PhD degree in Molecular Medicine

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples “Federico II”

Faculty of Medicine

Settore disciplinare: MED/04

Functional characterization of the endocytic protein Epsin3 in breast cancer

Iavarone Claudia

IFOM-IEO Campus, Milan

Matricola n. R08909

Supervisor: Prof. Pier Paolo Di Fiore

IFOM-IEO Campus, Milan

Anno accademico 2012-2013

TABLE OF CONTENTS

FIGURE AND TABLE INDEX

TABLE OF ABBREVIATIONS

AJ	Adherent junction
AML	Acute myeloid leukemia
AP2	Adaptor protein 2
Cbl	Casitas B-lineage lymphoma
CCP	Clathrin coated pit
CME	Clathrin-mediated endocytosis
CSC	Cancer stem cell
DCIS	Ductal carcinoma in situ
DSL	Delta Serrate Lag2
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal transition
ENTH	Epsin N-terminal homology
EPN1	Eps15-interacting protein 1
EPN2	Eps15-interacting protein 2
EPN3	Eps15-interacting protein 3
EPS15	Epidermal growth factor receptor substrate 15
ER	Estrogen receptor
ERK	Extracellular signal regulated kinase
FISH	Fluorescence in situ hybridization
GPCR	G protein-coupled receptor
HER2	Human epidermal growth factor receptor 2
HGF	Hepatocyte growth factor
HIF1a	Hypoxia-induced factor 1 alpha
Hip1	Huntingtin-interacting protein 1
IDC	Invasive ductal carcinoma
IF	Immunofluorescence
IHC	Immunohistochemistry
IL-2Rβ	Interleukin 2 receptor type β
KD	Knock-down
KI	Knock-in
KO	Knock-out
MAPK	Mitogen-activated protein kinase
MHC I	Major histocompatibility complex I
NCE	Non clathrin endocytosis
NICD	Notch intracellular domain
NSCLC	Non-small-cell lung cancer
PgR	Progesterone receptor
PI3K	Phosphoinositide 3-kinase
PIP2	Phosphatidylinositol-4,5- bisphosphate
PLCγ	Phospholipase C γ
PM	Plasma membrane
Q-PCR	Quantitative PCR

RTK	Tyrosine kinase receptor
SARA	Smad anchor for receptor activation
SD	Standard deviation
SEM	Standard error of the mean
shRNA	Short hairpin RNA
SNX1	Sorting nexin 1
TCGA	Tumor cancer genome atlas
TGFβ	Transforming growth factor β
TGFβR	Transforming growth factor β receptor
TK	Tyrosine kinase
TMA	Tissue microarray
UIM	Ubiquitin interacting motif
WB	Western blot

ABSTRACT

Endocytosis plays a critical role in the maintenance of cellular homeostasis. Due to its ability to resolve extracellular signals, the endocytic pathway controls many cellular processes, including transcription, proliferation, cell migration and cell fate determination. One relevant, unanswered, question regarding the role of endocytosis in the cell is whether subversion of the endocytic route is involved in the development of pathological conditions, such as cancer. This possibility is supported by studies showing altered expression of several endocytic proteins in human tumors. In a previous study in our lab, the endocytic protein Epsin 3 (Epn3) was found in a gene signature prognostic for metastatic breast cancer. Epn3 belongs to the Epsin family of endocytic proteins. Unlike the other Epsin members, Epn1 and Epn2, which are ubiquitously expressed, Epn3 is exclusively expressed in gastric cells and in wounded or pathological tissues exhibiting altered cell-extracellular matrix interactions. While Epn1 and Epn2 have been well characterized as endocytic adaptors, the exact function of Epn3 protein in endocytosis or signaling is largely unknown.

In the present study, we show that Epn3 is overexpressed in approximately 30% of breast tumors and that upregulation of this protein correlates with clinico-pathological parameters of aggressive disease. We also show that *EPN3* gene is amplified in human breast tumors and that this genetic alteration can occur independently from *HER2* amplification in these tumors. Through functional studies using *in vitro* and *in vivo* breast cancer models, we demonstrate that breast tumor cells harboring Epn3 amplification/overexpression are dependent on Epn3 deregulation for the maintenance of their tumorigenic potential. In agreement with a possible oncogenic role, overexpression of Epn3 in a breast tumor cell line

increases the tumorigenic potential *in vivo*. Of note, Epn3 overexpression is also able to induce, both in normal and tumor breast cells, transcriptional and morphological changes typical of an epithelial-to-mesenchymal transition (EMT), in a TGFβ–dependent manner; indeed, Epn3 overexpression induces upregulation of TGFβ ligands and receptors and inhibition of TGFβ signaling is able to revert Epn3-induced EMT. In addition, we show that Epn3 increases the capacity of normal mammary epithelial cells to form mammospheres *in vitro*, suggesting that Epn3 may contribute to the expansion of the breast cancer stem cell compartment.

Finally, we demonstrate that Epn3 is involved in E-cadherin internalization, by inducing its downregulation from the cell surface upon TGFβ-stimulation.

In conclusion, our data suggest a novel oncogenic role for Epn3 in human breast cancer, through its action as an endocytic protein on E-cadherin trafficking. Overexpression of Epn3 might enhance E-cadherin internalization, and consequently induce EMT activating the expression of mesenchymal proteins that promote cell invasion and lead to tumor aggressiveness.

INTRODUCTION

Emerging evidence suggests that endocytosis is a master regulator of cell signaling events that controls not only signal attenuation, but also the duration, intensity, integration and spatial restriction of the signals [1]. By exerting this regulatory role, endocytosis is believed to control many cellular processes including transcription, proliferation, cell migration and cell fate determination [2]. It is possible therefore that subversion of endocytic pathways could have a role in the development and progression of pathological conditions, such as cancer. Indeed, the altered expression of several endocytic proteins in human tumors argues in favor of such hypothesis (reviewed in [3]).

We recently uncovered a potential role for the endocytic protein Epsin3 (Epn3) in breast tumorigenesis. By performing a transcriptional profiling of human breast tumors from a cohort of surgical breast cancer patients with complete follow-up, we identified Epn3 in a transcriptional signature that is prognostic for metastatic breast cancer (our preliminary unpublished data). Increased *EPN3* mRNA levels were detected in 20% of all breast tumors analyzed and this percentage increased to 45% in the subgroup of poor prognosis tumors. Interestingly, the *EPN3* gene is located on the chromosome 17 which also harbors the *HER2* gene, one of the most clinically relevant and frequently amplified oncogenes in breast cancer.

Based on these initial observations, we hypothesized that Epn3 might be involved in the development of breast cancer and designed a study aimed at: i) analyzing Epn3 expression in human breast tumors and its association with clinico-pathological and prognostic markers in a large cohort of breast cancer patients; ii) characterizing the involvement of Epn3 in breast tumorigenesis

through functional studies in cell-based model systems; iii) investigating the molecular mechanisms underlying such role of Epn3.

Thus, in this introduction I will provide an overview of the current literature relevant to this project: in Chapter 1, I will present evidence arguing for the existence of a network between endocytosis, cell signaling and cancer; in Chapter 2, I will describe what is currently known about Epn3 and the epsin family; finally, in Chapter 3, I will explain why the identification of novel markers through genetic profiling is a useful approach to advance the clinical management of breast cancer.

1 The complex network of endocytosis, cell signaling and cancer

1.1 Endocytosis as a master regulator of cell signaling

Endocytosis has long been viewed as a cellular program through which eukaryotic cells internalize nutrients, membrane-associated molecules and viruses. However, this is a simplistic and primordial definition of the endocytic process. Recent data has shown that endocytosis is a master regulator of signaling circuits, since it has the ability to resolve signals coming from the extracellular environment in time and space. Indeed, endocytosis regulates signal transduction as well as morphogenetic aspects of normal cell physiology, including cell proliferation, cell adhesion and migration (reviewed in [1, 2, 4]).

1.1.1 “Who and how”: different entry routes for different cargoes

Multiple cell entry routes have been identified. Figure 1 illustrates the main endocytic pathways studied so far: phagocytosis, macropinocytosis, and micropinocytosis (which includes clathrin-dependent endocytosis, caveolin-dependent endocytosis and clathrin and caveolin-independent pathways). These internalization pathways differ in the cargoes they transport and in the protein machinery that is involved in the endocytic route.

Phagocytosis internalizes large particles (> 500nm), including bacteria and apoptotic cells [5] while macropinocytosis is responsible for fluid uptake [6]. Both processes are characterized by large actin-based rearrangements of the plasma membrane (PM) [7].

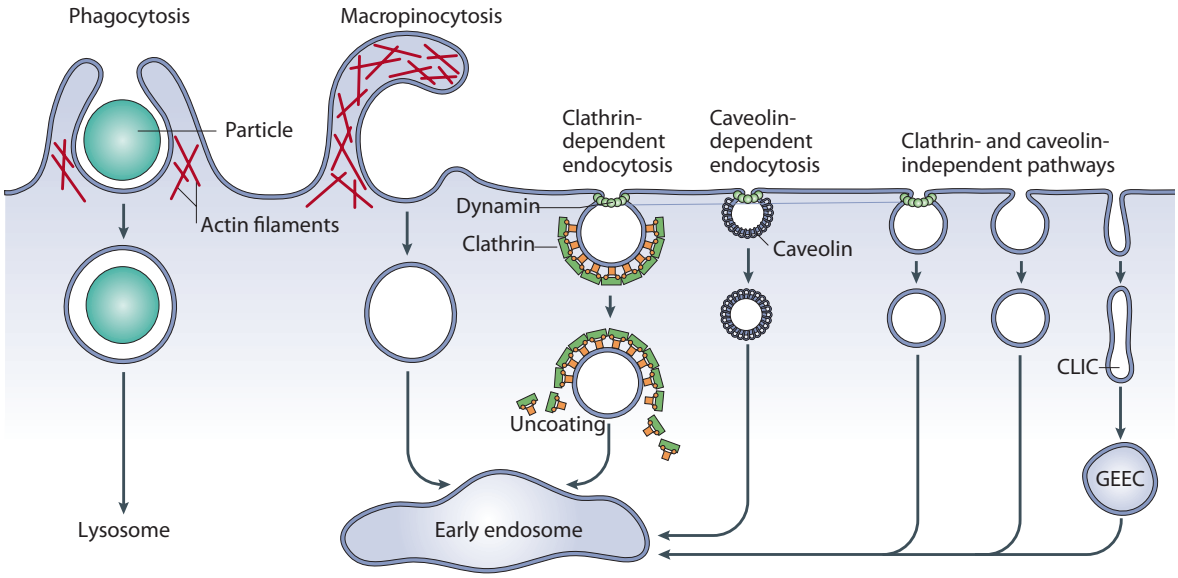


Figure 1 | Pathways of entry into cells.

The known endocytic pathways are shown. The major endocytic regulators are represented (i.e. actin filaments, clathrin, dynamin, caveolin). The intracellular trafficking of internalized molecules is illustrated (i.e. early endosome, lysosome) [adapted from [8]]

Micropinocytosis is, instead, characterized by smaller invaginations of the PM (<200 nm) and comprises clathrin-mediated endocytosis (CME) and non-clathrin-mediated endocytosis (NCE).

CME is the most well characterized endocytic pathway. This pathway is responsible for the uptake of nutrients, pathogens, growth factor receptors and PM associated molecules (reviewed in [9]). PM cargoes are recruited in clathrin-coated pits (CCPs) and adaptor molecules mediate this process by binding both clathrin and protein cargoes [10-12].

Several accessory proteins have been described to be associated with CCPs. The best characterized adaptor protein is the adaptor protein 2 (AP2), as it was the first accessory protein to have been described and it is the most abundant in cells.

Nevertheless, many other adaptors have been proposed to substitute AP2, such as Epsin and EPS15 (in the case of internalized epidermal growth factor receptor - EGFR [13, 14]) or β-arrestin (in the case of internalization of G protein-coupled receptors - GPCR [15]).

The term NCE is used to refer to a heterogeneous group of pathways that are insensitive to clathrin depletion, while they are dependent on cholesterol-rich PM domains [8, 16]. The molecular characterization of NCE is still limited and currently its classification relies on three major criteria: 1) dependency on dynamin for releasing vesicles; 2) presence of “coat-like” proteins involved in the membrane curvature, as in the case of caveolin-mediated endocytosis; 3) dependency on cargo-specific small GTPases, such as Arf-6 for the internalization of major histocompatibility complex I (MHCI) [17], RhoA for uptake of interleukin-2 receptor b (IL-2Rβ) [18] and, finally, Cdc-42 and GRAF-1, which regulate the endocytosis of fluid-phase markers through the CLIC/GEEC pathway [19].

Some receptors can be internalized through both CME and NCE routes with different signaling outcomes. The choice of internalization pathway is one of the many mechanisms used by endocytosis to control signaling output.

1.1.2 Endocytosis regulates cell signaling at different levels

Regulation of cell signaling through endocytosis can occur at different stages of the internalization process: before receptor entry at the PM, at the cell entry route selection step, and after endosome formation (Figure 2).

At the PM, several mechanisms are used by endocytosis to control receptor signaling (Figure 2A). The major mechanism responsible for the negative regulation of cell surface signals is ligand-induced internalization of signaling receptors: continuous ligand stimulation, including those ligands activating tyrosine

kinase receptors (RTKs) and GPCR, induces receptor internalization and degradation, thereby reducing the number of receptors from the PM.

Another mechanism used by endocytosis for signaling control is the recruitment and activation, exclusively at the PM, of specific molecular transducers. This approach is used during the internalization of RTKs, such as EGFR, and GPCRs (Figure 2A). In the case of EGFR, ligand binding induces receptor autophosphorylation that, in the presence of phosphatidylinositol 4,5-bisphosphate (PIP_2) at the PM, leads to either activation of phosphoinositide 3-kinase (PI3K) or phospholipase Cγ (PLCγ). The activation of these two signaling pathways is abrogated in the endosomes by constitutive lack of accessible PIP_2 in these intracellular compartments [20] (Figure 2A,a). GPCR-mediated signaling via PLC starts at the PM where trimeric G proteins reside and where G proteins can initiate and mediate signaling, due to the prevalent, if not exclusive, localization of PLC substrate, PIP_2, at the PM [21] (Figure 2A,b).

PIP_2 controls also the duration of signals coming from the PM by directly regulating the maturation of CCPs [22]. Indeed, in the case of RTKs and GPCRs internalization, cargo accumulation in CCPs was shown to be a highly regulated process that directly impacts on PM signaling [23].

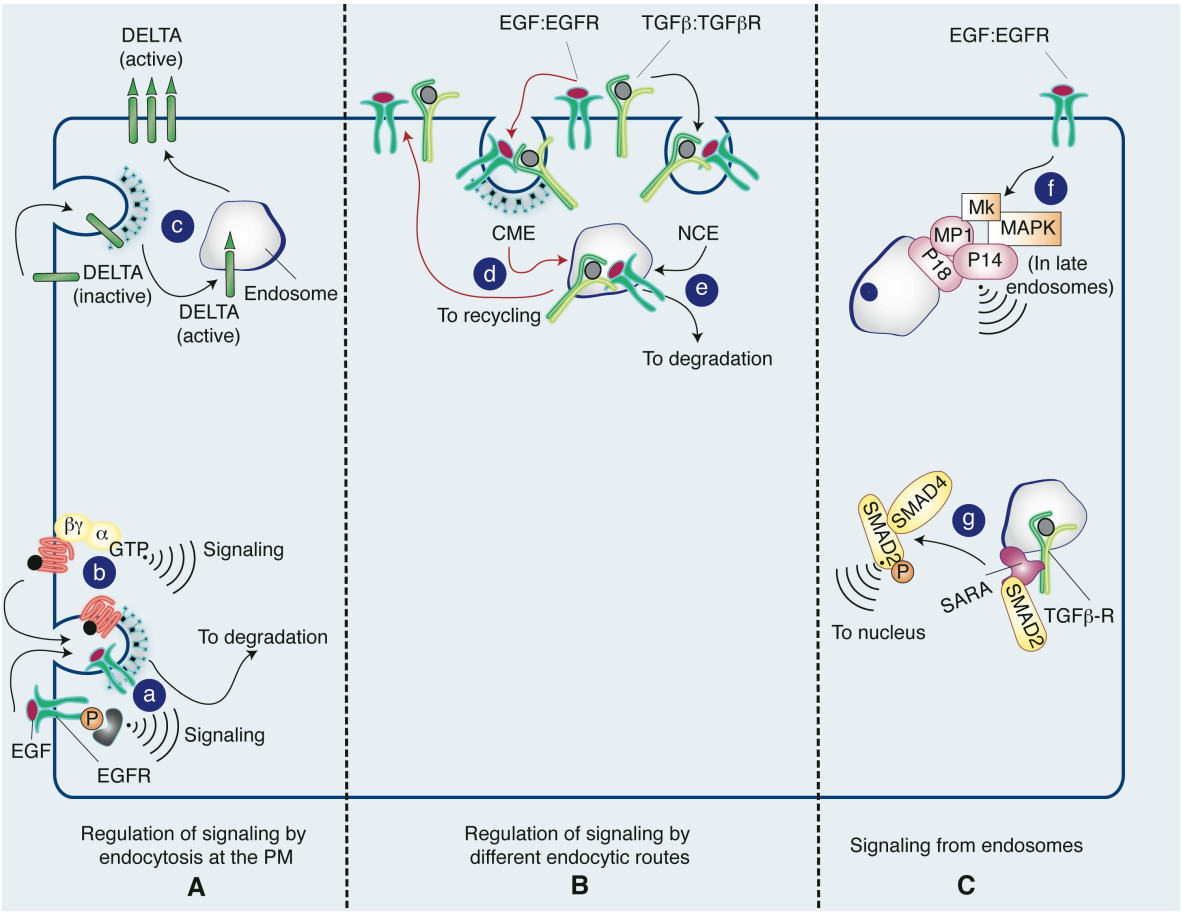


Figure 2 | Endocytosis regulates signaling at different levels.

A) Endocytosis controls signaling at the PM. B) Different routes regulate different biological output. C) Endocytosis can sustain signaling through the endosomes [adapted from [4]] (see text for details).

