through the canonical RSmad/Smad4 pathway exerts pleiotropic functions in early mouse embryo; these functions are cell lineage specific and strictly temporally controlled. The first known role of Nodal is to maintain the epiblast in an undifferentiated state through a self-sustaining positive feedback loop. At a strictly timed stage (E5.5), Nodal induces anterior visceral endoderm (AVE), a specialized subpopulation of visceral endoderm cells initially found at the distal tip of the embryo. These extraembryonic cells express specific Nodal target genes (such as *Cerl*, *Lim1* and many others). The AVE then rotates (E6.0) towards the prospective anterior pole of the embryo. AVE induction is the very first event by which Nodal\Smad signaling establishes the anterior pole of the antero-posterior axis in the mouse embryo; Nodal-dependent mesoderm induction, in facts, will occur later (E6.5) at the opposite side, marking the posterior pole (Figs. 2a, b).

At streak stages (from E6.5 on), Nodal has a major role in patterning of the nascent mesoderm. In particular, it is a gradient of Nodal activity that patterns the primitive streak along its antero-posterior axis, inducing the node at very hugh level and the rest of the primitive streak at lower level of signal (Dunn et al., 2004; Lowe et al., 2001; Lu and Robertson, 2004; Vincent et al., 2003).

Interestingly, the many Nodal functions in early mouse embryo are not uncoordinated or randomly executed. On the contrary, Nodal signaling seamlessly orchestrates the maintenance or restriction of embryonic pluripotency and establishes the body plan. The AVE played a central role in this mechanism as it starts to express *Cerberus* and *Lefty1* (Fig. 2b) that are secreted molecules that inhibit Nodal itself. These molecules serve two main goals:

1. they limit the effects of Nodal on itself; in facts, the primary stimulus to *Nodal* transcription is Nodal ligand itself through a strong feedforward loop. At the same time, Nodal induces its own antagonists, which in turn reduce *Nodal* expression;

2. they restrict the effect of Nodal on epiblast cells at the prospective posterior pole; with its migration, AVE creates a Nodal-free zone at the distal tip and at the

anterior side of the embryo, a fundamental condition for head development (Piccolo et al., 1999). On the other hand, the region of the epiblast that is more far from the AVE will be the zone where Nodal is less inhibited, and thus more expressed and active (thanks to the feedforward loop). It is in this region that Nodal induces the mesoderm master genes *T* (*Brachyury*), *Wnt3a*, *Eomes* and others (Arnold and Robertson, 2009), in a small group of epiblast cells; mesoderm ingression will mark the formation of primitive streak and thus the beginning of gastrulation.

TGFβ a double edged sword in cancer and metastasis progression

It is well established that several TGFβ responses remain fully operational in cancer cells and contribute positively to tumor invasiveness and metastasis (Fig. 3a). Metastasis is the cause of 90% of death from solid tumors; yet, the complexity of this process remains enigmatic. Several recent developments underlie, however, that for deeper understandings and clinical improvements it is mandatory to unveil, in mechanistic terms, the genetic and epigenetic changes that program the acquisition of distinct metastatic traits (Fidler, 2003; Gupta and Massague, 2006)

Mutant-p53 inhibits p63 anti-metastastic function

The output of TGFβ stimulation strongly depends on its cooperation with other proteins, such as Smad cofactors, and signaling cascades. Our group provided evidences that p53 family members are critical determinants for key TGFβ gene responses in different cellular and developmental settings (Cordenonsi et al., 2003).

More that 80% of p53 mutations in cancers are missense mutations, generating stable, but transcriptionally deficient proteins (Soussi and Lozano, 2005). It is believed that the potential advantages for the tumor to retain a mutated p53, instead of deleting it for good, stem from molecularly undefined “gain-of-function” properties that render mutant-p53 a dominant oncogene.

One mechanism for mutant-p53 function was proposed by our group in 2009 (Adorno et al., 2009). Mutant-p53 physically interacts with an anti-metastatic protein, p63, and inhibits its function. Theassembly of the complex required active Smad2,

that upon TGFβ stimulation, creates a platform for mutantp53 to interact with p63 in a ternary complex (Fig. 3b).

p63 is a member of p53 family expressed in many isoforms with different proprieties not completely understood. Tap63 isoform contains a transcription domain and can induces cell cycle arrest and apoptosis (Gressner et al., 2005). In presence of mutant-p53, TGFβ attains control over p63 and this unleashes TGFβ malignant effects. Knockdown of mutant-p53 in metastatic cancer cells does not affect the expression of the TGFβ invasive program but rather forestalls its phenotypic exploitation. Inactivation of p63 transforms non-invasive cells into malignant tumors and rescues metastasis ability in mutant-p53-depleted breast cancer cells. Conversely, addng back p63 inhibits metastasis.

In Adorno et al, 2009, were also identified two genes, Sharp1 and CyclinG2, differentially regulated by mutant-p53 that are the final downstream targets of the TGFβ/mutant-p53/p63 pathway. These were functionally validated *in vitro* as essential mediators of p63-mediated antagonism toward TGFβ responses. Traditional prognostic markers are able to confidentially assign prognosis to less than 50% of breast cancer patients. For the rest of the patients, new prognostic tools are required to assess the risk of metastasis and thus identify those that would benefit from adjuvant treatments (van 't Veer et al., 2002). Strikingly, in cancer patients, expression of Sharp1 and CyclinG2 represents a "minimal signature" with prognostic value independent from currently used clinical and histopathological variables. In spite of its simplicity, the minimal signature has predictive power comparable to more complex gene sets of predictors. The mechanism by which Sharp1 and CyclinG2 may act as metastasis suppressors *in vivo* remain ground for future studies. In the meantime, their use a diagnostic tools should be implemented for patients' stratification in the clinical laboratory.

Hypoxia signaling in cancer

The Hypoxia pathway

Hypoxia-inducible factor 1 (HIF-1) is a basic helix- loop-helix-PAS domain transcription factor composed by 2 subunits, HIF-1α and HIF-1β. HIF-1 controls the expression of hundreds of genes, but the list of target genes varies considerably from

one cell type to another; therefore, the complete HIF-1 transcriptome is likely to include thousands of genes. In particular, HIF-1α mediates developmental and physiological pathways that either deliver O_2 to cells or allow cells to survive O_2 deprivation. O_2 levels control HIF-1α protein stability through the PHD2-VHL pathway. The prolyl hydroxylase domain protein 2 (PHD2) utilizes O_2 under high oxygen levels to foster the hydroxylation of HIF-1α at the Pro402 and 564. These post-trascriptional modification are recognized by VHL, that recruits an E3 ubiquitin-protein ligase complex that ubiquitinates HIF-1α and this marks HIF-1α for degradation by the 26S proteasome (Fig. 3c) (Kaelin and Ratcliffe, 2008). Thus, in normoxic conditions HIF-1α is continuously synthesized and degraded but under hypoxic conditions PHD2 activity is reduced due to substrate limitation. Once the protein is stabilized, it dimerizes with HIF-1β (which is costitutively expressed), binds to cis-acting hypoxia response elements in target genes, and recruits coactivator proteins, which leads to increased transcription (Semenza, 2007).

Role of HIFs in tumor progression and metastasis

Immunohistochemical analysis of human cancer biopsies revealed increased levels (relative to surrounding normal tissue) of HIF-1α or HIF-2α protein (or both) in the majority of primary human cancers and their metastases. Moreover, increased levels of HIF-1α or HIF-2α protein correlate with increased patient mortality in many human cancers (Semenza).

Intratumoral hypoxia is a major mechanism underlying the increased levels of HIF-1α and HIF-2α in cancer and stromal cells. For example, the median PO_2 level measured in breast cancers was 10 mm Hg, as compared to 65 mm Hg in normal breast tissue (Vaupel, 2004). Other inducers of HIF-1α in the tumor micro-environment include reactive oxygen and nitrogen species, which also inhibit proteasomal degradation of HIF-1α (Dewhirst et al., 2008; Gao et al., 2007) Moreover activation of the phosphatidylinositol-3-kinase and MAP kinase pathways (either as a result of oncogenic mutation or increased signaling from receptor tyrosine kinases and G-protein-coupled receptors) increases HIF-1α synthesis, primarily through the action of mTOR (Laughner et al., 2001).

HIFs regulate many critical aspects of cancer biology, such as angiogenesis and glucose/energy metabolism, which controls oxygen delivery and use. Recent studies

implicate regulatory functions of hypoxia also in many steps of metastasis progression, as epithelial-mesenchymal transition (EMT), intravasation to the blood vessel walls, extravasation and finally dissemination at the distant organs (Fig. 3d).

Hypoxia pathway promotes each step of the metastasis progression. Extravasation may be promoted by hypoxia response factors VEGF, MMP1, and MMP2 in a similar way to intravasation. ANGPTL4, a key molecule for extravasation in lung, is upregulated by both TGFβ and hypoxia, suggesting a possible synergistic priming effect of the two pathways. Chemokine receptor CXCR4 plays a key role in metastatic homing of tumor cells to organs expressing high levels of its ligand SDF1 (Muller et al., 2001; Rey and Semenza, 2010; Wong et al., 2011; Zhang et al., 2011). Hypoxia response molecules that promote survival, invasion, and angiogenesis in the primary tumor may function similarly in the secondary site. In addition, hypoxia may upregulate proteins that mediate interactions with unique stromal cells in the secondary organ. Despite the different mechanisms of action in organ-specific metastasis, blocking HIF-1α significantly inhibited metastases to both bone and lung in animal models, establishing HIF-1α as a promising therapeutic target for treating metastasis at multiple sites.

Hypoxia and HIF levels do not stringently correlate in cancer: while most advanced cancers show increased HIFs protein levels, only a fraction show signs of hypoxic areas, suggesting that a substantial proportion of HIF regulation must derive from other mechanisms. Also, pVHL, a central player in the O2 pathway, is relevant in renal cancer, how about other solid tumors? Moreover, it should be remembered that cancer cell invasion in peritumoral stroma occurs in normoxia, calling hypoxia much into question, at least for initial disseminating steps. The disproportion between what we know (about hypoxia-HIF) and should know (hypoxia-independent HIF stabilization) is suggestive the presence of an alternative/oxygen-independent mechanisms that stabilize HIF in tumors.

PART 1

RESULTS

Ectodermin inhibits TGFβ signaling in mouse embryo

To investigate the role of Ecto *in vivo*, we generated *Ecto* conditional and germ-line knockout alleles. Mice heterozygous for the *Ecto* null mutation (*Ecto*+/-) were viable and fertile; however, homozygosity resulted in embryonic lethality. Indeed, embryos from heterozygote intercrosses were collected at different stages of gestation and *Ecto* mutants could be recovered at the expected Mendelian ratios at E5.5 through E7.5, but not at later stages. Morphological and histological analyses demonstrated that *Ecto* mutants display striking defects in embryonic polarity and tissue patterning. When compared to control littermates, at E6.5 *Ecto* mutants were smaller and lacked a clear distinction between epiblast and Extraembryonic Ectoderm (EXE) (Fig. 4a). We then monitored cell viability, but wild-type and mutant embryos showed comparable proliferation rates and apoptosis (Fig. 4b and data not shown). Wild-type embryos formed mesoderm as a consequence of gastrulation; in contrast, *Ecto* mutants could readily be identified by the undivided proamniotic cavity and the lack of a primitive streak (Figs. 4a-c). Defective mesoderm formation was confirmed by *in situ* hybridization at early streak stage examining the expression of markers, such as *T*, *Eomes* and *Wnt3* (Fig. 4c). These data indicate that *Ecto* is required for mouse gastrulation. As the pregastrulation development of extraembryonic tissues relies on the activity of early acting Nodal/Smad4 signaling (Arnold and Robertson, 2009), we tested if defects in *Ecto* mutants initiated with abnormal extraembryonic development.

Expression of AVE markers at E5.5 was strikingly upregulated in *Ecto* mutants: when these markers were barely detectable in wild-type littermates, signals of the Nodal targets *Cerberus-like* (*Cerl*) and *Lefty1* mRNAs were massive in knockout embryos, becoming rapidly saturated in an abnormally broad AVE domain (Fig. 5a). While in E6.5 wild-type embryos AVE markers are usually restricted to an anterior narrow stripe of cells, in *Ecto* mutants robust *Lim1* expression was vastly expanded around the epiblast (Fig. 5a). These results support a model in which Nodal induces AVE and Ecto restrains this Nodal function.

Ectodermin is required cell autonomously to restrain Nodal responsiveness

Genetic evidences indicate that AVE responds to Nodal ligands emanating from the epiblast (Lu and Robertson, 2004). Thus, we next challenged the notion that AVE expansion is caused by a cell-autonomous enhanced Smad responsiveness as opposed to being secondary to increased ligand expression/availability in the epiblast (Fig. 5b). The latter hypothesis is based on the Nodal feedforward autoregolatory loop, which could lead to an increase in Nodal ligand availability. To this end, we made use of the paternally-inherited *Sox2-Cre* transgene, recombining the *Ecto* conditional allele in the epiblast lineage specifically (*Sox2-Cre; Ecto fl/-* embryos, hereafter *Ecto-EpiKO* or *EpiKO*). In these *EpiKO* mutants, a genetically wild-type AVE did not display any of the abnormalities characterizing the *Ecto* germline mutants, as *Cerl* mRNAs were comparable in localization and intensity to wild-type embryos (Fig. 5c). In line, at E5.5, *Nodal* is expressed normally in *Ecto* mutant embryos (Fig. 5d) and, by immunofluorescence, Smad2 phosphorylation is comparable between wild-type and *Ecto* mutants (Fig. 5e). Thus, *Ecto* is required cell autonomously to restrain Nodal responsiveness, downstream of Nodal production and Smad2 phosphorylation.

Loss of Ectodermin causes a Smad4-dependent AVE expansion

The AVE phenotype of *Ecto-/-* embryos is unprecedented and is opposite to those reported for *Nodal*, *Smad2* and *Smad4* knockouts (Brennan et al., 2001; Waldrip et al., 1998; Yang et al., 1998). Hence, we investigated the genetic relationships between *Ecto* and its biochemical target *Smad4* (Dupont et al., 2009). We analyzed embryos from crosses of mice carrying the floxed alleles for the two genes (*Ecto fl/-* and *Smad4 fl/-*) that were undergoing zygotic deletion in the *CAG-Cre* maternal background. *Ecto fl/-; CAG-Cre* embryos were indistinguishable from *Ecto* germline

homozygous mutants (compare Fig. 6a with Fig. 5a) and *Smad4 fl/-;CAG-Cre* phenocopied morphologically the previously reported defects of the null-allele (Yang et al., 1998). Extending these studies, we found that Smad4 is dispensable for VE (as revealed by the detection of the *AFP* marker, Fig. 6b), but required for *Cerl* and *Lim1* induction. Remarkably, embryos double mutants for Smad4 and Ecto were indistinguishable from Smad4 mutants (Fig. 6a). Thus, Smad4 is an obligate mediator of Ecto activity.

The phenotype of Ectodermin mutant can be rescued lowering Nodal signaling

Data presented so far suggest that disruption of the Ecto/Smad4 inhibitory axis leads to excessive Nodal responsiveness in AVE. If so, this should be rebalanced by a concomitant reduction of the Nodal dosage. To this end, we combined *Ecto* mutant with a strong *Nodal* hypomorph (Norris et al., 2002) (*Nodal*Δ*600*/-), leading to a remarkable rescue of AVE patterning (Fig. 7a). This indicates that the AVE phenotype in *Ecto* mutants is due to enhanced Nodal signaling (Fig. 7b). As previously reported (Norris et al., 2002), low levels of Nodal expression in *Nodal*Δ*600*/- were per se often insufficient for AVE rotation; notably, however, in the *Ecto-/-; Nodal*Δ*600*/- compound mutants the AVE invariably rotates (Fig. 7c). Thus, lowering the Nodal dosage in *Ecto* mutants compensates an exalted Smad intracellular responsiveness and viceversa.

These results suggests that, in vivo, the net activity of Nodal/TGFβ is the result of two components: extracellular ligand availability and negative control of over Smad responsiveness; the loss of negative control of responsiveness in *Ecto* mutants profoundly alters embryonic patterning, but these defects are normalized by reduction of Nodal ligand availability.

Lack of mesoderm in *ecto* mutants is due to increased Nodal and defective EXE development

Next, we characterized molecularly the development of the other extraembryonic tissue in *Ecto* mutants. The extraembrionic markers *Eomes*, *Cdx2* and *BMP4* were undetectable in E5.5 *Ecto*-/- embryos (Fig. 8a and data not shown). This represents a cell autonomous requirement as *Ecto-EpiKO* embryos displayed normal EXE development (marked with *BMP4* in situ, Fig. 8b). Lack of EXE in *Ecto* mutants is paradoxically similar to the phenotype of *Nodal* mutants (Brennan et al., 2001);

however, in the case of *Nodal*, this is secondary to defective epiblast patterning as Nodal sustains *Oct4* and *FGF4* transcription in the epiblast (Guzman-Ayala et al., 2004; Lu and Robertson, 2004; Mesnard et al., 2006). In contrast, *FGF4* and *Oct4* are normally expressed in *Ecto* mutants (Fig. 8b and data not shown). Strikingly, Nodal attenuation rescued the EXE phenotype of *Ecto* mutants, as *Eomes* and *BMP4* transcripts were invariably rescued in combined *Ecto-/-; Nodal*Δ*600/-* or *Ecto-/-; Nodal+/-* embryos (Fig. 8c and data not shown). Taken together, these data strongly suggest that *Ecto* protects the extraembrionic ectoderm from excessive Nodal signaling.

By losing the EXE, *Ecto* mutants are deprived of an essential source of mesoderm inducing and patterning signals (Arnold and Robertson, 2009); at the same time, they display a massive expression of Nodal antagonists, such as *Cerberus* and *Lefty*. This raises questions on what is the primary cause of defective mesoderm in *Ecto* mutants. Remarkably, attenuation of Nodal signaling in compound *Ecto/Nodal* mutants also rescues mesoderm development, as revealed by transcription of the pan-mesodermal markers *Eomes* and *T* at the early gastrula stage (Fig. 9). Importantly, while the combination *Ecto-/-; Nodal*Δ*600/-* rescues EXE, mesoderm and AVE, compound *Ecto-/-; Nodal+/-* could rescue EXE and mesoderm but not AVE (compare red boxed pictures in Fig. 9), revealing that lack of mesoderm in *Ecto* mutants is primarily due to lack of EXE, and that this may be uncoupled from exalted AVE activity. A further complicating issue is the fact that AVE and EXE development might be linked, as the EXE has also been proposed to secrete AVE inhibiting factors (Rodriguez et al., 2005; Yamamoto et al., 2009). Is then the AVE expansion observed in *Ecto* mutants due to loss of EXE? Our results suggest that this is not the case, as *Ecto-/-; Nodal +/-* embryos display rescued EXE in the presence of a still expanded AVE (Fig. 9, red box). Thus, the two events seem uncoupled, and expanded AVE in *Ecto* mutants is primarily due to enhanced Nodal responsiveness of the visceral endoderm.

In the Epiblast Ecto patterns the primitive strike through the regulation of Smad4 activity

The *Sox2-Cre; Ecto fl/-* embryos (*Ecto-EpiKO*) allow to study more directly the role of *Ecto* in the epiblast, bypassing its early requirements in extraembryonic tissues, and likely allowing us to monitor post-gastrulation defects. It seemed

interesting, since we monitored the localization of Ecto protein by immunofluorescence, and we found that it is expressed at higher levels in epiblast nuclei than in extraembryonic cells (Fig. 10a). This differential enrichment is particularly sharp at E5.5 but declines as development proceeds, with Ecto becoming ubiquitous during gastrulation. Previous work established that a gradient of Nodal activity patterns the primitive streak (Dunn et al., 2004; Lowe et al., 2001; Vincent et al., 2003). Because Ecto is required for Smad4 monoubiquitination, preventing the incorporation of Smad4 into Smad transcriptional complexes (Dupont et al., 2009); we tested whether patterned Ecto activity contributed to regionalize Smad4 activity in vivo. By means of antiubiquitin immunoprecipitations from embryonic lysates, we noticed that endogenous Smad4 is preferentially ubiquitinated in the proximal, but not distal half of the gastrulating embryo (Fig. 10b).

This raised the possibility that *Ecto* works to prevent excessive, if not precocious, Nodal/Smad4 signaling in proximal epiblast. Strikingly, we found that approximately a third of the *Ecto-EpiKO* embryos displayed at streak stages an expanded *FoxA2* expression (Fig. 10c). These embryos appeared smaller and lacked an overtly elongated streak, and likely failed to undergo proper gastrulation. At later stages, *Ecto-EpiKO* embryos displayed a severely abnormal morphogenesis, highlighted by an expanded node and aberrant morphogenetic movements, in some cases leading to the growth within the amniotic cavity of a column of epiblast cells contiguous to the node rim. When node-derivatives were analyzed molecularly, surviving *Ecto-EpiKO* embryos showed expansion of the Node (marked by *FoxA2* staining, Fig. 10d), an almost radial expansion of the definitive endoderm marker *Cerl* (Fig. 10e), as well as duplications of Node and anterior axial mesendoderm tissues (*Shh* in situ, Fig. 10f). Together, the data suggest that *Ecto* is essential to orchestrate intensity of Nodal/Smad4 responses for proper primitive streak development (Fig. 10g).

DISCUSSION

The TGFβ cascade is a fundamental player in mammalian development and adult tissue homeostasis. Although TGFβ ligands are widely expressed in vertebrate tissues, they can elicit their effects in a strict temporally and spatially controlled manner. For the signal to reach only the appropriate cells and with the correct intensity, mechanisms must be in place to determine where and when cells must not respond to TGFβ. Yet, little is known on factors that shape Smad responsiveness in vivo (Schmierer and Hill, 2005). In the mammalian embryo, the TGFβ-related Nodal ligand signals through Smads to act as morphogen; depending on the cellular context and developmental timing, Nodal signaling seamlessly orchestrates the maintenance or restriction of embryonic pluripotency and establishes the body plan (Tam and Loebel, 2007). However, whether extracellular gradients of Nodal ligand exist in mouse embryos is unclear; with this work, we explored the possibility that the morphogenetic properties of Nodal may rely on negative Smad regulation.

The mechanisms by which cells clear nuclear Smad activity are starting to emerge (Itoh and ten Dijke, 2007; Wrana, 2009). However, these mechanistic results are largely uncoupled by sound in vivo validation, as in the case of Smurfs protein: these molecules were originally identified as intracellular inhibitors of the TGFβ cascade (Zhu et al., 1999), but the corresponding knockout embryos have a phenotype linked to the planar cell polarity pathway (Narimatsu et al., 2009). With our work, we now provide the first genetic evidences indicating that key to the morphogenetic properties of Nodal is a negative, cell-autonomous Smad regulation operated by the Smad4 Ubiquitin-ligase Ectodermin (Dupont et al., 2005; Dupont et al., 2009). Even more surprisingly, our data unveil that the extracellular modulation of TGFβ ligands becomes irrelevant in the presence of unleashed intracellular Smad4 activity.

Adding a further twist, we found that Ecto activity on Smad4 ubiquitination appears spatially regulated: a ubiquitination-mediated patterning mechanism is shown to be essential to generate the "landscape" of TGFβ responsiveness, playing an equal if not superior role to extracellular ligand distribution. Thus, during gastrulation, the Nodal "morphogenetic gradient" is actually not a simple extracellular gradient of ligands, but is mainly mediated by negative Smad regulation.

Analysis of *Ecto* knockouts showed how loss of *Ecto* invariably "upgrades" Nodal responses to a high-threshold of activity in all the tissues of the early mouse

embryo: this was true for known Nodal responses, such as Anterior Visceral Endoderm (AVE) induction, but also unveiled novel functions of Nodal in the differentiation of the trophoblast lineage. In vivo, the net activity of Nodal/TGFβ is the result of two components: extracellular ligand availability and, unprecedently, negative control of over Smad responsiveness; the sole loss of this second component in *Ecto* mutants is sufficient to profoundly alter embryonic patterning.

Adding a further twist, we found that Ecto activity appears spatially regulated: at the endogenous level, we found Smad4 ubiquitinated, and thus inactive, in the proximal embryo (posterior mesoderm, whose development is in fact *Smad4*-independent); in contrast, Smad4 is free of ubiquitin, and thus fully active, in the most anterior segment of the primitive streak, where the Node will form in a *Smad4*-dependent way.

Our previous studies indicated a function for Ectodermin as a general inhibitor of both TGFβ and BMP responses (Dupont et al., 2009; Dupont et al., 2005). It was therefore surprising to observe mainly Nodal-related phenotypes in *Ecto* mutant embryos. We reasoned that this might simply reflect the high degree of overlap between Nodal and BMP signaling in early mouse embryos (Ben-Haim et al., 2006; Di-Gregorio et al., 2007; Yamamoto et al., 2009); more intriguingly, this might highlight an intrinsic difference in their regulatory logic, with BMP signaling being mostly regulated in the extracellular space, and Nodal relying more on intracellular Smad regulation. Indeed, unleashing BMP activity leads to phenotypes profoundly different from those here described, as Chordin and Noggin double mutants embryos show reduced AVE and axial mesoderm (Bachiller et al., 2000).

More than ten years of mouse knockouts for pathway activators (Smads and TGFβ receptors) have been essential to know what are the consequences of losing TGFβ responses; however, lack of appropriate genetic tools for Smad inhibitors left us blind on other equally central questions, namely, where, when and why cells must not respond to TGFβ. So far, no intracellular inhibitor of TGFβ signaling has received genetic support in corresponding knockout animals. This study shows how the *Ecto* mutants can be used to explore the uncharted territories of enhanced Nodal signaling during mouse embryonic development.

PART 2

RESULTS

Hypoxia pathway is downstream of TGFβ in controlling metastasis progression

We recently identified in breast cancer cells two genes, *Sharp1* and *CyclinG2*, that are downstream of p63 (Adorno et al., 2009), a key suppressor of cell migration and metastasis in epithelial tumors (Adorno et al., 2009; Muller et al., 2009; Su et al.). *Sharp1* and *CyclinG2* levels are correlated to those of *p63* in primary TNBC (Pearson coefficients *r* are, respectively, 0,251 and 0,205; Fig. 11a), prompting us to investigate their prognostic value in this class of tumors. The analysis was carried out on a cohort of 250 primary TNBC extracted from a collection of 1555 primary breast cancers, combining data from 8 independent clinically annotated gene-expression datasets (see Methods and Tables 1, 2). We defined two groups of TNBC with respectively high and low levels of *Sharp1* and *CyclinG2* expression. Crucially, by univariate Kaplan-Meier analyses, the TNBC group with low-*Sharp1*/*CyclinG2* expression displayed a significantly higher probability to develop metastasis and reduced survival (Figs. 10b, c). The fact that expression of just two genes could be prognostic in TNBC suggested us a functional relevance, beside their utility as markers.