At the PM, endocytosis can also regulate ligand accessibility. One good example of this kind of regulation is the Notch receptor system. Notch receptors bind ligands of the DSL family (Delta, Serrate and Lag2), which are transmembrane proteins. After binding, two proteolitic cleavages occur in the Notch-presenting cell, referred as signal-receiving cell. These cleavages release the intracellular domain of Notch (NICD), allowing its translocation to the nucleus where it promotes the transcription of target genes. Endocytosis seems to be

necessary to activate Notch signaling both in signal-receiving and in signal-sending cells (reviewed in [24, 25]). In the signal-receiving cell, Notch endocytosis is essential for a second cleavage that occurs in the endosomes [26]. In the signal-sending cell, DSL internalization produces a mechanical pulling force that induces a conformational change in Notch and leads to Notch internalization and activation in the single-receiving cell [27]. Recent evidence in *Drosophila melanogaster*, shows that Delta endocytosis and recycling is essential to achieve high ligand concentration at the PM, necessary to activate a robust Notch signaling response [28] (Figure 2A,c).

At the internalization step, many signaling receptors can select the preferred cell entry route. EGFR, transforming growth factor-β receptor (TGFβR), Notch and Wnt receptors are some of the receptors that can be internalized either through CME or NCE. Of note, internalization of a receptor through different endocytic routes results in different final biological outputs (Figure 2B). For example, internalization of TGFβR through CME leads to downstream signaling (phosphorylation and activation of Smad2 and Smad3 proteins) and recycling of the receptors back to the PM (Figure 2B). However, if TGFβR is internalized via NCE, the recptor is preferentially targeted for ubiquitination and degradation, a process that involves the recruitment of Smad7 and the E3 ligase SMURF [29] (Figure 2B,e).

While in the case of TGFβR it is not clear the mechanism that regulates the entry through a specific endocytic pathway, in the case of EGFR, the choice of internalization pathway depends on the dose of ligand and the ubiquitination state of the receptor [13, 30, 31] (Figure 2B, d).

At low EGF dose, EGFR enters the cell via CME, which leads to recycling and sustained signaling in the presence of high EGF concentration, the receptor

gets heavily ubiquitinated and is preferentially committed to NCE and targeted for degradation, thus leading to signal extinction [13, 30, 31].

The last stage of signaling control occurs at the endosomes. Internalized cargo proteins are trafficked to intracellular compartments, called early endosomes, from which they can be either routed to lysosomes, where they will be degraded, or recycled back to the PM, where they can continue their signaling activities. Both processes depend on the action of proteins belonging to the Rab small GTPase family; in particular, Rab11 guides the cargoes back to PM while Rab7 regulates the sorting to lysosomes and degradation (reviewed in [32]).
Besides this role in regulating protein trafficking, growing evidence suggests that early endosomes function as signaling platforms that can regulate duration, amplitude and specificity of the downstream signals (reviewed in [2, 4, 33]) (Figure 2C).

Endosomes can influence signaling by sustaining signals coming from the PM. Many RTKs remain bound to their ligands along the endocytic route and thus active in the endosomal compartments. In the case of EGFR, it has been shown that almost the entire ERK-MAPK activation cascade can be detected in the endosome [34, 35]. This could be explained by the presence of scaffold proteins, such as P18, P14, that anchor MAPK and ERK proteins to the endosomal membrane, facilitating their activation [36, 37] (Figure 2C, f).

Canonical GPCR signaling was also recently found to continue in the endosomal compartment; indeed, it has been demonstrated that the β2-adrenoreceptor, a prototypical GPCR, can signal both at PM and in the early endosomes contributing to the overall cyclic AMP cellular response within several minutes after internalization [38].

Endosomes can also support signaling by allowing the assembly of specific signaling complexes. One example of this mechanism occurs with TGFβ signaling. Onece internalized, TGFβ receptors are routed to endosomes where scaffold FYVE-containing protein SARA (Smad anchor for receptor activation) can recruit Smad2 to be phosphorylated by TGFβ receptors. This initiates the TGF signaling cascade that continues with the formation of Smad2-Smad4 complex and its translocation to the nucleus, where activation of the transcription of target genes occurs [39, 40] (Figure 2C,g).

Finally, endosomes are able to regulate the spatial restriction of signals to prevent its uniform distribution throughout the cell. Indeed, early endosomes can redirect signaling molecules to specialized areas of the PM. Spatial restriction of the signals is necessary for the execution of a number of polarized functions, including directed cell migration and epithelial-cell polarization [41].

During chemotactic migration, cells must orient themselves in the right direction and their intracellular machinery must be coordinated in order to generate propulsive forces [42, 43]. In mammalian cells, motogenic stimuli, such as hepatocyte growth factor (HGF), activate clathrin- and Rab5-mediated endocytosis of Rac proteins (small GTPases that relay signals from cell surface to actin cytoskeleton). Rac activation occurs on early endosomes, where the RacGEF Tiam1 is also recruited. This leads to recycling of Rac to specific regions of the PM where actin-based migratory protrusions are formed [44] (Figure 3).

The spatial regulation of the signals is also important to regulate the epithelial cell polarity and this is achieved through the trafficking of adhesion molecules, such as E-cadherin. E-cadherin, the prototypical epithelial cadherin, is localized at the lateral surface of epithelial cells and represents the major component of the cell-cell adherent junctions (AJs) [45].

E-cadherin can be internalized through CME and NCE and, thus, it can be recycled back to cell surface or committed to degradation (reviewed in [46]). One possibility is that the decision to recycle or degrade E-cadherin is executed by the endosome-retention machinery. Indeed, it has been shown that EGF-induced E-cadherin internalization into early endosomes containing the sorting nixing SNX1 retrieves E-cadherin from degradation while promoting its efficient recycling, and the re-establishment of epithelial adhesion [47] (Figure 3).

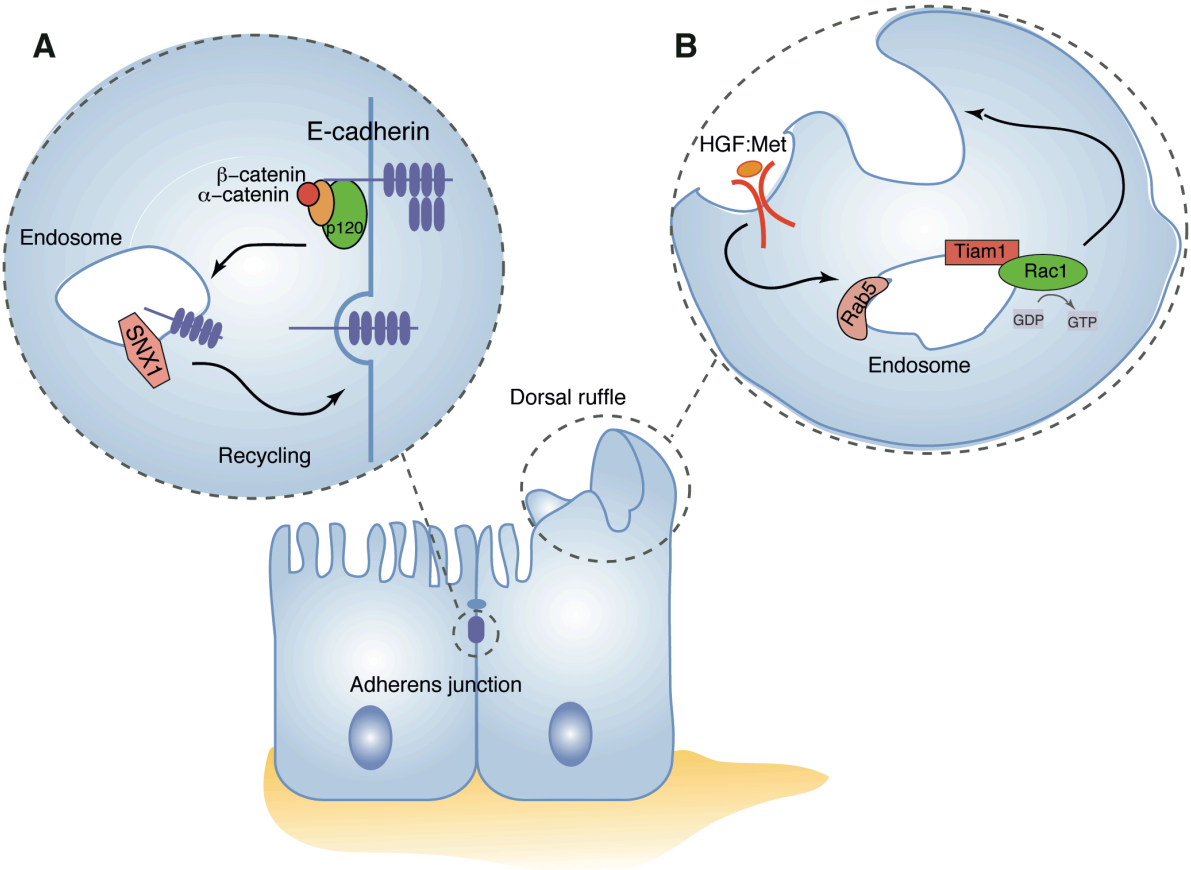


Figure 3 | Endosome sorting regulates cell signaling

A) Epithelial cell–cell contacts contain the adherent junctions protein complex of E-cadherin, p120 catenin (p120), α- catenin and β-catenin. Upon E-cadherin internalization, the recruitment of SNX1 in the endosomes leads to recycling. B) During growth factor stimulation, some cells form apical protrusions called dorsal

ruffles. Endocytic recycling of Rac1 to these dorsal ruffles requires the Rac1 GEF Tiam1 on Rab5 endosomes in response to HGF [adapted from [48]]

1.2 Endocytosis and cancer

The tight and bidirectional connection between endocytosis and cellular signaling suggests that alterations of the endocytic programs might have an important role in several human pathologies, including cancer.

In recent years, intense cancer research originated a wider and deeper knowledge of the hallmarks of cancer. These include self-sufficiency in growth signals, unlimited replicative potential, resistance to cell death signals, and activation of invasion and metastatic programs (reviewed in [49]). Evidence suggests that endocytosis is embedded in all these cellular aspects. Indeed, endocytosis regulates growth signaling either in a negative fashion, by inducing degradation of the growth factor receptors, or in a positive manner by sustaining the signals in the endosomes (reviewed in [4]). Furthermore, endocytosis controls the spatial restriction of signals needed for directed cell movement adopted by metastatic cells, by regulating internalization and recycling of adhesion molecules, including cadherins and integrins (reviewed in [41]).

A subversion of these endocytic functions might play a role in the tumorigenic process. Growing evidence suggests that altered vesicular trafficking of growth factor receptors and adhesion molecules can have a causative role in tumors. Moreover, deregulated expression of various endocytic proteins in human cancers has been reported (reviewed in [3, 4, 50]).

1.2.1 Starting from the “top”: alterations of PM cargoes – the RTK receptor family example

PM cargoes belonging to the RTK receptor family are frequently mutated in human tumors (reviewed in [48]). These mutations can induce changes in the vesicular trafficking of cargoes and affect their downstream signaling. RTK receptors, like EGFR and Met, sustain proliferative signals in the cell. The major mechanism that the cell adopts to restrict prolonged signals deriving from RTK receptors is their internalization, ubiquitination and consequent degradation. However, mutations in RTKs either increase the kinase activity of the receptors or affect the ability of the receptors to be properly ubiquitinated and downregulated (reviewed in [50]). The final outcome is, thus, sustained signaling that cannot be attenuated, which unbalances cellular homeostasis.

The EGFR is involved in the transmission of cell differentiation, proliferation and survival signals. Somatically acquired mutations in the tyrosine kinase (TK) domain of EGFR have been found in non-small-cell lung cancer (NSCLC) patients [51]. These mutations increase the kinase activity of the receptor [51, 52] and enhance selective transduction of survival signals [51]. EGFR TK mutations are found in 10-30% of NSCLC patients and are predictive of significant clinical response to the kinase inhibitor gefitinib [51, 52]. Of note, these EGFR mutants present in NSCLC are constitutively internalized and are continuously trafficked to recycling endosomes, which provides direct evidence of a causative role of aberrant trafficking of EGFR in EGFR-mediated oncogenesis [53].

Another oncogenic EGFR mutant, named EGFRvIII has been found in glioblastoma [54]. This mutant is characterized by an in-frame deletion in its extracellular domain and by its incapacity to bind ligands. This mutant form of EGFR transforms fibroblast and enhances tumorigenic potential of cancer cells

both *in vitro* and *in vivo* ([55, 56]. Moreover, EGFRvIII is inefficiently degraded: although it can still be internalized, EGFRvIII is recycled back to PM rather than being delivered to lysosomes [57].

Met, the receptor for HGF is implicated in growth, survival and dissemination of various human cancers (reviewed in [58]). Mutations in the kinase domain of Met were originally found in human papillary renal carcinoma. These mutants exhibit constitutive kinase activity and promote tumorigenesis *in vitro* and *in vivo* [59, 60]. Recently, it was shown that these mutants are tumorigenic not only because they are more active but also because they are continuously internalized/recycled and cannot be degraded, which leads to increased Met signaling in the endosomes [61]. In this study, the authors show that blocking of endocytosis *in vivo* is able to inhibit the tumorigenic potential of Met mutants, and, thus, provide strong supporting evidence for a link between changes in receptor endocytosis and tumorigenesis [61].

One important regulator of RTKs signaling is the E3 ubiquitin ligase Cbl (Casitas B-lineage lymphoma) (reviewed in [62]). Indeed, Cbl acts by ubiquitinating RTKs, which induces receptor internalization and targets it for degradation. The oncogenic forms of Cbl described in the literature are characterized by loss of the E3 ligase activity. This leads to impaired degradation of RTKs and sustained signaling. Recent studies have found Cbl mutations in acute myeloid leukemia (AML) that cause impaired internalization and degradation of the Flt3 receptor, a RTK known to regulate hematopoietic progenitor cell homeostasis. Of note, Cbl mutants lead to an increased activity of Flt3 and downstream signaling of PI3K and STAT pathways [63, 64]. All these observations point toward a new mechanism of altered RTKs signaling in human cancers.

1.2.2 Moving toward the “inside machine”: alterations of endocytic proteins

A wide range of endocytic proteins shows altered expression in human cancers (reviewed in [3, 50]). A high number of such alterations has been found in recent years, but whether there is a causal link between aberrant regulation of endocytosis, resulting from the altered expression of such proteins, and tumorigenesis remains to be determined.

In this section, I will give an overview of what is known about deregulated expression of Huntingtin-interacting protein 1 (Hip1), Rab25, Numb, and hypoxia-induced factor α (HIFα), for which altered expression has been implicated in tumorigenesis.

Hip1 is a cofactor protein involved in CME [65]. Hip1 overexpression has been observed in various tumors [66-68]. Even though the genetic liaison underlying the changes in Hip expression levels in tumors is not yet known, there is biological evidence that deregulation of Hip expression might be causal for cancer. Indeed, in prostate cancer, Hip1 protein overexpression promotes survival of cancer cells and correlates with prostate cancer progression and metastasis [69]. Of note, Hip1 overexpression is able to transform NIH/3T3 fibroblast cells and its tumorigenic potential is linked to a deregulation of receptor trafficking, such as EGFR. Moreover, EGFR expression is elevated in tumors showing upregulation of Hip1 [66-68].

Rab25 is another prominent example of an endocytic protein whose expression is altered in tumors. Rab25 belongs to the Rab11 subfamily of small GTPases that regulate the recycling of internalized cargoes back to the PM [32]. Elevated expression of Rab25 is found in ovarian and breast cancer and it is associated with decreased disease-free survival or overall survival, respectively [70]. Forced overexpression of Rab25 increases tumorigenesis *in vitro* and *in vivo*

of ovarian and breast cancer cells [70]. Rab25 has been reported to directly interact with $\alpha 5\beta 1$ integrin promoting its relocalization at the pseudopodial tips at the cell front, thus increasing the ability of tumor cells to invade the ECM and form metastasis [71, 72]. The overexpression of Rab25 results from gene amplification, which provides a more solid genetic basis to propose a causal role in tumorigenesis [70].

In some tumors, downregulation of endocytic proteins is also observed. This is the case of the protein Numb. Numb is a cell fate determinant, which controls cell fate by both asymmetrical partitioning itself and by antagonizing Notch activity [73]. In mammalian cells, Numb works also as an endocytic protein since it binds the major components of the CCPs (such as AP2) and co-traffics with internalizing receptors [74]. Numb protein expression is severely decreased in breast and lung tumors due to protein degradation and this alteration correlates with increased Notch activity [75, 76]. Moreover, Numb regulates p53 levels by forming a tricomplex with p53 and Hdm2 that results in the inhibition of the p53 ubiquitination. As a result, Numb-negative breast tumors show decreased p53 levels [77]. The genetic alterations underlying Numb deregulated expression in tumors are not yet known. However, it is clear that perturbation of Numb expression has a causative role in cancer, which provides an additional link between altered expression of endocytic proteins and tumorigenesis.

HIF1α is a transcription factor involved in the cellular response to hypoxia. Indeed, in the absence of oxygen HIFα is stabilized and activates the expression of some hypoxia-response genes, such as vascular endothelial growth factor (reviwed in [78]). It has been shown that HIF1α in renal carcinoma cells inhibits the transcription of RABAPTIN-5 gene, a critical Rab5 effector [79], which impairs Rab5-mediated early endosome fusion and delays the endocytic pathway. As a

consequence, the resident time of activated EGFR in endosomes is prolonged and signaling leading to cell proliferation and survival is sustained [79]. In agreement with this mechanism, tumor hypoxia and HIF1α overexpression generally correlate with an aggressive phenotype and poor patient prognosis [79, 80]. Importantly, primary kidney and breast tumors with strong hypoxic signatures show significantly lower expression of rabaptin-5 RNA and protein [79].

2 Epsin proteins

Epn3 belongs to the epsin family of endocytic adaptor proteins, which is involved in important physiological processes, first and foremost, the coordination of endocytosis and signaling. While the other members of the family have been well characterized as endocytic adaptors, the exact role of Epn3 in endocytic processes is still unknown. Preliminary results from our lab argue for the involvement of Epn3 in breast cancer, since the *EPN3* gene was identified in a signature associated with aggressive breast cancer. These observations prompted us to investigate the role of Epn3 in physiology and cancer. In this chapter, I will provide an overview of the current knowledge on epsin proteins with a particular focus on Epn3 and its biological functions.

2.1 Overview of the epsin protein family

In mammals, the epsin family is composed of 3 members, named epsin1 (Epn1), epsin2 (Epn2) and Epn3.

The first family member to be identified was the rat epsin1 due to its ability to bind the accessory endocytic protein Eps15 (indeed, epsin stands for Eps15 interacting protein1) [81]. In addition to Eps15, Epn1 was found also to bind the clathrin adaptor protein AP2 and to colocalize with clathrin coats [81]. In a different study, Epn1 was isolated as binding partner of intersectin [82], another endocytic factor, and POB1, a component of RalGTPase network [83]. The sequence analysis of Epn1 revealed that it is closely related to the *Xenopus* mitotic phosphoprotein MP90 [81].

Soon after, the second member of the family, Epn2, was identified. Epn2 shows a high degree of sequence homology to Epn1 and a similar cellular localization and expression pattern [84]. Both Epn1 and Epn2 are ubiquitously

expressed in all body tissues, although they are enriched in the brain. The two proteins show a general cytoplasmic localization in the cell with a pronounced accumulation in puncta on the PM [81, 84].

The third member of the family, Epn3, was identified as an extracellular matrix-induced protein in human keratinocytes. Due to the high sequence homology between Epn3 and the two other members of the epsin family, the protein was classified as belonging to the epsin family and named Epn3 [85]. Nevertheless, Epn3 is distinct from the other epsin proteins, since it has a different expression profile in normal tissues. Indeed, Epn3 is not expressed or is expressed at low levels in physiological conditions and in normal tissues, apart from migrating keratinocytes [85]. However, high expression of Epn3 was found in gastric parietal cells and in some cases of gastric cancer [86].

In vertebrates, another epsin-like protein, also known as epsinR or Epn4, displays partial sequence homology with the other epsins, mostly in the N-terminal domain (see below) but it possesses different features. Indeed, epsinR is localized at the endosome membrane and it is mainly involved in the formation of clathrin coated vesicles from the Golgi network and in retrograde transport [87, 88].

Epsin proteins are evolutionary well conserved with homologues in the lower species. *Saccharomyces cerevisiae* contains at least two epsin paralogs, Ent1 and Ent2, while one epsin orthologue, liquid facets (LqF), is present in *Drosophila*. A *Xenopus* epsin was identified as mitotic phosphoprotein in oocytes (MP90) (reviewed in [89]).

2.1.1 Structure and domain organization of epsins

All members of the epsin family share a common domain organization (Figure 4).

At the N-terminus, epsins contain a highly conserved domain of about 150 amino acids known as the Epsin N-Terminal homology (ENTH) domain. This domain has been proposed to represent a separately folded protein module due to its structural characteristics [90]. The crystal structure of mammalian ENTH domain revealed a superhelical fold comprising seven α helices [91, 92].

The ENTH domain binds PIP_2, which is enriched in endocytic sites within the PM. The binding of ENTH to PIP_2 contributes to membrane deformation occurring at the initial steps of CCP formation [93, 94]. Furthermore, this domain is also able to bind to the promyelocytic leukemia zinc finger transcriptional factor (PLZF), thus suggesting a possible role of epsin proteins in transcriptional regulation [91]. Consistent with this, epsin proteins have been shown to shuttle between nucleus and cytoplasm [91, 95]. Nevertheless, the exact role of this interaction is not clear yet.

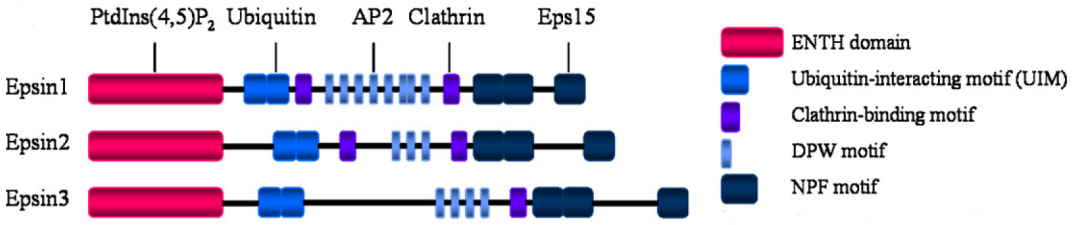


Figure 4 | Domain architecture of human epsin family members.

Different protein domain and their interactors (on the top) are indicated. For a detailed description of each domain see text. [adapted from [94]].

Several ubiquitin-interacting motifs (UIMs) are located adjacent to the ENTH domain. UIMs are responsible for the monoubiquitination of epsins as well as the recognition of ubiquitinated cargos [13, 96, 97].

The central part of epsins includes the presence of multiple DPW (Asp-Pro-Trp) motifs through which epsins bind AP2 complex [81], flanked by clathrin binding consensus sequences [98]. The number of these motifs and the affinity for these two endocytic components can vary among epsin members: indeed, Epn3 shows a low affinity to the clathrin binding sequence while the other members display a higher affinity [85, 86].

The C-terminus of the proteins comprises NPF (Asn-Pro-Phe) repeats that are required for the binding to EH-containing proteins, including Eps15, POB1 and intersectin [81-83, 99, 100]

2.1.2 Biological functions of epsins

Epsin proteins fulfill different roles in the endocytic process by acting as adaptor proteins that link surface receptors to the major components of endocytic machinery, i.e. AP2, clathrin and Eps15.

Many studies have demonstrated a clear role for Epn1 and Epn2 in the internalization process of various cargoes, while the exact function of Epn3 in endocytosis and/or signaling is still missing. Therefore, the biological functions described below are referred only to the first two members of the family.

Genetic studies conducted in *Drosophila melanogaster* demonstrated that the epsin parologue, liquid facet (LqF), is essential for the endocytosis of the Notch ligands DSL in the signal-sending cell [101, 102]. Furthermore, epsin-dependent endocytosis of DSL produces a mechanical pulling force that drives Notch conformational change necessary for Notch internalization and activation in the single-receiving cell [27]. Since endocytosis of both DSL ligands and Notch is essential to activate Notch-dependent signaling, epsin contributes to Notch pathway activation.

A detailed study of double Epn1/Epn2 knock-out (KO) mice revealed that these proteins are essential for vitality, since embryos died during gestation, and that Epn1 and Epn2 are redundant, since single KO pups were viable [103]. Interestingly, development defects observed in double KO embryos recapitulate those produced by inhibition of Notch signaling [103].

In mammals, it has been widely demonstrated that Epn1, together with Eps15, plays a role in both CME and in NCE of EGFR. In the latter case, this is mediated by the ability of Epn1 to bind ubiquitinated EGFR via UIM [13]. Besides EGFR, Epn1 and Epn2 also bind the ubiquitinated form of vascular endothelial growth factor receptor (VEGRF), and, by doing so, allow its internalization and degradation [104]. As a consequence, loss of Epn1 and Epn2 in endothelial tissue results in excessive VEGFR signaling that leads to disorganized vasculature, leaky tumor angiogenesis and tumor growth retardation [104].

Complementary to their endocytic role, epsins are involved in many other cellular processes. For instance, epsin proteins are able to regulate small GTPases and, consequently, actin-based motility and polarity.

In yeast, the ENTH domain interacts with Cdc42 GTPase-activating proteins (Cdc42 GAPs) regulating polarized cell growth. This interaction is necessary to sustain cell viability [105].

In mammals, epsin proteins exert important functions in the regulation of Rho GTPase signaling. Specifically, epsin interacts through its ENTH domain with Ral-binding protein 1 (RalBP1) and this interaction leads to cell migration in fibrosarcoma cell lines [106].

Finally, a number of studies conducted in mammalian cells, revealed that epsin proteins are subjected to Cdk1-dependent phosphorylation that renders them incompetent for the binding to AP2 and Eps15 and, thus, unable to play their

endocytic role [107, 108]. Furthermore, epsin proteins regulate mitotic membrane morphology and spindle orientation during mitosis through its ability to induce membrane curvature [109]. Therefore, the endocytic function of epsins and its role in mitosis are mutually exclusive and tightly regulated.

2.2 The case of Epn3: what makes the difference

The protein Epn3 shows the same domain architecture and a high degree of sequence identity when compared to the other two members of the family. Indeed, Epn3 contains the highly conserved ENTH domain with more than 90% homology. Epn3 also contains multiple DPW and NPF motifs to bind AP2 and Eps15. Differently from Epn1 and Epn2, Epn3 exhibits a low affinity clathrin binding [85] (Spradling et al JBC 2001).

This similarity of sequence and domain organization suggests that Epn3 could behave analogously to the other epsins. Indeed, like the other members, Epn3 shows a cytoplasmic localization with a pronounced accumulation in perinuclear region. Furthermore, like Epn1 and Epn2, Epn3 co-localizes with clathrin coated-structures and is able to shuttle between cytoplasm and nucleus [85, 86]. Finally, together with the other epsins, Epn3 has been shown to be involved in membrane fission and clathrin-coated vesicle formation [110].

The characteristic that makes Epn3 divergent from the other epsins is its different tissue expression pattern. Unlike Epn1 and Epn2, which are ubiquitously expressed, Epn3 is prevalently expressed in gastric parietal cells [86] and in wounded or pathological epithelial tissues exhibiting altered cell-extracellular matrix interaction [85]. In particular, Epn3 expression was found to be upregulated in migrating keratinocytes in cutaneous wounds, but not in differentiating cells [85].

In gastric cells, Epn3 high expression levels are found in the apical caniculi where it co-localizes with clathrin. These cells are specialized in the control of the stomach lumen acidification via exo-endocytosis of vesicles containing H/K ATPase [86].

All these data suggest a putative role of Epn3 in the endocytic process. However, the exact function of this atypical epsin in endocytosis, signaling and/or cell migration is still unknown.

3 Breast cancer

The project presented here is based on the preliminary results derived from a gene-expression analysis of a small cohort human breast tumor samples showing the association between expression of the endocytic protein Epn3 and poor prognosis in breast cancer patients. These preliminary data strongly argued for a role of Epn3 as a breast cancer marker with potential to be used in the clinical management of the disease. In light of this data, we decided to investigate the involvement of Epn3 in breast cancer initiation.

Breast cancer is the most commonly diagnosed cancer and the main cause of cancer-related deaths in women worldwide. Despite the fact that state-of-the-art diagnostic and therapeutic tools allow to treat breast cancer patients at the very early stage of disease, disease recurrence and metastasis are responsible for the most part of breast cancer-related deaths. This situation is mainly due to our incapacity to predict breast cancer evolution. Breast cancer is a complex and heterogeneous disease, made up of distinct clinical, morphological and molecular entities [111] and our understanding of the mechanisms underlying its progression is still limited. The molecular profile of breast cancer should help to understand how tumors can differ among patients and to predict the clinical outcome and the response to specific therapy. A major need in the clinical management of breast cancer patients is, therefore, the identification and validation of molecular signatures or individual genes as prognostic markers able to predict the risk of disease recurrence and that can be used to better stratify the patients and develop more rationale therapies. Epn3 might be one of such markers.

3.1 Histological and molecular classification of breast cancer

Pathologists, on the basis of their microscopic observations, have classified breast cancer in non-invasive or invasive types [112, 113]. Non-invasive breast cancer, also referred as *in situ*, is characterized by the presence of cancer cells restricted to the basement membrane; the most common type of non-invasive breast cancer is the Ductal Carcinoma *in situ* (DCIS) [112]. On the other hand, the invasive breast cancer occurs when cancer cells infiltrate the surrounding tissues. Invasive tumors show a higher risk to form metastasis. The invasive ductal carcinoma (IDC) represents the most frequent type, comprising the 70-80% of all the breast cancer cases [113].

Although useful, the histological classification of breast cancer does not take into account the molecular features of the disease. To investigate the molecular differences between breast tumors, researchers have relied on gene expression microarrays. Such approach consents the simultaneous expression analysis of thousands of genes in breast tumors, and thus allows creating molecular profiles of the analyzed breast cancers.

Perou's group was the first to provide a molecular classification of breast cancer, using cDNA microarray of 38 breast cancer cases [111]. An unsupervised analysis was used for grouping several breast tumors according to their biological characteristics regardless their clinical and prognostic variables. The classification divides breast cancer into 4 main groups, named Luminal-A, Luminal-B, HER2 and Basal-like. The identification of these molecular classes was validated by several subsequent studies [114-116]. After this initial classification, new molecular subtypes have been identified, such as claudin-low [117] and normal-like [118].

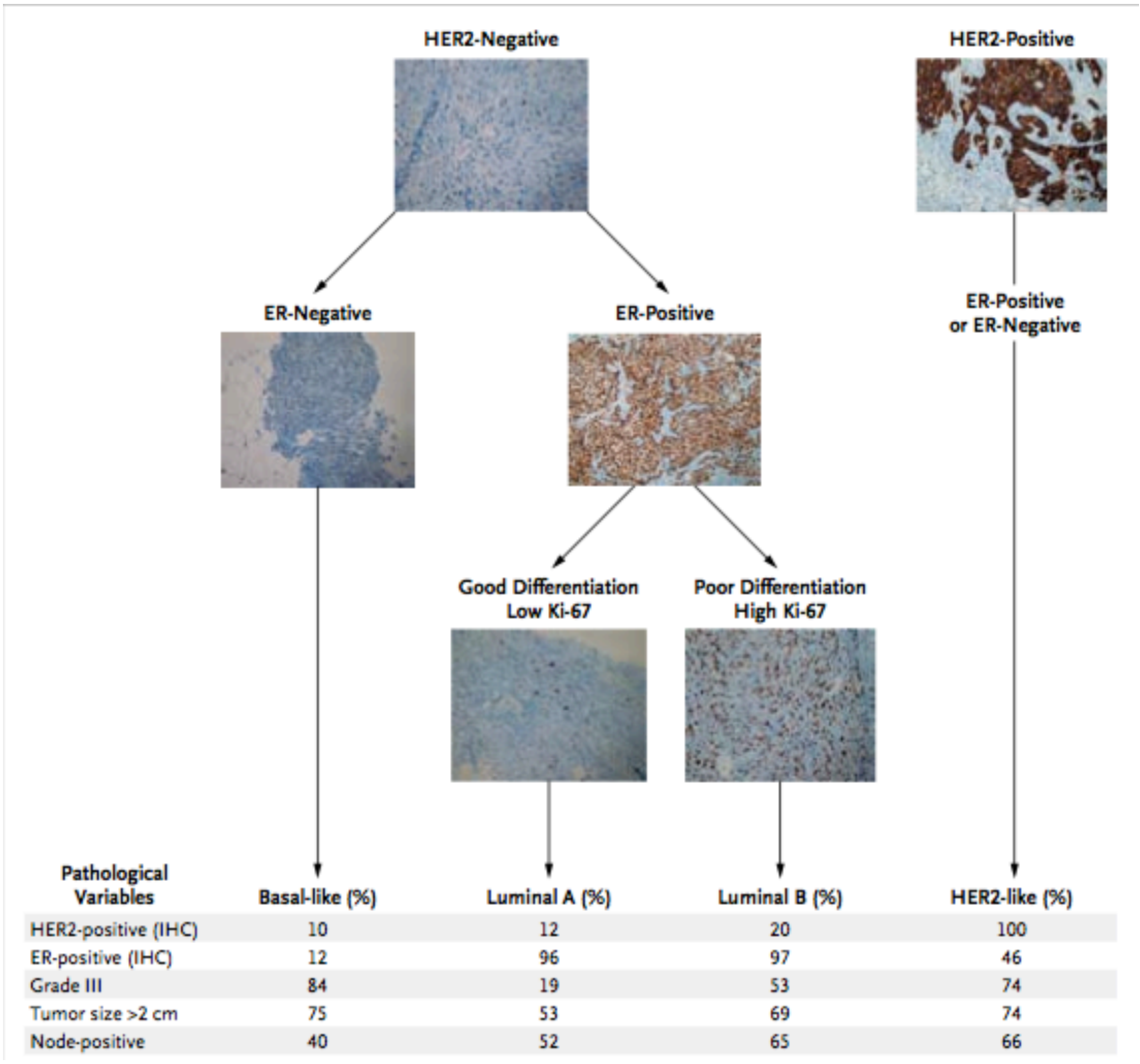
The main four molecular breast cancer subtypes are briefly described below (Figure 5).

Luminal A breast carcinoma represents 50-60% of all breast cancer cases. This type of tumor is characterized by high expression levels of estrogen receptor (ER) and progesterone receptor (PgR), absence of HER2 expression, low proliferation rate and low grade. Patients with this type of breast carcinoma have a good prognosis [114].

Luminal B type makes up 10-20% of all breast cancers. Luminal B carcinoma has a more aggressive behavior, higher histological grade, higher proliferative index and worse prognosis than Luminal A [111]. Luminal A and B tumors express ER, but due to their differences in prognosis, there is a strong need to find other biomarkers able to distinguish between these two subtypes.

The HER2 subtype represents 15-20% of all breast cancer cases and is characterized by high expression of HER2 and other genes associated with HER2 amplicon located on chromosome 17. These tumors are highly proliferative, they have high histological grade and have p53 mutations. From the clinical point of view, HER2 subtype is characterized by a poor prognosis [116].

Basal-like breast carcinoma comprises 10-20% of breast cancer cases and is more frequent in a premenopausal age [119]. One of the most relevant features of this tumor subtype is the absence of expression of all the three key receptors in breast cancer: ER, PgR and HER2. Therefore, these tumors are also called “triple negative”. The basal-like subtype is characterized by high rate of p53 mutations and occurs commonly in BRCA1 germline mutations carriers. All these characteristics result in the poor prognosis of basal-like carcinoma patients.