To gain insights into the mechanisms by which Sharp1 and CyclinG2 are linked to malignant progression we asked whether their expression could be linked to other known tumorigenic pathways by Gene Set Enrichment Analyses (GSEA, see Methods). Specifically, we searched for statistical association between low *Sharp1*/*CyclinG2* expression with other gene "signatures" that register elevated

activity of specific signaling pathways in tumor samples. These included signatures for Ras, Wnt/TCF4 or TGFβ signaling, as well as signatures typifying mutations in BRCA or p53, transcriptional activity of Hypoxia-Inducible-Factors (HIFs), disrupted cell cycle/S-phase entry, proteasomal components (Table 4). To broaden our search parameters, we also employed gene sets extracted from the Molecular Signature database (see Methods) for a total of 254 classifiers. Interestingly, at the top of the list were signatures of TGFβ activity and p53 mutation (Fig. 11d), validating in clinical data the convergence of mutant-p53 and TGFβ for *Sharp1*/*CyclinG2* regulation previously shown in the well-established TNBC cellular model MDA-MB-231 (hereafter MDA-231) (Adorno et al., 2009). That said, our attention was attracted by the second highest hit, namely, the association between low *Sharp1*/*CyclinG2* and high HIF activity (Fig. 11d). Hypoxia-inducible factors have been involved in multiple aspects of malignancy in several tumor types (Gordan and Simon, 2007; Kaelin and Ratcliffe, 2008; Semenza); in breast cancer, a large body of clinical data shows that high levels of HIF-1α or HIF-2α in tissue samples is linked to poor prognosis and high patients' mortality (Semenza, 2010). Several evidences strongly support the role of HIFs also in TNBC. In fact, in our coohort of TNBC, a signature of HIF activity predicts metastasis proclivity and survival in our cohort of TNBC (Figs. 11e, f). We confirmed the role of HIF in migration *in vitro* in MDA-231 cells that, depleted of HIF-1α (Fig. 12a) or HIF-2α (Fig. 12c), have a reduced transwell migration triggered by TGFβ (Fig. 12b, d)

Sharp1 inhibts HIFs transcriptional responses

Are Sharp1 or CyclinG2 causal for HIFs repression? Interestingly, a physical association between overexpressed HIF-1α and Sharp1 has been previously noted in transfected COS7 cells (Sato et al., 2008). We found a robust physical association between HIF-1α and Sharp1 at endogenous levels in independent TNBC cell lines, such as MDA-231, Hs578T and SUM159 (Figs. 12e, f and data not shown). Conversely, no interaction was detected between HIF-1α and CyclinG2 (data not shown); moreover, at difference with Sharp1 (see below) we did not detect any inhibitory effect of CyclinG2 overexpression on the activation of HIF targets (data not shown), prompting us to focus on Sharp1 for further analyses.

Sharp1 belongs to the family of bHLH protein that reconizes the E box sequence (5'-CANNTG-3'), found in the transcriptional regulatory region of a

number of genes. Of the basic helix-loop-helix (bHLH) proteins binding to the E box sequence, class B of bHLH proteins, BHLHB2 (also referred to as the DEC1/Eip1/SHARP-2/Stra13/Clast5) and BHLHB3 (also referred to as the DEC2/SHARP-1/SHARP1), are transcription factors that contain a unique orange domain. These transcription factors repress the transcription of target genes not only via binding to the E box sequence but also via protein-protein interactions with other transcription factors. Both the *BHLHB2* and *BHLHB3* genes are widely expressed in both embryonic and adult tissues (Liu et al.; Sato et al., 2008). Their gene expressions are regulated in a cell type-specific manner by various extracellular stimuli, such as growth factors, serum starvation, hypoxia, hormones, nutrient, cytokines, light and infection.

Sharp1 levels appear elevated in the TNBC non-metastatic MCF10Atk1 (or MII) cells compared to more aggressive MDA-231 (Fig. 13a). Taking advantage of this differential Sharp1 expression, we reasoned that, if Sharp1 opposes HIF activity, gain-of-Sharp1 in MDA-231 should blunt HIF transcriptional responses; and, conversely, loss-of-Sharp1 in MII cells should enhance HIF-dependent responses. A main layer for HIF-1α and HIF-2α regulation occurs in response to microenvironmental oxygen levels (Kaelin and Ratcliffe, 2008). Under hypoxia, HIFs levels are stabilized to activate a key set of target genes, such as *Angiopoietin-like4 (ANGPTL4)*, *LOXL2* and *miR-210,* that have been involved in multiple steps of the metastastic cascade of ER-negative breast cancer cells(Buffa et al.; Padua et al., 2008; Wong et al.). As expected, these genes were upregulated by hypoxia in control MDA-231 (Fig. 13b, lane 3), but gain-of-Sharp1 dampened these inductions (Fig. 13b, lane 4). Similarly, inductions of HIF targets related to tumor metabolism, such as *carbonic anhydrase IX (CAIX)*, *hexokinase2 (HK2)* and *Pyruvate dehydrogenase kinase1 (PDK1)*, or the autophagy and stress-regulators *BNIP3* and *NDRG1* (Kaelin and Ratcliffe, 2008), were also inhibited by Sharp1 (Fig. 13b). Thus, sustaining Sharp1 expression is sufficient to oppose HIF responses. Conversely, in MII cells, loss-of-Sharp1 strongly cooperated with a pulse of hypoxia to upregulate the HIF targets *VEGFA*, *CAIX*, *ANGPTL4*, *HK2*, and *PDK1* (Fig. 13c). This indicates that Sharp1 is required to limit HIF activity and is an endogenous buffer against the effects of hypoxia.

We then wished to investigate in a more unbiased and genome-wide level if Sharp1 and hypoxia-inducible factors control a significantly overlapping set of genes. For this, transcriptomic profiles were obtained from cells stably expressing either shGFP or shRNAs against HIF-1α/HIF-2α and from cells overexpressing Sharp1. We identified two independent lists of genes differentially expressed upon Sharp1-overexpression or HIFs-depletion (Fig. 13d and for validation see Fig. 13e). Upon comparison, we found a highly statistically significant overlap between the two lists (Fischer's test, $p<10^{-73}$), supporting the notion that Sharp1 is a global inhibitor of HIF-1α/-2α gene responses. From these microarrays, we generated a list of genes repressed by Sharp1 (Table 4) and found that this signature has prognostic value in TNBC datasets: as shown in Fig. 13f, tumors expressing high levels of such genes (i.e., low Sharp1 activity signature, red line) display increased propensity to distant metastasis than tumors expressing low levels of the same signature (i.e. those retaining high Sharp1 activity, blue line). Cox multivariate analyses revealed that the signature of Sharp1-repressed genes did not add prognostic information when combined to a signature of high HIFs activity (p value = 0.4434); a similar result was obtained from the combination of the *Sharp1/CyclinG2* and HIF activity signatures (p value = 0.3072), at least suggesting that the prognostic value of Sharp1 activity is contained in that one of HIFs.

We then investigated the functional relevance of the Sharp1-HIF axis for malignant behavior of TNBC cells. *In vivo*, both HIFs depletion and gain-of-Sharp1 opposed lung metastasis of MDA-231 cells after tail-vein injection in recipient mice (Fig. 14a). Sharp1-mediated inhibition was partially rescued by overexpression of PA-HIF-1α, a constitutively active and pVHL-insensitive HIF-1α (Kaelin and Ratcliffe, 2008). Of note, gain-of-Sharp1 phenocopies that one of p63 in the same experimental set-up (data not shown). In agreement with the role of Sharp1 as HIF inhibitor, experimental tumors derived from Sharp1-expressing cells displayed a reduction of the HIF-targets CAIX and Glut1 to levels comparable to tumors emerging from HIFs-depleted cells (Fig. 14b)

The above experiments suggest that by antagonizing HIFs, Sharp1 may control migratory and invasive cell behaviors in tumors. To further substantiate this conclusion, we monitored the relevance of Sharp1 for HIF-dependent cell migration *in vitro*. Loss-of-Sharp1 leads to increased transwell migration in MII cells and this

effect was rescued by concomitant depletion of HIF-1α (Fig. 13c, for siRNA validation see Fig. 14d). As shown in Fig. 14e, similar results were obtained in MDA-231 cells depleted of mutant-p53 (a treatment that unleashes endogenous p63 activity and raises Sharp1 levels (Adorno et al., 2009)). To extend the validity of this epistasis, gain-of-HIFs opposes the anti-migratory effects of gain-of-Sharp1 in other TNBC cell lines, such as SUM159 and Hs578T. Sharp1-expressing cells were unable to migrate in a wound-healing assay, and this could be rescued by overexpression of PA-HIF-1α or PA-HIF-2a (Figs. 14f, g). These responses occurred in absence of significant effect on cell growth (data not shown).

Sharp1 promotes HIFs protein degradation independently by oxygen levels

The data presented so far establish Sharp1 as a physiological inhibitor of HIF function. We next sought to determine the mechanism of this inhibition. A main layer of HIF regulation occurs at the level of protein stability (Kaelin and Ratcliffe, 2008). Strikingly, Sharp1 overexpression greatly reduced endogenous HIF-1α protein levels, a result confirmed in three independent TNBC cell lines (MDA-231, SUM159 and Hs578T) (Figs. 15a, b). Sharp1 was equally effective against HIF-2α (Fig. 15c). Intriguingly, the effect of Sharp1 was independent from oxygen levels, as downregulation of HIF-1α protein occurred at high efficiency in cells cultured in both normoxic and in hypoxic conditions (Figs. 15a, c). Notably, in experimental tumors derived from orthotopically injected MDA-231 cells, Sharp1 also downregulated HIF-1α protein levels to those of cells expressing shHIFs (Fig. 15d).

We then tested if endogenous Sharp1 is relevant as inhibitor of HIF-1α protein levels. For this, we depleted Sharp1 from MII cells and MDA-231 cells. Sharp1-depletion induced robust HIF-1α stabilization, a finding confirmed with independent Sharp1 siRNAs (Figs. 15e, f). In agreement with Sharp1 expression being regulated by TAp63 in TNBC cell lines (Fig. 15g), effective downregulation of endogenous HIF-1α protein could be observed after raising p63 activity in MDA-231 cells, as obtained by transfecting an expression vector for TAp63α, or by unleashing endogenous TAp63α activity upon knockdown of mutant-p53 (Fig. 15i, h). Conversely, loss of p63 phenocopies loss-of-Sharp1 in MII cells (Fig. 15j).

At this point, we wondered how Sharp1 impinges on HIF protein levels. Although oxygen-dependent/pVHL-mediated degradation of HIF-1α or HIF-2α has received considerable attention (Kaelin and Ratcliffe, 2008), it is also clear that HIFs

are unstable proteins degraded by the proteasome even under hypoxic conditions (Kong et al., 2007; Uchida et al., 2004). Yet, this aspect of HIF regulation has remained only partially understood. We thus tested if the proteasome is required for Sharp1 effects. Indeed, treatments of MDA-231 or HEK293T cells with proteasome inhibitors effectively opposed Sharp1-mediated HIF-1α degradation, resulting in accumulation of unmodified and polyubiquitinated HIF-1α isoforms (Fig.16a).

We then tested if Sharp1 promoted degradation of HIF-1α by acting upstream of the prolyl-hydroxylases PHDs or of pVHL-mediated ubiquitination. For this, we first assayed whether a hydroxylation-mutant HIF-1α resists to Sharp1 and, second, the effectiveness of Sharp1 overexpression in cells lacking pVHL activity. We found that Sharp1 was able to downregulate the levels of wild-type and hydroxylation-mutant/PA-HIF-1α (Fig. 16b), and could effectively operate in VHL-mutant renal cell carcinoma (RCC4) cells (Fig. 16c). In conclusion, Sharp1 curtails HIF-1α activity independently from the pVHL pathway.

Ubiquitin-ligases other than pVHL, such as RACK1 (Liu et al., 2007), have been reported to promote oxygen-independent instability of HIFs. To test whether ubiquitination is involved in Sharp1-mediated HIF-1α degradation, we used cells carrying a temperature-sensitive mutant of the E1 ubiquitin-activating enzyme (UBE1) (Monney et al., 1998). As shown in Fig. 15d, Sharp1 is equally active in permissive (34°C) and non-permissive (39°C) conditions. As positive control for this assay (Chowdary et al., 1994), mdm2-mediated degradation of p53 is inhibited in non-permissive conditions (Fig. 16d). This suggests that Sharp1 operates independently of HIF-1α ubiquitination. Consistently, overexpression of Sharp1 does not increase polyubiquitination of HIF-1α, differing from the effects of transfected VHL (Fig. 16e).

As alternative mechanism for proteasomal-dependent degradation other than ubiquitination, we hypothesized that Sharp1 may present HIF-1α to the proteasome. The proteasome has been previously shown to bind and degrade specific short-lived proteins in an ubiquitin-independent manner (Asher et al., 2005; Sdek et al., 2005), matching the effects of Sharp1. Interestingly, unmodified HIF-1α can bind directly the 20S α4 subunit of the proteasome (Cho et al., 2001). By co-immunoprecipitations (co-IP), both HIF-1α and Sharp1 associate with the 20S proteasomal subunit (Fig. 16f).

To capture in living cells if Sharp1 is instrumental for HIF-1α recognition by the proteasome, we stabilized protein complexes by treating HEK293T and MDA-231 cells with the bifunctional and cell permeable crosslinker DSP prior to cell lysis and co-IP. As shown in Figs. 16g and 16h, expression of Sharp1 increased association of HIF-1α with the 20S subunit. Given the requirement of Sharp1 for HIF-1α instability, we then tested the relevance of endogenous Sharp1 for the formation of the HIF-1α/20S complex. In co-IP assays, HIF-1α failed to efficiently associate to the 20S subunit in lysates of Sharp1-depleted cells, indicating that Sharp1 is essential for efficient recognition of HIF-1α by the 20S proteasome (Fig. 16i).

Is the capacity to bind HIF and the proteasome instrumental for Sharp1 function? To address this question we first dissected the structural requirements of Sharp1 to dock HIF-1α to the 20S by comparing different Sharp1 deletion constructs using co-IP assays. The N-terminal bHLH domain of Sharp1 (b, see diagram in Fig. 17a) was required and sufficient for proteasomal association (Fig. 17b); as for HIF association, the bHLH is sufficient, albeit not essential (Fig. 17c). The sole bHLH domain is capable to trigger HIF-1α and HIF-2α instability in a proteasome-dependent manner, thus recapitulating the effects of full-length Sharp1 (Figs. 17d and data not shown). Interestingly, a bHLH-*deleted* version of Sharp1 (Sharp1-Δb) retained HIF-1α association, but lost proteasomal recognition. We proposed that the bHLH domain of Sharp1 is required to bring HIF-1α close to the 20S proteosome, inducing its degradation (Fig. 17e). We conclude from these results that Sharp1 needs to associate with both HIFs and the proteasome to cause HIF degradation.

DISCUSSION

Sharp1 is a metastasis suppressor downstream of p63 in TNBC

The ability of p53 to switch from friend to foe depending on its mutational status presented striking analogies with the switch in TGFβ responsiveness reported in many cancers: TGFβ is a tumor suppressors of early neoplasms but, at later stages of the diseases, it becomes a potent prometastatic factor, fostering cancer cell migration, epithelial-mesenchymal transition and stromal invasion (Derynck et al., 2001).

Prompted by the cooperation between wild-type p53 and Smad in growth arrest responses, we previously studied the role of mutant-p53 in TGFβ responsiveness (Adorno et al., 2009). We found a new signaling pathway that instills metastatic proclivity to epithelial cancer cells: TGFβ -->Smad/mutant-p53 --| p63 --> metastasis protection. Downstream of this pathway two genes, Sharp1 and CyclinG2, were identified. However, the mechanism by which Sharp1 and CyclinG2 may act as metastasis suppressors *in vivo* was not known.

In the present work we clarify how Sharp1 acts as metastasis suppressor, providing a mechanistic link between p63, Sharp1 and HIFs activities in Triple negative Breast Cancer (TNBC). We found that in TNBC Sharp1 is able to inhibits HIFs-driven metastasis progression.

By analyzing expression levels in cohort of 250 TNBC, we confirmed that expression of Sharp1 and CyclinG2 correlates with p63 expression and that their expression levels can predict survival and metastasis proclivity in TNBC patients. Bioinformatic analysis of these tumors revealed that the Hypoxia pathway is differentially expressed in tumors with low Sharp1\CyclinG2 activity, and infact signature of Hypoxia activity can equally well split the cohort of TNBC according to metastasis proclivity and survival. Sharp1 signature does not add prognostic information to HIF signature, suggesting that they may operate in the same pathway in vivo.

Sharp1 inhibits HIFs-induced migration in TNBC cells lines and in experimental tumors

We confirmed the role of HIF as metastasis inducer in a cell migration assays *in vitro* in multiple TNBC cells lines. As far as human breast cancer is concerned, both HIF-1α and HIF-2α have been connected to metastasis in breast cancer patients (suggesting perhaps a redundant/cumulative function) (Wong et al., 2011). Others have also claimed that in breast cancer HIF-1α inhibition alone is critical for tumor progression (Lu et al., 2010). In mouse models of breast cancer, there is good genetic evidence is on the relevance of HIF-1α, as conditional knockout in mammary cells of HIF-1α is sufficient to block metastasis in the Polyoma Middle T mouse model (Liao et al., 2007). Beside this, in our assay we coul not appreciate any overt differences between HIF-1α and HIF-2α: infact HIF-2α is required for cell migration and it is regulated by Sharp1 in the same way of HIF-1α.

Sharp1 shows metastasis-suppressor activities also *in vivo*: after tail vein injections, cells overexpressing Sharp1 colonize the lung less efficinetly, in a manner similar to cells depleted of HIFs. Expanding the biological relevance on cell migration and local invasiveness of the primary tumor we show that Sharp1 is instrumental to limit cell invasion to the skin in experimental tumors. Sharp1-tumors then display severely reduced expression of CAIX, the best validated read-out of HIF activity in breast cancer, as well as GLUT1 and HIF-1α (Fig. 13).

Sharp1 is an adaptor protein between HIFs and the 20S-α4 subunit

We show that Sharp1 and HIF-1α binds at endogenous levels in multiple cells lines and that Sharp1 degrades HIFs in a proteasome depedent manner. Moreover Sharp1 increases the binding of HIF1α to the proteasome. In our model Sharp1 binds to HIFs with almost two domains while the bHLH domain is clearly necessary for the binding to the α4 subunit of the proteasome. Conversely Sharp1 deleted of its proteosomal binding domain (ΔbHLH) stabilizes HIF by preventing HIF association to the proteasome. These data suggest that Sharp1-proteosome recognition is causal for Sharp1-induced HIF degradation (Fig. 16).

The mechanism by which Sharp1 leads to HIFs protein degradation in these breast cancer contexts is independent from the oxygen/pVHL pathway as it occurres very efficiently in hypoxia (in which the pathway is turned off. This conclusion is also

in line with the results obtained with genetically modified RCC4 pVHL-/- cells) were we show that Sharp1 mediated degradation is also independent from ubiquitination: cells that are temperature sensitive mutants for the universal E1 Ubiquitin conjugating enzyme Sharp1 is still active to degrade HIF protein; moreover overexpression of Sharp1 does not lead to upregulation of ubiquitinated HIF (as expected in case a Ubiquitin ligase were involved). Many other short living protein, as p53 and p73, are degradated also by an ubiquitin-independent manner (Asher et al., 2005).

In this work we propose that Sharp1 is an essential cellular determinant of the intrinsic instability of HIF proteins, independetly from the oxygen levels.

It is now well known that many tumors shift to glycolytic metabolism even at high oxygen levels, a phenomenon called "Warburg effect" (Dewhirst, 2007). Moreover, is now accepted that intratumoral hypoxia is a cycling phenomena because tumors induce angiogenesis, even if with low efficiency. Cancer cells are then subjected to both normoxia and hypoxia, in a periodic manner (Dewhirst et al., 2008). For these reason it is clear that the regulation of HIF dependent from the oxygen levels cannot explain the entire scenario. Sharp1, acting in parallel to oxygen-dependent mechanisms for HIFs degradation, can be considered a buffer against the effects of hypoxic fluctuations in breast cancer. In tumors where Sharp1 is shut down (by TGFβ/mutant-p53/p63 patway), HIF is stabilized and can starts its trascriptional program, even in normoxic conditions.

HIFs activation represents a final common event in cancer pathogenesis; thus, an issue deserving attention in future studies is what signals and factors other than p63 and TGFβ can regulate Sharp1 activity, stability or expression. Finally, the present work opens interesting scenarios on the possibility to therapeutically enhance Sharp1 expression/activity in order to blunt malignancy of TNBC subtypes otherwise refractory to current therapies.

EXPERIMENTAL PROCEDURES

Genotyping

Offsprings were genotyped by PCR on genomic tail DNA extracted by standard procedures (see Table A for oligo sequences). After in situ, individual embryos were manually dissected with a tungsten wire (FineScienceTools) to eliminate the EXE and ectoplacental cone, thus avoiding maternal contaminations. Epiblast/VE tissues were lysed overnight at 55°C with mild agitation in 10mM Tris/HCl pH=8.0, 50mM KCl, 2mM MgCl2, 0,3% Tween-20, 0,5% NP-40 supplemented with fresh proteinaseK (Invitrogen, 1:40). Lysis volume was adjusted according to the stage: E5.5 20µl, E6.5 40µl. After vortexing, proteinaseK was inactivated 10min at 95°C, quenched on ice, and samples were centrifuged 10min/4°C/10000rcf. 4ul of the fresh supernatants were used for each PCR reaction using EX-Taq polymerase (Takara). For detection of the *Ecto*allele in embryos of early stages, nested PCR was employed when necessary.

Generation of *Ecto* knockout and conditional alleles

To generate the Ecto/Tif1γ targeting vector, a genomic clone spanning exons 2, 3 and 4 was used (Yan et al., 2004). Briefly, a loxP flanked (floxed) PGK-Neo cassette was inserted within the first intron, and a third loxP site was inserted within the fourth intron. The targeting fragment was electroporated into 129/Sv H1 ES cells as described previously (Cammas et al., 2000). After selection, neomycin-resistant ES clones were expanded, and their genomic DNA was screened by PCR. Positive clones were further validated with Southern blotting analysis with two independent probes (not shown). ES cells bearing the correctly targeted allele were injected into C57BL/6

blastocysts to produce chimeric offspring. These were back-crossed with C57BL/6 mice, and their offspring was genotyped by PCR. Mice heterozygous for the targeted allele were then crossed with CMV-Cre transgenic mice (Dupe et al., 1997), and the offspring was analyzed by PCR to identify animals with either complete recombination of the loxP sites (null allele, *Ecto-*) or lacking of the PGK-Neo cassette due to recombination of the first and second loxP sites (conditional allele, Ecto fl). Cre-negative Ecto+/- and Ecto fl/fl mice were subsequently kept on a C57BL/6 background for phenotypic analyses. Animal care was in accordance with our institutional guidelines.

Generation of Ecto-EpiKO and compound Ecto-/-;Smad4-/- and Ecto-/-;NodalΔ600/- embryos

To obtain epiblast-specific *Ecto* knockout embryos, *Sox2-Cre; Ecto+/-* males were crossed with *Ecto fl/fl* females. In this setup the *Sox2-Cre* transgene selectively deletes the floxed alleles in ICM/epiblast cells (Hayashi et al., 2002). Embryos were genotyped after in situ hybridization for *Ecto fl*, *Ecto+*, *Ecto-*, and *Cre* alleles. Embryos were scored as mutants in the presence of *Cre*, *Ecto fl*, *Ecto-* and absence of *Ecto+* alleles. To obtain embryos homozygous null for both *Ecto* and *Smad4*, *Ecto fl/fl;Smad4 fl/fl* (Bardeesy et al., 2006) males were crossed with *CAG-Cre; Ecto+/-; Smad4+/-* females. In this setup, the Cre protein supplied by the mother within the oocyte completely recombinates the paternal floxed alleles after fertilization, irrespective of transgene trasmission ((Sakai and Miyazaki, 1997) and our unpublished observations), raising the expected frequency of compound null embryos to 25%. Embryos were genotyped after in situ hybridization for *Ecto fl* (recognizing also the *Ecto+* allele), *Ecto-*, *Smad4 fl* (recognizing also the *Smad4+* allele), and *Smad4-* alleles. Embryos were scored as compound mutants in the presence of *Ecto-* and *Smad4-* and in the absence of *Ecto fl* and *Smad4 fl* alleles. To obtain *Ecto-/-* embryos with reduced Nodal signaling, *Ecto+/-; Nodal+/-* (lacZ allele (Collignon et al., 1996)) mice were crossed with *Ecto+/-; Nodal+/Δ600* ((Norris et al., 2002) mice. Embryos were genotyped after in situ hybridization for *Ecto+*, *Ecto-*, *lacZ* and *NodalΔ600* alleles.

In situ hybridization

Mouse embryos were staged based on their morphology, considering the morning of the vaginal plug as E0.5. Embryos were manually dissected in ice-cold DEPC-treated phosphatebuffered saline (PBS) and fixed overnight in PBS 4% PFA at 4°C, dehydrated (for storage) rehydrated through methanol series. Whole-mount in situ hybridizations were performed according to http://www.hhmi.ucla.edu/derobertis/ (Xenopus ISH protocol), with minor modifications to ensure efficient genotyping after staining: Day1, post-fixing after proteinaseK treatment was done with 4% PFA only, 1 hour at 4°C; Day3, washes were done with PBS 0,5% Goat Serum (GS, Invitrogen), without AP1 incubation before BM-Purple staining (Roche), and without post-fixation. Embryos were mounted in 80% glycerol and photographed with a Leica DMR microscope equipped with a Leica DC500 camera. For each experiment, at least 5 embryos of every genotype were analyzed.

Immunofluorescence and histology

For immunostaining, embryos were fixed overnight in PBS 4% PFA supplemented of phosphatase inhibitors (Sigma) at 4°C, dehydrated and rehydrated through methanol series. Embryos were permeabilized with two washes in PBS 0,5% NP40, 20min at 4°C, followed by one wash in PBS 0,3% TRITON X100, 20min at RT. After two washes in PBS 0,1% TRITON X100 (PBT), 15min at RT, embryos were blocked with two washes in PBT 10% GS, 1 hour at RT, and incubated overnight with rabbit anti-Ecto primary antibody (Sigma HPA004345, 1:75) in PBT 10% GS or rabbit mAb anti-phospho-Smad2 (CST-3108, 1:50) in PBT 3% BSA. The following day, embryos were washed twice in PBT 2% GS, 15min at 4°C, and five more times in PBT 2% GS, 1 hour at 4°C. Secondary Alexa555 goat anti-rabbit antibody (1:200) was incubated overnight in PBT 5% GS. The third day, embryos were washed five times in PBT, 15min at RT, mounted in 80% glycerol and photographed with a Nikon Eclipse E600 confocal microscope equipped with a Biorad Radiance2000 camera/laser scanning system. Nuclear localizations were confirmed by colocalization with YOYO1 staining (Invitrogen). For histological analysis, deciduae were collected in PBS, fixed in Bouin's overnight, dehydrated, and embedded in paraffin. Serial sections were cut at 6μm and stained with Hematoxylin and Eosin according to standard procedures. Similar procedures were applied to obtain sections of embryos after *in situ*.

Migration and Invasion Assays

For wound healing assays, cells were seeded at low confluency (15%) in 6-well plates in complete medium. For MDA-231 and SUM-159, starting from day 3, cells were starved in medium without serum, for 24 hours. At day 4, confluent cells monolayer was scraped with a P200 tip to obtain a wound in each well, the medium was replaced with fresh serum-free medium with 5ng/ml TGFβ1. For HS578T, starting from day 3, cells were starved in medium without serum, at day 4 confluent cells monolayer was scraped with a P200 tip to obtain a wound in each well, the medium was replaced with either with complete fresh medium or with serum-free medium. After 16-24 hours the cells were fixed with 4% PFA and photographed. For the quantification, we counted cells migrated inside the scratch (pictures were taken at time 0 as reference) in 5 different fields of the wound.