Pathological Variables	Basal-like (%)	Luminal A (%)	Luminal B (%)	HER2-like (%)
HER2-positive (IHC)	10	12	20	100
ER-positive (IHC)	12	96	97	46
Grade III	84	19	53	74
Tumor size >2 cm	75	53	69	74
Node-positive	40	52	65	66

Figure 5 | Molecular class of breast cancer with clinico-pathological features.

HER2 status, estrogen receptor (ER) and proliferation index (Ki67) are represented for each molecular subtype. They determined by immunohistochemical (IHC) analysis. [adapted from [116]]

3.2 Clinical management of breast cancer patients

Despite the fact that in the last decade the technological development allowed advances in diagnosing and treating breast cancer patients, breast cancer is still the leading cause of cancer-related women deaths [120]. Indeed, our ability to treat advanced and metastatic breast cancers is still limited and this results in a high death rate for breast cancer at these stages.

Current routine clinical management of breast cancer patients involves the evaluation of clinico-pathological and biological markers that are used to determine prognosis and response to therapies. The classical markers include clinical and histological factors, such as tumor size, axillary lymph node status [121], tumor grade [122], but also biological factors, such as proliferative index [123], HER2 [124] and hormone receptors status [125]. However, disease recurrence in breast cancer patients is high and clinicians do not have the necessary tools to predict it. Indeed, optical predictors of metastases have yet to be identified.

Targeted therapies have been developed during the last years, including endocrine therapy (i.e.tamoxifen) to treat hormone receptors positive patients, and anti-HER2 therapy (i.e.trastuzumab) that is specific for HER2 positive patients. For instance, trastuzumab is the first-line treatment for HER2-positive metastatic breast cancer patients, in combination with adjuvant chemotherapy agents [126]. Nevertheless, only 30% of the patients harboring HER2 overexpression respond to this treatment regimen, while approximately 70% of patients may have resistance to trastuzumab (reviewed in [127]).

Thus, the clinical management of breast cancer patients remains complicated, mainly because we still do not understand completely the mechanisms of tumor progression. We need a deeper knowledge of tumor biology

and microenvironment to be able to provide patients with better and individualized care. An individualized approach to breast management has the potential to offer clear clinical benefits as well as the ability to bring more tailored and cost-effective strategies to breast cancer patients. Currently, the more promising approach to provide a real improvement in the clinical management of breast cancer patients relies on the identification of specific biomarkers to predict disease progression.

3.3 Predictive and prognostic gene signatures in breast cancer

The application of high-throughput technologies, such as microarrays, to gene expression analysis of tumors has allowed the generation of prognostic and predictive gene signatures that are useful to predict the outcome and the response to therapy of breast cancer patients (reviewed in [128]).

The first prognostic breast cancer gene signature was identified in the work of *Van't Veer et al* in 2002 [129]. In this study, researchers used cDNA microarray analysis of primary breast tumors and identified a list of 70 genes able to discrimate between good and bad prognosis. This "poor prognosis" signature was found to be strongly predictive of disease outcome, as defined by the presence of distant metastasis within 5 years [129, 130]. A subsequent study by Minn and colleagues screened for genes associated with the ability of breast cancer cells to form metastases [131]. Through the analysis of gene expression data, a signature associated with the formation of breast cancer lung metastases and poor outcome was identified [131]. During the last years, several prognostic and predictive gene signatures have been generated but only few of them are really used to stratify breast cancer patients.

One of those signatures is represented by the Mammaprint test. Mammaprint was the first commercialized microarray-based multigene assay for

breast cancer (Agendia). This test comprises 70 genes and was generated from node-negative breast cancer patients who did not receive any adjuvant therapy [132]. The different expression of the 70 genes, which is evaluated by the assay, allows the stratification of breast cancer patients into good or poor prognosis groups. The Breast Cancer International Group has launched a clinical trial, named MINDACT (Microarray In Node negative Disease may Avoid ChemoTherapy) to determine whether this molecular test can identify low-risk patients who can be spared from cytotoxic therapy [133].

An alternative approach to Mammaprint assay is the Oncotype DX assay. This assay measures the expression of 21 genes in formalin-fixed paraffin-embedded tumor samples through RT-PCR. Oncotype DX assay has shown to predict the risk of recurrence in node-negative, ER positive breast cancer patients who had received treatment with tamoxifen [134].

Although the results of the introduction of these technologies in the clinical management of breast cancer patients are quite promising, it is still not clear if these gene expression-profiling techniques will be capable of replacing the traditional methods based on histopathological assessments. An alternative approach could be the development of signatures comprising a limited number of genes with validated antibodies suitable for IHC detection of their respective proteins. This approach, together with the development of more sophisticated and automated imaging methods, would hypothetically provide uniform, objective and more comparable information for routine clinical practice.

MATERIALS AND METHODS

1 Tissue microarray (TMA) studies

1.1 Patient selection and study design.

Case-control study group: we established a case-control study on a large series of breast cancer patients (n = 349) who had undergone surgery at the European Institute of Oncology (IEO) in Milan for the removal of primary breast cancer between 1994 and 1997. Disease recurrence (any relapse and distant relapse) was within 18 years (median 9.2 years and 10.6 years respectively). For some patients not all information was available. (See Table 1 for further details on patient characteristics)

Validation study group: this group included 916 breast cancer patients who underwent surgery at IEO for the removal of primary breast cancer between 1997 and 2000. For some patients not all information was available. (See Table 2 for further details on patient characteristics).

Written informed consent for research use of biological samples was obtained from all patients, and the research project was approved by IEO's Institutional Ethical Committee.

1.2 Analysis of Epn3 expression in breast cancers by immunohistochemistry on TMA

Tissue sample preparation, IHC and TMA-IHC analyses were performed in collaboration with the Molecular Pathology Unit at the IEO, Milan.

Normal (when available) and tumor areas (diameter 0.6 mm) from each biopsy sample, previously identified on haematoxylin-eosin stained sections, were

removed from the paraffin donor blocks and deposited on the recipient block using a custom-built precision instrument (Tissue Arrayer - Beecher Instruments, Sun Prairie, WI 53590, USA). Two-µm sections of the resulting recipient block were cut, mounted on glass slides, and used for IHC.

TMAs were analyzed for Epn3 protein expression by IHC. TMA sections were routinely processed, placed for 30 minutes in 0.25 mM EDTA at 95°C for antigen retrieval and incubated for 3 hours with the VI31 anti-Epn3 monoclonal antibody (1:30 0000, produced in-house); bound antibody was revealed using the EnVision Plus/HRP detection system (DAKO) and diaminobenzidine as a chromogenic substrate. TMA sections were finally counterstained with hematoxylin and mounted. Positive and negative controls were included in each experiment and only clear staining of the tumor cell nuclei was considered positive for Epn3 expression. A semiquantitative approach was used to generate a score for each tissue core, ranging from 0 to 3 according to the signal intensity. Scores 1 (weak), 2 (moderate) and 3 (strong) were assigned when at least 30% of tumor cells in the sample were positive.

Tumors showing IHC scores > 1.0 were assigned to the Epn3-HIGH group, whereas those with IHC scores ≤ 1.0 were considered as the Epn3-LOW group.
ER and PgR proteins, measured by IHC on whole tissue sections, were retrieved from histopathology reports. The rate of proliferation was measured by determining the percentage of nuclei in which labeled antigen Ki-67, a marker of cell division, was expressed (MIB1 antibody DAKO, Cytomation). HER2 expression was measured by IHC on TMAs, processed as previously described, using an anti-HER2 polyclonal antibody (1:160, DAKO, Cytomation). HER2 overexpression was evaluated according to the scoring system recommended by the DAKO HercepTest: score 0, no staining or membrane staining in <10% of the

tumor cells; score 1, barely perceptible membrane staining in >10% of the tumor cells; score 2, weak-to-moderate staining of the entire membrane in >10% of the tumor cells; score 3, strong staining of the entire membrane in >10% of the tumor cells. Scores of 2 and 3 were considered to represent overexpression.

1.3 Statistical analysis

Association between the clinical/pathological features of the tumors and Epn3 expression was evaluated by the Fisher's exact test. In the breast cancer validation cohort, logistic regression was used to assess the association between Epn3 expression and the presence or absence of events (relapse or death) from the date of surgery to the date of the event or the date of last follow-up. Follow-up was updated to 2012. Kaplan-Meier plots were carried out by means of the proportional hazards Cox-model. All *P*-values were two- sided. A *P*-value equal to or less than 0.05 was considered significant. All statistical analyses were carried out using SAS statistical software (SAS Institute, Inc., Cary, NC).

1.4 FISH analysis

FISH on breast TMAs and breast cell lines was performed in collaboration with the Molecular Pathology Unit at the IEO in Milan.

DNA probes were labeled with a fluorescent dye (Cy3-dUTP or Green-dUTP) that is incorporated into the DNA by nick translation. The following reaction mixture was used: 3µl Buffer 10X, 0.6µl dAGC, 0.3µl dUTP/Cy-3, 3µl ß-mercaptoethanol, 0.3µl DNA polymerase, 6µl DNase (1:700 in H2O), and 5 µg DNA in 30 µl final volume. The reaction mix was incubated at 16°C for 2 hours. The probe was precipitated using the following reaction mixture: 3µl salmon sperm DNA, 10µl Cot-1 DNA, 1/10 volume NaAcetate, 3 volumes of cold 100% ethanol. The reaction was then placed at -80°C for 15min (or at - 20°C for 30 min) and then

centrifuged at 13000 rpm for 20 min at 4°C. The supernatant was removed and the pellet dried and resuspended in 15 µl of hybridization mix (for 15ml: 7.5 ml ultrapure formamide, 6 ml dextran solphate 25%, 1.5 ml 20 x SSC). The mixture was then incubated with shaking for 10 min at room temperature.

To prepare the cells, the cell suspension was centrifuged at 1500 rpm for 10 min at room temperature. The supernatant was removed and the cell pellet was resuspended in 10 ml hypotonic solution (0.075M KCL) and incubated at 37°C for 18 min. Cells were then fixed in 3:1 methanol:acetic acid.

For the probe hybridyzation, probes, dissolved in hybridization mix, were placed on pretreated slides, and covered with a coverslip and sealed with rubber cement. Slides were placed in a Hybrite machine (Vysis) and incubated at 73°C for 3 min and then at 37°C overnight. Slides were then washed 3 times in 0.1x SSC for 5 min each wash, then incubated with DAPI (DAPI in 2 X SSC) for 5 min.

2 Reagents

Human recombinant TGFβ1 was from Tebu-bio. TGFβ kinase inhibitor LY2109761 was purchased from Selleckbio.

2.1 Antibodies

For western blot, anti-Epn3 VI31 antibody (epitope: aa residues 464-483, *Homo sapiens*) and anti-Epn1/2 (epitope: aa 249-401, *Homo sapiens*) were produced in-house through the Antibody Facility. Anti-vinculin, anti-tubulin and anti-Flag were from Sigma. Anti-E-cadherin, anti-N-cadherin and anti-Vimentin were from Transduction BD. Anti-Snail, anti-P-Smad2, anti-P-Smad3 were from Cell Signaling. For immunofluorescence, anti-E-cadherin was from Transduction BD. For E-cadherin internalization assay, anti-Ecadherin (HECD-1) was from Abcam.

2.2 Constructs and plasmids

The pSicoR lentiviral vectors (Addgene), developed by Ventura *et al*. [135], were used for constitutive shRNA expression. Vectors were engineered to express shRNA specifically targeting Epn3 expression (shEpn3#1, shEpn3#2) or luciferase (shLuc) and mismatch (shMis) as a controls. The pSicoR vector expresses the puromycin resistance gene as selection marker.

The pBABE retroviral vector was used to generate a construct to overexpress Epn3 in mammalian cells. The human Epn3 coding sequence was previously cloned downstream the Flag epitope in pBABE vector. The pBABE vector expresses the puromycin resistance gene as selection marker.

The pLVX-puro lentiviral vector was used to generate a construct to overexpress Epn3 in mammalian cells. The human Epn3 coding sequence, already present in the lab, was inserted by digestion with the restriction enzymes (New England Biolabs) and ligation into the pLVX-puro vector from pCDNA vector already present in the lab. pLVX contains a puromycin resistance gene

2.3 RNAi oligos

For stable Epn3 KD the following shRNA oligo sequences were used:

shEpn3#1:Fwd-TGCGAGAACCTCTACACCATTTCAAGAGAATGGTGTAGAGG
TTCTCGCTTTTTTC

shEpn3#1:Rev-TCGAGAAAAAAGCGAGAACCTCTACACCATTCTCTTGAAAT
GGTGTAGAGGTTCTCGCA

shEpn3#2: FWD-TGAGCTAGAAACTGAACGCCTTCAAGAGAGGCGTTCAGTTT
CTAGCTCTTTTTTC

shEpn3#2: Rev-TCGAGAAAAAAGAGCTAGAAACTGAACGCCTCTCTTGA
AGGCGTTCAGTTTCTAGCTCA

shOligoMis: Fwd- TGAGCGAACCGATACACTATTTCAAGAGAATAGTGTA

TCGGTTCGCTCTTTTTTC

shOligoMis: Rev- TCGAGAAAAAAGAGCGAACCGATACACTATTCTCTTG

AAATAGTGTATCGGTTCGCTCA

For transient KD of TGFβ receptors the following oligos from Qiagen were used:

TGFBR1 smart pool:

GAGAAGAACGUUCGUGGUU

UGCGAGAACUAUUGUGUUA

GACCACAGACAAAGUUAUA

CGAGAUAGGCCGUUUGUAU

TGFBR2 smart pool:

CAACAACGGUGCAGUCAAG

GACGAGAACAUAACACUAG

GAAAUGACAUCUCGCUGUA

CCAAUAUCCUCGUGAAGAA

2.4 Q-PCR

For Q-PCR experiments the Taqman chemistry (Applied Biosystems) was used. In the table below, the Taqman assays employed are listed:

Gene symbol	Assay ID	Ref Seq
EPN3	hs00978957_m1	NM_017957.1
CDH1	hs00170423_m1	NM_004360.3
CDH2	hs00169953_m1	NM_001792.3
VIM	hs00185584_m1	NM_003380.3
TGFB1	Hs00998133_m1	NM_000660.4
TGFB2	hs00234244_m1	NM_003238.3
TGFBR1	hs00610320_m1	NM_004612.2
TGFBR2	hs00234253_m1	NM_003242.5
18S	Hs99999901_s1	X03205.1

3 Cloning techniques

3.1 Agarose gel electrophoresis

DNA samples were loaded on 0.8%-2% agarose gels along with DNA markers (1 kb DNA Ladder, NEB). Gels were made in TAE buffer containing Gel Red (Biotium), according to manufacturer's instructions, and run at 80 V until desired separation was achieved. DNA bands were visualized under a UV lamp.

3.2 Minipreps

Individual colonies were used to inoculate 2 ml LB (containing the appropriate antibiotic) and grown overnight at 37°C. Bacteria were transferred to Eppendorf tubes and centrifuged for 5 minutes at 16,000*xg* using a 5415 R centrifuge. Minipreps were performed with the Wizard Plus SV Minipreps Kit (Promega) following manufacturer's instructions. The plasmids were eluted in 50 µl nuclease free H2O.

3.3 Diagnostic DNA restriction

Between 0.5 and 5 µg DNA were digested for 2 hours at 37°C with 10-20 units of restriction enzyme (New England Biolabs). For digestion, the volume was made up to 20- 50 µl with the appropriate buffer and distilled H2O.

3.4 Large scale plasmid preparation

Cells containing transfected DNA were expanded into 250 ml cultures overnight. Plasmid DNA was isolated from these cells using the Qiagen Maxi-prep kit according to manufacturer's instructions.

3.5 Transformation of competent cells

An aliquot of competent cells TOP10 (Invitrogen) were thawed on ice for approximately 10 minutes prior to the addition of plasmid DNA. Cells were incubated with DNA on ice for 30 minutes and then subjected to a heat shock for 45 seconds at 42°C. Cells were returned to ice for an additional 5 minutes. Then, 300 µl of SOC was added and the cells were left at 37°C for a further 60 minutes before plating them onto agar plates with the appropriate antibiotic. Two plates for each reaction were used, one plated with 100 µl of the transformed bacterial cells and the other one with the rest. Plates were incubated overnight at 37°C.

4 Cell culture procedures

4.1 Cell lines

All human breast cell lines were from the American Type Culture Collection (ATCC). The MDA-MB-361, BT474, MCF7 and SKBR3 cell lines were cultured in DMEM medium (from Lonza), supplemented with 10% fetal bovine serum (FBS, HyClone) and 4 mM L- glutamine (Euroclone). The HCC1569, HCC1428, ZR7530, ZR751 and AU565 cell lines were cultured in RPMI-1640 medium (from Lonza), supplemented with 10% FBS and 4 mM L- glutamine. The BT483 cell line was cultured in RPMI-1640 medium supplemented with 20% FBS and 0.01 mg/ml insulin (Sigma). MDA-MB-175 cell line was cultured in Leibovitz's L-15 medium

(Invitrogen, Life Science Technologies) supplemented with 10% FBS. MDA-MB-436 cell line was cultured in Leibovitz's L-15 medium supplemented with 10% FBS, 10 µg/ml insulin, 16 µg/ml glutathione (from Sigma). The MCF10A cell line was cultured in a 1:1 mixture of DMEM and Ham's F12 medium (Gibco, Life Technologies), supplemented with 5% Horse Serum (Invitrogen), 20 ng/ml human epidermal growth factor (EGF; Invitrogen), and 50 ng/ml cholera toxin, 10 µg/ml insulin and 500 ng/ml hydrocortisone (from Sigma). The UACC-812 cell line was cultured in Leibovitz's L-15 medium supplemented with 20% FBS, 2 mM L-glutamine and 20 ng/ml human EGF. All cells were cultured at 37°C in a humidified atmosphere containing 5% CO2.

4.2 Retroviral and lentiviral infections

Retroviruses were packaged by transfection of either pBABE, pBABE-Flag-Epn3, pBABE-Flag-Epn1 or pBABE-mTwist vectors in phoenix cells. Twenty-four hours (24 h) after transfection, the supernatant was discarded and replaced with 5 mL of fresh medium concentrated to 5 ml for each 10 cm plate and, 24 h later, the supernatant was collected, filtered through a 45 µm filter and added to MCF10A cells (40%-50% confluent), which were then incubated at 37°C for 8 h. MCF10A cells were passaged 24 h after infection and puromycin (2 µg/ml) was added to the culture.

Lentiviruses were packaged by transfection of either pSicoR, pSicorR-shLUC, pSicoR-shEpn3#1, pSicoR-shEpn3#2, pLVX, or pLVX-Epn3 together with the helper vectors in 293T cells. After transfection (24 h), the supernatant was concentrated to 5 ml for each plate; 24 h later, supernatant was collected, filtered through 45 µm filter and added to the cells. Viral supernatant was overlayed on the MCF7, BT474 or HCC1569 cells at 40% confluence and incubated at 37°C for 8 h.

Cells were passaged 24 h after infection and puromycin (1.5 µg/ml) was added to the culture.

4.3 Transient transfections

RNAi transfections were performed using Lipofectamine RNAimax reagent from Invitrogen, according to manufacturer's instructions. Cells were subjected to double transfection, treated with 8 nM RNAi oligos and analyzed 4 days after transfection. The negative control siRNA used was All Stars from Qiagen.

4.4 mRNA extraction and cDNA synthesis

mRNA from control and test cell lines was extracted using the RNeasy kit from Qiagen, according to the manufacturer's protocol.

Single stranded cDNA synthesis was performed using the SuperScriptII reverse trancriptase (Invitrogen) following manufacturer's instructions. Briefly, 1 µg of total RNA was mixed with 250 ng of random primers in RNase-free water and then incubated at 70°C for 5 min. Following the incubation, 10X reaction buffer, dNTPs mix (0.5 mM final concentration), and 1 µl of reverse transcriptase were added to the mix (20 µl final volume) and the reaction was incubated at 42°C for 1 hour. Finally, the reaction was inactivated by heating at 70°C for 15 min.

4.5 Immunofluorescence studies

Cells were plated on glass coverslips. Cells were fixed in 4% paraformaldehyde (in Pipes Buffer) for 10 min, washed with PBS and permeabilized in PBS 0.1% Triton X-100 for 10 min at room temperature. To prevent non-specific binding of the antibodies, cells were incubated with PBS in the presence of 5% BSA for 30 min. The coverslips were incubated with primary antibodies diluted in PBS 0.2 % BSA. After 1 hour of incubation at room temperature, coverslips were washed 3 times with PBS. Cells were then incubated for 30 min at room temperature with the

appropriate secondary antibody Cy3 (Amersham), Alexa 488-conjugated (Molecular Probes).

After three washes in PBS, coverslips were mounted in a 90% glycerol solution containing diazabicyclo-(2.2.2)octane antifade (Sigma) and examined under a wild-field immunofluorescence microscope (Leica). Images were further processed with the Adobe Photoshop software (Adobe) or with Image J to merge the images of the single channels.

5 Protein procedures

5.1 Cell lysis

After washing with PBS 1X, cells were lysed in RIPA directly in the cell culture plates using a cell-scraper and clarified by centrifugation at 16,000 *xg* for 10 min at 4^{o}C using a 5415 R centrifuge. Protein concentration was measured by the Bradford assay (Biorad) following manufacturer's instructions.

5.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Gels for resolution of proteins were made from a 30%, 37,5:1 mix of acrylamide: bisacrylamide (Sigma). As polymerization catalysts, 10% ammonium persulphate (APS) and TEMED were used.

5.3 Western blot (WB)

Desired amounts of proteins were loaded onto 1-1.5 mm thick SDS-PAGE gels for electrophoresis (Biorad). Proteins were transferred in western transfer tanks (Biorad) to nitrocellulose (Schleicher and Schnell) in 1X Western transfer buffer (supplemented with 20% methanol) at 30 V overnight or 100 V for 1 hour for small gels and at 30 V overnight or 0.8 A for 2 hours for large gels. Ponceau staining

was used to determine the efficiency protein transfer onto the filters. Filters were blocked for 1 hour (or overnight) in 5% milk in TBS supplemented with 0.1% Tween (TBS-T). After blocking, filters were incubated with the primary antibody, diluted in TBS-T 5% milk, for 1 hour at room temperature, followed by three washes of five minutes each in TBS-T. Filters were then incubated with the appropriate horseradish peroxidase-conjugated secondary antibody diluted in TBS-T for 30 min. After the incubation with the secondary antibody, the filter was washed 3 times in TBS-T (5 minutes each) and the bound secondary antibody was revealed using the ECL method (Amersham).

6 Biological assays

6.1 Soft agar assay

For soft agar assays, 1% base bottom agar and 0.4% top agarose (GIBCO) solutions were prepared. A stock solution of 5% agarose was prepared by dissolving 5g of agarose in 100 ml of PBS and aliquots of 30 ml were prepared and sterilized by autoclave and store at 4^{o}C. At the moment of the experiment, the stock solution of agarose was dissolved in microwave and then mixed with culture medium at the final concentration of 1% and 4 ml were added in each 6-multiwell of a plate (bottom agar layer). The bottom agar layer was incubated at 4^{o}C. Meanwhile cells were trypsinized and counted and a cellular suspension of 25 000 cells/ml was prepared. Top agar layer was prepared by mixing 46ml of culture medium and 4 ml of stock agar solution (5%). Then 1 ml of the cellular suspension was added to 3ml of top agar solution and the total 4ml of cells and agar were plated in 6-multiwell plates and incubated for 3 weeks. After this period, colonies were counted at the phase-contrast microscope (magnification 4x).

6.2 *In vivo* xenografts assay

Six to 8 week-old NOD/SCID IL2R gamma-chain null female mice were injected in the inguinal mammary fat pad with 300 000 BT474 and 1 000 000 HCC1569 cells

For the preparation of the cells, cells were trypsinized and resuspended in sterilized PBS. Cells were centrifuged at 1200 rpm for 5 minutes and cellular pellet was resuspended in 10ml of PBS and counted. The desired total number of cells was centrifuged again in at 1200 rpm for 5 minutes and cells were in 40 µL of a 1:1 Matrigel-PBS solution and immediately injected into the inguinal mammary fat pad of NOD/SCID IL2R gamma-chain null female mice. Mice were monitored by hand-palpation for tumor development. Tumor growth was measured by using a vernier caliper and applying the standard formula: tumor volume = (a x b2)/2. Mice were sacrificed when tumors reached a dimension of 1.5 – 2 mm3. Tumors were explanted and processed for cell extraction.

6.3 Invasion assay

Invasion assays Growth Factor Reduced Matrigel invasion chambers placed on 24-multiwell plates from BD were used. Cells were trypsinized, resuspended in culture medium and counted. 40000 cells were resuspended in 500ul of culture medium and added to the top chamber of the transwells and incubated at 37°C for 24 h. After 24h of incubation, transwells were washed with PBS and fixed with PFA 4% for 20 minutes. Transwell filters were stained with crystal violet for 20 minutes and then extensive washes with H_2O were performed. Stained membranes were used to take photographs of cells and then were dissolved in 100ul of 10% acetic acid and absorbance at 595 nm was measured at spectrophotometer.

6.4 Mammosphere-forming assay

Mammosphere forming assays were performed as described in [136]. 6-multiwell plates were coated with 1,2% Poly-HEMA (2-hydroxyethylmethacrylate) (Sigma) solution and sored at 4°C. Cells were trypsinized, resuspended in PBS and centrifuged at 1200 rpm for 5 minutes. Pellet was resuspended in PBS and PKH-26 dye (Sigma) diluted 1:20000 and incubated for 5minutes at room temperature. Cells were centrifuged and resuspended in stem cell medium and counted. A cellular suspension of 20000 cells/ml was prepared and 2m were plated in 6-multiwell plates in quadruplicates (4000 cells/well), and incubated at 37°C for one week. After one the mammospheres were observed at fluorescent microscope and only mammospheres containing one PKH26-labelled cell inside the spheres were counted. The no-labelled mammospheres and spheres containing more than one PKH26-positive cells where considered respectively spheres originating from differentiated/precursor cells and cellular aggregates.

6.5 E-cadherin internalization assay

E-cadherin internalization assay was performed using an anti-Ecadherin antibody (HECD-1) from Abcam, recognizing the extracellular domain of E-cadherin. Briefly, MCF10A cells were plated on coverslip at about 30% of confluency. The day after cells were starved overnight by incubating them with culture medium containing only serum (horse serum 5%). After starvation, cells on coverslips were washed with PBS and incubated with the anti-E-cadherin antibody (2μg/ml) for 1h at 4°C.

After 1h incubation, time zero cells were fixed with PFA 4% for 10 minutes. For the remaining samples, PBS washes were performed and cells were incubating with culture medium containing serum and TGFβ (5ng/ml) for 90 and 180 minutes. After incubation cells were fixed with PFA 4% for 10 minutes. Immunofluorescence was performed. Coverslips were washed with PBS and

permeabilized in PBS 0.1% Triton X-100 for 10 min at room temperature. To prevent non-specific binding of the antibodies, cells were incubated with PBS in the presence of 5% BSA for 30 min.

The coverslips were incubated with primary polyclonal anti-Epn3 antibody (rabbit, produced in house) diluted in PBS 0.2 % BSA. After 1 hour of incubation at room temperature, coverslips were washed with PBS. Cells were then incubated for 1h at room temperature with secondary anti-mouse antibody Cy3 (Amersham) to detect E-cadherin and anti-rabbit Alexa 488-conjugated (Molecular Probes) to detect Epn3. After washes in PBS, coverslips were mounted in a 90% glycerol solution containing diazabicyclo-(2.2.2)octane antifade (Sigma) and examined under a confocal immunofluorescence microscope (Leica). Images were further processed with the Adobe Photoshop software (Adobe) or with Image J to merge the images of the single channels.

AIMS AND RATIONALE OF THE STUDY

The starting point for the research presented in this thesis was provided by a preliminary gene expression analysis conducted in our lab to identify novel genes that might contribute to poor prognosis in breast cancer. In this preliminary study, the endocytic protein Epn3 was found to be part of a gene signature associated with metastatic relapse in breast cancer patients. Based on these preliminary data, we hypothesized that Epn3 might have a pathogenetic role in this malignancy. Given the emerging data supporting a crosstalk between endocytosis, signaling and cancer (see Introduction, sections 1.1 and 1.2), it is possible that Epn3 exerts an oncogenic activity through an endocytosis-based mechanism (although we do not exclude that a non-endocytic function might also be relevant to the putative role of Epn3). Indeed, several endocytic proteins are abnormally expressed in human cancers and, for some of them, a causative link between altered trafficking and tumorigenesis has been demonstrated. Therefore, a hot topic in the endocytic field is to demonstrate whether deregulation of internalization processes can promote tumor initiation. In light of this scenario, we were interested in clarifying the role of Epn3 in breast tumorigenesis, which could provide additional evidence to support a strong connection between endocytosis and cancer.

An additional relevant aspect that emerged from our preliminary study was the potential prognostic value of Epn3 in breast cancer that, if confirmed, might contribute to a more refined clinical management of breast cancer patients. Indeed, the identification of novel biomarkers and new therapeutic targets to better stratify breast cancer patients seems to be the major need in the clinical management of this type of cancer. Advanced technologies have allowed the generation of gene expression profiles that have prognostic power and can bring

therapeutic benefits to patients. Nevertheless, to allow the development of proper patient stratification tools that can be easily used in the clinical setting the identification of novel prognostic biomarkers from these signatures is currently needed.

In line with our rationale, the first aim of the study presented in this thesis was to assess the clinical value of Epn3 as a prognostic marker. To this end, we performed IHC analysis on large cohorts of breast cancer patients to evaluate the correlation of Epn3 with classical clinical parameters of breast cancer and its ability to predict clinical outcome and patient survival (See Results, section 1).

The second aim was to validate the role of Epn3 in breast tumorigenesis by exploiting *in vitro* and *in vivo* assays to study cell transformation and tumorigenic potential upon Epn3 ablation or overexpression (see Results, section 2).

Finally, the third goal was to investigate the molecular mechanisms underlying Epn3 function in breast cancer and normal cells to uncover whether Epn3 exerts its oncogenic properties through an endocytic function, which could interfere with one or more signaling pathways (see Results, section 3).

Results from the proposed study should allow us to define whether Epn3 overexpression has a causal role in breast cancer, as well as provide valuable insights into the molecular mechanisms through which Epn3 could contribute to breast tumorigenesis. In a future perspective, our results could also lead to the identification of novel therapeutic targets and to a better clinical management of breast cancer patients.

RESULTS

1 Analysis of Epn3 expression in human breast tumors

1.1 Epn3 expression correlates with aggressive clinical/pathological parameters in invasive breast tumors

In a previous preliminary study performed in our lab with the aim of screening for novel genes contributing to poor prognosis in breast cancer, a transcriptional profile associated with metastatic relapse in breast cancer patients was derived (unpublished data). In this study, the Affimetrix GeneChip platform was used to perform gene expression analysis of tumors from 70 patients with node-negative primary breast carcinomas. This patient cohort included surgical patients who developed metastases within 5 years (29 cases) and patients who continued to be disease-free (41 cases) after 12 years. The analysis of the genes differentially expressed between these two groups generated a gene signature prognostic for metastasis, in which *EPN3* was included. *EPN3* was overexpressed in 20% of the breast tumors analyzed and in 45% of the poor-prognosis cases. These preliminary observations argued for a possible role for Epn3 as prognostic marker in breast cancer and prompted us to further investigate this hypothesis.

To assess the clinical relevance of Epn3 expression in breast cancer, we investigated the correlation between Epn3 expression in breast cancer patients and clinical/pathological parameters and disease outcome. To this purpose, we decided to analyze larger patient cohorts by immunohistochemistry (IHC) on breast tumor tissue microarray (TMAs), already available in our lab.

To define the parameters for Epn3 quantification, we evaluated the expression of Epn3 on normal breast tissues (n=31) using a specific anti-Epn3

antibody ("VI31" monoclonal antibody) that was produced in-house since commercial anti-Epn3 antibodies were not available to perform a comprehensive Epn3 protein expression analysis. Normal breast tissue samples showed very low levels of Epn3 expression (median score=0.5), which lead us to define tumors with high Epn3 expression as those displaying an expression score >1 (Figure 6)

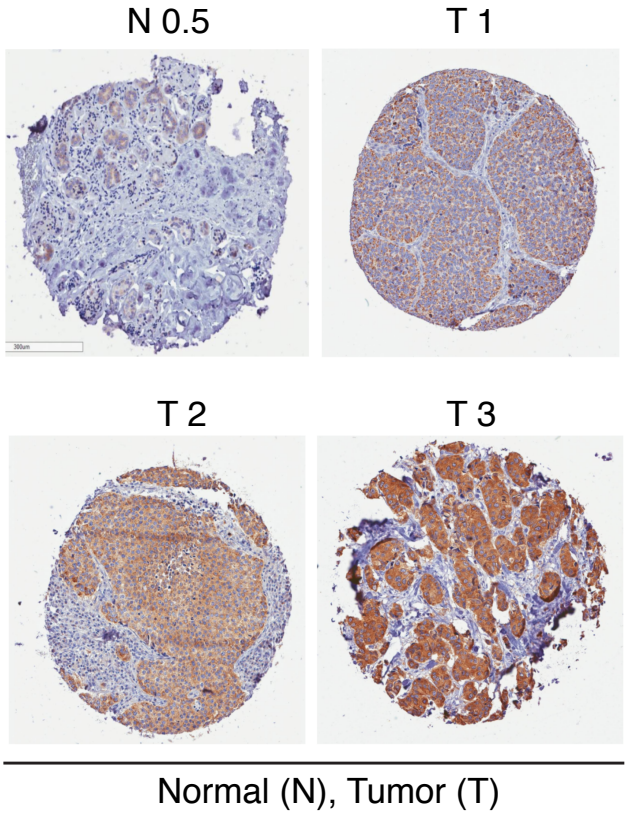


Figure 6 | Immunohistochemical analysis of Epn3 expression in human breast tumor samples

Immunohistochemical (IHC) analysis of Epn3 expression was performed on formalin-fixed paraffin-embedded breast tumor and normal tissue sections (2μm) using the VI31 anti-Epn3 antibody produced in-house.
A representative image of a normal breast tissue section expressing low levels of Epn3 (score=0.5) is shown. Three representative images of tumor samples showing different levels of Epn3 expression (scores from 1 to 3) are shown. Tumors with high Epn3 expression were defined as those displaying an expression score >1. According to the score assigned, tumors were classified as EPN3-high and EPN3-low. Scale bar 300μm

Epn3 protein levels were then analyzed by IHC on 2 independent case collections: one ‘case control’ group – composed of 349 patients who developed an event, including loco-regional relapse and distant metastasis after the first breast cancer surgery (also referred to as training set,
Table 1) – and one ‘validation’ group – comprising 916 consecutive cases (our test set, Table 2). In the case-control cohort, Epn3 expression was high in approximately 30% of tumors (92 out of 298 analyzable cases). Elevated Epn3 levels significantly correlated with many clinical and pathological parameters indicative of aggressive disease, such as tumor size (pT), node positive status, high grade (GRADE), PgR negative status, high proliferative index (Ki67) and HER2 positive status (Table 3). Moreover, there was a significant association with the more aggressive molecular subtypes, such as Luminal B and triple negative (TNBC) tumors [111, 137] (Table 3). Finally, high expression of Epn3 correlated with any event of disease relapse, distant metastasis and death (Table 3).

To refine our analysis, we analyzed the expression of Epn3 on the consecutive group (Table 2). In this group, we confirmed the association between high Epn3 levels and high grade, hormone receptor negative tumors (ER and PgR), high Ki67 and positive HER2 status as well as the different molecular subtypes (Luminal B, Luminal B-HER2+, HER2+ and triple negative) (Table 4).

Altogether these results indicate the Epn3 expression is associated with prognostic markers of aggressive breast cancer, disease recurrence and survival, suggesting a clinical value for Epn3 detection in breast cancer.