For transwell migration assays, MDA-231 cells were transfected with siRNAs at low confluency (15%) in 10cm dishes. After 8 hours, the medium was changed to no serum medium. At day 3, cells from each plate were seeded in 24-well PET inserts (Falcon, 8.0μm pore size) in medium without serum, with 0.1% BSA and either 10μM SB505124 or 5ng/ml TGFβ1 (as indicated) both in top and bottom chambers. Before plating, inserts and 24-well plates had been incubated 2 hours with serum free medium with 0.1% BSA. Each sample was plated in triplicate (500000 cells/insert). After overnight incubation, cells in the upper side of the top chamber were removed with a cotton swab and the migrated cells (in the lower side of the filter) fixed in 4% PFA. To visualize the cells, the filter was stained with 0.5% Crystal Violet and photographed, then total number of migrated cells were counted.

Stably transfected MCF10A-MII cells were starved overnight (in Assay medium: DMEM/F12 2% HS freshly supplemented with 10μg/ml Insulin, 8.5ng/ml Cholera Toxin, 500ng/ml hydrocortisone). The day after they were trypsinized, counted, resuspended in Assay medium to a concentration of $0.5x10^6$ cells/ml and 400μl (200000 cells/insert) were plated on the top of 24-well PET inserts (Millicell, Millipore). We plated 4 replicas for each sample. After 5 hours, the cells were fixed and stained as above. For the quantification, 4 fields of each filter were then photographed and the surface covered by cells was measured using ImageJ. Each experiment was repeated at least two times with consistent results.

In vivo Assays

Mice were housed in Specific Pathogen Free (SPF) animal facilities and treated in conformity with approved institutional guidelines (U. Padua). For xenograft tumor assay, cells were resuspended in 100μl of PBS and injected in the tail vein of SCID female mice, aged-matched between 5 and 7 weeks. We injected 10 mice for each sample ($3x10^5$ cells for each mouse). After four weeks, animals were sacrificed and lungs removed and fixed. Serial sections of the lungs were cut at a distance of 70 μm from each other, stained with Hematoxylin and Eosin (H&E) and processed for human cytokeratin expression. For orthotopic tumor assays, $8x10^5$ cells (resuspended in 100μl of DMEM) were injected in the fourth inguinal mammary fat pad. After 24 days, primary tumors were removed and analyzed for expression of HIF-1α and its targets by immunohistochemistry. After 30 days from the primary tumor removal, mice lungs were assayed for the presence of metastatic lesions.

Immunohistochemistry

Immunohistochemical cytokeratin staining was performed on formalin-fixed, paraffin-embedded tissue using an indirect immunoperoxidase technique (Bond Polymer Refine Detection; Vision BioSystems, UK). Sections mounted on silanized slides will be dewaxed in xylene, dehydrated in ethanol, boiled in 0.01 M citrate buffer (pH 6.0) for 20 min. in a microwave oven and then incubated with 3% hydrogen peroxide for 5 min. After washing with PBS, they were be incubated in 10% normal BSA for 5 min, followed by incubation for 45 min with monoclonal mouse anti-human Cytokeratin, AE1/AE3 (1:200, BioGenex). After washing, sections were incubated with labeled polymer (Bond Polymer Refine Detection) and diaminobenzidine. The sections were then counterstained with hematoxylin, dehydrated, cleared, and mounted.

For HIF-1α, CAIX and Glut1 staining in experimental primary tumors, tumor masses were fixed in 4% paraformaldehyde, paraffin-embedded and cut in 5μm-thick sections. Sections were re-hydrated and then antigen retrieval was performed by incubation with citrate buffer 0.01M pH6 at 95°C for 20 min. Slides were blocked in 5% BSA for at least 1 hour and then incubated with the primary antibody diluted in BSA 5% over night at 4°C. Primary antibody dilutions were: anti-CAIX (1:1000; Novus Biologicals NB100-417), anti-GLUT1 (1:500; Santa Cruz sc-7903), anti-HIF-

1α (rabbit, 1:100, Sigma-Aldrich Prestige HPA001275). After incubation with primary antibody, sections were washed and incubated for at least 1 hour at room temperature with Alexa dyes-conjugated secondary Ab (diluted 1:1000 in BSA 5%). Tissues were counterstained with DAPI for 15 min. at room temperature (1:10000, Sigma-Aldrich) to visualize nuclei. The specificity of each staining procedure was confirmed by replacing the primary antibodies with the specific isotype control.

Coimmunoprecipitation Assays

To study endogenous Sharp1/HIF-1α binding, the cells were treated with 800μM $CoCl_2$ (Sigma) for 24 hours. One confluent 15cm dish of MDA-231 cells was lysed by sonication in 25mM Hepes (pH 7.8), 5mM EDTA, 100mM KCl, 0.1% NP40, 10% glycerol freshly supplemented with protease inhibitor cocktail (Roche), phosphatase inhibitor cocktail II (Sigma), 1mM DTT, 10μM MG-132 (Sigma) and 6mM $MgCl_2$. Prior to immunoprecipitation, proteinA-bound Sepharose beads were incubated overnight with Sharp1 antibody (custom antibody, Bethyl) in PBS with 5% BSA and 0.05% CHAPS at 4°C. We then added 6mM $MgCl_2$ to the extracts before immunoprecipitation with protein-A Sepharose at 4°C for 4 hours. After three washings in binding buffer, co-purified proteins were analysed by Western Blotting.

For domain mapping, ubiquitination assays and HIF-1α/α4-subunit interations cells were harvested in Ub-lysis buffer: 50mM Hepes (pH 7.8), 200mM NaCl, 5mM EDTA, 1% NP-40, 5% glycerol, freshly complemented with 1mM DTT (except for experiments with DSP treatment), protease inhibitor cocktail, phosphatase inhibitor cocktail II. For ubiquitination assays, HEK293T cells were transfected with the indicated plasmids at the following doses: Sharp1 (2.4μg/10cm dish), VHL (2.4μg/10cm dish), Flag-HIF-1α (600ng/10cm dish), HA-ubiquitin (5μg/10cm dish), pMix.2Δ7-lux (600ng/10 cm dish) for normalization. Total amount of tranfected DNA was kept constant by adding pCSP1 plasmid. Before harvesting, cells were treated for 8 hours with proteasome inhibitor mix: 20μM MG-132, 20μM MG-115 (Sigma), 20μM lactacystin (Sigma). Cells were harvested in Ub-lysis buffer supplemented with 10μM MG-132. Lysates were immunoprecipitated at 4°C for 4 hours with anti-Flag M2 affinity resin (Sigma) that had been previously blocked overnight in PBS with 2% BSA and 0.05% CHAPS at 4°C. After immunoprecipitation the beads were washed three times (2 minutes rotating at room temperature) with Washing buffer (50mM

Hepes pH 7.8, 500mM NaCl, 1mM EDTA, 1% NP-40, 5% glycerol, 2.5mM $MgCl_2$). Free and ubiquitinated HIF-1α forms were then analyzed by Western Blotting.

For HIF-1α/α4-subunit interaction after endogenous Sharp1 depletions, HEK293T cells were transfected with control or Sharp1 siRNA at day 1 and the medium changed after 8 hours. The day after cells were transfected with Flag-HIF-1α plasmid (600ng/10cm dish) and pMix.2Δ7-lux (600ng/10 cm dish). At day 4 cells were lysed and luciferase assays were used to normalize for transfection efficiency. Normalized extracts were brought to the same volume with Ub-Lysis buffer before immunoprecipitation with anti-Flag M2 affinity resin (same IP conditions as above). For the mapping of α4-subunit-interacting domain of Sharp1, HEK293T cells were transfected with 1μg/10cm dish of either wild-type or mutant forms of Flag-Sharp1. For Sharp1/HIF-1α interaction experiments, HEK293T cells were transfected with V5-HIF-1α (2μg/10cm dish), Flag-Sharp1 (wt or mutants, 4μg/10cm dish), pMix.2Δ7-lux for normalization (600ng/10 cm dish) and harvested in the presence of 10μM MG-132. Cell extract was then subjected to immunoprecipitation using anti-V5 resin (Sigma) using the same conditions of Flag-M2 resin.

To detect HIF-1α/20Sα4-subunit binding in the presence of overexpressed Sharp1, HEK293T cells were transfected with Flag-HIF-1α (1μg/10cm dish) and Sharp1 (4μg/10cm dish). Before harvesting, cells were washed twice with large volumes of pre-warmed PBS and then incubated at room temperature for 30 minutes with freshly resuspended 2.5 mM DSP (Thermo). The crosslinking reaction was stopped by adding Tris pH7.5 to a final concentration of 50mM for 10 min. After sonication, the extracts were processed, immunoprecipitated with Flag-M2 resin as above and finally washed three times with 50mM Hepes (pH 7.8), 500mM NaCl, 1mM EDTA, 1% NP-40, 5% glycerol, 2.5mM $MgCl_2$. To allow protein visualization of the copurified protein, the beads were incubated with 50mM DTT at 37°C for 30 minutes before boiling in SDS protein sample buffer for 5 minutes. To visualize HIF-1α/20Sα4 complexes in MDA-231 cells, control and Sharp1-overexpressing cells were treated with DSP as above and endogenous HIF-1α was immunoprecipitated using 5μg of rabbit polyclonal antibody from Abcam (ab2185, or total rabbit IgG as control) in Ub-Lysis buffer at 4°C for 4 hours. The beads were then washed three times with 50mM Hepes (pH 7.8), 500mM NaCl, 1mM EDTA, 1% NP-40, 5% glycerol, 2.5mM $MgCl_2$ and the amine-crosslinking was reversed by treating the

beads with 50mM DTT at 37°C for 30 minutes before boiling in SDS protein sample buffer for 5 minutes.

Transient transfections and stable transduction

siRNAs (10 pmoles/cm^2) were transfected using Lipofectamine-RNAiMax reagent (Invitrogen), a list of the siRNAs is provided in Table 8. For transient overexpression studies, DNA plasmids were transfected using LT1 reagent (Mirus) except for ts20 cells that were transfected with JetPEI reagent (Polyplus).

Stable overexpression and silencing were obtained by transducing MDA-231 cells with retroviral or lentiviral vectors. The efficiency of knockdown or overexpression was controlled by qPCR or with Western Blotting.

Plasmids

pCDNA3.1Zeo-mSHARP1, pFlagCMV6b-mSharp1 were a gift from Y.Kato. Flag-HIF-1α wt was a gift from S.K. Libutti. pGL2-hVEGFpromoter was a gift from Y. Maeda. pRRLsin.ppts.hCMV.gfppre, pMDG.VSVG, pCMV8.74, pMD2env were gift from L. Naldini. pGL3b-Mix.2Δ7promoter-luciferase is minimal promoter sequence of Mix.2 gene (-46, +1) cloned upstream of luciferase cds(Cordenonsi et al., 2003), HA-Ubiquitin was described in(Dupont et al., 2009). pLKO.1puro was obtained from Addgene (#8453).

HA-hHIF-1α wt, HA-hHIF-1α P402A/P564A and HA-hHIF-2α wt coding sequences were subloned in pEBB from original pCDNA3-vectors (Addgene #18949, #18955, #18950 respectively). Flag epitope was substituted with V5-tag in Flag-HIF-1α wt plasmid to generate V5-hHIF-1α wt. For stable overexpression hHIF1α P402A/P564A and hHIF-2α P405A/P531A were subcloned into pRRL; mSharp1 and TAp63α were cloned into pLPCX vector. For transient overexpression, mSharp1 and HA-VHL were cloned into pRK5. HA-VHL was subcloned from Addgene construct #19234.

To create the constructs for HIF-1α and HIF-2α stable knock-down in MDA-231, the interfering sequences B for HIF-1α and A for HIF-2α (see Table 8) were cloned into pLKO.1 vector according to manifacturer's protocol. For MCF10A-MII cells experiments, Sharp1 interfering sequence A (or control sequence) was cloned in pSuper-Retro-puro vector, while HIF-1α sequences A and B (or control sequence) were cloned in pSuper-Retro-hygro vector to allow double selection with puromycin

and hygromycin. Each contruct was controlled by sequencing.

Cell lines and Treatments

HEK293T and HEK293Tgp cells were cultured in DMEM 10%FBS, MDA-231 cells in DMEM/F12 10%FBS, MCF10Tk1 (MII) cells in DMEM/F12 5%HS freshly supplemented with 10μg/ml Insulin, 20ng/ml EGF, 8.5ng/ml Cholera Toxin, 500ng/ml hydrocortisone. SUM-159PT were kindly provided by Dr. S. Ethier and were cultured in F12 5%FBS freshly supplemented with 5μg/ml Insulin and 1μg/ml hydrocortisone. Hs578T cells were obtained from ICLC and mainteined in DMEM 10%FBS freshly supplemented with 10μg/ml Insulin. Control shGFP-MDA and shp53-MDA-231 cells have been previously described(Adorno et al., 2009).

RCC4 and RCC4/VHL cells were a kind gift from Dr. Kaelin and were cultivated in DMEM 10%FBS. The temperature-sensitive ts20 Balb/C 3T3 clone A31 cell line was a gift from Dr. C. Borner and was mainteined in DMEM 10%FBS medium at the permissive temperature of 34°C. For the ubiquitin-independent degradation of HIF-1α, ts20 cells were first incubated for 8 hours at the non-permissive temperature of 39°C (to inactivate E1 enzyme), then transfected with the indicated plasmids and placed again at 39°C for additional 48 hours. A parallel experiment was run at 34°C as control.

For TGF-β treatment, cells were starved for 24 hours in serum-free medium and then either treated with 5ng/ml TGF-β1 or with serum-free medium for 4 hours before harvesting. To visualize otherwise undetectable HIF-1α levels, SUM159 and Hs578T were cultivated at 1% O_2 for 24 hours before harvesting.

Western Blotting and Antibodies

To monitor endogenous and overexpressed protein regulations, cells were harvested in Ub-lysis buffer freshly complemented with 1mM DTT, protease inhibitor cocktail, phosphatase inhibitor cocktail II and 20μM MG-132. Proteins were loaded according to Bradford quantification (for endogenous protein) or according to luciferase quantification (for overexpressed protein).

We used the following antibodies: anti-Sharp1 (Bethyl, custom), anti-HIF-1α (Becton Dickinson, #610958), anti-HIF-2a (Novus Biologicals, NB100-12), anti-HA (Covance, HA.11), Flag-M2 (Sigma), V5-HRP (Sigma), anti-GAPDH (Millipore, MAB374), anti-p53 (Santa Cruz Biotechnology, DO-1), anti-α4 subunit (Enzo Life

Sciences, MCP34), anti-PAI1 (Becton Dickinson, 612025).

Breast cancer data collection and processing

We retrieved breast cancer datasets from Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo), containing informations about patients' clinical outcome and gene expression data obtained with Affymetrix HG-U133A arrays (Table 1).

Duplicate samples between distinct datasets were removed, and the datasets were renamed/modified as follows:

- *Stockholm*(Pawitan et al., 2005) was re-named as *KI_Stockholm* (Karolinska Institutet Stockholm);
- *EMC-286*(Wang et al., 2005) and *EMC-58*(Minn et al., 2007) were merged to create *EMC-344* (Erasmus Medical Center);
- *MSK*(Minn et al., 2005) was renamed as *MSKCC* (Memorial Sloan-Kettering Cancer Center);
- *Uppsala-Miller*(Miller et al., 2005), *Ivshina-Miller*(Ivshina et al., 2006), and *Loi*(Loi et al., 2007; Loi et al., 2008; Loi et al.)have been split into *KI_Uppsala*, comprising all 258 unique patients from the Uppsala University Hospital, and *OXF*, composed of 178 samples collected at the John Radcliffe Hospital in Oxford and formerly part of GSE6532;
- *Sotiriou* dataset of Ref (33)(Sotiriou et al., 2006) is entirely included in GSE6532;
- *Desmedt*(Desmedt et al., 2007) was re-named as *TRANSBIG* (after the consortium of cancer centers where samples have been collected);
- *Schmidt*(Schmidt et al., 2008) was re-named as *Mainz* (Johannes Gutenberg University in Mainz);
- *Veridex* dataset (Zhang et al., 2009) was not modified.

The re-organization of the original datasets resulted in a meta-dataset comprising 1555 unique samples (Table 2).

To standardize clinical information among the various datasets, we redefined the outcome descriptions based on the clinical annotations of each individual study. Specifically, we defined two major types of events, i.e., *metastasis* and *survival*, and censored all outcomes at 7 years. *Metastasis* is associated to the metastatic spread and includes the terms *recurrence free survival*, *metastasis free survival*, *distant*

metastasis free survival, *distant recurrence*, and *time to distant metastasis*. *Survival* is intended as death because of cancer and includes *overall survival*, *disease free survival*, and *disease specific survival*.

TNBCs have been defined among the 349 ER- samples using the clinical annotations of PR and HER2, where available, or the expression level of ERBB2, when the status of HER2 and PR were not annotated. Specifically, we labeled as HER2- all those samples with an expression signal of ERBB2 probe-set (216836_s_at) lower than the 99th percentile of ERBB2 signal distribution in clinical samples annotated as HER2-. This threshold has been selected to correctly classify 98% of clinically annotated TNBC. This resulted in a cohort of 250 primary TNBCs.

Since raw data (e.g., CEL files) were available for all samples, expression values of the meta-dataset were generated from fluorescence signals using the robust multi-array average procedure (RMA). Specifically, intensity levels have been background adjusted, normalized using quantile normalization, and $\log_2$ expression values calculated using median polish summarization (Irizarry et al., 2003).

To identify two groups of tumors with either high or low levels of *Sharp1/CyclinG2* we used the classifier described in Ref(Adorno et al., 2009). Briefly, we defined a classification rule based on summarizing the standardized expression levels of Sharp1 and CyclinG2 into a combined score with zero mean. Tumors were then classified as *Sharp1/CyclinG2 Low* if the combined score was negative and as *Sharp1/CyclinG2 High* if the combined score was positive. This classification was applied to $\log_2$ expression values obtained using RMA on the meta-dataset described above. The same classifier was used to identify two groups of tumors with either high or low HIFs-activity summarizing the standardized expression levels of genes in the *HIFs* signature (Table 4) into a combined score with zero mean. Tumors were classified as *HIFs* if the combined score was negative and as *HIFs High* if the combined score was positive. These classifications were applied to $\log_2$ expression values obtained using RMA on the meta-dataset described above.

Survival analysis

To evaluate the prognostic value of the *Sharp1/CyclinG2* and *HIFs* signatures, we estimated, using the Kaplan-Meier method(Kalbfleisch and Prentice, 1980), the probabilities that patients would remain free of metastatic events (*metastasis*) or of death (*survival*). To confirm these findings, the Kaplan-Meier curves were compared

using the log-rank or Mantel-Haenszel test(Harrington and Fleming, 1982). P-values were calculated according to the standard normal asymptotic distribution. When comparing *Sharp1/CyclinG2 signature High* and *Sharp1/CyclinG2 signature Low* groups, the group with low *Sharp1/CyclinG2* mRNA levels displayed a significantly higher probability (at a significance level $\alpha=5x10^{-2}$) to develop metastasis and reduced survival. When comparing *HIFs High* and *HIFs Low* groups, the group with high-*HIFs* activity displayed a significantly higher probability (at a significance level $\alpha=5x10^{-2}$) to develop metastasis and reduced survival. Survival analysis was performed using *survcomp* package of R and the Kaplan-Meier plots were drawn using the *km.coxph.plot* function of the same package. Comparisons between Kaplan-Meyer curves were carried out using the log-rank test of the *surv_test* function (*coin* R package), i.e., testing the null hypothesis of no difference against the two-sided alternative.

Multivariate analysis using a Cox proportional-hazards model

The analysis of the risk of metastasis for the TNBC samples was conducted using Cox proportional-hazards regression modeling. In particular, we examined the relationship between metastasis free survival and the *Sharp1/CyclinG2 signature* predictor and the available clinical variables, i.e., tumor diameter and age. We fitted Cox proportional-hazards regression model first by using clinical variables only (Model 1), and then adding the *Sharp1/CyclinG2 signature predictor* (Model 2).

Over-representation analysis

Over-representation analysis was performed using Gene Set Enrichment Analysis and a Fisher's exact test. In both approaches, gene sets were taken from the Molecular Signatures Database (http://www.broadinstitute.org/gsea/msigdb/index.jsp) and from previously published gene signatures. In particular, we investigated whether low Sharp1 and CyclinG2 expressions were associated with elevated expression of TGF-β, HIFs, mutant-p53 (Miller et al., 2005), Wnt/TCF4 (van de Wetering et al., 2002), E2F3 (Bild et al., 2006), H-Ras (Bild et al., 2006) and of 248 gene sets derived from the BioCarta pathway database. GSEA software (http://www.broadinstitute.org/gsea/index.jsp) (Subramanian et al., 2005) was applied on $\log_2$ expression data of TNBC samples classified as *Sharp1/CyclinG2 signature High* or as *Sharp1/CyclinG2 signature Low*. GSEA returned 3 gene sets up-regulated

in phenotype *Sharp1/CyclinG2 signature Low* (Enrichment Score ES<0) and significantly enriched at FDR<25% when using Signal2Noise as metric and 10,000 permutations of phenotype labels. As an alternative approach, we tested the over-representation of signaling pathways in genes up-regulated in samples with low *Sharp1/CyclinG2* expression using a Fisher's exact test. Briefly, genes up-regulated in *Sharp1/CyclinG2 signature Low* TNBCs have been identified using SAM algorithm coded in the *samr* R package (Tusher et al., 2001). In SAM, we estimated the percentage of false positive predictions with 1000 permutations and set the q-value threshold at 0.01. SAM comparison of gene expression profiles in *Low Sharp1/CyclinG2* and *High low Sharp1/CyclinG2* TNBC samples resulted in 2804 unique genes (3303 annotated probe set) up-regulated in *Low Sharp1/CyclinG2* samples. Considering these S overexpressed single-symbol-annotated genes, the over-representation of pre-defined pathway signatures has been calculated as the hypergeometric probability of having α genes for a specific pathway in S, under the null hypothesis that they were picked out randomly from the N total genes of the microarray. Over-representation has been quantified using the one-sided Fisher's exact test coded in the *phyper* function of R *stats* package and p-values adjusted for false discovery rate with the *p.adjust* function (R *stats* package). The Fisher's test indicated that the HIF gene signature (Table 4) is significantly enriched at FDR<5% in genes up-regulated in samples with *Low Sharp1/CyclinG2*.

Statistical analysis

We tested if *Sharp1/CyclinG2* and *HIFs* signatures are two independent predictors of metastasis free survival using a multivariate analysis. We fitted Cox proportional-hazards regression models first by using the *HIFs* score (Model 1), and then adding the *Sharp1/CyclinG2 signature* score (Model 2). Model 1 and Model 2 may be compared to assess whether the *Sharp1/CyclinG2 signature* adds additional prognostic information over the *HIFs* signature. The difference between the residual deviances of models constructed without (i.e., Model 1) and with the *Sharp1/CyclinG2 signature* (i.e., Model 2) was equal to 1.0427 and did not exceed the .95 quantile of the chi-square distribution with one degree of freedom (p-value = 0. 3072). Thus, there is no evidence that the *Sharp1/CyclinG2 signature* adds prognostic value to the model constructed using the *HIF* signature score.

Statistical analysis for correlations and comparison between experimental groups were made with R and Prism software (Graphpad). For correlation studies, Pearson's correlation of genes expression values in the TNBC meta-dataset has been computed using the *cor* of R. The statistical significance of the linear dependence between the expression levels of *Sharp1*, HIFs target genes, and loading controls has been assessed with the *cor.test* function of R, testing the null hypothesis that the Pearson's correlation is zero.

RNA and Microarray analysis

For qRT-PCR analysis, total RNA was purified with Trizol (Invitrogen) and Poly(A)$^+$-RNA was retrotranscribed with M-MLV Reverse Transcriptase (Invitrogen) and oligo-d(T) primers. Real-time PCR were perfomed on a RotorGene 3000 (Corbett) using the FastStart SYBR Green Master Mix (Roche), for standard RT-PCR please refer to (Adorno et al., 2009). The primers are listed in Table 9.

For Microarray analysis, four replicas of MDA-shGFP, MDA-shHIF and MDA-Sharp1 cells were prepared. Total RNA was extracted using Trizol (Invitrogen) according to the manufacturer's instructions. Four biological mRNA replicates for each group were hybridized on Affymetrix GeneChip Human Genome HG-U133 Plus 2.0 arrays. Sample preparation for microarray hybridization was carried out as described in the Affymetrix GeneChip® Expression Analysis Technical Manual. All data analyses were performed in R using Bioconductor libraries and R statistical packages. Probe level signals have been converted to expression values using RMA(Irizarry et al., 2003). Raw data are available at Gene Expression Omnibus GSE33950.

Differentially expressed genes have been identified using the Rank Product method coded in the *RankProd* R package(Breitling et al., 2004) with default parameters. Rank Product is a non-parametric statistic that detects items consistently highly ranked in a number of lists as, e.g., genes that are consistently up-regulated in replicate experiments. To identify genes whose expression is modified by Sharp1-overexpression or HIF-depletion, we compared the expression profile MDA-shHIF or MDA-Sharp1 cells with MDA-shGFP and selected those transcripts whose estimated percentage of false positive predictions (i.e., False Discovery Rate, FDR), was ≤0.05. This selection was further refined setting the lower limit for fold change induction (or reduction) to 1.5. Using this combined filter, we identified 1918 and 1700 probe-sets

differentially regulated by Sharp1-overexpression and HIF-depletion, respectively. Comparing these two independent lists of probes, we found a highly statistical significant overlap of 234 probe-sets ($p<e^{-72}$ using *phyper* function of R *stats* package for one-sided Fisher's exact test), 207 of which fully annotated in the *hgu133plus2.db* Bioconductor package (Tables 6 and 7).

Sweave documents for these analyses are attached as "Microarray Sweave.pdf" and "Microarray Sweave.Rnw".

Sharp1 signature

To derive a *Sharp1-activity signature* (*Sharp1* signature) we selected transcripts down-regulated in MDA-Sharp1 cells (as compared to MDA-shGFP at FDR≤0.05 and fold change ≤-1.5), negatively correlated with Sharp1 expression levels in the TNBC dataset, and not present in the *HIF* signature (Table 4). This combined filter resulted in a 27-gene signature. The same classifier used for the *Sharp1/CyclinG2* and *HIFs* signatures was used to identify two groups of tumors with either high or low Sharp1 activity summarizing the standardized expression levels of genes in the *Sharp1-activity* signature (Table 4) into a combined score with zero mean. Tumors were classified as *Sharp1-activity Low* if the combined score was negative and as *Sharp1-activity High* if the combined score was positive. These classifications were applied to log2 expression values obtained using RMA on the meta-dataset described above. When comparing *Sharp1-activity High* and *Sharp1-activity Low* groups, the group with high-*Sharp1* activity displayed a significantly higher probability (at a significance level $\alpha=5x10^{-2}$) to develop metastasis. Survival analysis was performed as previously described. The *Sharp1-activity* signature did not add prognostic value to the model constructed using the *HIFs* signature score. Indeed, the difference between the residual deviances of *HIFs* models constructed without (i.e., Model 1) and with the *Sharp1-activity* signature score (i.e., Model 2) was equal to 0.0688 and did not exceed the .95 quantile of the chi-square distribution with one degree of freedom (p-value = 0.7931).

REFERENCES

Adorno, M., Cordenonsi, M., Montagner, M., Dupont, S., Wong, C., Hann, B., Solari, A., Bobisse, S., Rondina, M.B., Guzzardo, V., *et al*. (2009). A Mutant-p53/Smad complex opposes p63 to empower TGFbeta-induced metastasis. Cell *137*, 87-98.

Aprelikova, O., Chandramouli, G.V., Wood, M., Vasselli, J.R., Riss, J., Maranchie, J.K., Linehan, W.M., and Barrett, J.C. (2004). Regulation of HIF prolyl hydroxylases by hypoxia-inducible factors. J Cell Biochem *92*, 491-501.

Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nature reviews *10*, 91-103.

Asher, G., Tsvetkov, P., Kahana, C., and Shaul, Y. (2005). A mechanism of ubiquitin-independent proteasomal degradation of the tumor suppressors p53 and p73. Genes Dev *19*, 316-321.

Bild, A.H., Yao, G., Chang, J.T., Wang, Q., Potti, A., Chasse, D., Joshi, M.B., Harpole, D., Lancaster, J.M., Berchuck, A., *et al*. (2006). Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature *439*, 353-357.

Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett *573*, 83-92.

Brennan, J., Lu, C.C., Norris, D.P., Rodriguez, T.A., Beddington, R.S., and Robertson, E.J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. Nature *411*, 965-969.

Buffa, F.M., Camps, C., Winchester, L., Snell, C.E., Gee, H.E., Sheldon, H., Taylor, M., Harris, A.L., and Ragoussis, J. microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. Cancer Res *71*, 5635-5645.

Cho, S., Choi, Y.J., Kim, J.M., Jeong, S.T., Kim, J.H., Kim, S.H., and Ryu, S.E. (2001). Binding and regulation of HIF-1alpha by a subunit of the proteasome complex, PSMA7. FEBS Lett *498*, 62-66.

Chowdary, D.R., Dermody, J.J., Jha, K.K., and Ozer, H.L. (1994). Accumulation of p53 in a mutant cell line defective in the ubiquitin pathway. Mol Cell Biol *14*, 1997-2003.

Cordenonsi, M., Dupont, S., Maretto, S., Insinga, A., Imbriano, C., and Piccolo, S. (2003). Links between tumor suppressors: p53 is required for TGF-beta gene responses by cooperating with Smads. Cell *113*, 301-314.

Desmedt, C., Piette, F., Loi, S., Wang, Y., Lallemand, F., Haibe-Kains, B., Viale, G., Delorenzi, M., Zhang, Y., d'Assignies, M.S., *et al*. (2007). Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res *13*, 3207-3214.

Dewhirst, M.W., Cao, Y., and Moeller, B. (2008). Cycling hypoxia and free radicals regulate angiogenesis and radiotherapy response. Nat Rev Cancer *8*, 425-437.

Dunn, N.R., Vincent, S.D., Oxburgh, L., Robertson, E.J., and Bikoff, E.K. (2004). Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. Development *131*, 1717-1728.

Dupont, S., Mamidi, A., Cordenonsi, M., Montagner, M., Zacchigna, L., Adorno, M., Martello, G., Stinchfield, M.J., Soligo, S., Morsut, L., *et al.* (2009). FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination. Cell *136*, 123-135.

Dupont, S., Zacchigna, L., Cordenonsi, M., Soligo, S., Adorno, M., Rugge, M., and Piccolo, S. (2005). Germ-layer specification and control of cell growth by Ectodermin, a Smad4 ubiquitin ligase. Cell *121*, 87-99.

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

Gao, P., Zhang, H., Dinavahi, R., Li, F., Xiang, Y., Raman, V., Bhujwalla, Z.M., Felsher, D.W., Cheng, L., Pevsner, J., *et al.* (2007). HIF-dependent antitumorigenic effect of antioxidants in vivo. Cancer Cell *12*, 230-238.

Gordan, J.D., and Simon, M.C. (2007). Hypoxia-inducible factors: central regulators of the tumor phenotype. Curr Opin Genet Dev *17*, 71-77.

Gressner, O., Schilling, T., Lorenz, K., Schulze Schleithoff, E., Koch, A., Schulze-Bergkamen, H., Lena, A.M., Candi, E., Terrinoni, A., Catani, M.V., *et al.* (2005). TAp63alpha induces apoptosis by activating signaling via death receptors and mitochondria. EMBO J *24*, 2458-2471.

Gupta, G.P., and Massague, J. (2006). Cancer metastasis: building a framework. Cell *127*, 679-695.

Guzman-Ayala, M., Ben-Haim, N., Beck, S., and Constam, D.B. (2004). Nodal protein processing and fibroblast growth factor 4 synergize to maintain a trophoblast stem cell microenvironment. Proc Natl Acad Sci U S A *101*, 15656-15660.

Harrington, D., and Fleming, T. (1982). A class of rank test procedures for censored survival data. Biometrika *69*.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics *4*, 249-264.

Itoh, S., and ten Dijke, P. (2007). Negative regulation of TGF-beta receptor/Smad signal transduction. Current opinion in cell biology *19*, 176-184.

Ivshina, A.V., George, J., Senko, O., Mow, B., Putti, T.C., Smeds, J., Lindahl, T., Pawitan, Y., Hall, P., Nordgren, H., *et al.* (2006). Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. Cancer Res *66*, 10292-10301.

Kaelin, W.G., Jr., and Ratcliffe, P.J. (2008). Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. Mol Cell *30*, 393-402.

Kalbfleisch, J., and Prentice, R. (1980). The statistical analysis of failure time data. Wiley Publication.

Kong, X., Alvarez-Castelao, B., Lin, Z., Castano, J.G., and Caro, J. (2007). Constitutive/hypoxic degradation of HIF-alpha proteins by the proteasome is independent of von Hippel Lindau protein ubiquitylation and the transactivation activity of the protein. J Biol Chem *282*, 15498-15505.

Laughner, E., Taghavi, P., Chiles, K., Mahon, P.C., and Semenza, G.L. (2001). HER2 (neu) signaling increases the rate of hypoxia-inducible factor 1alpha (HIF-1alpha) synthesis: novel mechanism for HIF-1-mediated vascular endothelial growth factor expression. Mol Cell Biol *21*, 3995-4004.

Liu, Y., Sato, F., Kawamoto, T., Fujimoto, K., Morohashi, S., Akasaka, H., Kondo, J., Wu, Y., Noshiro, M., Kato, Y., *et al.* Anti-apoptotic effect of the basic helix-loop-helix (bHLH) transcription factor DEC2 in human breast cancer cells. Genes Cells *15*, 315-325.

Liu, Y.V., Baek, J.H., Zhang, H., Diez, R., Cole, R.N., and Semenza, G.L. (2007). RACK1 competes with HSP90 for binding to HIF-1alpha and is required for O(2)-independent and HSP90 inhibitor-induced degradation of HIF-1alpha. Mol Cell *25*, 207-217.

Loi, S., Haibe-Kains, B., Desmedt, C., Lallemand, F., Tutt, A.M., Gillet, C., Ellis, P., Harris, A., Bergh, J., Foekens, J.A., *et al*. (2007). Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. J Clin Oncol *25*, 1239-1246.

Loi, S., Haibe-Kains, B., Desmedt, C., Wirapati, P., Lallemand, F., Tutt, A.M., Gillet, C., Ellis, P., Ryder, K., Reid, J.F., *et al*. (2008). Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC Genomics *9*, 239.

Loi, S., Haibe-Kains, B., Majjaj, S., Lallemand, F., Durbecq, V., Larsimont, D., Gonzalez-Angulo, A.M., Pusztai, L., Symmans, W.F., Bardelli, A., *et al*. PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer. Proc Natl Acad Sci U S A *107*, 10208-10213.

Lowe, L.A., Yamada, S., and Kuehn, M.R. (2001). Genetic dissection of nodal function in patterning the mouse embryo. Development *128*, 1831-1843.

Lu, C.C., and Robertson, E.J. (2004). Multiple roles for Nodal in the epiblast of the mouse embryo in the establishment of anterior-posterior patterning. Dev Biol *273*, 149-159.

Massague, J. (2000). How cells read TGF-beta signals. Nature reviews *1*, 169-178.

Mesnard, D., Guzman-Ayala, M., and Constam, D.B. (2006). Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning. Development *133*, 2497-2505.

Miller, L.D., Smeds, J., George, J., Vega, V.B., Vergara, L., Ploner, A., Pawitan, Y., Hall, P., Klaar, S., Liu, E.T., *et al*. (2005). An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. Proc Natl Acad Sci U S A *102*, 13550-13555.

Minn, A.J., Gupta, G.P., Padua, D., Bos, P., Nguyen, D.X., Nuyten, D., Kreike, B., Zhang, Y., Wang, Y., Ishwaran, H., *et al*. (2007). Lung metastasis genes couple breast tumor size and metastatic spread. Proc Natl Acad Sci U S A *104*, 6740-6745.

Minn, A.J., Gupta, G.P., Siegel, P.M., Bos, P.D., Shu, W., Giri, D.D., Viale, A., Olshen, A.B., Gerald, W.L., and Massague, J. (2005). Genes that mediate breast cancer metastasis to lung. Nature *436*, 518-524.

Monney, L., Otter, I., Olivier, R., Ozer, H.L., Haas, A.L., Omura, S., and Borner, C. (1998). Defects in the ubiquitin pathway induce caspase-independent apoptosis blocked by Bcl-2. J Biol Chem *273*, 6121-6131.

Muller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M.E., McClanahan, T., Murphy, E., Yuan, W., Wagner, S.N., *et al*. (2001). Involvement of chemokine receptors in breast cancer metastasis. Nature *410*, 50-56.

Muller, P.A., Caswell, P.T., Doyle, B., Iwanicki, M.P., Tan, E.H., Karim, S., Lukashchuk, N., Gillespie, D.A., Ludwig, R.L., Gosselin, P., *et al*. (2009). Mutant p53 drives invasion by promoting integrin recycling. Cell *139*, 1327-1341.

Narimatsu, M., Bose, R., Pye, M., Zhang, L., Miller, B., Ching, P., Sakuma, R., Luga, V., Roncari, L., Attisano, L., *et al*. (2009). Regulation of planar cell polarity by Smurf ubiquitin ligases. Cell *137*, 295-307.

Norris, D.P., Brennan, J., Bikoff, E.K., and Robertson, E.J. (2002). The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo. Development *129*, 3455-3468.

Padua, D., Zhang, X.H., Wang, Q., Nadal, C., Gerald, W.L., Gomis, R.R., and Massague, J. (2008). TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. Cell *133*, 66-77.

Pawitan, Y., Bjohle, J., Amler, L., Borg, A.L., Egyhazi, S., Hall, P., Han, X., Holmberg, L., Huang, F., Klaar, S., *et al*. (2005). Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. Breast Cancer Res *7*, R953-964.

Piccolo, S., Agius, E., Leyns, L., Bhattacharyya, S., Grunz, H., Bouwmeester, T., and De Robertis, E.M. (1999). The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. Nature *397*, 707-710.

Rey, S., and Semenza, G.L. (2010). Hypoxia-inducible factor-1-dependent mechanisms of vascularization and vascular remodelling. Cardiovasc Res *86*, 236-242.

Rodriguez, T.A., Srinivas, S., Clements, M.P., Smith, J.C., and Beddington, R.S. (2005). Induction and migration of the anterior visceral endoderm is regulated by the extra-embryonic ectoderm. Development *132*, 2513-2520.

Sato, F., Bhawal, U.K., Kawamoto, T., Fujimoto, K., Imaizumi, T., Imanaka, T., Kondo, J., Koyanagi, S., Noshiro, M., Yoshida, H., *et al*. (2008). Basic-helix-loop-helix (bHLH) transcription factor DEC2 negatively regulates vascular endothelial growth factor expression. Genes Cells *13*, 131-144.

Schmidt, M., Bohm, D., von Torne, C., Steiner, E., Puhl, A., Pilch, H., Lehr, H.A., Hengstler, J.G., Kolbl, H., and Gehrmann, M. (2008). The humoral immune system has a key prognostic impact in node-negative breast cancer. Cancer Res *68*, 5405-5413.

Schmierer, B., and Hill, C.S. (2005). Kinetic analysis of Smad nucleocytoplasmic shuttling reveals a mechanism for transforming growth factor beta-dependent nuclear accumulation of Smads. Molecular and cellular biology *25*, 9845-9858.

Sdek, P., Ying, H., Chang, D.L., Qiu, W., Zheng, H., Touitou, R., Allday, M.J., and Xiao, Z.X. (2005). MDM2 promotes proteasome-dependent ubiquitin-independent degradation of retinoblastoma protein. Mol Cell *20*, 699-708.

Semenza, G.L. (2007). Hypoxia-inducible factor 1 (HIF-1) pathway. Sci STKE *2007*, cm8.

Semenza, G.L. (2010). Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. Oncogene *29*, 625-634.

Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., Nordgren, H., Farmer, P., Praz, V., Haibe-Kains, B., *et al*. (2006). Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst *98*, 262-272.

Soussi, T., and Lozano, G. (2005). p53 mutation heterogeneity in cancer. Biochem Biophys Res Commun *331*, 834-842.

Su, X., Chakravarti, D., Cho, M.S., Liu, L., Gi, Y.J., Lin, Y.L., Leung, M.L., El-Naggar, A., Creighton, C.J., Suraokar, M.B., *et al*. TAp63 suppresses metastasis through coordinate regulation of Dicer and miRNAs. Nature *467*, 986-990.

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

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

Uchida, T., Rossignol, F., Matthay, M.A., Mounier, R., Couette, S., Clottes, E., and Clerici, C. (2004). Prolonged hypoxia differentially regulates hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha expression in lung epithelial cells: implication of natural antisense HIF-1alpha. J Biol Chem *279*, 14871-14878.

van 't Veer, L.J., Dai, H., van de Vijver, M.J., He, Y.D., Hart, A.A., Mao, M., Peterse, H.L., van der Kooy, K., Marton, M.J., Witteveen, A.T., *et al*. (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature *415*, 530-536.

van de Wetering, M., Sancho, E., Verweij, C., de Lau, W., Oving, I., Hurlstone, A., van der Horn, K., Batlle, E., Coudreuse, D., Haramis, A.P., *et al*. (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell *111*, 241-250.

van Uden, P., Kenneth, N.S., and Rocha, S. (2008). Regulation of hypoxia-inducible factor-1alpha by NF-kappaB. Biochem J *412*, 477-484.

Vaupel, P. (2004). The role of hypoxia-induced factors in tumor progression. Oncologist *9 Suppl 5*, 10-17.

Vincent, S.D., Dunn, N.R., Hayashi, S., Norris, D.P., and Robertson, E.J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. Genes Dev *17*, 1646-1662.

Wakefield, L.M., and Roberts, A.B. (2002). TGF-beta signaling: positive and negative effects on tumorigenesis. Current opinion in genetics & development *12*, 22-29.

Waldrip, W.R., Bikoff, E.K., Hoodless, P.A., Wrana, J.L., and Robertson, E.J. (1998). Smad2 signaling in extraembryonic tissues determines anterior-posterior polarity of the early mouse embryo. Cell *92*, 797-808.

Wang, Y., Klijn, J.G., Zhang, Y., Sieuwerts, A.M., Look, M.P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M.E., Yu, J., *et al*. (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet *365*, 671-679.

Wong, C.C., Gilkes, D.M., Zhang, H., Chen, J., Wei, H., Chaturvedi, P., Fraley, S.I., Wong, C.M., Khoo, U.S., Ng, I.O., *et al*. (2011). Hypoxia-inducible factor 1 is a master regulator of breast cancer metastatic niche formation. Proc Natl Acad Sci U S A *108*, 16369-16374.

Wrana, J.L. (2009). The secret life of Smad4. Cell *136*, 13-14.

Wright, C.W., and Duckett, C.S. (2009). The aryl hydrocarbon nuclear translocator alters CD30-mediated NF-kappaB-dependent transcription. Science *323*, 251-255.

Yamamoto, M., Beppu, H., Takaoka, K., Meno, C., Li, E., Miyazono, K., and Hamada, H. (2009). Antagonism between Smad1 and Smad2 signaling determines the site of distal visceral endoderm formation in the mouse embryo. J Cell Biol *184*, 323-334.

Yang, M.H., Wu, M.Z., Chiou, S.H., Chen, P.M., Chang, S.Y., Liu, C.J., Teng, S.C., and Wu, K.J. (2008). Direct regulation of TWIST by HIF-1alpha promotes metastasis. Nat Cell Biol *10*, 295-305.

Yang, X., Li, C., Xu, X., and Deng, C. (1998). The tumor suppressor SMAD4/DPC4 is essential for epiblast proliferation and mesoderm induction in mice. Proc Natl Acad Sci U S A *95*, 3667-3672.

Zhang, H., Wong, C.C., Wei, H., Gilkes, D.M., Korangath, P., Chaturvedi, P., Schito, L., Chen, J., Krishnamachary, B., Winnard, P.T., Jr., *et al*. (2011). HIF-1-dependent expression of angiopoietin-like 4 and L1CAM mediates vascular metastasis of hypoxic breast cancer cells to the lungs. Oncogene.

Zhang, Y., Sieuwerts, A.M., McGreevy, M., Casey, G., Cufer, T., Paradiso, A., Harbeck, N., Span, P.N., Hicks, D.G., Crowe, J., *et al*. (2009). The 76-gene signature defines high-risk patients that benefit from adjuvant tamoxifen therapy. Breast Cancer Res Treat *116*, 303-309.

Zhu, H., Kavsak, P., Abdollah, S., Wrana, J.L., and Thomsen, G.H. (1999). A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. Nature *400*, 687-693.

FIGURE

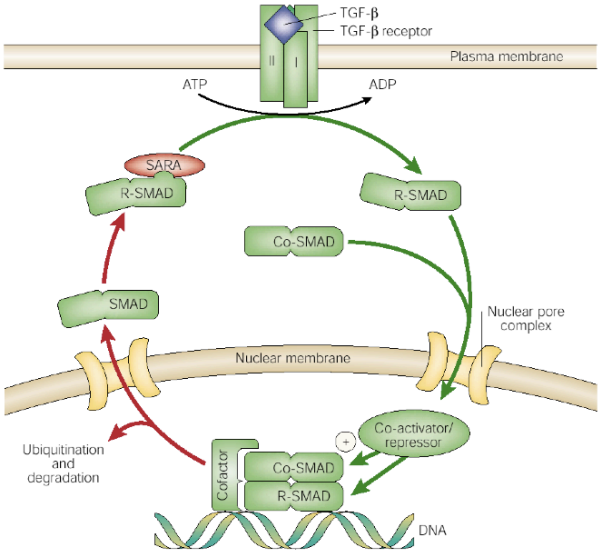
Figure 1

a. The TGFβ cascade. TGFβ ligands bind the serine-threonine kinase receptors on the cell surface; this event activate the receptors that in turn phosphorylate and thus activate receptor Smads (RSmads), which are the first intracellular signal mediators. The R-Smads bind Smad-4, the Co-Smad, and this multimer translocates into the nucleus; here, specific target genes transcription will be regulated (Massague, 2000).

b. The proposed cycle of monoubiquitination/deubiquitination of Smad4. The two enzymes involved, Ectodermin/Tif1γ/TRIM33 and FAM/Usp9x trigger the opposing reactions, respectively in the nucleus and in the cytoplasm. This cycle allows the cells to maintain a reliably transduction of rapidly varying signals from the extracellular space (Wrana, 2009).

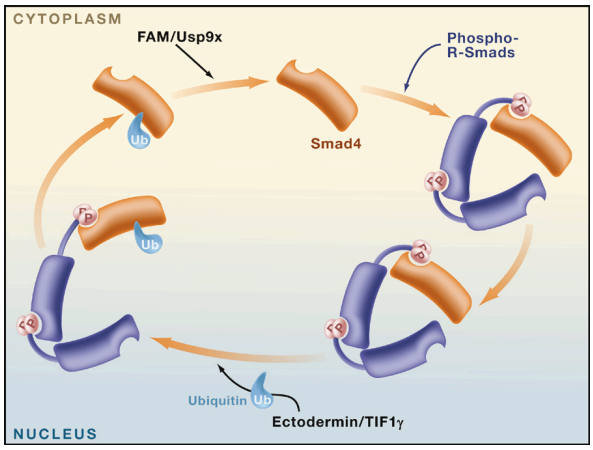
c. Proposed model of Ecto binding to the hystone and DNA to disrupt the active complex between Smad4 and Rsmads

a



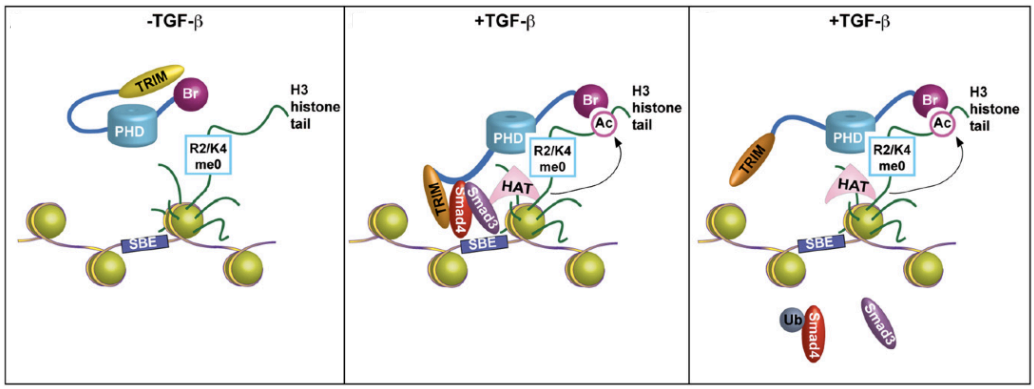
from Massaguè, 2002

b



from Wrana et al, 2009

c



from Agricola et al, 2011

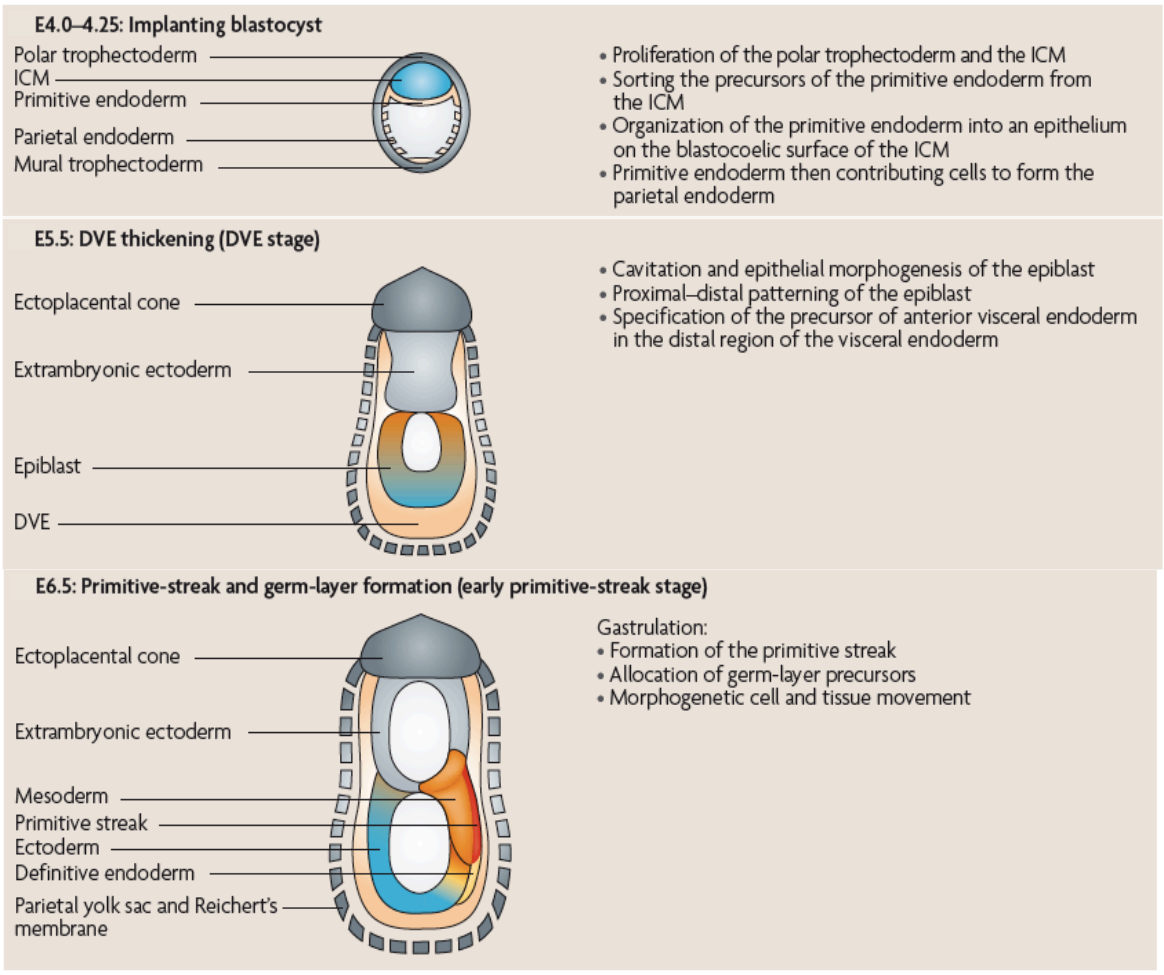
FIGURE 1

Figure 2

a. The blastocyst is composed by three differentiated lineages, namely trophectoderm (grey), primitive endoderm (yellow) and inner cell mass (light blue). At around E5.5 the egg cylinder has formed, thanks to the elongation of the blastocyst; this originates the proximo-distal axis. The epiblast has undergone cavitation, originating the proamniotic cavity. Around E6.5 gastrulation starts: at a discrete place in the proximal epiblast, mesoderm is induced. The primitive streak begins the elongation towards the distal tip of the egg cylinder: this elongation originates the first axis of the embryo proper, the antero-posterior axis. Modified from Tam and Loebel, 2007.

b. Events that requires Nodal during early mouse development. The principal TGFβ ligand of the early mouse embryo is Nodal; by E5.0 it is expressed widely in the epiblast (violet) and in thesurrounding visceral endoderm (not shown in this scheme for clarity reasons). At E5.5, Nodal gives positive signal both to the distal visceral endoderm, where the AVE (red) is induced, and to the trophoblast, where the trophoblast stem cells (TS, yellow) start growing and differentiating. At E6.5 mouse gastrulation begins; AVE rotates towards the prospective anterior side, and mesoderm induction happens at the opposite pole of the epiblast (light blue). Nodal is a fundamental cue for mesoderm induction; other positive signals come from the trophoblast compartment. Mesoderm induction originates an elongated structure known as primitive streak (PS) that is patterned mainly by a Nodal morphogenetic field (depicted as a gradient from dark blue (anterior PS) to light blue (posterior PS).