CASE-CONTROL Cohort (N = 349)			
Parameter	**Group**	**N**	**%**
Histotype	Ductal	290	83,09
	Lobular	50	14,33
	Mixed/Other	9	2,58
Age at Surgery	<50	171	49
	≥50	178	51
pT	pT1	178	51,44
	pT2	142	41,04
	pT3	18	5,2
	pT4	8	2,3
Grade	G1	56	18,42
	G2	132	43,42
	G3	116	38,16
Ki67 Status	LOW	116	34,42
	HIGH	221	65,58
HER2 Status	NEG	227	88,22
	POS	37	11,78
Nodal Status	NEG	153	43,84
	POS	196	56,16
ER Status	NEG	106	31,45
	POS	231	68,55
PgR Status	NEG	140	41,54
	POS	197	58,46
Subtype	LUMINAL A	88	28,85
	LUMINAL B	129	42,3
	HER2 POS	23	7,54
	LUM B HER2 POS	10	3,28
	TNBC	55	18,03
Any Event	ANY	196	56,16
	NONE	153	43,84
Distant event	DISTANT	244	69,92
	NONE	105	30,09
Status	Alive	233	66,76
	Dead	116	33,24

Table 1 | Clinical and pathological information of the case-control dataset of breast cancer patients

The clinical and pathological information of the patients included in the case-control TMA is reported. These patients have undergone surgery at the European Institute of Oncology (IEO) between 1994 and 1997. Disease recurrence (any and distant relapse) was within 18 years. Any Relapse: loco-regional relapse, distant metastasis; Distant Relapse: distant metastasis; secondary primary cancer and contralateral breast cancer were not considered as a primary tumor related event. For some patients not all information was available.

CONSECUTIVE Cohort (N=916)			
Cohort	**Group**	**N**	**%**
Histotype	Ductal	728	79,48
	Lobular	102	11,14
	Mixed/Other	86	9,39
Age at Surgery	<50	410	44,76
	≥50	506	55,24
pT	pT1	85	9,28
	pT2	756	82,53
	pT3	63	6,88
	pT4	12	1,31
Grade	G1	68	7,74
	G2	340	38,68
	G3	471	53,58
Ki67 Status	LOW	169	18,47
	HIGH	746	81,53
HER2 Status	NEG	797	88,95
	POS	99	11,05
Nodal Status	NEG	306	33,89
	POS	597	66,11
ER Status	NEG	199	21,72
	POS	717	78,28
PgR Status	NEG	310	33,84
	POS	606	66,16
Subtype	LUMINAL A	162	18,1
	LUMINAL B	508	56,76
	HER2 POS	49	5,47
	LUM B HER2 POS	50	5,59
	TNBC	126	14,08
Any Event	ANY	272	29,69
	NONE	644	70,31
Distant event	DISTANT	245	26,75
	NONE	671	73,25
Status	Alive	664	72,49
	Dead	252	27,51

Table 2 | Clinical and pathological information of the consecutive cohort of breast cancer patients

The clinical and pathological information of the patients included in the consecutive cohort is reported. These patients have undergone surgery at the European Institute of Oncology (IEO) between 1997 and 2000. Disease recurrence (any and distant relapse) was within 18 years. Any Relapse: loco-regional relapse, distant metastasis; Distant Relapse: distant metastasis;

secondary primary cancer and contralateral breast cancer were not considered as a primary tumor related event. For some patients not all information was available.

CASE-CONTROL COHORT (N = 349*)				
		EPN3 LOW	**EPN3 HIGH**	
All Patients (298*)		206	92 (30.87%)	Pvalue
Histotype	Ductal	163	86 (**34.54%)**	
	Lobular	35	5 (12.50%)	**0.0053**
pT	1	116	38 (24.68 %)	
	2-3-4	89	53 (**37.32%**)	**0.0185**
Nodal Status	NEG	102	24 (19.05 %)	
	POS	104	68 (**39.53 %**)	**0.0002**
GRADE	G1	45	6 (11.76 %)	
	G2	80	29 (26.61 %)	
	G3	54	52 (**49.06 %**)	**<0.0001**
ER	POS	144	56 (28 %)	
	NEG	57	33 (36.67 %)	0.1388
PgR	POS	128	43 (25.15 %)	
	NEG	73	46 (**38.66 %**)	**0.0141**
Ki67	LOW	82	13 (13.68 %)	
	HIGH	119	76 (**38.97 %**)	**<0.0001**
HER2	NEG	181	70 (27.89 %)	
	POS	15	17 (**53.13 %**)	**0.0036**
LUMINAL A		68	9 (11.69%)	
LUMINAL B		77	44 (36.36%)	
LUMINAL B HER2 POS		6	3 (33.33%)	
HER2 POS		7	12 (63.16%)	
TNBC		34	16 (32%)	**<0.0001**
Distant Relapse	NO	162	47 (22.49 %)	
	YES	44	45 (**50.56 %**)	**<0.0001**
Any Relapse	NO	139	30 (17.75 %)	
	YES	67	62 (**48.06 %**)	**<0.0001**
Status	Alive	157	44 (21.89 %)	
	Dead	49	48 (**49.48 %**)	**<0.0001**

Table 3 | Correlation of Epn3 expression and clinico-pathological parameters in the case-control cohort of 349 breast cancer patients

Epn3 expression was measured by IHC on TMA using the anti-Epn3 VI31 antibody. Tumors with high Epn3 status (EPN3 HIGH) display an expression score >1, while low Epn3 status (EPN3 LOW) tumors show an expression score <1. Asterisks indicate that the number of scored cases (298) is lower than the total number of cases (349). Scoring of samples was not possible when individual cores detached from the slides during experimental manipulation and when clinical information was not available.

CONSECUTIVE COHORT (N = 916*)				
		EPN3 LOW	**EPN3 HIGH**	
All Patients (762*)		479	283 (37.14%)	P value
Histotype	Ductal	365	258 (**41.41%)**	
	Lobular	62	13 (17.33%)	
	Other/Mixed	52	12 (18.75%)	**<0.0001**
pT	1	51	22 (30.14 %)	
	2-3-4	428	261 (37.88%)	0.1929
Nodal Status	NEG	170	84 (33.07 %)	
	POS	299	199 (39.96%)	0.0651
GRADE	G1	47	5 (9.62 %)	
	G2	209	67 (24.28 %)	
	G3	199	205 (**50.74 %**)	**<0.0001**
ER	POS	395	201 (33.72 %)	
	NEG	84	82 (**49.40 %**)	**0.0002**
PgR	POS	328	172 (34.40 %)	
	NEG	151	111 (**42.37 %**)	**0.0306**
Ki67	LOW	116	16 (12.12 %)	
	HIGH	362	267 (**42.45 %**)	**<0.0001**
HER2	NEG	445	225 (33.58 %)	
	POS	28	55 (**66.27 %**)	**<0.0001**
LUMINAL A		113	16 (12.40%)	
LUMINAL B		270	162 (37.50%)	
LUMINAL B HER2 POSITIVE		12	29 (70.73%)	
HER2 POSITIVE		16	26 (61.90%)	
TNBC		61	47 (43.52%)	**<0.0001**

Table 4 | Correlation of Epn3 expression and clinico-pathological parameters in the consecutive cohort of 916 breast cancer patients

Epn3 expression was measured by IHC on TMA using anti-Epn3 VI31 antibody. Tumors with a high Epn3 status (EPN3 HIGH) display an expression score >1, while low Epn3 status (EPN3 LOW) tumors show en expression score <1. Asterisks indicate that the number of scored cases (762) is lower than the total number of cases (916). Scoring of samples was not possible when individual cores detached from the slides during experimental manipulation and when clinical information was not available.

1.2 Epn3 expression is associated with high risk of breast cancer recurrence and poor disease outcome

Since Epn3 overexpression is associated with indicators of aggressive breast cancer, we decided to investigate whether Epn3 status could be a prognostic marker able to predict disease recurrence and survival in breast cancer patients.

In the case-control group, we found that Epn3 overexpression is significantly correlated with high risk of disease recurrence (any or distant relapse) and death (Table 3).

To validate the prognostic potential of Epn3 observed in the case-control cohort, we performed Kaplan-Meier analysis on the consecutive cohort of 916 patients (Table 2). Tumor patients were stratified on the basis of their Epn3 expression levels in high (score >1) and low (score <1) Epn3 status. Kaplan-Meier curves of disease-free survival and overall survival revealed that high Epn3 status correlates with increased risk of disease relapse ($p=0.001$) and decreased overall survival ($p=0.0028$) (Figure 7)

Since HER2 is one of the most amplified oncogene in breast cancer strongly associated with poor outcome [124, 138] (see also Discussion, section 1.1) we tested whether Epn3 can retain its prognostic value independent of HER2 status. We analyzed the association between Epn3 status and survival in the subpopulation of HER2-negative patients present in the consecutive cohort. Also in this case, our results confirmed an inverse correlation between Epn3 expression and disease-free survival ($p=0.0003$) and overall survival ($p=0.0018$) (Figure 8).

Together, our data identify Epn3 overexpression as a new prognostic marker in breast cancer. This marker significantly correlates with high risk of disease recurrence and fatal outcome. Importantly, Epn3 retains its prognostic value independently of HER2 status in breast cancer patients.

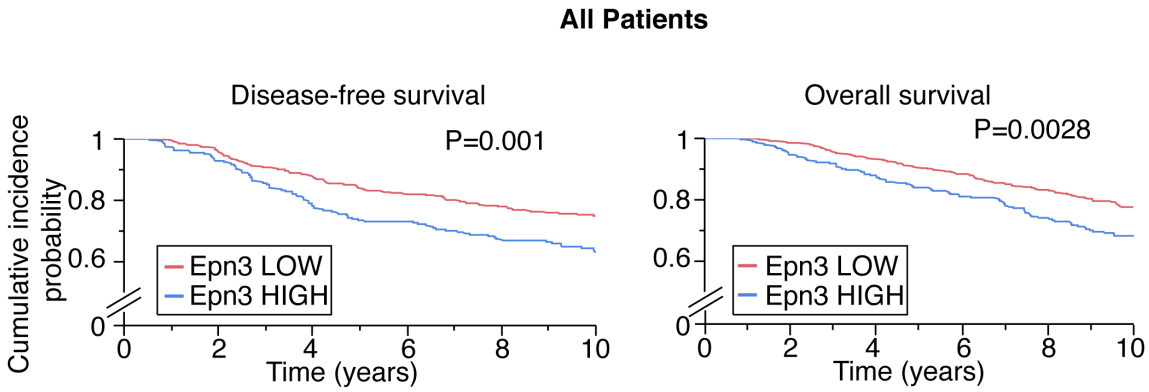


Figure 7 | Cumulative incidence probability of disease-free survival and overall survival in the consecutive cohort

Kaplan-Meier plots of the 10-years cumulative incidence probability of disease-free survival and overall survival according to low (LOW, score <1, red line) and high (HIGH, score >1, blue line) Epn3 expression measured by IHC in the consecutive cohort of 916 breast cancer patients. P-values (P) are indicated.

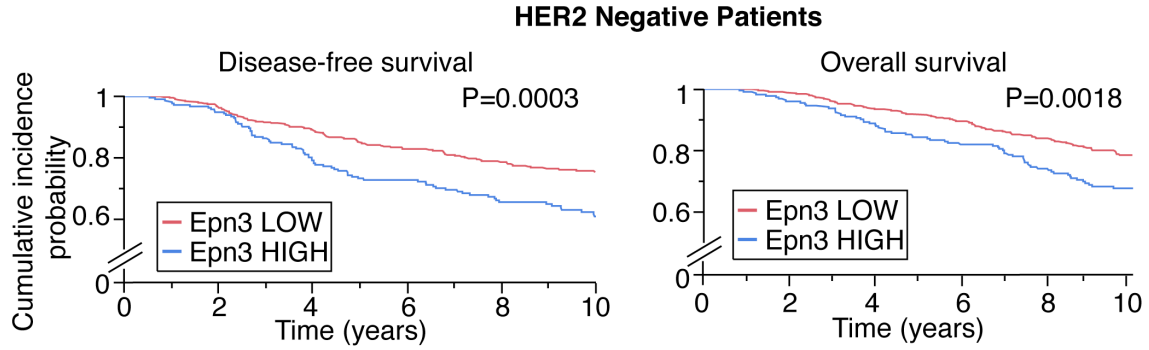


Figure 8 | Cumulative incidence probability of disease-free survival and overall survival in HER2-negative patients of the consecutive cohort

Kaplan-Meier plots of the 10-years cumulative incidence probability of disease-free survival and overall survival according to low (LOW, score <1, red line) and high (HIGH, score >1, blue line) Epn3 expression measured by IHC in HER2-negative patients of the consecutive cohort of 916 breast cancer patients. P-values (P) are indicated.

1.3 The *EPN3* gene is amplified in human breast tumors

Gene amplification is one of the molecular mechanisms that lead to oncogene activation [139]. To assess whether the *EPN3* gene is amplified in breast tumors, we performed dual color interphase fluorescence *in situ* hybridization (FISH) analysis on human breast tumor biopsies. The *EPN3* gene is located on chromosome 17, a chromosome very relevant to breast tumorigenesis since one of the most frequently amplified breast oncogenes, *HER2,* is located there. Of note, *EPN3* and *HER2* are located in the 17q12-21 region, which is frequently rearranged in breast cancers, approximately 12M base pairs apart from each other [140, 141] (Figure 9A). Thus, we were also interested in understanding whether these genes were co-amplified in the analyzed tumors.

We used a BAC clone encompassing the *EPN3* gene and a *HER2* gene probe to detect *EPN3* and *HER2* gene amplification, respectively, in a subset of 281 breast tumors. We observed *EPN3* gene amplification in approximately 18%, and *HER2* amplification in 15% of the total samples analyzed (Figure 9B). Despite the observation that co-amplification of the two genes occurred to some extent, our results show that *HER2* and *EPN3* amplifications are independent events. Indeed, 66% of *EPN3*-amplified tumors did not show *HER2* amplification while 60% of *HER2*-amplified cases did not display *EPN3* amplification. These results are in line with previous data showing that these two amplicons can be amplified independently [142].

These data indicate that *EPN3* is amplified in a subset of breast tumors and, together with the observation that Epn3 overexpression has a potential prognostic value in breast cancer patients - including those with HER2-negative status - strongly suggest that the involvement of Epn3 in cancer progression could be both dependent and independent from HER2 status.

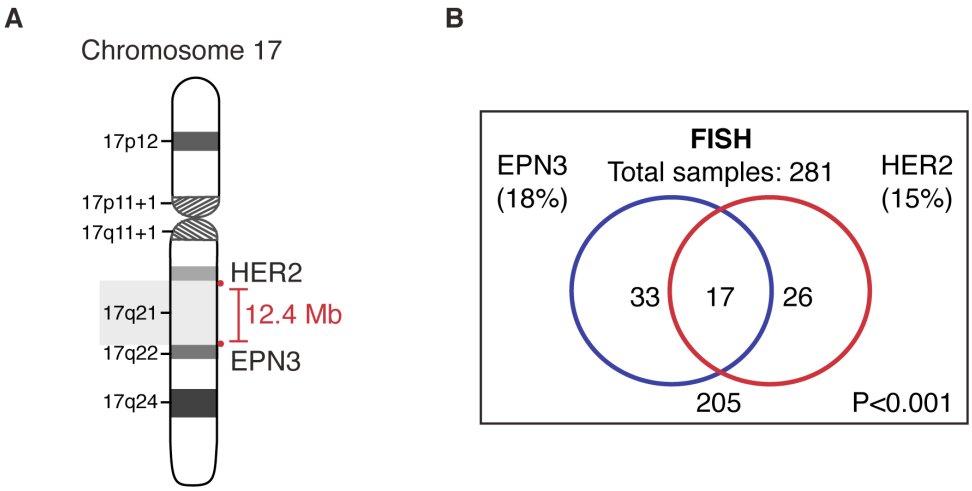


Figure 9 | Analysis of *EPN3* gene amplification in human breast tumor samples

A) Scheme of chromosome 17. The *EPN3* gene is located in chromosome 17, 12.4 Mb from the HER2 gene. B) Representative scheme of *EPN3* and *HER2* gene-amplified samples as assessed by FISH analysis in 281breast tumor samples. Venn diagram shows the overlap between EPN3-FISH and HER2-FISH data. Both *EPN3* and *HER2* genes were considered amplified when the ratio between each specific gene and the chromosome 17 centromere copy number was >2.25. Significance of the co-amplification was calculated with Fisher't.

2 Characterization of the role of Epn3 in breast tumorigenesis

To identify suitable cell model systems to investigate the role of Epn3 in breast tumorigenesis, we analyzed the expression of the protein in a panel of commercially available human normal and tumor breast cell lines, both by western blot (WB) and quantitative PCR (Q-PCR) analysis. We observed that Epn3 is highly expressed in two breast tumor cell lines, BT474 and MDA-MB361, in comparison to the normal breast cell lines, MCF10A and HMEC, both at the protein and mRNA level (Figure 10A,B).

We then assessed whether Epn3 overexpression in the BT474 and MDA-MB361 cell lines was due to gene amplification, as observed in human breast tumor samples (Figure 10). We performed dual color FISH analysis on metaphase chromosomes, using a centromeric probe for chromosome 17 and a BAC clone encompassing the *EPN3* gene, and observed that *EPN3* is indeed amplified in both BT474 and MDA-MB361 cells (Figure 11).

Based on these results, we selected different cell lines to be used as model systems in the analysis of the involvement of Epn3 in breast tumorigenesis: i) the BT474 breast tumor cell line, which overexpresses Epn3 due to gene amplification, was chosen to study the dependence of tumorigenic phenotypes on *EPN3* locus alteration; ii) the normal breast epithelial cell line MCF10A and the breast tumor cell lines HCC1569 and MCF7 - which express low levels of Epn3 and do not show *EPN3* amplification – were selected to uncover the transforming potential of Epn3 overexpression.

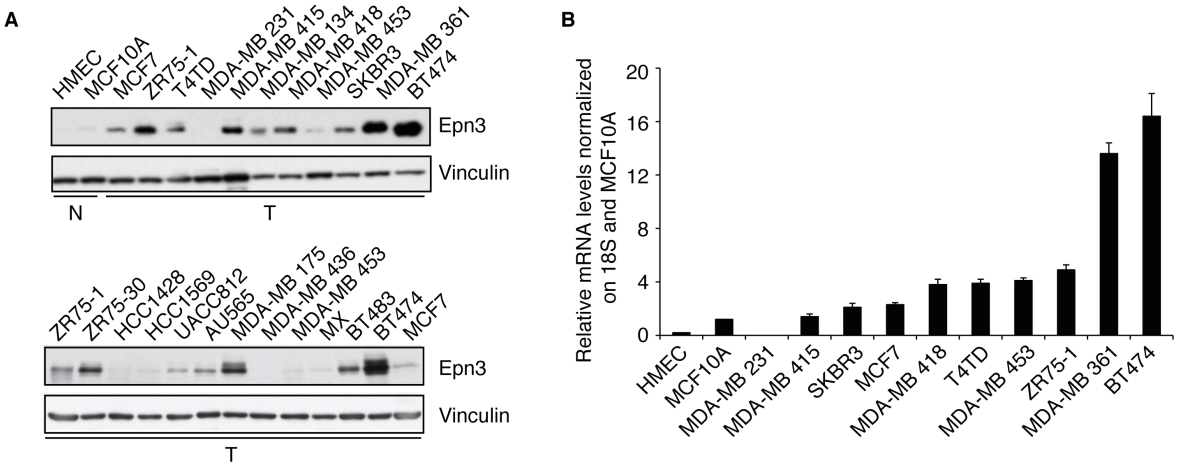


Figure 10 | Analysis of Epn3 expression in a panel of human breast normal and tumor epithelial cell lines

A) Epn3 expression in human normal mammary epithelial cell lines (N) and human breast tumor epithelial cell lines (T) was evaluated by western blot (WB) analysis using VI31 antibody. Vinculin was used as loading control. B) Expression of *EPN3* transcript levels was examined with Q-PCR analysis in normal mammary cell lines (N) and some of the breast tumor (T) cell lines. Results were normalized to the housekeeping gene 18S and are expressed as fold-change relative to MCF10A cells. Results represent the mean ± SD of an experiment performed in triplicate.

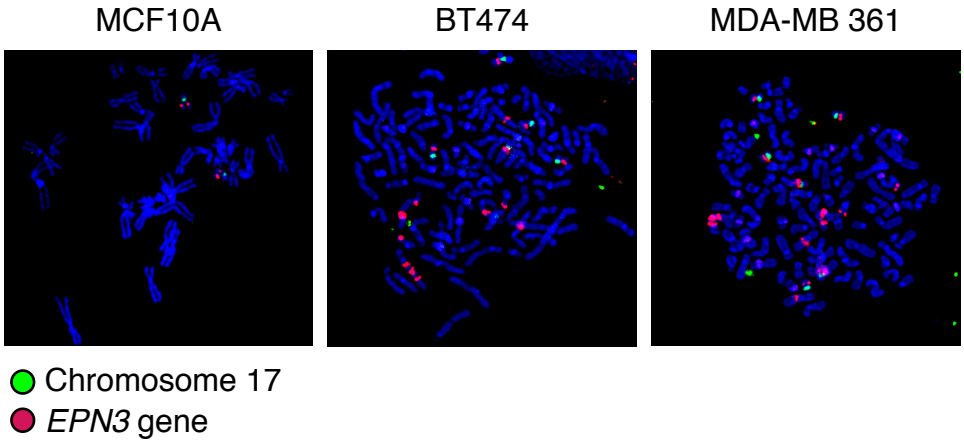


Figure 11 | Analysis of *EPN3* gene amplification in breast cell lines

EPN3 gene amplification evaluated by FISH analysis on metaphase chromosomes in MCF10A, BT474 and MDA-MB 361 cells. A centromeric probe for chromosome 17 (green) and a BAC clone for *EPN3* gene (red) were used. DAPI was used for nuclei staining. *EPN3* gene was considered amplified when the ratio *EPN3*/Chr.17 was >2-fold. At least three nuclei per sample were analyzed. Representative images at 100x magnification are shown.

2.1 Functional characterization of Epn3 ablation in breast tumor cells carrying *EPN3* gene amplification

2.1.1 Epn3 is required for *in vitro* anchorage-independent growth of breast tumor BT474 cells

To shed light on the role of Epn3 in breast carcinogenesis, we stably silenced Epn3 expression in BT474 cells, and analyzed the consequences of this genetic manipulation on tumorigenic phenotypes.

We used a lentiviral vector system encoding two short-hairpin RNA (shRNA) oligos targeting Epn3 (shEpn3#1 and shEpn3#2) to stably knockdown (KD) Epn3 in BT474 cells. As controls, we used a shRNA targeting luciferase (shLUC) or a shRNA mismatch sequence (shMis) (see Materials and Methods). Both Epn3-shRNA yielded efficient ablation of Epn3 expression as assessed by WB analysis (Figure 12A). We performed a soft agar assay and we found that Epn3 ablation in BT474 cells strongly impaired the capacity of these cells to grow in an anchorage-independent manner, causing a reduction in the number and size of colonies formed in semi-solid medium (~2-fold reduction in BT474 Epn3-KD cells in comparison to control cells, $p<0.05$) (Figure 12B)

To provide a direct evidence of the specific role of Epn3 overexpression driven by gene amplification in the tumorigenic capacity of breast tumor cells, we decided to silence Epn3 in MCF7 breast tumor cells, which show basal levels of Epn3 without any alteration of *EPN3* locus (Figure 10), in order to perform a comparative phenotytpic analysis between BT474 and MCF7 Epn3-KD cells.

We stably silenced Epn3 expression in MCF7 cells by infecting these cells with the same lentiviral systems used in BT474 cells. As for BT474 cells, we obtained a good Epn3 KD efficiency in MCF7 cells, as assessed by WB analysis (Figure 13A). However, contrary to the results obtained in BT474, we found that

ablation of Epn3 in MCF7 did not affect colony formation in the soft agar assay (Figure 13B).

These results indicate that Epn3 ablation selectively impairs the anchorage-independent growth of cells harboring *EPN3* gene amplification, such as BT474, while it does not significantly affect the growth of tumor cells with physiological Epn3 levels.

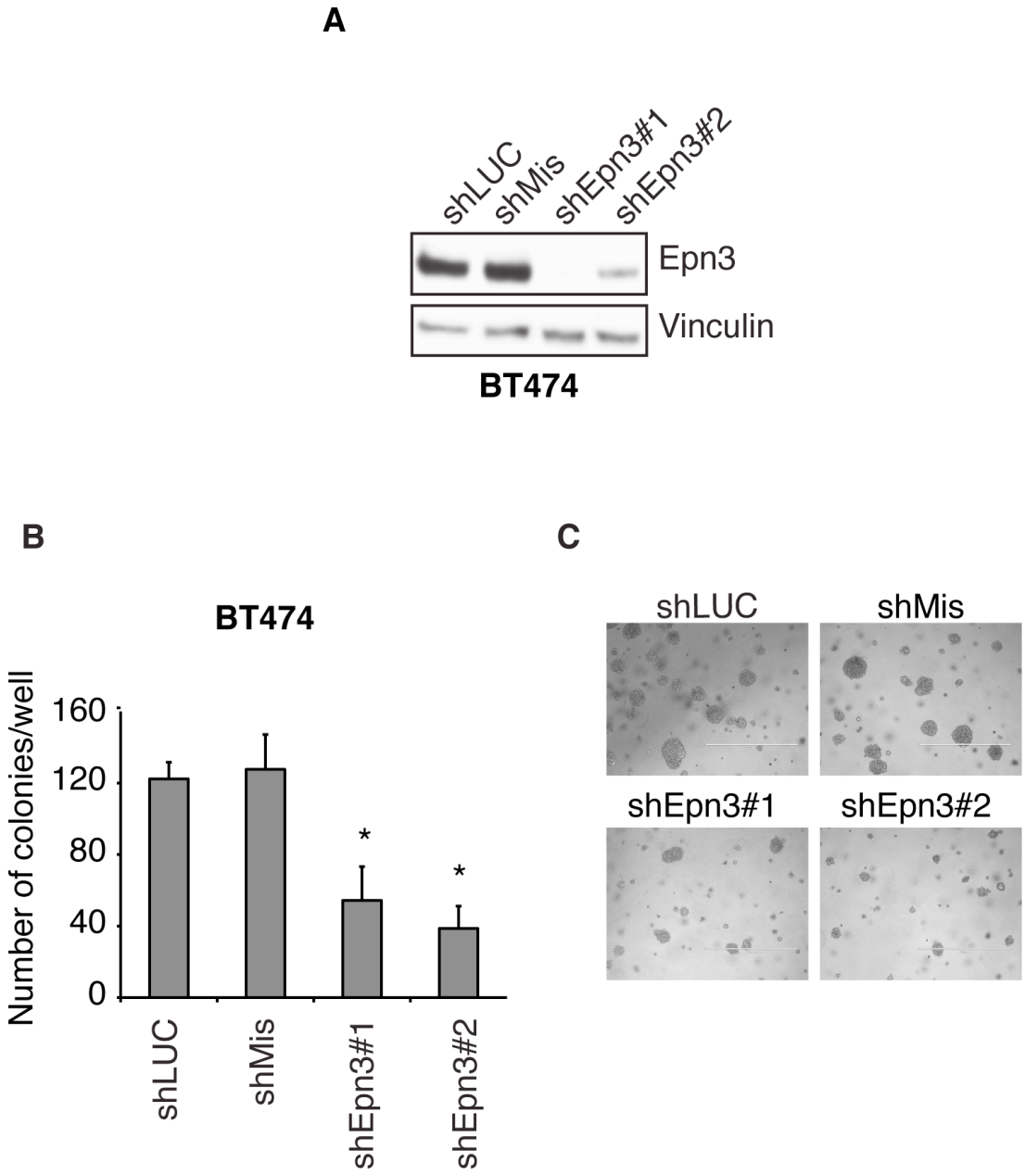


Figure 12 | Effect of Epn3 ablation on anchorage-independent growth of BT474 cells

BT474 cells were infected with lentiviral vectors (pSICOR) expressing shRNAs targeting Epn3 (shEpn3#1, shEpn3#2). Lentiviral vectors encoding a shRNA targeting luciferase (shLUC) and a mismatch shRNA (shMis) were used as negative controls. A) Efficiency of stable silencing in BT474 was assessed by WB analysis using VI31 antibody. Vinculin was used as loading control. B) The anchorage-independent growth of BT474-infected cells was assessed using the soft agar assay. Cells (25,000) were plated in 6-well culture plates in a medium containing 0.3% agar. After 3 weeks in culture, colonies were counted by bright-field microscopy at 4x magnification. Bar graphs report mean colony number per well ± SD of an experiment performed in triplicate and are representative of 3 independent experiments. Asterisks indicate mean p-value < 0.05 compared to shLUC and shMis. C) Representative bright-field microscopy images of soft-agar colonies formed by BT474-infected cells are shown. Magnification 4x. Scale bar 100 µm.

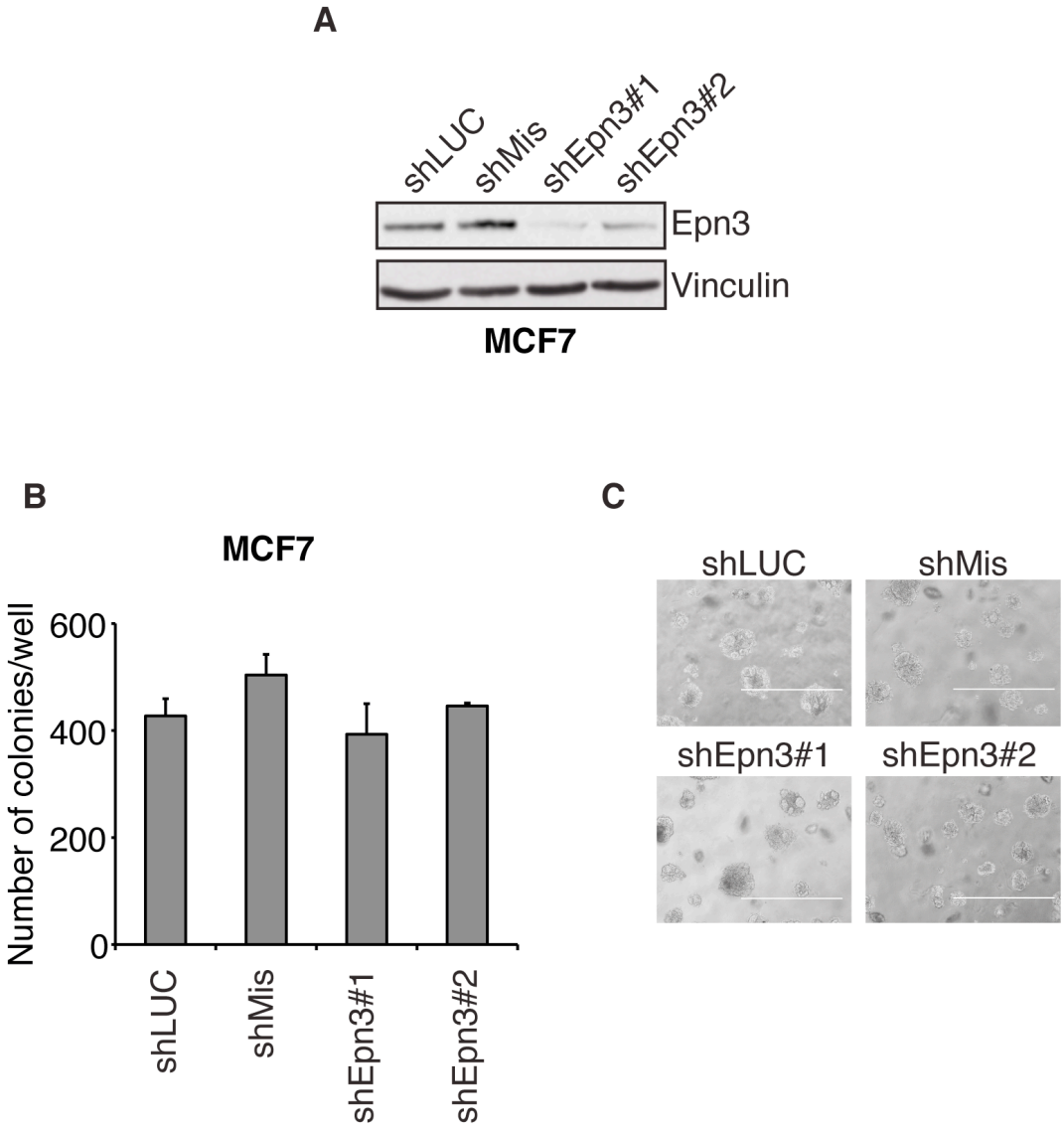


Figure 13 | Effect of Epn3 ablation on anchorage-independent growth of MCF7 cells

MCF7 cells were infected with lentiviral vectors (pSICOR) expressing shRNAs targeting Epn3 (shEpn3#1, shEpn3#2). Lentiviral vectors encoding a shRNA targeting luciferase (shLUC) and a mismatch shRNA (shMis) were used as negative controls. A) Efficiency of stable silencing in MCF7 cells was assessed by WB analysis using VI31 antibody. Vinculin was used as loading control. B) The anchorage-independent growth of MCF7-infected cells was assessed using the soft agar assay. Cells (25,000) were plated in 6-well culture plates in a medium containing 0.3% agar. After 3 weeks in culture, colonies were counted by bright-field microscopy at 4x magnification. Bar graphs report mean colony number per well ± SD of an experiment performed in triplicate and are representative of 3 independent experiments. C) Representative bright-field microscopy images of soft-agar colonies formed by MCF7-infected cells are shown. Magnification 4x. Scale bar 100 μm.

2.1.2 Epn3 is required for *in vivo* tumorigenic growth of breast tumor BT474 cells

Based on the *in vitro* results, we next investigated the effects of Epn3 ablation on the *in vivo* tumorigenic potential of BT474 cells.

To this aim, $3x10^5$ BT474 Epn3-KD cells were orthotopically injected into the inguinal mammary fat pad of NOD/SCID female immunocompromised mice. As a control, we injected BT474 control ($3x10^5$ BT474 shLUC) cells into the contralateral mammary gland. Eleven weeks post-injection, tumors reached a palpable size (~$100mm^3$) and were monitored every 7 days for another 3 weeks. Remarkably, tumors derived from Epn3-KD BT474 cells grew slower and were ~2-fold smaller in comparison with tumors generated by control BT474 cells (Figure 14A,B). To assess the Epn3 expression levels on the tumors extracted from the mammary gland, we dissociated them into cells and performed WB analysis. We observed that the efficiency of Epn3 ablation in tumors was the same as the one obtained in parental Epn3 KD cells (Figure 14C).

Our results show that BT474 cells are dependent on altered Epn3 expression for the maintenance of their malignant phenotype.