a


E4.0–4.25: Implanting blastocyst
Polar trophectoderm
ICM
Primitive endoderm
Parietal endoderm
Mural trophectoderm
• Proliferation of the polar trophectoderm and the ICM
• Sorting the precursors of the primitive endoderm from the ICM
• Organization of the primitive endoderm into an epithelium on the blastocoelic surface of the ICM
• Primitive endoderm then contributing cells to form the parietal endoderm
E5.5: DVE thickening (DVE stage)
Ectoplacental cone
Extrambryonic ectoderm
Epiblast
DVE
• Cavitation and epithelial morphogenesis of the epiblast
• Proximal–distal patterning of the epiblast
• Specification of the precursor of anterior visceral endoderm in the distal region of the visceral endoderm
E6.5: Primitive-streak and germ-layer formation (early primitive-streak stage)
Ectoplacental cone
Extrambryonic ectoderm
Mesoderm
Primitive streak
Ectoderm
Definitive endoderm
Parietal yolk sac and Reichert's membrane
Gastrulation:
• Formation of the primitive streak
• Allocation of germ-layer precursors
• Morphogenetic cell and tissue movement

b

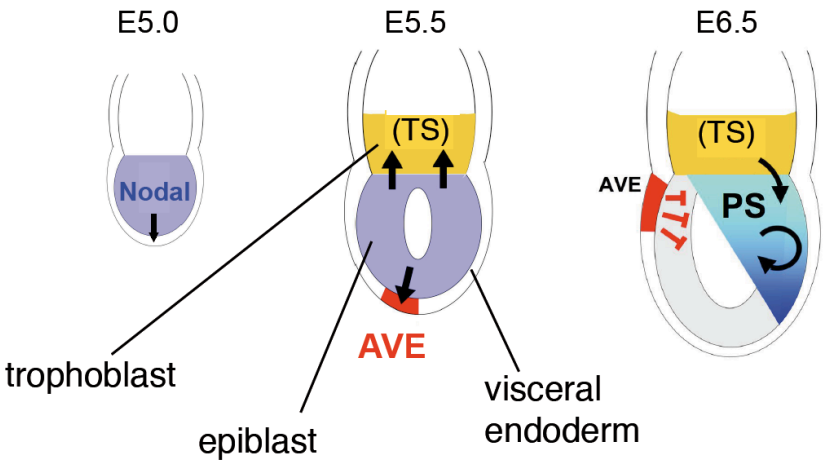

E5.0
E5.5
E6.5
Nodal
(TS)
AVE
trophoblast
epiblast
visceral endoderm
(TS)
AVE
PS

FIGURE 2

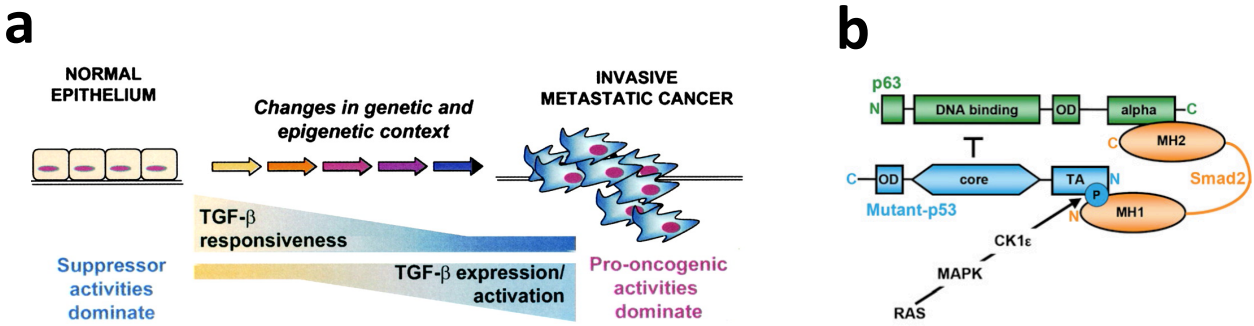
Figure 3

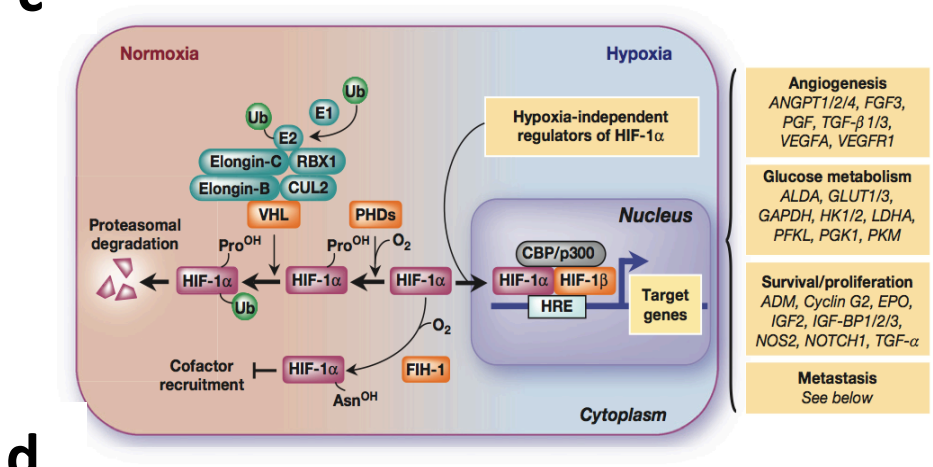
a. Scheme of TGFβ as tumor suppressor and then tumor promoting factor.

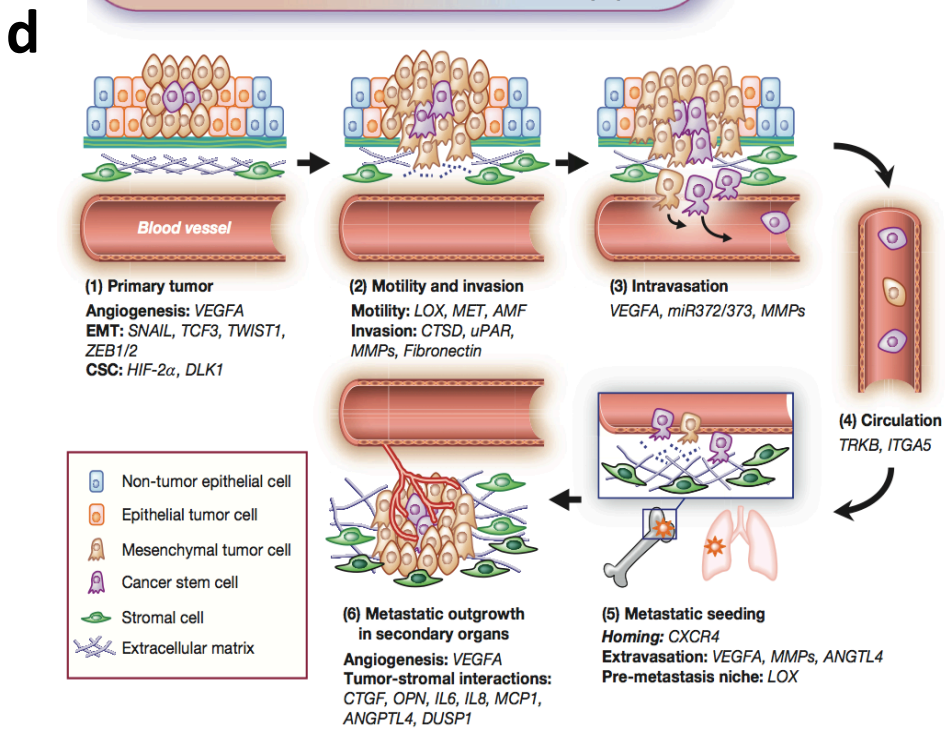
b. Scheme of the binding between Smads, mutant p53 and p63 induced by K-ras induced phophorilation of p53.

c. Hypoxia pathway and list of target genes regulated by HIF involved in tumor progression.

d. Model of metastastasis formation and progression with lists of HIFs target genes involved in each step.







From Lu et al, 2010

FIGURE 3

Figure 4

a. Immunofluorescent staining of wild-type (WT) and Ecto-/- (-/-) embryos with anti-Ecto antibody. Ecto is ubiquitously expressed, displaying a strong enrichment in epiblast cells at E5.5.

b. Hematoxilin and Eosin staining of sections of wild-type (WT) and Ecto-/- embryos within intact decidual tissues at early-streak stage. Note the absence of primitive streak formation (arrowhead) and how the embryo lacks a distinction between epiblast (epi.) and extraembryonic ectoderm (EXE). Scale bars correspond to 10μm.

c. Above, immunofluorescent stainings for the mitotic marker phospho-Histone3 (P-H3, red channel) on wild-type and *Ecto* mutant embryos at E5.0. YOYO1 serves as nuclear counterstain (green channel).

below, quantitations of P-H3 positive cells in the embryonic (Em.) and extraembryonic (Ex.) portions of wild-type (n=17) and *Ecto* mutant (n=5) embryos. The number of P-H3 positive cells in each embryo was determined based on z-stack confocal images covering the whole embryo. Data are given as mean + SD.

c. E6.5 *Ecto* mutant embryos do not express the pan-mesodermal marker *T* (also known as *Brachyury*), the extraembryonic ectoderm and mesoderm marker *Eomes*, and the early mesoderm marker *Wnt3*. Lateral views, anterior to the left.

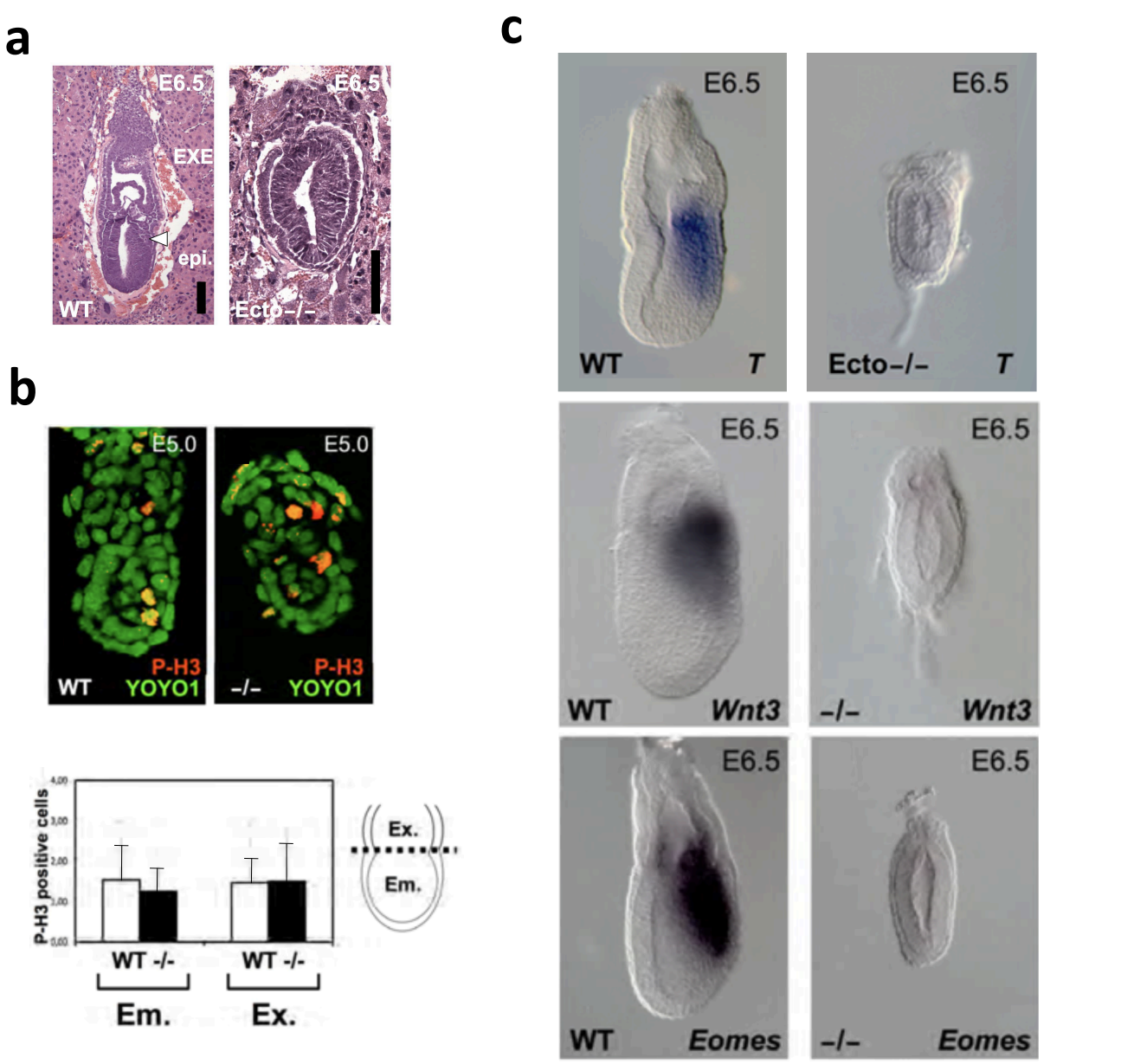

a
E6.5
EXE
epi.
WT
E6.5
Ecto−/−
b
E5.0
P-H3
YOYO1
WT
E5.0
P-H3
YOYO1
−/−
P-H3 positive cells
WT -/-
WT -/-
Em.
Ex.
Ex.
Em.
c
E6.5
WT
T
E6.5
Ecto−/−
T
E6.5
WT
Wnt3
E6.5
−/−
Wnt3
E6.5
WT
Eomes
E6.5
−/−
Eomes

FIGURE 4

Figure 5

a. Above, at early pregastrulation stages, AVE is strongly expanded in *Ecto* mutants, as assayed by expression of the Nodal/Smad targets *Cerl* and *Lefty1*.

Below, as development proceeds, the AVE of *Ecto*-deficient (-/-) embryos further expands, encircling the epiblast. Lateral views, anterior to the left, of early-streak stage embryos (wild type, WT, and *Ecto* knockout, -/-) stained for *Lim1* are shown in the upper part of the panel. Transverse (or optical) sections of early-streak stage embryos stained for the same markers are provided in the bottom part (anterior to the left). Note on the right part that while in wild-type embryos *Lim1* stains both the AVE and the posterior primitive streak, in *Ecto* mutants the AVE is much broader and the mesodermal expression domain of *Lim1* is lost.

b. A schematic representation of the model. The earliest *Nodal*-dependent phenomenon of mouse embryogenesis, AVE induction, is much exaggerated in *Ecto*-/- (-/-) embryos.

c. Panels show in situ for *Cerl* at pre-streak stage in wild-type and *Sox2Cre;Ecto fl/- (EpiKO)* embryos, namely, in embryos where *Ecto* is inactivated in epiblast cells, but not in extraembryonic tissues.

d. Nodal is normally expressed in Ecto mutants at E5.5, but it is rapidly downregulated as development proceeds.

e. Smad2 is normally activated in Ecto mutants, as assayed by immunofluorescence for phospho- Smad2 (P-Sm2, red channel). Merged images with nuclear counterstain are also shown (YOYO1, green channel).

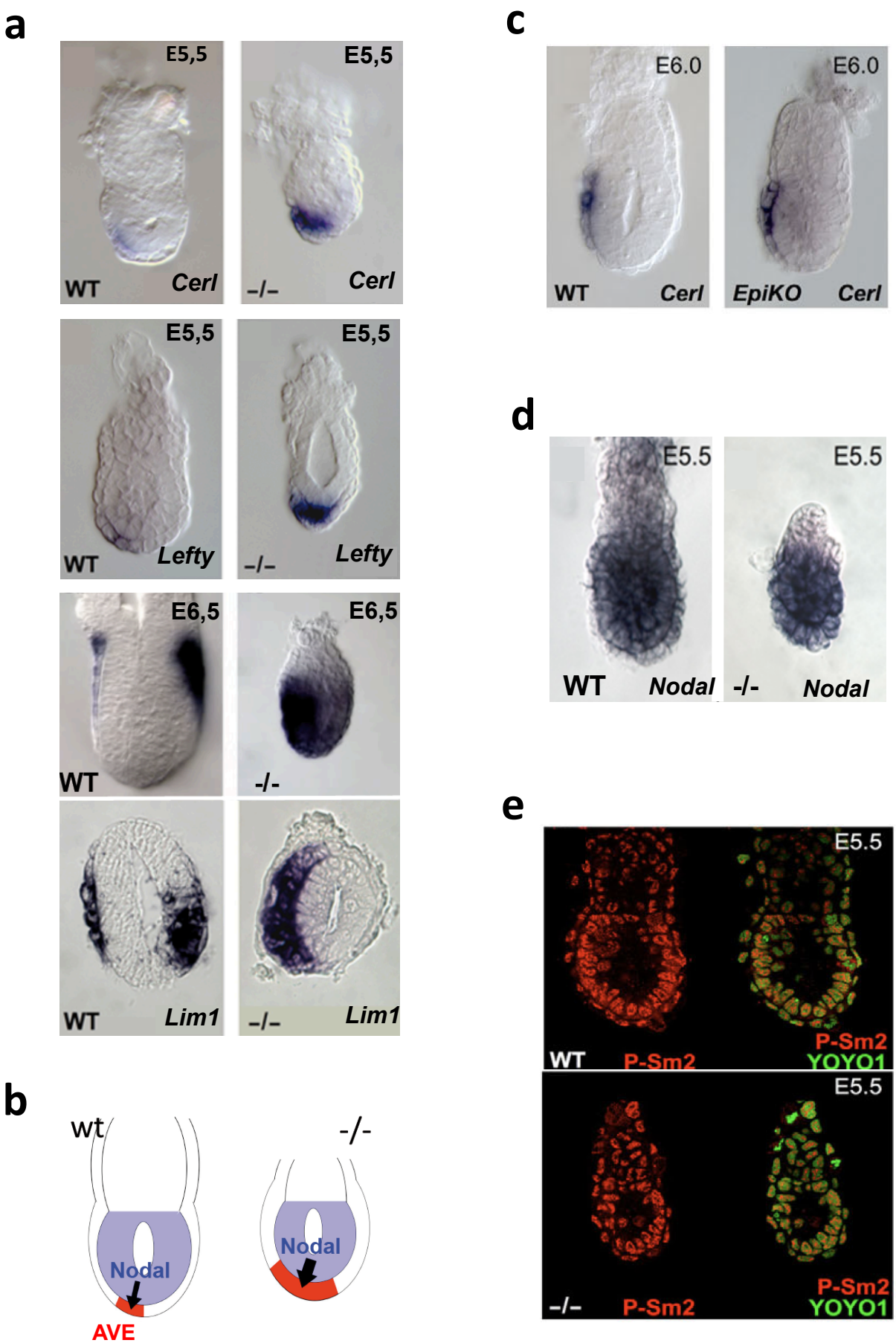

a
E5,5
WT
Cerl
E5,5
-/-
Cerl
E5,5
WT
Lefty
E5,5
-/-
Lefty
E6,5
WT
E6,5
-/-
WT
Lim1
-/-
Lim1
b
wt
-/-
Nodal
Nodal
AVE
c
E6.0
WT
Cerl
E6.0
EpiKO
Cerl
d
E5.5
WT
Nodal
E5.5
-/-
Nodal
e
E5.5
WT
P-Sm2
P-Sm2
YOYO1
E5.5
-/-
P-Sm2
P-Sm2
YOYO1

FIGURE 5

Figure 6

a. Excessive AVE formation in *Ecto* mutants is dependent on Smad4 activity. In situ hybridization for the AVE markers *Cerl* and *Lim1* in wild-type (WT), *Ecto*-/-, *Smad4*-/- and *Ecto*/*Smad4* double mutant (*Ecto*-/- *Smad4*-/-) embryos. AVE expansion is observed in *Ecto* mutants but not in embryos also lacking *Smad4*.

b. Visceral Endoderm (VE) specification occurs normally in *Smad4*-/- embryos. In situ hybridization for the visceral endoderm (VE) marker alpha-feto protein (*AFP*) on E6.0 wild-type (WT) and *Smad4*-/- embryos. *AFP* staining in *Smad4* deficient embryos serves as control that the visceral endoderm is correctly specified even if it cannot be induced to AVE in the presence or absence of *Ecto*.

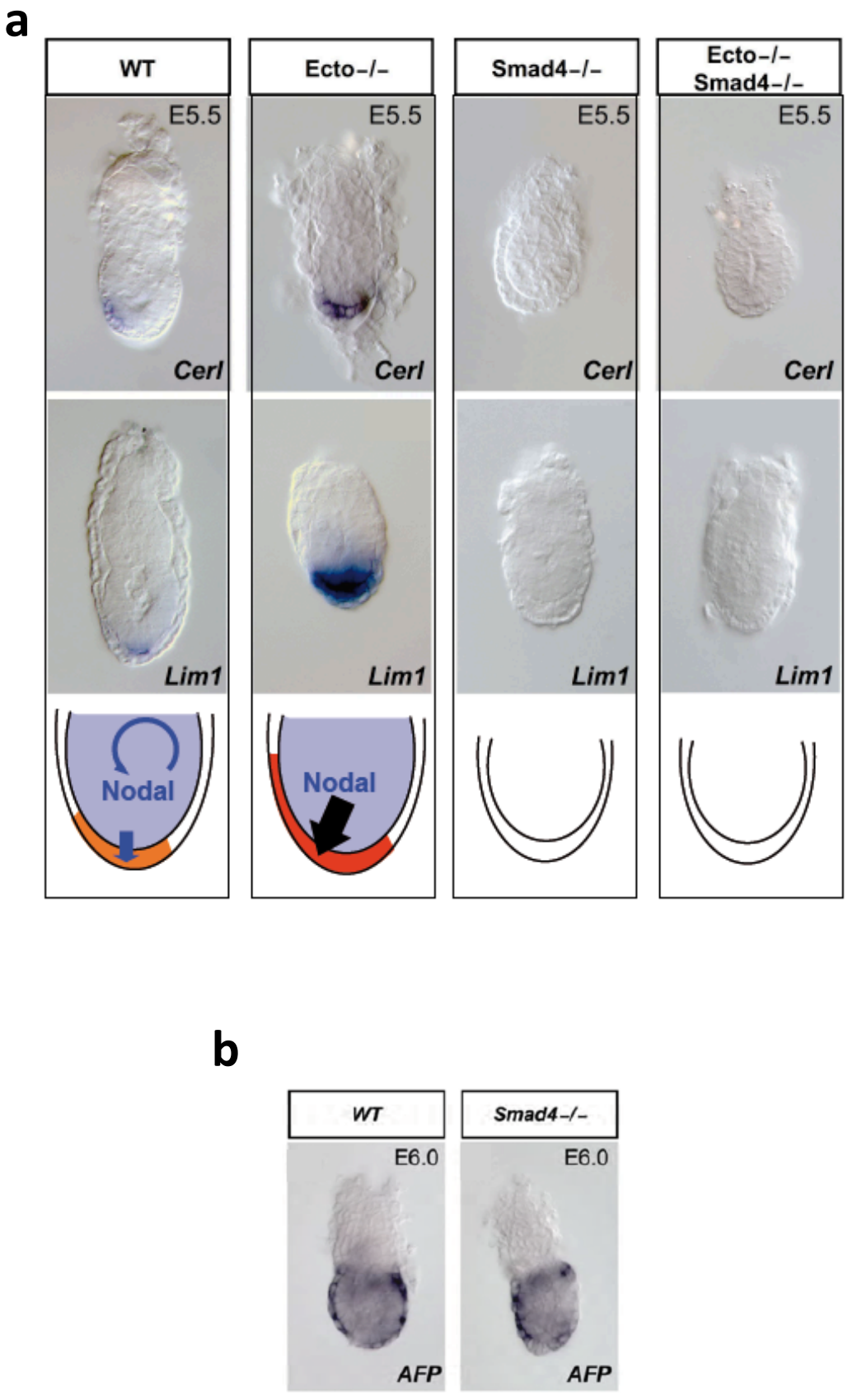

a
WT
Ecto−/−
Smad4−/−
Ecto−/−
Smad4−/−
E5.5
Cerl
Lim1
Nodal
b
WT
Smad4−/−
E6.0
AFP

FIGURE 6

Figure 7

a. Reduction of Nodal signaling by the combined use of null (*Nodal-*) and hypomorphic (*NodalΔ600*) alleles neutralize AVE expansion in *Ecto* mutants, as assayed by in situ hybridization for *Lim1* at pre-streak stage. Note how decreased *Nodal* also rescues the overall morphology and size of the *Ecto* mutants. Lateral views, anterior to the left.

b. The epiblast markers *FGF4* and *Oct4* are normally expressed in *Ecto* mutants (Ecto-/-).

c. *Lim1* in situ hybridization at early streak stage in wild-type (WT), *Nodal*Δ*600*/- and compound *Nodal*Δ*600*/-;*Ecto*-/- embryos. Note how reduction of Nodal signaling in *Nodal*Δ*600*/- embryos impairs rotation of the AVE (white arrowhead) toward the anterior pole of the embryo, and how this rotation is restored by inactivation of *Ecto* in *Nodal*Δ*600*/-; *Ecto*-/- embryos. Lateral view, anterior to the left.

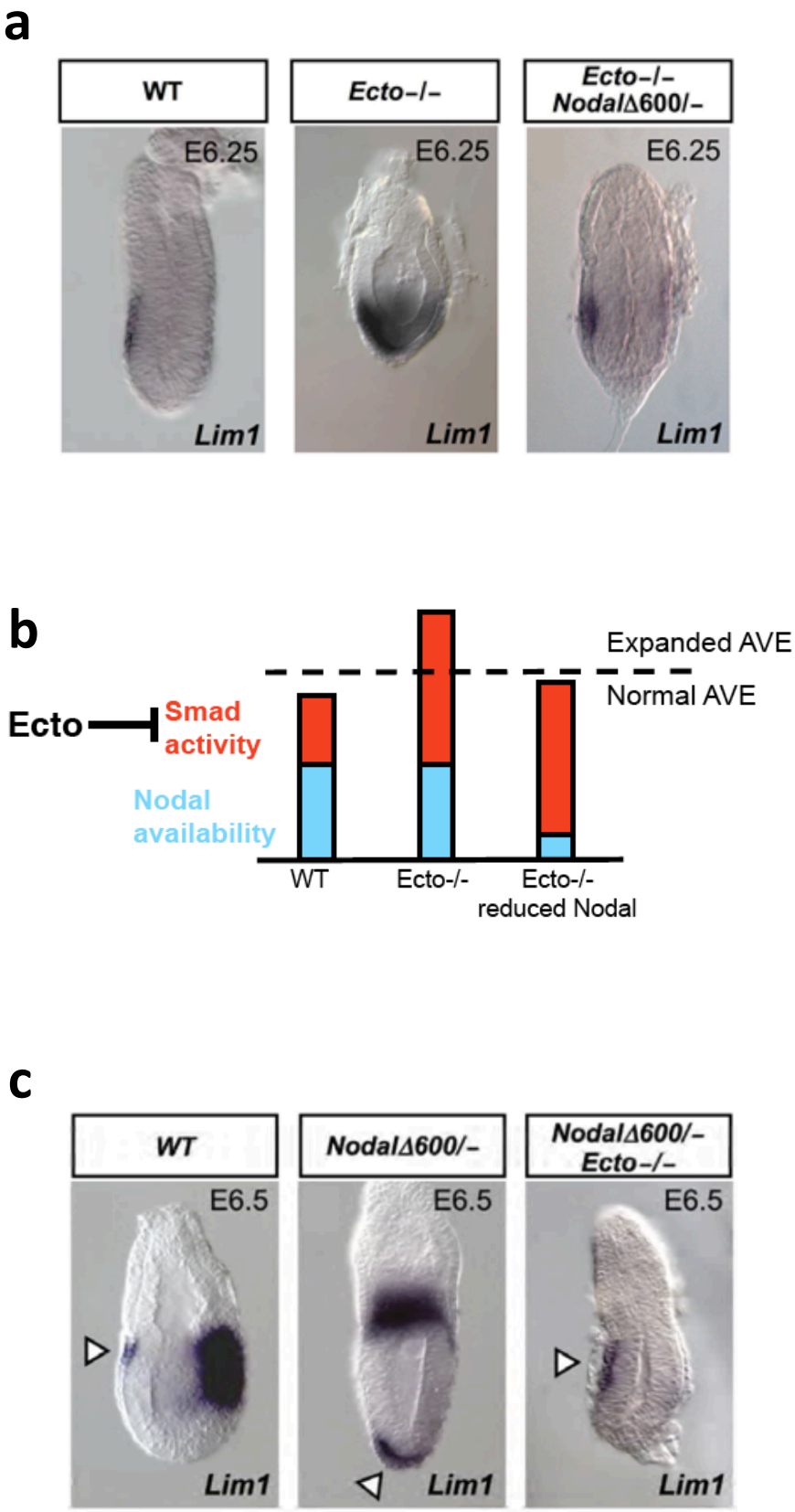

a
WT
Ecto−/−
Ecto−/−
NodalΔ600/−
E6.25
E6.25
E6.25
Lim1
Lim1
Lim1
b
Expanded AVE
Normal AVE
Ecto
Smad
activity
Nodal
availability
WT
Ecto-/-
Ecto-/-
reduced Nodal
c
WT
NodalΔ600/−
NodalΔ600/−
Ecto−/−
E6.5
E6.5
E6.5
Lim1
Lim1
Lim1

FIGURE 7

Figure 8

a. *Ecto* mutants (Ecto-/-) lack expression of extraembrionic ectoderm markers as *Eomes* at E5.5.

b. The epiblast marker *Oct4* is normally expressed in *Ecto* mutants (Ecto-/-).

c. *Ecto* acts cell-autonomously within the extraembryonic tissues to maintain EXE fates. Panels show in situ for *BMP4* in wild-type and *Sox2-Cre;Ecto fl/- (EpiKO*) embryos, namely, in embryos where *Ecto* is inactivated in epiblast cells, but not in extraembryonic tissues.

d. Reduction of *Nodal* dosage rescues EXE formation in *Ecto* mutant embryos, as assayed by *BMP4* expression. Note that EXE compartment is rescued both in the compound mutants *Ecto-/-; Nodal*+/- and *Ecto-/-; Nodal*Δ*600*/-.

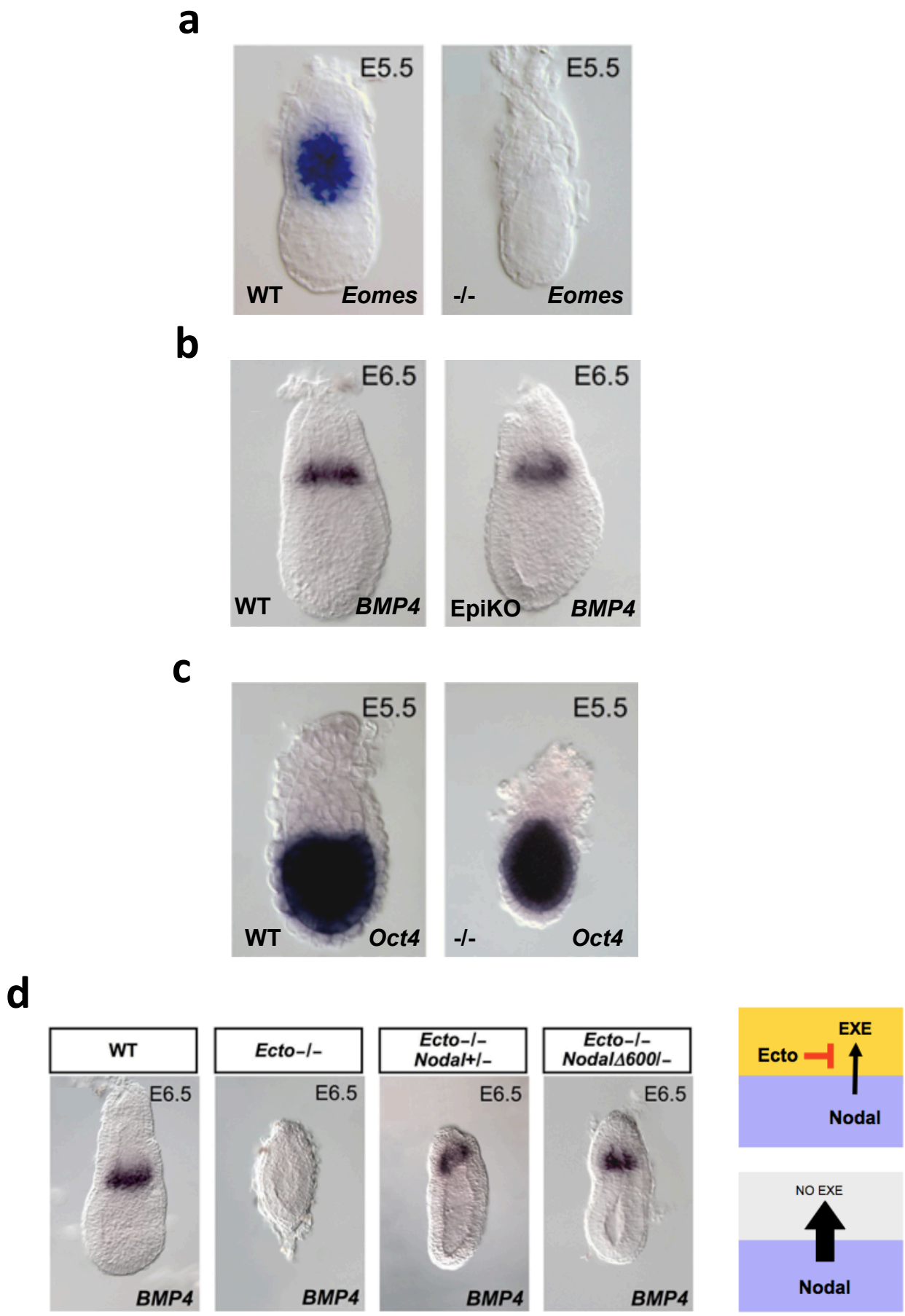

a
E5.5
WT
Eomes
E5.5
-/-
Eomes
b
E6.5
WT
BMP4
E6.5
EpiKO
BMP4
c
E5.5
WT
Oct4
E5.5
-/-
Oct4
d
WT
Ecto-/-
Ecto-/-
Nodal+/-
Ecto-/-
NodalΔ600/-
E6.5
BMP4
E6.5
BMP4
E6.5
BMP4
E6.5
BMP4
EXE
Ecto
Nodal
NO EXE
Nodal

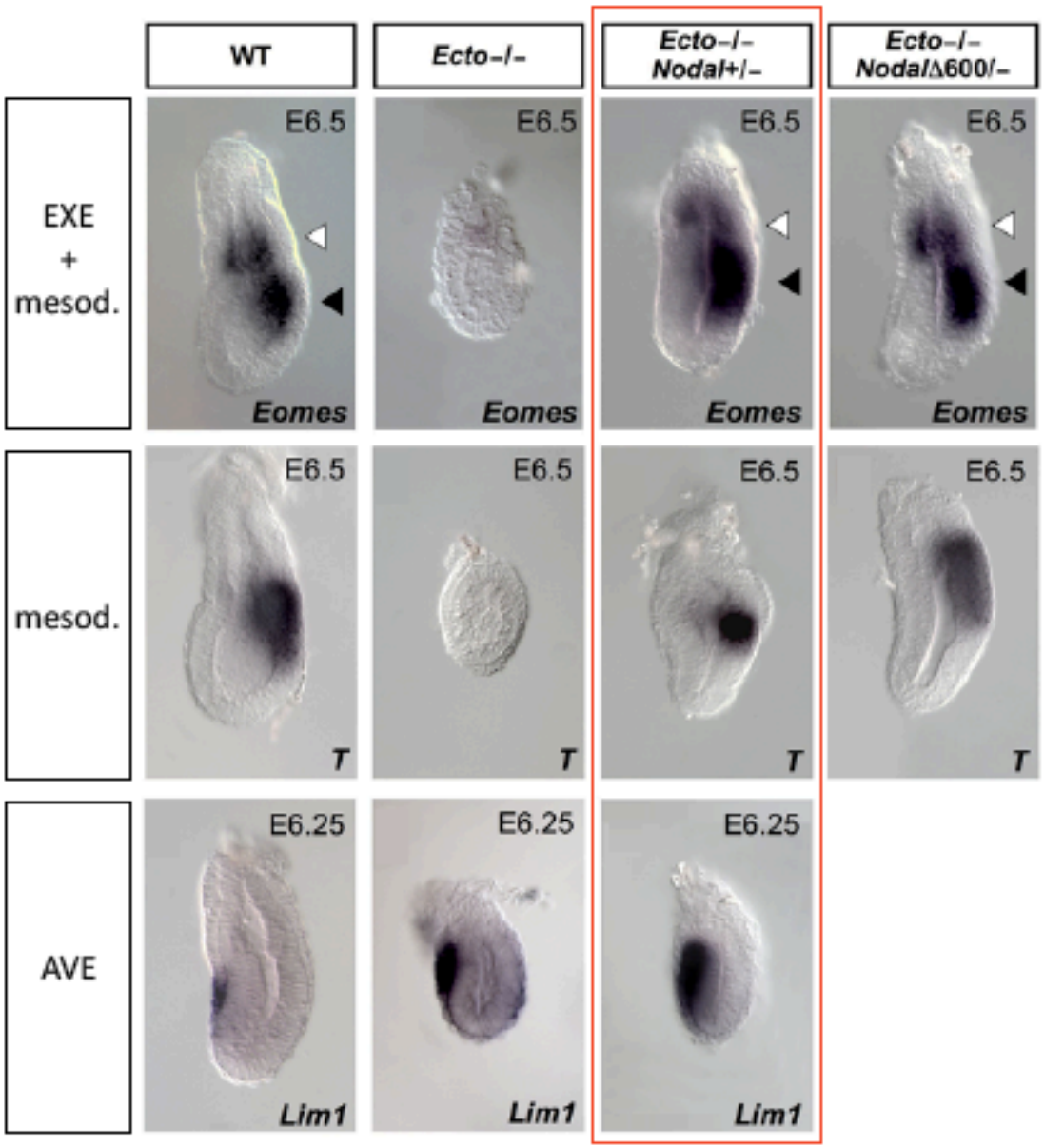
FIGURE 8

Figure 9

First two rows: reduction of *Nodal* rescues mesoderm formation in *Ecto* mutant embryos, as assayed by *Eomes* expression and *T*. Lateral views, anterior to the left.

Third row: analysis of *Lim1* expression in AVE of wild-type (WT), *Ecto* mutants and *Ecto-/-; Nodal+/-* embryos. Note how *Ecto-/-; Nodal+/-* embryos (red box) show already rescued mesoderm and EXE development (white arrowheads), but still display expanded AVE, suggesting that lack of mesoderm in *Ecto* mutants is primarily due to lack of EXE-derived inducing signal.

The table below shows how EXE and mesoderm markers are coupled throughout the mutants' series, whereas AVE seems uncoupled in *Ecto-/-; Nodal+/-* embryos. Tick marks stand for normal tissue; X marks stand for impaired tissue; arrows up stand for expanded tissue.



EXE	✓	✗	✓	✓
mesod.	✓	✗	✓	✓
AVE	✓	↑↑	↑↑	✓

FIGURE 9

Figure 10

a. Immunofluorescent staining of wild-type (WT) and *Ecto*-/- (-/-) embryos with anti-Ecto antibody. Ecto is ubiquitously expressed, displaying a strong enrichment in epiblast cells at E5.5.

b. Immunoprecipitation of ubiquitinated-Smad4 from dissected Distal (D) and Proximal (P) portions of the embryonic cup at E7.0. Upper panel shows anti-Smad4 (Sm4) immunoblot of anti- Ubiquitin (Ub) immunoprecipitates. Co. is anti-Ubiquitin antibody alone. Lower panels show immunoblotting of embryonic whole extracts.

c. Expression of the anterior mesoderm marker *FoxA2* encompasses the whole primitive streak in *Sox2-Cre;Ecto fl*/- (*EpiKO*) streak stage embryos. Lateral and posterior views of a wild-type embryo (WT), and lateral and posterior views of an *EpiKO* embryo are provided. Dashed lines indicate the boundary between extraembryonic (EXE) and embryonic (epi.) tissues.

d. Expansion of the node in *EpiKO* embryos. Pictures show a close-up of the distal region of sibling embryos stained for the node marker *FoxA2*, taken from the anterior.

e. In situ hybridizations for the definitive endoderm marker *Cerl* at E7.5. Lateral views, anterior to the left. Below are shown transverse sections of the corresponding embryos, taken at the level of white lines.

f. *EpiKO* embryos display a widened and duplicated anterior node, as assayed by expression of *Shh* at E8.5/9.0. Pictures show a close-up of the node region, anterior to the top. Below are shown corresponding whole embryo lateral views. White arrowheads show the direction from which closeups were taken.

g. representative model of the Nodal gradient on the primitive strike.

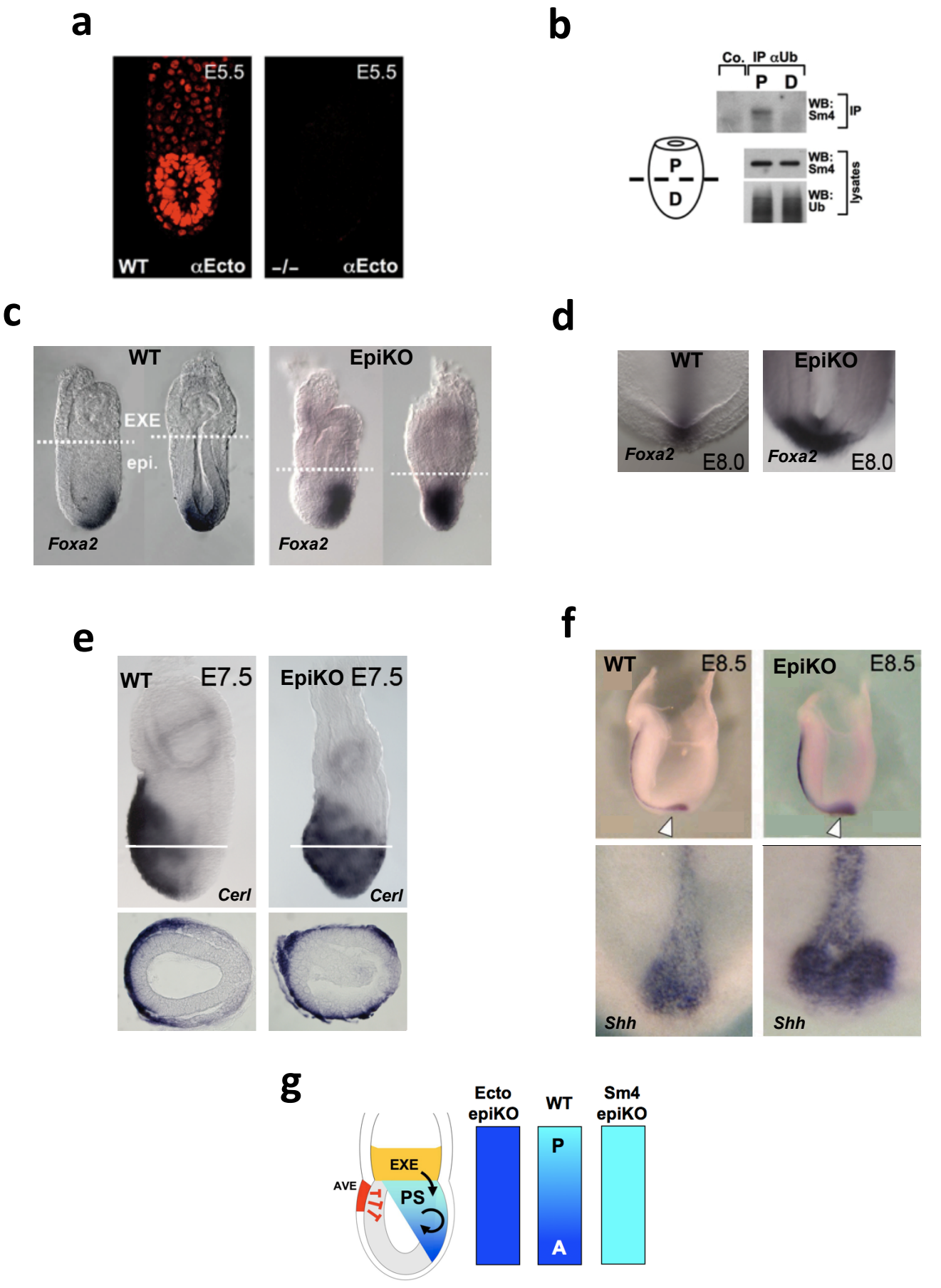

a
E5.5
WT
αEcto
E5.5
-/-
αEcto
b
Co.
IP αUb
P
D
WB: Sm4
IP
P
D
WB: Sm4
WB: Ub
lysates
c
WT
EXE
epi.
Foxa2
EpiKO
Foxa2
d
WT
Foxa2
E8.0
EpiKO
Foxa2
E8.0
e
WT
E7.5
Cerl
EpiKO E7.5
Cerl
f
WT
E8.5
EpiKO
E8.5
Shh
Shh
g
Ecto epiKO
WT
Sm4 epiKO
EXE
AVE
PS
TTY
P
A

FIGURE 10

Figure 11

a. Correlation between expression values of *Sharp1* and *p63* (left) or *CyclinG2* and *p63* (right) in the cohort of primary TNBC.

b. Kaplan-Meier graph representing the probability of metastasis-free survival in triple negative breast cancer patients (TNBC) stratified according to high or low expression levels of *Sharp1* and *CyclinG2* (*Sharp1*/*CyclinG2* signature). The log-rank test p value reflects the significance of the association between the ‘‘High’’ group and longer metastasis-free survival.

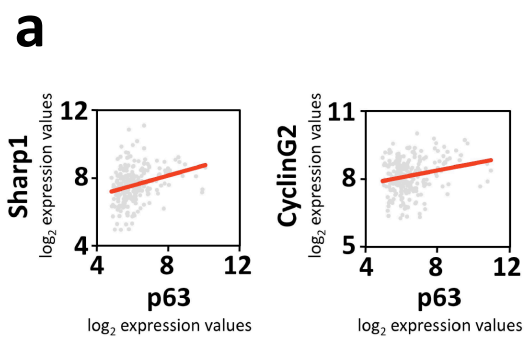
c. Univariate Kaplan-Meier and log-rank test p value reflect the association between low *Sharp1*/*CyclinG2* expression levels and reduced survival.

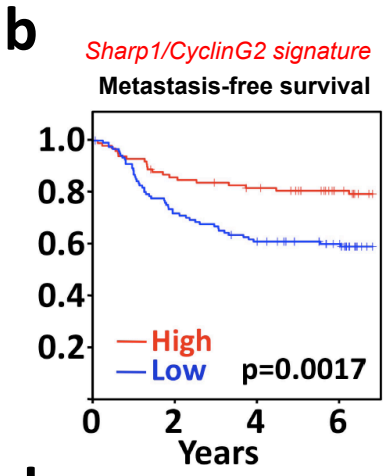
d. GSEA analysis for association between high or low *Sharp1*/*CyclinG2* expression values and gene sets denoting the activation of specific signaling pathways. In bold (red box) are indicated the signatures that reach statistical significance. The table also shows a selection of the other classifiers that did not reach statistical significance. Enrichment in patients with Low *Sharp1* and *CyclinG2* expression (Enrichment Score, ES<0) was considered statistically significant with a False Discovery Rate (FDR q-val) < 0.25.

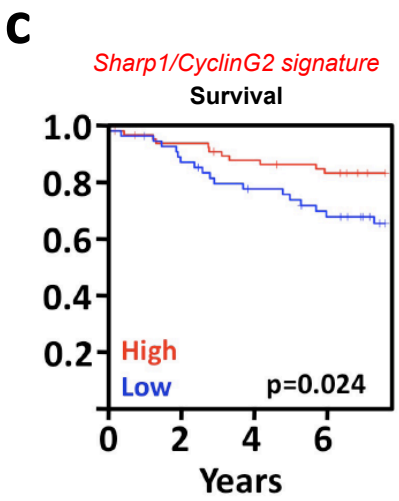
e. Kaplan-Meier curve of metastasis-free survival in TNBC patient datasets stratified

according to the expression levels of genes activated by HIFs (as in Table 4). The log-rank test p-value indicates that tumors with higher expression of HIFactivated genes are significantly associated with poor prognosis.

f. Survival probability of TNBC cancer patients with High or Low HIFs signature.

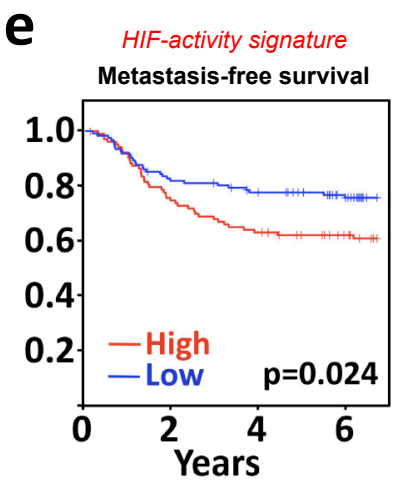






d

Pathway	ES	FDR q-val
mutant-p53	***-0.58***	***0.05***
HIF	***-0.47***	***0.19***
TGFβ	***-0.43***	***0.15***
Proteasome	-0.33	0.34
Brca	-0.32	0.37
Cell cycle	-0.31	0.33
TCF4	-0.27	0.44
Ras	-025	0.46
E2F3	-0.19	0.69



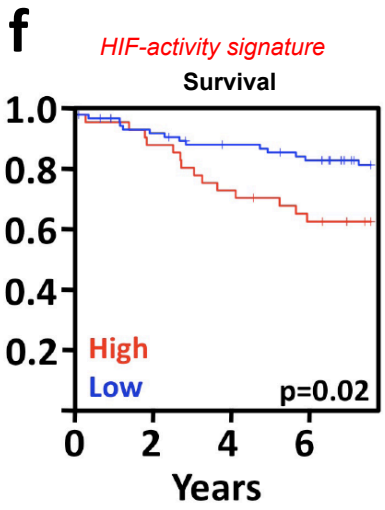


FIGURE 11

Figure 12

a and **b**. efficacy of siRNA mediated knock-downs of indicated proteins as assayed by Western blot.

b and **d**. Depletion of endogenous HIF-1α (**b**) or HIF-2α (**d**) inhibits motility of TNBC cells as assayed by TGFβ-triggered transwell migration experiments with MDA-231.

e. Coimmunoprecipitation (Co-IP) of endogenous Sharp1 with endogenous HIF-1α, from extracts of MDA-231 and Hs578T cells. In the Sharp1 immunoblot, the asterisk points to a background band due to cross-reaction of IgGs.

f. HEK293T cells were transfected with HA-HIF-2α either with empty vector or with Flag-Sharp1; lysates were then immunoprecipitated with Flag-M2 resin. Coimmunoprecipitated HA-HIF-2α was then visualized by Western Blotting

.

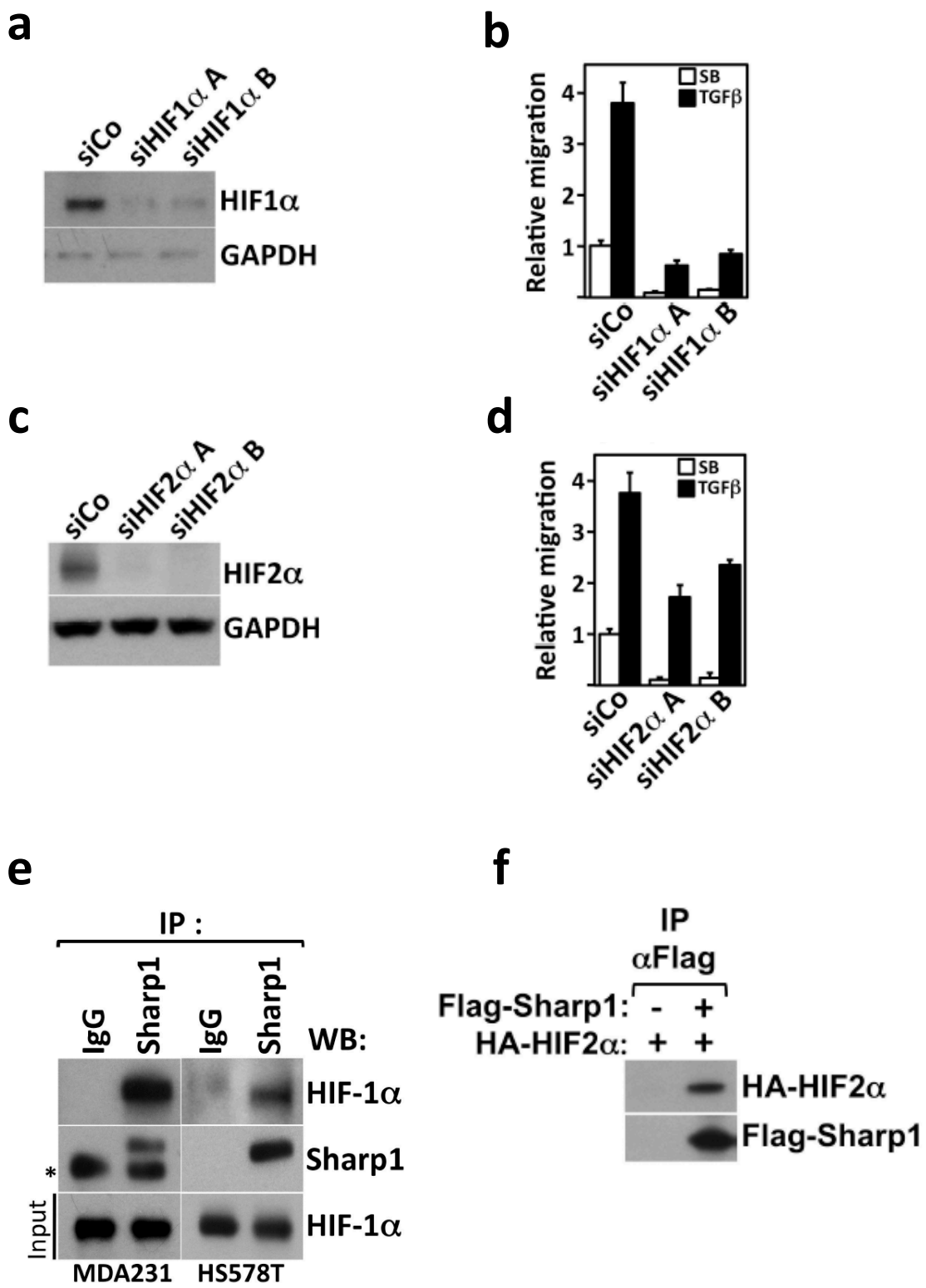

a
siCo
siHIF1α A
siHIF1α B
HIF1α
GAPDH
b
Relative migration
SB
TGFβ
siCo
siHIF1α A
siHIF1α B
c
siCo
siHIF2α A
siHIF2α B
HIF2α
GAPDH
d
Relative migration
SB
TGFβ
siCo
siHIF2α A
siHIF2α B
e
IP :
IgG
Sharp1
IgG
Sharp1
WB:
HIF-1α
Sharp1
*
Input
HIF-1α
MDA231
HS578T
f
IP
αFlag
Flag-Sharp1: - +
HA-HIF2α: + +
HA-HIF2α
Flag-Sharp1

FIGURE 12

Figure 13

a. above, qPCR analysis of *Sharp1* mRNA levels in MDA-231 and MII cells. The expression of *Sharp1* in MDA-231 is set to 1.

Belov, Western Blot analysis of p63 protein in MDA-231 and MII. Bars show mean and SD.

b. qPCR analyses of selected HIFs targets in MDA-231 cells stably overexpressing empty vector (Control) or Sharp1, and incubated in low (1%) or normal (20%) oxygen levels to modulate HIFs activity. Expression levels are relative to *GAPDH*, data are normalized to lane 1. Data are presented as mean and SD.

c. qPCR analyses of selected HIF targets in MII cells transfected with the indicated siRNAs (Sharp1 siRNA A, Table 8), and incubated in low (1%) or normal (20%) oxygen levels. Data are presented as mean and SD.

d. Heat-map indicate the fold-expression levels of the genes that were coherently downregulated (green) or upregulated (red) in cells overexpressing *Sharp1* or depleted of HIFs, as measured by microarray analysis once compared to control shGFP-MDA-231 cells (see Methods).

e, qPCR validation of selected Sharp1/HIF target genes identified in these microarray analysis. Bars show mean and SD.

f. Kaplan-Meier graphs representing the metastasis-free survival of TNBC patients stratified according to Sharp1 activity signature. p-value was calculated with the log-rank test.

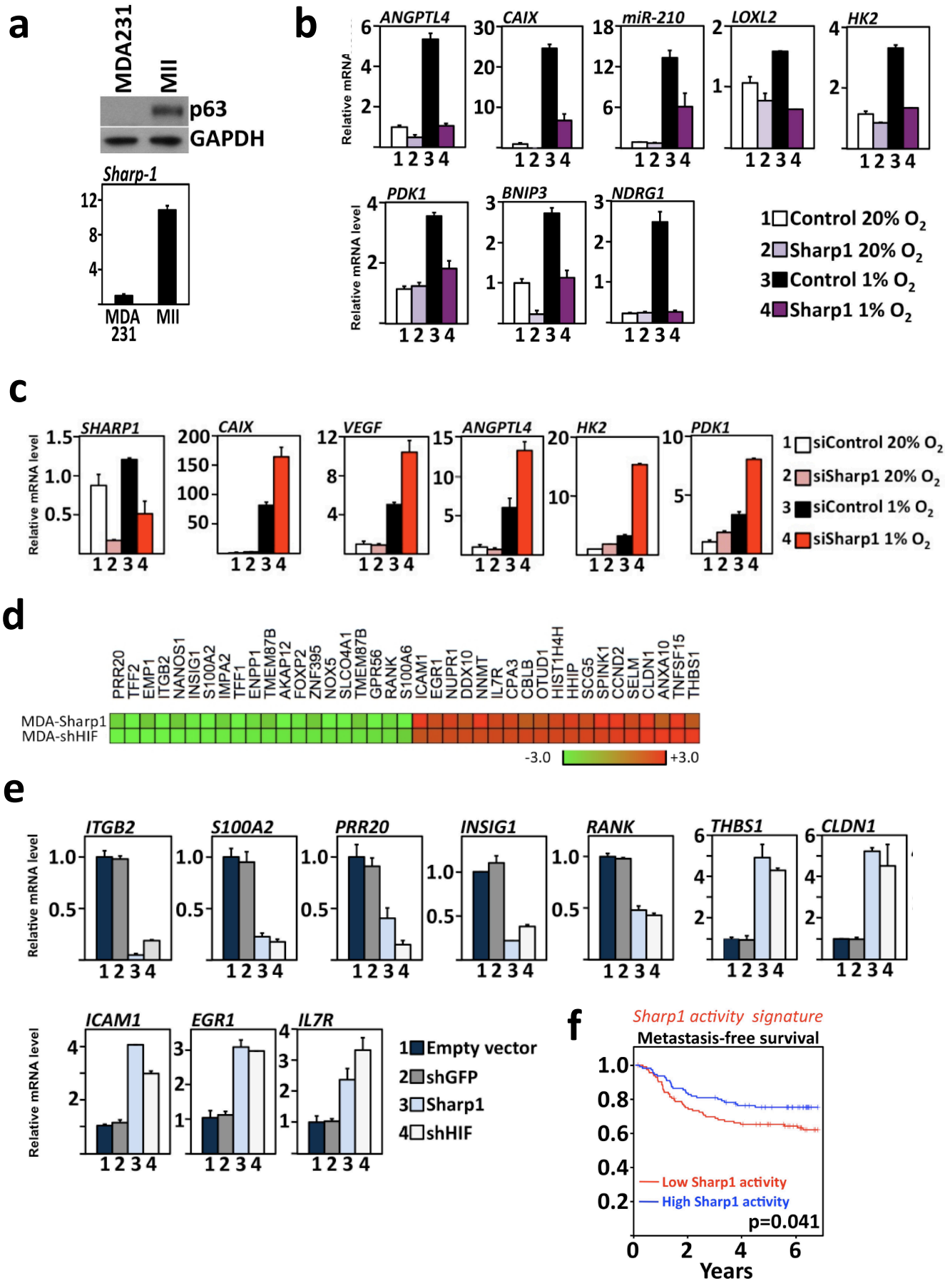

a
MDA231
MII
p63
GAPDH
Sharp-1
MDA 231
MII
b
ANGPTL4
CAIX
miR-210
LOXL2
HK2
PDK1
BNIP3
NDRG1
Relative mRNA level
1 Control 20% O2
2 Sharp1 20% O2
3 Control 1% O2
4 Sharp1 1% O2
c
SHARP1
CAIX
VEGF
ANGPTL4
HK2
PDK1
Relative mRNA level
1 siControl 20% O2
2 siSharp1 20% O2
3 siControl 1% O2
4 siSharp1 1% O2
d
PRR20
TFF2
EMP1
ITGB2
NANOS1
INSIG1
S100A2
IMPA2
TFF1
ENPP1
TMEM87B
AKAP12
FOXP2
ZNF395
NOX5
SLCO4A1
TMEM87B
GPR56
RANK
S100A6
ICAM1
EGR1
NUPR1
DDX10
NNMT
IL7R
CPA3
CBLB
OTUD1
HIST1H4H
HHIP
SCG5
SPINK1
CCND2
SELM
CLDN1
ANXA10
TNFSF15
THBS1
MDA-Sharp1
MDA-shHIF
-3.0
+3.0
e
ITGB2
S100A2
PRR20
INSIG1
RANK
THBS1
CLDN1
ICAM1
EGR1
IL7R
Relative mRNA level
1 Empty vector
2 shGFP
3 Sharp1
4 shHIF
f
Sharp1 activity signature
Metastasis-free survival
Low Sharp1 activity
High Sharp1 activity
p=0.041
Years

FIGURE 13

Figure 14

a. Lung colonization assay of mice intravenously injected with the metastatic TNBC cell line MDA-231 cells as described. Plot shows the number of the metastatic foci per section (2 sections/lung) (n>8) (*** p<0.0001 based on Student's t-test). On the right: representative immunohistochemical stainings for human cytokeratins.

b. Sharp1 inhibits HIF activity *in vivo* in the primary tumor. CAIX and Glut1 protein levels were analysed in tumors emerging after fat pad injection. Cells were scored according to the intensity of the signal. Representative immunofluorescent images are shown on the right, green for CAIX and Glut-1, blue for DAPI. The bars indicate mean and s.e.m.

c. Transwell migration assay of MII cells stably expressing the indicated shRNAs. Plot shows the quantification of the area covered by the migrated cells, relative to the first lane that was set to 1. In right panels: representative images.

d. Control of protein and mRNA knockdowns used in Fig. 13c.

e. Transwell migration assay in MDA-231 cells. Cells expressing shRNA targeting GFP (lanes 1 and 2) or a shRNA targeting mutant-p53 (lanes 3-8), were transiently transfected with the indicated siRNAs. Graphs show the quantification of cells that passed through the filter, relative to the first lane that was set to 1. Data are presented as mean and S.D.

f. Above, wound-healing assays of multiple TNBC cell lines: monolayers of MDA-231, SUM159 and Hs578T cells stably transfected with the indicated plasmids were scratched with a pipette tip and cells allowed to migrate.

Below, quantification of the cells migrated inside the wound is shown in the plot (***: p<0.0001; based on Student's t test). The plot presents mean and s.e.m., the experiments were repeated at least three times with consistent results.

g. Scratch assays of MDA-231 cells expressing *Sharp1* in absence and presence of HIF-2α.

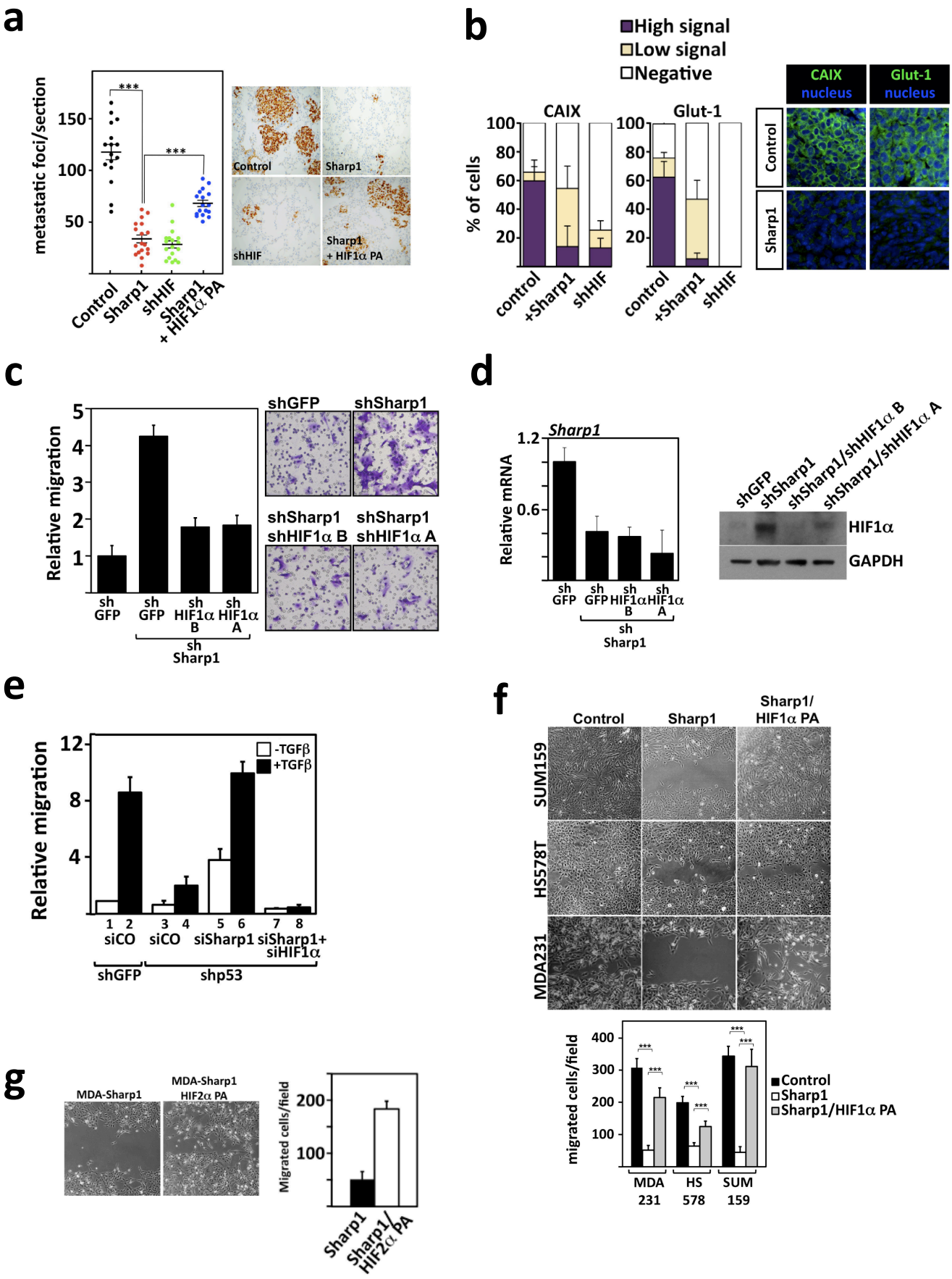

a
metastatic foci/section
150
100
50
***
***
Control
Sharp1
shHIF
Sharp1 + HIF1α PA
Control
Sharp1
shHIF
Sharp1 + HIF1α PA
b
High signal
Low signal
Negative
CAIX
Glut-1
% of cells
100
80
60
40
20
control
+Sharp1
shHIF
CAIX nucleus
Glut-1 nucleus
Control
Sharp1
c
Relative migration
5
4
3
2
1
shGFP
shSharp1
shSharp1 shHIF1α B
shSharp1 shHIF1α A
sh GFP
sh GFP
sh HIF1α B
sh HIF1α A
sh Sharp1
d
Sharp1
Relative mRNA
1.2
0.6
sh GFP
sh GFP
sh HIF1α B
sh HIF1α A
sh Sharp1
shGFP
shSharp1
shSharp1/shHIF1α B
shSharp1/shHIF1α A
HIF1α
GAPDH
e
Relative migration
12
8
4
-TGFβ
+TGFβ
1 2
siCO
3 4
siCO
5 6
siSharp1
7 8
siSharp1+ siHIF1α
shGFP
shp53
f
Control
Sharp1
Sharp1/ HIF1α PA
SUM159
HS578T
MDA231
migrated cells/field
400
300
200
100
***
***
***
***
***
***
Control
Sharp1
Sharp1/HIF1α PA
MDA 231
HS 578
SUM 159
g
MDA-Sharp1
MDA-Sharp1 HIF2α PA
Migrated cells/field
200
100
Sharp1
Sharp1/ HIF2α PA

FIGURE 14

Figure 15

a. Western blot analysis of cell extracts from MDA-231 shows that stable Sharp1 overexpression downregulates HIF-1α protein levels both in normoxic conditions (20% O_2) and upon low oxygen conditions (1% O_2, 24 hours). Control cells were transduced with an empty vector.

b. Sharp1 overexpression downregulates HIF-1α protein levels in two additional TNBC cell lines (SUM159 and Hs578T).

c. Sharp1 inhibits HIF-2α protein stability in MDA-231 cells. Cell lysates from control and Sharp1-overexpressing cells were analyzed for endogenous HIF-2α protein expression, GAPDH serves as loading control.

d. Fluorescent IHC of HIF-1α protein in tumors arising from orthotopically injected control and Sharp1-overexpressing MDA-231. Sharp1 efficiently blunts HIF-1α protein levels *in vivo*. Representative images are shown on the left panels.

e. Relevance of endogenous Sharp1 in MII cells. After 6 hours of hypoxia (1% oxygen) HIF-1α protein levels raise (compare lanes 1 and 2); lane 3: depletion of Sharp1 (siRNAA) greatly increases HIF1a stabilization.

f. Relevance of endogenous Sharp1 in MDA-231 cells as assayed by anti-HIF1α western blots in control and Sharp-depleted cells (siRNA-A). In this set-up, cells were cultured in normoxic conditions.

g. Sharp1 mRNA levels are decreased by three independent siRNAs targeting endogenous *p63* (A, B, C) in MDA-231 and SUM-159.

h. Western blot analysis of cell extracts from MDA-231 cultured under hypoxia (1% oxygen for 24 hours) in order to better visualize changes in HIF-1α protein level. Note that HIF-1α protein levels are reduced in cells depleted of mutant-p53 (shp53),

i. or upon overexpression of *TAp63*α compared to empty vector.

j. Western blot analysis of MII cell extracts shows enhanced HIF-1α protein levels in cells transfected with siRNA against p63 after 6 hours of hypoxia (1% oxygen).

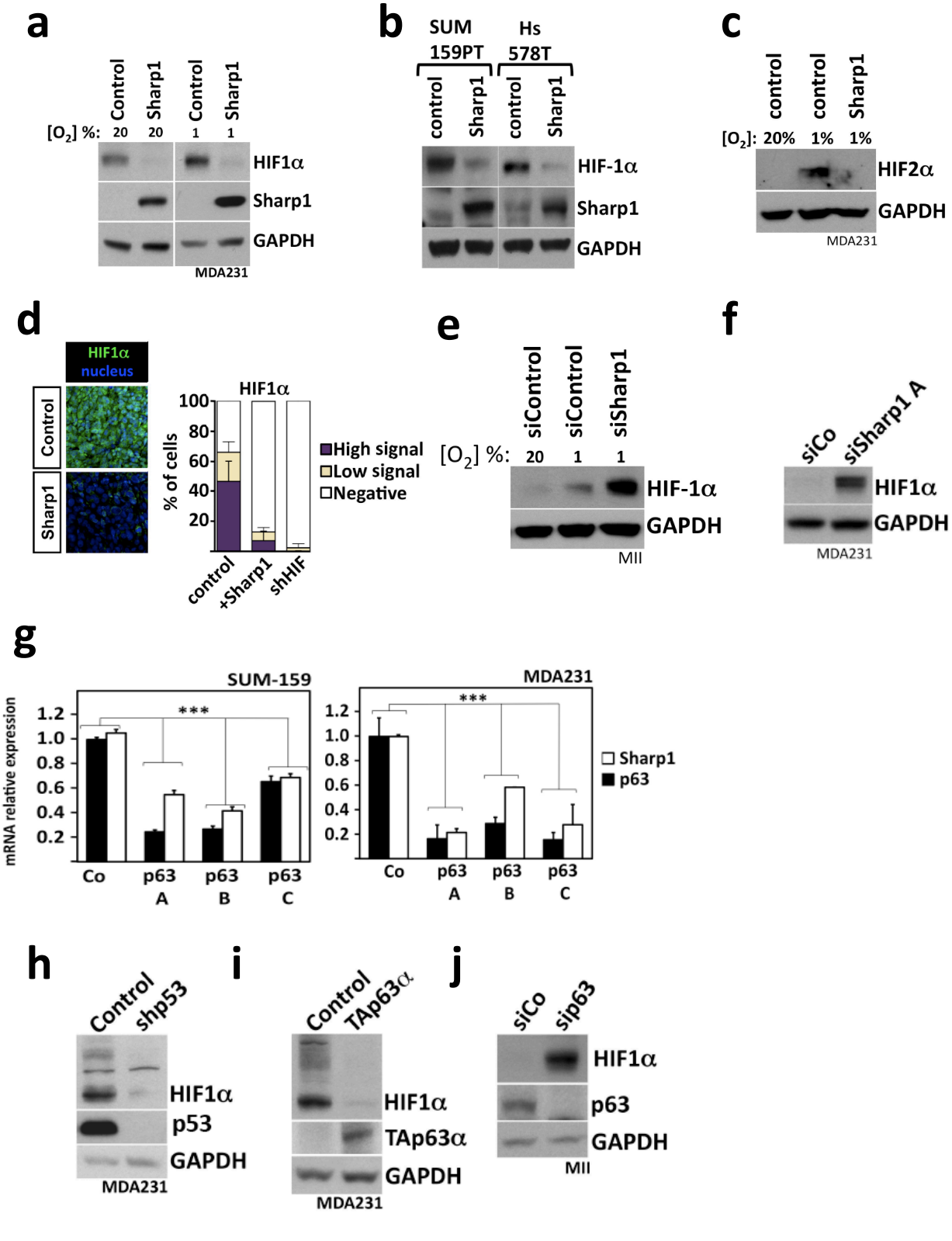

a
Control
Sharp1
Control
Sharp1
[O2] %: 20 20 1 1
HIF1α
Sharp1
GAPDH
MDA231
b
SUM 159PT
Hs 578T
control
Sharp1
control
Sharp1
HIF-1α
Sharp1
GAPDH
c
control
control
Sharp1
[O2]: 20% 1% 1%
HIF2α
GAPDH
MDA231
d
HIF1α nucleus
Control
Sharp1
HIF1α
% of cells
100
80
60
40
20
High signal
Low signal
Negative
control
+Sharp1
shHIF
e
siControl
siControl
siSharp1
[O2] %: 20 1 1
HIF-1α
GAPDH
MII
f
siCo
siSharp1 A
HIF1α
GAPDH
MDA231
g
SUM-159
MDA231
mRNA relative expression
1.2
1.0
0.8
0.6
0.4
0.2
***
Co
p63 A
p63 B
p63 C
Sharp1
p63
h
Control
shp53
HIF1α
p53
GAPDH
MDA231
i
Control
TAp63α
HIF1α
TAp63α
GAPDH
MDA231
j
siCo
sip63
HIF1α
p63
GAPDH
MII

FIGURE 15

Figure 16

a. Sharp1 expression leads to proteasomal-dependent degradation of endogenous HIF-1α in HEK293T cells.

b. Raising Sharp1 leads to reduction of wild-type (wt) or oxygen/pVHL-insensitive mutant (PA) HIF-1α levels. Note the presence of residual PA-HIF-1α protein remaining after Sharp1 overexpression (lane 4).

c. Western blots for the indicated antibodies of cell lysates from RCC4 cells, either null for pVHL or pVHL-reconstituted RCC

d. BALBc/3T3ts20 cells, bearing a temperature-sensitive UBE/E1, were transfected with HIF-1α, alone or with increasing doses of Sharp1. p53 western blot ensures the efficient inhibition of ubiquitin-dependent pathways. The asterisk indicates an aspecific bands at high molecular weight.

e. HEK293T cells were transfected with the indicated plasmids and immunoprecipitated with Flag-M2 resin. Non-Ub (free) HIF-1α is visualized at the bottom of the gel.

f. Extracts form HEK293T cells overexpressing the Flag-tagged proteins were subjected to anti-Flag immunoprecipitation (IP), and tested for association to endogenous α4 subunit of the 20S proteasome by western blot.

g. HEK293T cells were transfected with Flag-tagged HIF-1α alone or together with untagged Sharp1, and incubated with the cell permeable DSP crosslinker before harvesting. Extracts were subjected to anti-Flag immunoprecipitation (IP), and coprecipitating endogenous α4 subunit of the 20S proteasome was detected by immunoblotting after de-crosslinking of the lysate.

h. Endogenous HIF-1α was immunoprecipitated in control or Sharp1-overexpressing MDA-231 cells after DSP treatment, and co-precipitating endogenous α4 subunit of the 20S proteasome was detected by immunoblotting.

i. HEK293T cells were transfected with the indicated siRNAs and with Flag-tagged HIF-1α (+). After harvesting, extracts were subjected to anti-Flag immunoprecipitation (IP), the co-precipitating endogenous 20S α4 proteasome subunit and endogenous Sharp1 proteins were detected by immunoblotting.

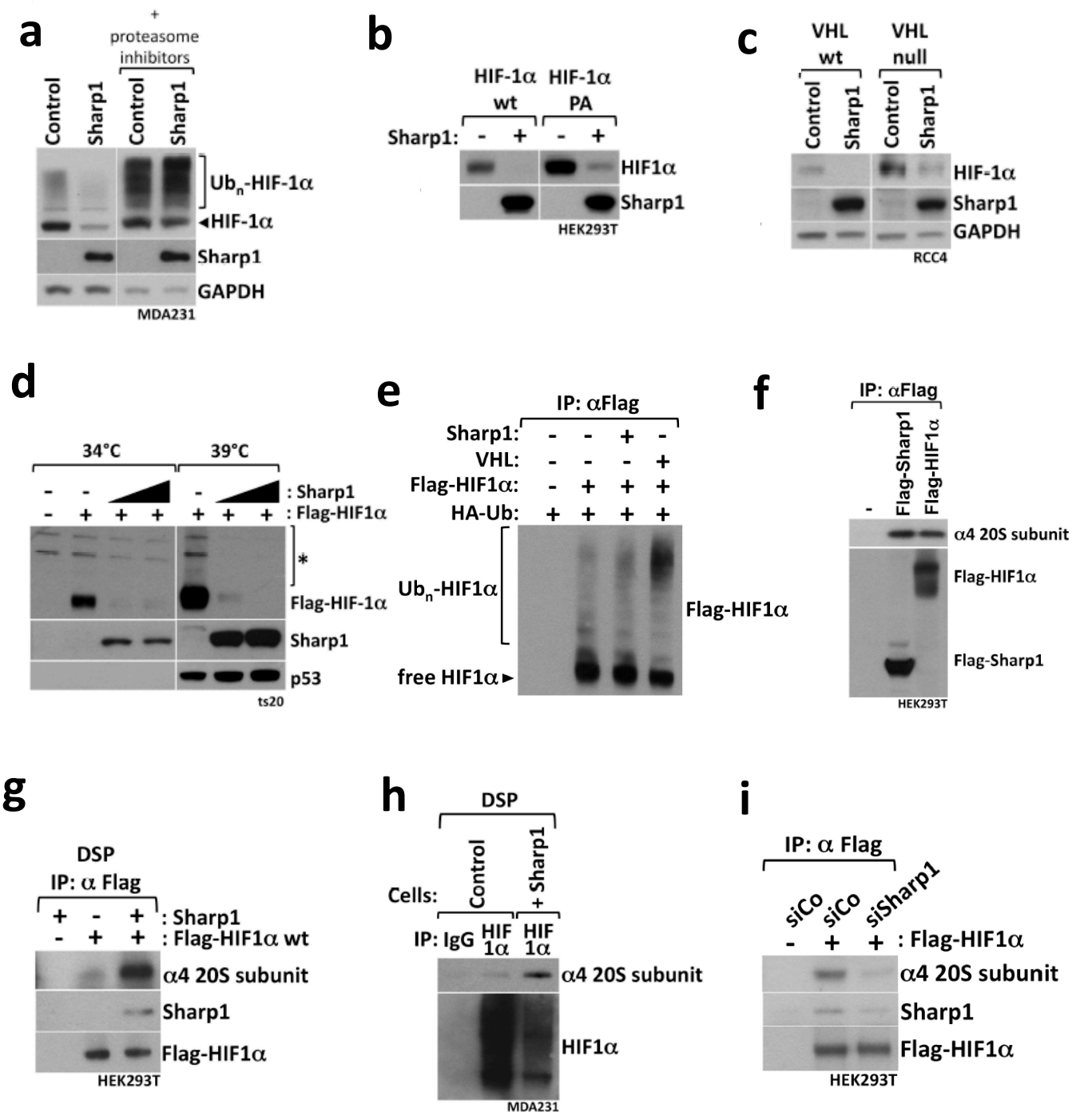

a
+ proteasome inhibitors
Control
Sharp1
Control
Sharp1
Ub$_n$-HIF-1α
◂HIF-1α
Sharp1
GAPDH
MDA231
b
HIF-1α wt
HIF-1α PA
Sharp1: - + - +
HIF1α
Sharp1
HEK293T
c
VHL wt
VHL null
Control
Sharp1
Control
Sharp1
HIF-1α
Sharp1
GAPDH
RCC4
d
34°C
39°C
: Sharp1
: Flag-HIF1α
*
Flag-HIF-1α
Sharp1
p53
ts20
e
IP: αFlag
Sharp1: - - + -
VHL: - - - +
Flag-HIF1α: - + + +
HA-Ub: + + + +
Ub$_n$-HIF1α
free HIF1α ►
Flag-HIF1α
f
IP: αFlag
-
Flag-Sharp1
Flag-HIF1α
α4 20S subunit
Flag-HIF1α
Flag-Sharp1
HEK293T
g
DSP
IP: α Flag
+ - + : Sharp1
- + + : Flag-HIF1α wt
α4 20S subunit
Sharp1
Flag-HIF1α
HEK293T
h
DSP
Cells:
Control
+ Sharp1
IP: IgG
HIF 1α
HIF 1α
α4 20S subunit
HIF1α
MDA231
i
IP: α Flag
siCo
siCo
siSharp1
- + + : Flag-HIF1α
α4 20S subunit
Sharp1
Flag-HIF1α
HEK293T

FIGURE 16

Figure 17

a. Diagrams of the domains of Sharp1 and corresponding deletion mutants (ΔA, Δb, b) used in co-immunoprecipitation experiments with HIF-1α and the α4 subunit of the 20S proteasome. On the right: summary of the results.

b. Sharp1 requires its bHLH domain to interacts with endogenous α4-20S subunit. HEK293T cells were transfected with wildtype or domain deletion mutants of Flag-Sharp1; lysates were then immunoprecipitated with Flag-M2 resin. Coimmunoprecipitated α4-20S subunit was then visualized by Western Blotting.

c. Flag-Sharp1 deletion constructs were transfected alone or together with V5-HIF-1α vector in HEK293T cells. Cells lysates were then immunoprecipitated with Flag-V5 resin and immunoblotted for copurified Flag-Sharp1 isoforms. Sharp1 bHLH domain contributes to HIF-1α binding.

d. Sharp1 requires its α4-interacting domain to promote HIF-1α degradation. HEK293T cells were transfected with Flag-HIF-1α together with empty vector (-), with wild-type Sharp1 (wt), or with the deletion mutants of Sharp1. Cells were then either untreated or treated with a proteasome inhibitors for 8 hours as indicated.

e. Scheme representing the model that we propose in wich Sharp1 through the bHLH domain can connect HIF-1α to the 20S proteasome.

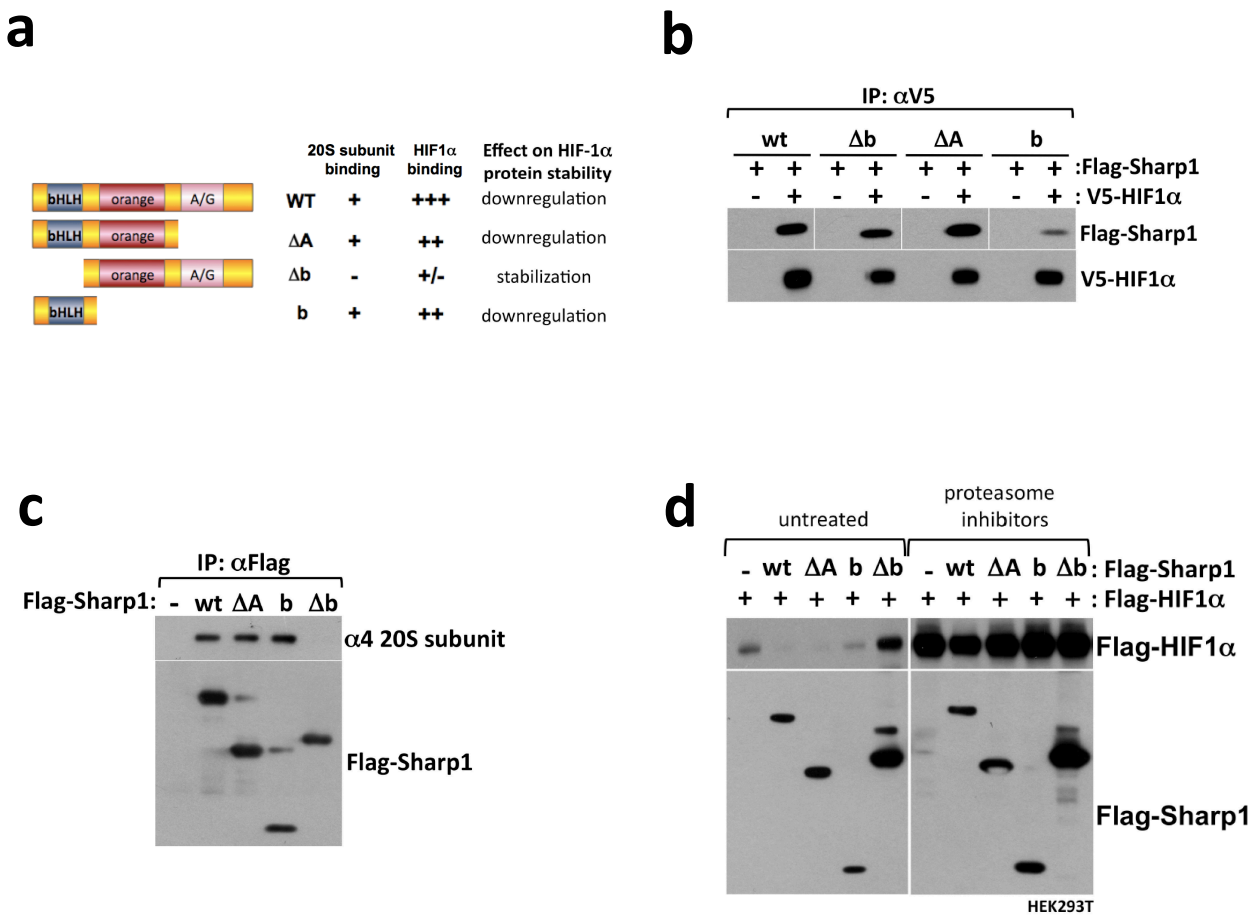

a
20S subunit binding
HIF1α binding
Effect on HIF-1α protein stability
bHLH
orange
A/G
WT + +++ downregulation
ΔA + ++ downregulation
Δb - +/- stabilization
b + ++ downregulation
b
IP: αV5
wt Δb ΔA b
:Flag-Sharp1
: V5-HIF1α
Flag-Sharp1
V5-HIF1α
c
IP: αFlag
Flag-Sharp1: - wt ΔA b Δb
α4 20S subunit
Flag-Sharp1
d
untreated
proteasome inhibitors
- wt ΔA b Δb - wt ΔA b Δb
: Flag-Sharp1
: Flag-HIF1α
Flag-HIF1α
Flag-Sharp1
HEK293T
e
cytoplasm
nucleus
bHLH
A/G
HIFs
Sharp1
Sharp1

FIGURE 17

TABLES

Table A:

PCR oligonucleotides used for genotyping of mouse offspring and embryos. The expected length of the amplified fragment is indicated on the right-most column (BAND).

GENE	ALLELE primer	Nucleotide sequence	BAND (bp)
ECTO	***WT and floxed***		**wt=180 fl=250**
	Ecto-F	CCACTGAGTCACCACTCTAGCACTCAGACG	
	Ecto-R	TCTTAAAAGCTCTGGAGGAACGTCGTCAGG	
	Null		**300**
	EctoNULL-F	GCGTGAAGAGCTTGATTTAGAATACATACACC	
	EctoNULL-R	CATTCAAAGACGAAGAGGAGGAGGTGAAGC	
	Null nested		**180**
	EctoNULL-int-F	TGAGAAAAGTGTACTTTCATCCATTT	
	EctoNULL-int-R	CGTCAGGACTGATTGCACAG	
SMAD4	***WT and Floxed***		**wt=249 fl=280**
	Sm4fl L1	GGGAAGCAAACAAAACAAGTC	
	sm4fl R1	CAGAACAGATGTGCAATCGAA	
	WT		**148**
	Swt LA1	ATACGCGCTTGGGTAGATCTTATGAACAGC	
	Swt RA1	AGGAGTGCAGTTGGAATGTAAAGGTGAAGG	
	Null		**280**
	S- LB1	CAGCATGCTCTGAGCTCACAATTCTCC	
	S- RB1	CATAACAGACTATGTATCAGGGTTAGTGTGG	
NODAL	***Null (LacZ)***		**200**
	Lacz_L2	GTAGGTTTTCCGGCTGATAAATAAGG	
	Lacz_R3	ATAACGACATTGGCGTAAGTGAAGC	
	Δ600		**283**
	Δ600_L1	AATGGTAGAGGGTAAAAAGGGTAGGG	
	Δ600_R2	GAACTTTCTAGAAGGGAAGTGAACTGG	
CRE	*Cre-R*	AGGTTCGTTCACTCATGGA	**218**
	Cre-F	TCGACCAGTTTAGTTACCC	

Table 1: Original breast cancer dataset downloaded from GEO

Study	Affymetrix platform	Samples	Data source	References
Stockholm	HG-U133A	159	GEO GSE1456	Pawitan et al., 2005
EMC-286	HG-U133A	286	GEO GSE2034	Wang et al., 2005
EMC-58	HG-U133A	58	GEO GSE5327	Minn et al., 2007
MSK	HG-U133A	82	GEO GSE2603	Minn et al., 2005
Uppsala-Miller	HG-U133A	236	GEO GSE3494	Miller et al., 2005
Ivshina-Miller	HG-U133A	249	GEO GSE4922	Ivshina et al., 2006
Loi	HG-U133A	327	GEO GSE6532	Loi et al., 2007; Loi et al., 2008; Loi et al., 2010
Sotiriou	HG-U133A	187	GEO GSE2990	Sotiriou et al., 2006
Desmedt	HG-U133A	198	GEO GSE7390	Desmedt et al., 2007
Schmidt	HG-U133A	200	GEO GSE11121	Schmidt et al., 2008
Veridex	HG-U133A	136	GEO GSE12093	Zhang et al., 2009

Table 2: Breast cancer re-organized cohorts comprised in the meta-dataset analyzed in this study.

Cohort	Affymetrix platform	Samples	Data source	References
KI_Stockholm	HG-U133 A	159	GSE1456	Pawitan et al., 2005
EMC-344	HG-U133A	344	GSE2034; GSE5327	Wang et al., 2005; Minn et al., 2007
MSKCC	HG-U133A	82	GSE2603	Minn et al., 2005
KI_Uppsala	HG-U133A	258	GSE3494 - GSE4922 - GSE6532	Loi et al, 2008; Ivshina et al, 2006; Miller et al, 2005
OXF	HG-U133A	178	GSE6532	Ivshina et al., 2006
TransBIG	HG-U133A	198	GSE7390	Desmedt et al., 2007
Mainz	HG-U133A	200	GSE11121	Schmidt et al., 2008
Veridex	HG-U133A	136	GSE12093	Zhang et al., 2009

Table 3: Multivariate analysis of the risk of metastasis for the Triple Negative Breast Cancer dataset using a Cox proportional-hazards model.

In Model 1, tumor size covariate has statistically significant coefficient at α=0.05. However, when the *Sharp1/CyclinG2 signature* is included (Model 2), affiliation to group 'Low', keeping constant all other covariates, significantly increases the hazard of metastasis by a factor of 2.157 on average.

Model 1: Multivariate analysis using clinical variables only.

Model 1 was obtained using n=183 observations and its residual deviance (i.e., minus twice the partial log likelihood) is equal to 546.4102.

Variable	Hazard Ratio	Hazard ratio 95% confidence interval	p-value
Age class > 60 (vs. Age class < 40)	0.963	(0.445 - 2.085)	0.920
Age class 40-60 (vs. Age class < 40)	0.782	(0.407 - 1.505)	0.460
Tumor size (cm)	1.207	(1.042 - 1.399)	0.012

Model 2: Multivariate analysis using clinical variables and the minimal signature.

Model 2 was obtained using n=183 observations and its residual deviance (i.e., minus twice the partial log likelihood) is equal to 539.7222.

Variable	Hazard Ratio	Hazard ratio 95% confidence interval	p-value
Age class > 60 (vs. Age class < 40)	0.843	(0.388 - 1.835)	0.670
Age class 40-60 (vs. Age class < 40)	0.786	(0.409 - 1.514)	0.470
Tumor size (cm)	1.213	(1.045 - 1.409)	0.011
Group low (vs. Group high)	2.157	(1.168 - 3.984)	0.014

Model 1 and Model 2 may be compared to assess whether the *Sharp1/CyclinG2 signature* adds additional prognostic information over the clinical variables. In particular, the value of adding the *Sharp1/CyclinG2 signature* to Model 1 is obtained subtracting the residual deviance of Model 1 from the one of Model 2 and testing this difference against a chi-square distribution with one degree of freedom. The difference between the residual deviances of the model constructed without the *Sharp1/CyclinG2 signature* (i.e., Model 1) and the model including the *Sharp1/CyclinG2 signature* (i.e., Model 2) is equal to 546.4102 – 539.7222 = 6.6879 and exceeds the .95 quantile of the chi-square distribution with one degree of freedom (p-value = 0.0097). As such, the *Sharp1/CyclinG2 signature* is a significant predictor of metastasis-free survival, adding new prognostic information beyond the one provided by the standard clinical predictors.

Clinical Predictor	Difference of residual deviances	p-value
Age class	6.5583	0.0104
Tumor size	6.8927	0.0087

Table shows how the *Sharp1/CyclinG2 signature* adds prognostic value not only to the multivariate model but also to any model constructed using any single clinical predictor. Indeed, the difference between the residual deviance of the model obtained using a single clinical variable plus the *Sharp1/CyclinG2 signature* (e.g., tumor diameter+ *Sharp1/CyclinG2 signature*) and the residual deviance of the model obtained using only a clinical variable, is significant for each clinical predictor.

Table 4: Gene Signatures used for GSEA, metastasis and survival analysis.

For references to the other Gene Signatures analyzed see Methods description.

TGF-β signature
ADAM19
ANGPTL4
CALCR
CCDC99
CD3EAP
CDH19
CDK5R1
CHST11
CHST3
COL4A1
CYP24A1
CYP26B1
DDB2
DNAJB2
EGR4
EHD1
EPHB2
ETS2
GADD45B
GPR17
GPR87
HBEGF
HSF2BP
IL11
KLF7
LIMS1
MCL1
MFHAS1
MYO10
NGF
PDGFB
PDLIM4
PLAU
PLXNA2
PPP1R13L
RUSC2
SERPINE1
SGK1
SLC7A1
SLC7A5
SLC7A6
SMOX
SPEN
SPHK1
SRRD
TGFBI
TMC7
TUBA4A
UCK2
VDR
WNT5B

HIF activity signature
VEGFA
PDK1
BNIP3
NDRG1
HK2
PDPK1
BHLHE40
DDIT4
INSIG1
PGK1
PFKP
SAP30
SLC2A3
SLC2A1
ALDOC
GLRX
ENO2
CENPF
MAFF
ADM
BIRC2
WSB1

Sharp1-activity signature
AGTBP1
CHN2
COBL
DSC2
EPS8L2
F2RL1
GBE1
GPR56
HSF2
IFIT3
IGF2BP3
IMPA2
ITGB2
LPIN1
LYZ
ME1
NOX5
PLCE1
RAPGEF5
S100A3
SH2D3A
SLC5A3
SLCO4A1
SREBF1
TGFA
WWTR1

Table 5. Enrichment of signaling pathways: genes up-regulated in Sharp1/CyclinG2 low samples.

Differentially expressed genes have been identified using SAM. Setting q-value<1% resulted in 2804 genes up-regulated in Sharp1/CyclinG2 signature low samples. Enrichment has been determined using a Fisher's test on all 254 signaling pathways of GSEA. p-values have been corrected using BH procedure.

Pathway	p-value	BH p-value
mutant-p53	<0.00001	<0.00001
Ras	<0.00001	<0.00001
TGFb	<0.00001	0.0015
MYC	0.0001	0.006
HIFs	0.0002	0.007
Src/RPTP	0.0002	0.007
TCF4	0.001	0.031
Proteasome	0.003	0.099
P27	0.007	0.1614
E2F3	0.007	0.161

Table 6 : Genes Downregulated in MDA-shHIF and MDA-Sharp1 cells

Genes were selected from microarray data according to Fold Change < -1.5 (compared to MDA-shGFP cells) and a pfp (percentage of false positive prediction) < 0.05

	MDA-shHIF	**MDA-Sharp1**
Genesymbols	*FoldChange*	*FoldChange*
PRR20	-3.85	-1.96
TFF2	-3.45	-11.11
EMP1	-3.23	-1.64
ITGB2	-3.12	-4.55
NANOS1	-2.86	-1.72
INSIG1	-2.78	-2.63
S100A2	-2.78	-4.55
IMPA2	-2.63	-1.92
TFF1	-2.63	-6.25
ENPP1	-2.56	-2.22
TMEM87B	-2.56	-1.64
AKAP12	-2.44	-3.85
FOXP2	-2.44	-2.50
ITGB2	-2.44	-2.70
ZNF395	-2.44	-1.64
INSIG1	-2.38	-2.70
NOX5	-2.38	-6.25
SLCO4A1	-2.38	-10.00
TMEM87B	-2.38	-1.67
GPR56	-2.27	-4.00
RANK	-2.22	-2.17
COL12A1	-2.17	-1.67
OLFML2A	-2.17	-2.94
SHANK2	-2.17	-3.03
ZNF395	-2.13	-1.69
AFAP1L1	-2.08	-1.50
AHNAK2	-2.08	-1.92
PLLP	-2.08	-2.63
WDR66	-2.08	-4.00
DSC2	-2.04	-2.17
SLCO4C1	-2.04	-1.67
ZNF395	-2.04	-1.69
IRAK3	-2.00	-2.27
USP12	-1.96	-1.50
AGTPBP1	-1.92	-2.04
FOXP2	-1.92	-2.78
ME1	-1.89	-2.08
SLC2A13	-1.89	-1.79
FSCN1	-1.85	-1.52
LIFR	-1.85	-1.52
LPIN1	-1.85	-1.75
PDCD4	-1.85	-1.54
SH2D3A	-1.85	-3.57
AKAP12	-1.82	-3.33

CALHM3	-1.82	-1.82
IFIT3	-1.82	-1.67
INADL	-1.82	-1.61
LOC645431	-1.82	-3.45
PAQR5	-1.82	-2.86
S100A6	-1.82	-2.94
FSCN1	-1.79	-1.50
SHANK2	-1.79	-3.23
ATPBD4	-1.75	-1.50
COL12A1	-1.75	-1.69
	MDA-shHIF	**MDA-Sharp1**
Genesymbols	*FoldChange*	*FoldChange*
INHBB	-1.75	-3.12
PDLIM2	-1.75	-1.50
PLCE1	-1.75	-1.54
ZNF385B	-1.75	-1.52
ACOX2	-1.72	-1.75
COL12A1	-1.72	-1.79
LOC100127888	-1.72	-7.14
MBD5	-1.72	-1.59
PLCE1	-1.72	-1.61
RAPGEF5	-1.72	-1.64
DSC2	-1.69	-1.92
IFIT3	-1.69	-1.56
NOX5	-1.69	-3.57
S100A3	-1.69	-2.27
SHANK2	-1.69	-3.33
ANKRD12	-1.67	-2.17
LOXL4	-1.67	-2.56
MAP7D3	-1.67	-1.52
TMCC3	-1.67	-3.33
LY75	-1.64	-2.00
LYZ	-1.64	-2.38
TFF3	-1.64	-3.23
F2RL1	-1.61	-2.17
IGF2BP3	-1.61	-1.79
SDR16C5	-1.61	-2.94
SREBF1	-1.61	-1.50
ZDHHC23	-1.61	-2.86
FLJ39632	-1.59	-1.54
MYO6	-1.59	-1.52
SUSD3	-1.59	-1.54
AGPAT9	-1.56	-2.50
DOPEY2	-1.56	-1.59
ENPP1	-1.56	-2.08
EPS8L2	-1.56	-1.69
FHOD3	-1.56	-1.75
INSIG1	-1.56	-2.63
QKI	-1.56	-1.56
TGFA	-1.56	-2.08
AKAP12	-1.54	-4.17
CAMK2D	-1.54	-1.96
CHN2	-1.54	-2.27
ENPP1	-1.54	-2.08
FAS	-1.54	-1.54
MYO6	-1.54	-1.89
SH3TC2	-1.54	-3.33
C13orf38	-1.52	-2.04
CDH19	-1.52	-1.54
CHD9	-1.52	-1.61
COBL	-1.52	-2.17
MEGF9	-1.52	-1.56
RICTOR	-1.52	-1.54
SPOCD1	-1.52	-3.23
TIMP3	-1.52	-2.63
ATP8A1	-1.50	-1.96
BNIP3	-1.50	-3.33
CA12	-1.50	-3.03
CCDC69	-1.50	-2.33
DUSP4	-1.50	-1.92
LFNG	-1.50	-1.89
NNT	-1.50	-1.82
PAG1	-1.50	-3.33
PLEC	-1.50	-1.50
SH3TC2	-1.50	-2.00
SLC5A3	-1.50	-1.50
WDR66	-1.50	-1.59

Table 7: Genes Upregulated in MDA-shHIF and MDA-Sharp1 cells

Genes were selected from microarray data according to Fold Change > +1.5 (compared to MDA-shGFP cells) and a pfp (percentage of false positive prediction) < 0.05

	MDA-shHIF	**MDA-Sharp1**
Genesymbols	*FoldChange*	*FoldChange*
THBS1	7.91	1.58
TNFSF15	3.37	5.79
ANXA10	3.26	1.51
CLDN1	2.90	3.78
LOC201651	2.86	3.85
SELM	2.64	2.08
CCND2	2.57	3.91
SPINK1	2.32	5.82
SCG5	2.29	1.98
HHIP	2.17	2.46
HIST1H4H	2.16	2.11
CBLB	2.13	1.53
OTUD1	2.13	1.75
CPA3	2.09	2.22
DMBT1	2.08	2.38
IL7R	2.06	2.12
DDX10	2.03	1.72
NNMT	2.03	7.74
NUPR1	2.01	2.41
SIPA1L2	2.01	3.34
IL7R	1.93	1.66
EGR1	1.91	1.92
OTUD1	1.90	1.64
CCND2	1.87	1.73
CLDN1	1.87	1.93
C11orf96	1.86	1.64
IFI27	1.86	1.50
COL4A5	1.84	1.78
MFI2	1.84	2.16
ANKRD55	1.82	1.65
HHIP	1.81	3.14
NNMT	1.81	7.07
CDK15	1.80	2.47
EGR1	1.79	2.46
NTN4	1.79	2.47
RAB15	1.79	2.26
SERPINB7	1.79	3.88
EFEMP1	1.78	3.25
EGR1	1.75	2.27
PLCXD2	1.74	2.64
C12orf39	1.73	11.98
DUSP10	1.73	2.25
HHIP	1.73	2.47
ICAM1	1.73	3.31

	MDA-shHIF	**MDA-Sharp1**
Genesymbols	*FoldChange*	*FoldChange*
COL11A1	1.72	2.85
ECM1	1.72	2.52
EFEMP1	1.71	2.60
FNDC3B	1.71	1.60
KBTBD8	1.71	2.11
ZNF503	1.69	1.78
IRX3	1.68	1.57
RAB15	1.68	2.06
FAM24B	1.66	1.54
RASSF2	1.66	1.52
ICAM1	1.65	2.39
SDSL	1.65	1.61
SIPA1L2	1.64	2.50
COL1A1	1.63	1.95
PALM2	1.63	2.06
RHOJ	1.63	2.32
C10orf35	1.62	1.65
HIST1H2BD	1.61	1.58
SLC16A9	1.60	1.68
CDR2L	1.59	1.73
LOC645638	1.59	1.64
ARL14	1.58	1.54
HIST1H4H	1.58	1.72
ID3	1.58	4.75
TMEM150A	1.58	1.67
AADAC	1.57	2.60
DUSP10	1.57	2.52
ID2	1.57	1.81
RARB	1.57	2.81
UCA1	1.57	3.72
C9orf119	1.54	1.56
CTPS2	1.54	1.53
CYLD	1.54	1.81
LOX	1.54	1.75
HIST2H2BE	1.53	1.78
HIST1H2AC	1.52	1.54
RHOBTB3	1.52	1.55
CACHD1	1.51	1.78
PCYOX1L	1.51	1.78
ZBTB47	1.51	1.57
FSTL1	1.50	1.50
KDELR3	1.50	1.67
SPOCK1	1.50	1.51

Table 8: Sequences of siRNA and shRNA

Name	Sequence	Reference and Validation
Control sequence A	GCAAGCUGACCCUGAAGUU	-
Control sequence B	GAAGUAUUCCGCGUACGUU	-
HIF-1α sequence A	GAAUUCUCAACCACAGUGC	See Suppl. Fig. 4 and(Yang et al., 2008)
HIF-1α sequence B	CUGAUGACCAGCAACUUGA	See Suppl. Fig. 4 and(van Uden et al., 2008)
HIF-2α sequence A	GCGACAGCUGGAGUAUGAA	See Suppl. Fig. 4 and(Aprelikova et al., 2004)
HIF-2α sequence B	CAGCAUCUUUGAUAGCAGU	See Suppl. Fig. 4 and(van Uden et al., 2008)
HIF-1β siRNA	AAGACUCGUACUUCCCAGUUU	See Suppl. Fig. 11 and(Wright and Duckett, 2009)
p63 siRNA A	CAUCAUGUCUGGACUAUUU	see Suppl. Figure 16
p63 siRNA B	CGACAGUCUUGUACAAUUU	see Suppl. Figure 16 and Ref. (Adorno et al., 2009)
p63 siRNA C	UCUUGUUUGUCGCACCAUCTT	see Suppl. Figure 16
Sharp1 sequence A	GCUUUAACCGCCUUAACCG	see Suppl. Figure 15
Sharp1 sequence B	CGAGACGACACCAAGGAUA	see Suppl. Figure 15

Control sequence A was used in experiments with MDA-231 and SUM-159 cells, Control sequence B was used in experiment with MCF10A-MII.

Table 9: qPCR and RT-PCR Primer sequences

Name	Sequence
VEGFA	CCTTGCCTTGCTGCTCTACCTC
	TTCTGCCCTCCTCCTTCTGC
HK2	GTGCCCGCCAGAAGACATTA
	TGCTCAGACCTCGCTCCATT
ANGPTL4	TCCGCAGGGACAAGAACTG
	GCCGTTGAGGTTGGAATGG
CAIX	TGGCTGCTGGTGACATCCTA
	TTGGTTCCCCTTCTGTGCTG
BNIP3	TCTGCTGCTCTCTCATTTGCTG
	AGGTGCTGGTGGAGGTTGTC
NDRG1	TTGTGCGCAAGGCTGGAT
	AATGTGCTGGCGGTAGGTGT
LOXL2	CTGGACCCCATGAAGAATGT
	CTCTGGCTTGTACGCTTTCC
miR-210	GGGTGGGGGTAAGCGAAAT
	CTGGACACACAAGGAAAAGAAAG
PDK1	TCCTGTCACCAGCCAGAATG
	TCCTTTGCCTTTTCCACCAA
CLDN1	CTGTCATTGGGGGTGCGATA
	GCCTTGGTGTTGGGTAAGAGG
ITGB2	ACCAGCCCAGAGGTGACTGT
	CGGATGACAAACGACTGCTC
INSIG1	ATTCGTTCTTGGCTCCTTG
	TCACTATGGGGCTTTTCAGGA
PRR20	ACCTGTGGACACCTCGACT
	GCCTTGGTTCCTCCATGAGC
S100A2	CTGCCTTGCTCTCCTTCCTG

	GCTTGAACTTGTCGCCCTCT
RANK	TTTCCTCCACGGACAAATGC
	CACAGACGCGAAGAGAAGCA
THBS1	GAACGGGAAACCCTGTGAAG
	GCAGTCCTTGCCTCCAAACT
IL7R	TGGAGAAAGTGGCTATGCTCA
	GACATCTGGGTCCTCAAAAGC
EGR1	AGCAGCACCTTCAACCTCA
	GGTCTCCACCAGCACCTTCT
ICAM1	CACAGTCACCTATGGCAACG
	CTGGCTTCGTCAGAATCACG
BHLHB3	AGCCTACCGTCCCACAGATT
	CCTTGGTGTGTCGTCTCGTTTC
GAPDH	CCGAGAACGACACGGACAC
	GGCTCCTGCTTGATGGTGAC

Name	Sequence
TAp63 isoform	GTCCCAGAGCACACAGACAA
	GAGGAGCCGTTCTGAATCTG
ΔNp63 isoform	CTGGAAAACAATGCCCAGAC
	GGGTGATGGAGAGAGAGCAT
p63α isoform	GAGGTTGGGCTGTTCATCAT
	AGGAGATGAGAAGGGGAGGA
p63β isoform	AACGCCCTCACTCCTACAAC
	CAGACTTGCCAGATCCTGA
p63γ isoform	ACGAAGATCCCCAGATGATG
	GCTCCACAAGCTCATTCCTG