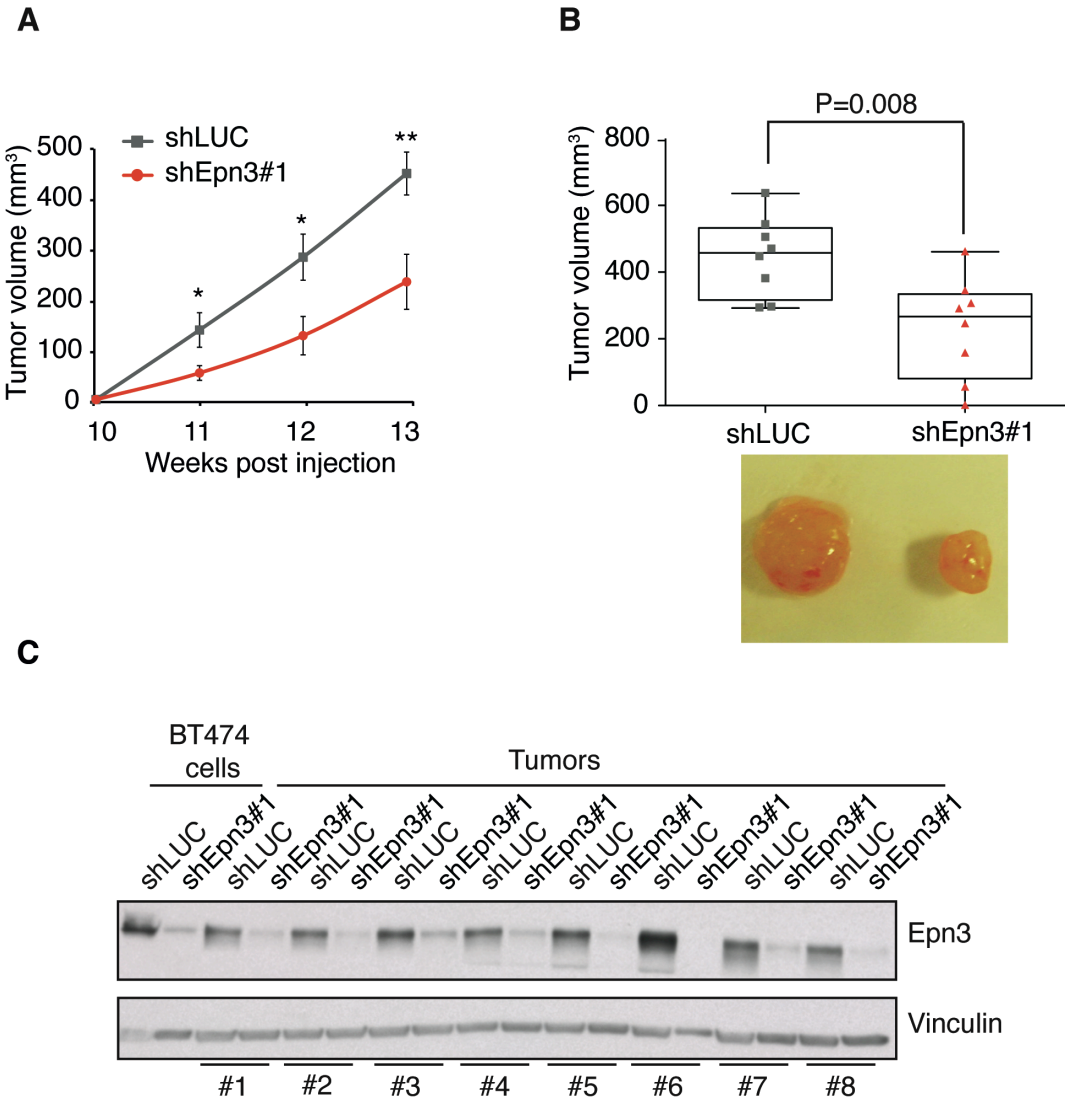


Figure 14 | Effect of Epn3 ablation on *in vivo* tumor growth of BT474 cells

Mammary tumors were generated by injecting 300 000 Epn3-KD cells (shEpn3#1) and control (shLUC) BT474 cells into the inguinal mammary fat pads of female NOD/SCID IL2 gamma-chain null mice. Tumors were grown for 13 weeks before being explanted. A) Kinetic evaluation of BT474 shLUC and shEpn3#1 tumor growth. Tumor volume was assessed by *in vivo* caliper measurements at the indicated time points after injection. Asterisks indicate mean p-value < 0.05 (*) and < 0.01 (**) compared to shLUC. B) The distribution of tumor volume of BT474 shLUC and shEpn3#1 is reported in a box-plot. Representative images of tumor

outgrowths are reported below the chart. C) The expression of Epn3 in mammary tumors was assessed by WB analysis using VI31 antibody. Vinculin was used as loading control. Results are representative of 2 experiments (n=8) and represent the mean± SEM.

2.2 Functional characterization of Epn3 overexpression in normal and tumor breast cell lines

2.2.1 Epn3 overexpression induces EMT and invasive phenotype in normal breast epithelial MCF10A cells

To investigate the molecular mechanisms through which Epn3 can induce breast tumorigenesis, we overexpressed Epn3 in normal MCF10A cells, which display low basal levels of Epn3 and show no alterations at the level of the *EPN3* locus. To this purpose, we used a retroviral vector system expressing Flag-tagged Epn3 (pBABE-Epn3). To make sure we could observe an Epn3-specific induced phenotype, we also transfected cells with Flag-tagged Epn1 (pBABE-Epn1), a member of the Epsin family that is not amplified in breast tumors (see Discussion, sections 1.2 and 4). Cells infected with empty vector (pBABE) were used as a negative control. We obtained high levels of Epn3 expression in the MCF10A cells, which were comparable to the levels observed in BT474 cells (Figure 15A, lanes *a*, *b* and *c*). Furthermore, we obtained the same level of Epn3 and Epn1 overexpression, as determined by anti-Flag WB analysis (Figure 15A, lanes *c* and *d*).

Before we were able to perform some tumorigenic assays (i.e. soft agar), we immediately observed that Epn3 overexpression induced a marked alteration of cell morphology in 2D-culture. Indeed, pBABE-Epn3 cells, but not pBABE-Epn1 cells, acquired a more elongated and spindle-like morphology compared to the

cobblestone-like epithelial morphology of pBABE cells (Figure 15B). These morphological changes were suggestive of EMT in pBABE-Epn3 cells. EMT is a complex cellular process in which polarized epithelial cells progressively lose their epithelial properties and acquire a mesenchymal phenotype characterized by the expression of specific mesenchymal genes, increased invasiveness and migration (reviewed in [143]).

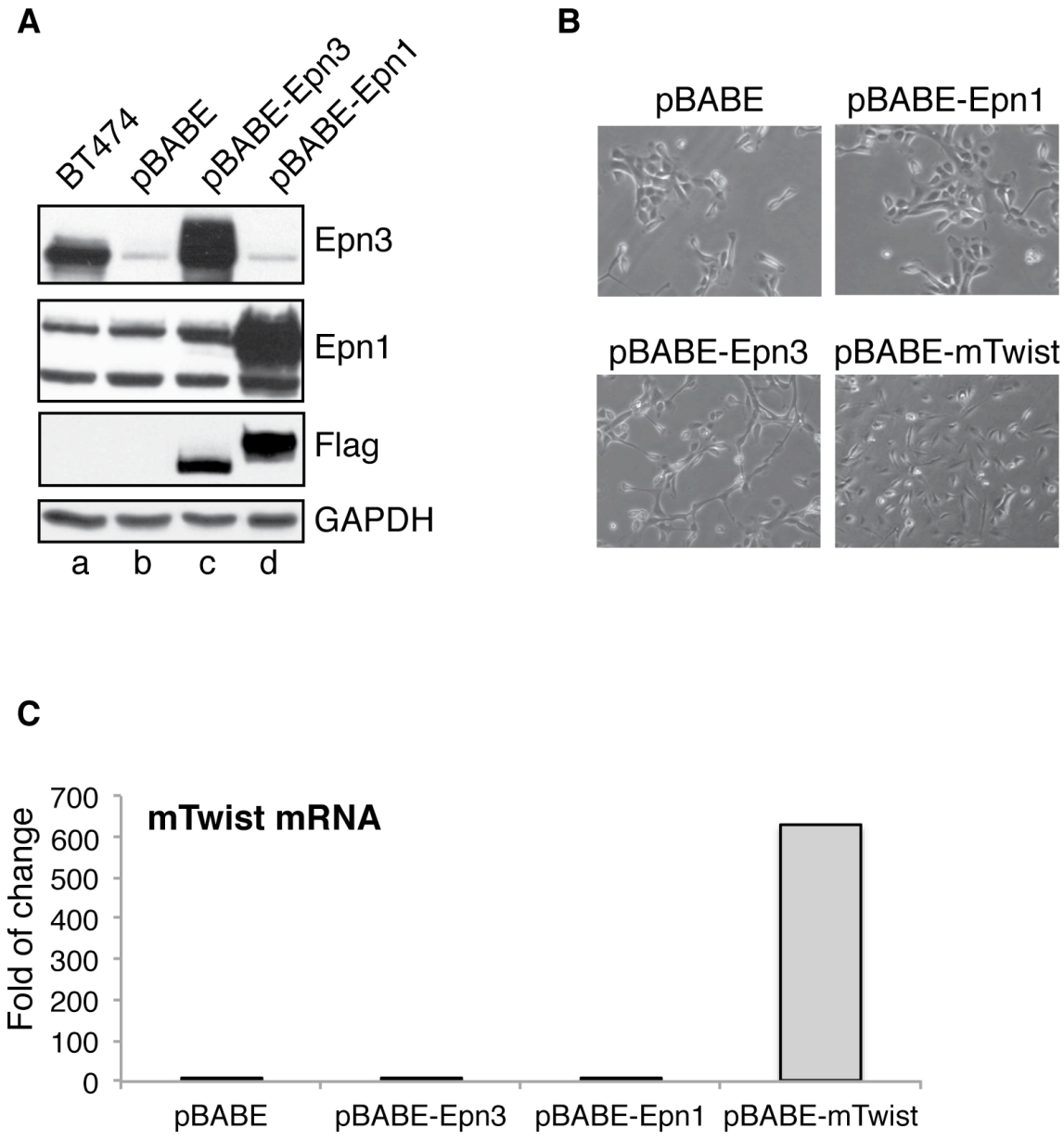


Figure 15 | Analysis of the effects of Epn3 overexpression in human normal mammary epithelial MCF10A cells

A) MCF10A cells were infected with retroviral vectors expressing Flag-Epn3 (pBABE-Epn3, *lane c*) or Flag-Epn1 (pBABE-Epn1, *lane d*). Cells infected with empty vector (pBABE, *lane b*) are used as negative control. Expression of Epn3, Epn1 (upper band) and Flag was examined by WB analysis. BT474 cells were used as positive control for Epn3 expression (*lane a*). GAPDH was used as a loading control. B) The morphology of MCF10A cells expressing control vector pBABE, pBABE-Epn1, pBABE-Epn3 or pBABE-mTwist was assessed by phase contrast microscopy. Representative images are shown. Magnification 4x. C) The overexpression of mTwist in MCF10A cells was evaluated by Q-PCR analysis. Data were normalized to the housekeeping gene GAPDH and are expressed as fold-change relative to pBABE cells. Results represent the mean ± SD of an experiment performed in triplicate.

To better investigate this phenotype, we generated MCF10A cells overexpressing the transcriptional factor Twist by transfecting them with retroviral vector expressing murine Twist (pBABE-mTwist). Twist is a well-known EMT-inducer gene and therefore represents a good positive control for induction of EMT phenotype [144]. We obtained high level of Twist overexpression in MCF10A as assessed by Q-PCR analysis (Figure 15C). By comparing pBABE-Epn3 and pBABE-mTwist cells at the morphological level, we observed a similar fibroblast-like shape in both cell types (Figure 15B). However, pBABE-mTwist cells were more scattered than pBABE-Epn3 cells suggesting a more advanced EMT phenotype.

To assess if this was the case, we analyzed the expression of known EMT markers, namely N-cadherin, Vimentin and E-cadherin. By performing WB and Q-PCR analysis, we observed higher levels of N-cadherin and vimentin expression in pBABE-Epn3 and pBABE-mTwist than in pBABE and pBABE-Epn1 cells (Figure 16A,B). On the contrary, E-cadherin protein expression was reduced in both pBABE-mTwist and pBABE-Epn3 cells (Figure 16A).

This “cadherin switch” is a known hallmark of EMT. Indeed, EMT is usually associated with a switch from E-cadherin to N-cadherin, which is expressed by mesenchymal cells, cancer cells and neural tissue (reviewed in [145]). Moreover, overexpression of N-cadherin is found in some carcinoma cells with decreased levels of E-cadherin; in these cells, N-cadherin acts as a weak intercellular adhesion molecules, promoting EMT progression and cell motility [146]. Of note, only Twist overexpression, but not Epn3, caused a reduction in the E-cadherin mRNA levels, suggesting that Twist and Epn3 may affect E-cadherin expression levels through distinct mechanisms.

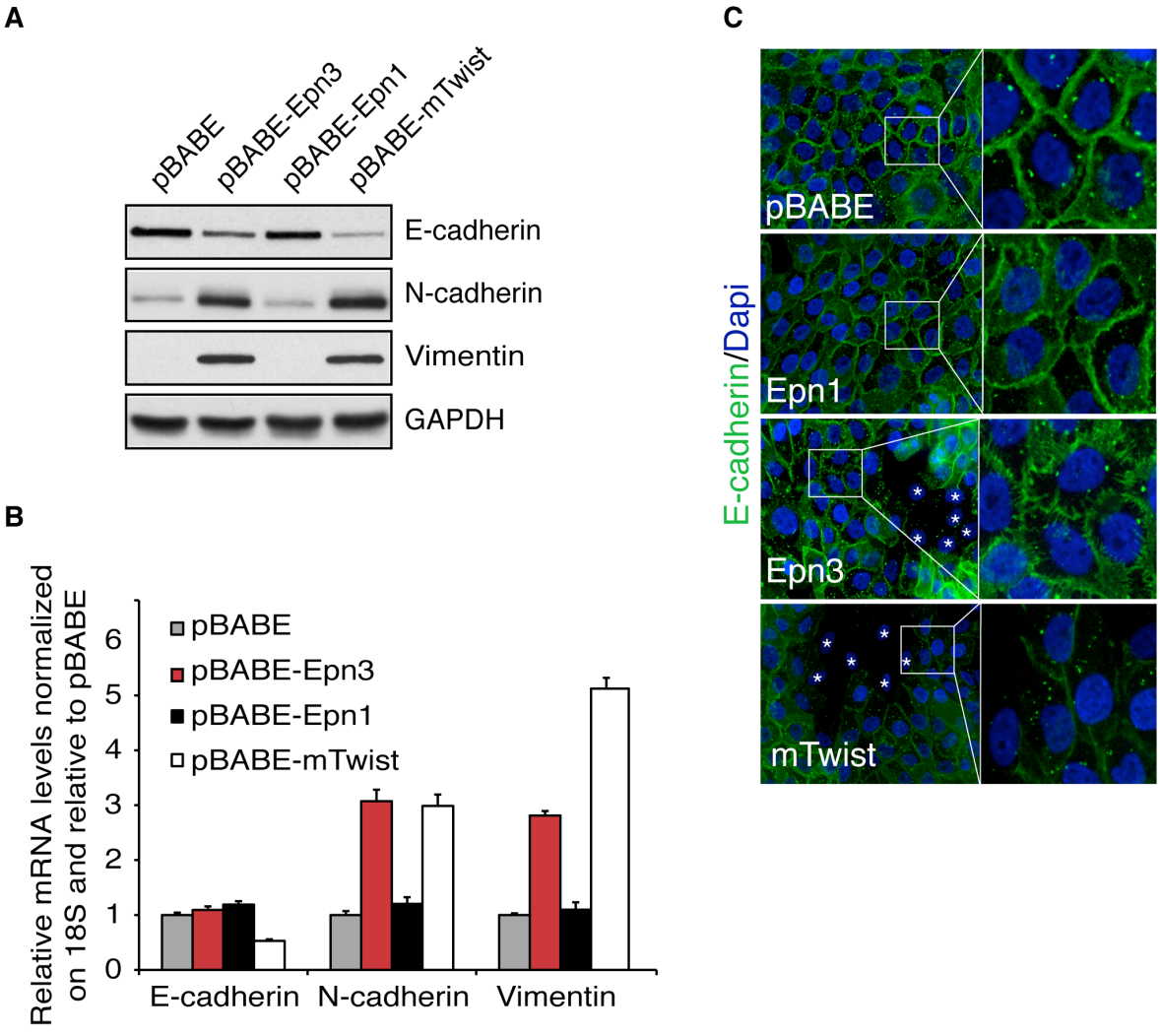


Figure 16 | Epn3 overexpression induces EMT phenotype in normal mammary epithelial MCF10A cells.

A) Expression of epithelial marker E-cadherin and the mesenchymal markers N-cadherin and vimentin was examined by WB in MCF10A cells expressing pBABE, pBABE-Epn1, pBABE-Epn3 or pBABE-mTwist. GAPDH was used as a loading control. B) Expression of epithelial E-cadherin, mesenchymal N-cadherin and Vimentin transcripts was examined by Q-PCR. Data were normalized to the housekeeping gene 18S and are expressed as fold-change relative to pBABE cells. Results represent the mean ± SD of an experiment performed in triplicate. C) Immunofluorescence staining of E-cadherin in MCF10A-infected cells cultured at confluency for 48 h. Green signal represents E-cadherin and blue signal represents nuclear staining by DAPI. White asterisks indicate cells lacking E-cadherin staining.

We next assessed E-cadherin expression at the cell-cell junctions by immunofluorescence (IF). We observed that E-cadherin expression was lost in a non-homogeneous way in both pBABE-Epn3 and pBABE-mTwist cells (Figure 16C). Moreover, upon Epn3 overexpression in the cells that still presented E-cadherin staining there was a marked rearrangement of E-cadherin at cell-cell junctions with the formation of comb-like structures with many radially oriented strands (Figure 16C), a phenotype that has been previously associated with a pre-EMT state [147]. On the contrary, E-cadherin appeared as a continuous belt between cells in pBABE and pBABE-Epn1 samples. These results suggest an alteration of cell-cell adherent junctions in pBABE-Epn3 cells that could be important for cell invasion and migration.

Since pBABE-Epn3 cells appear to have acquired characteristics of highly motile cells, we assessed their invasiveness. Cells were seeded on the top chamber of a transwell on a layer of Matrigel and cell invasion into this layer was assessed by crystal violet staining (see Material and Methods, section 6.3).

We observed that pBABE-Epn3 cells display ~50% increase of the capcacity of invading the Matrigel layer than pBABE control cells (Figure 17).

Taken together these data indicate that Epn3 overexpression in normal mammary epithelial MCF10A cells is able to induce an EMT-like phenotype characterized by changes in cell morphology, increased expression of mesenchymal markers, alteration of cell-cell junctions, and increased invasiveness.

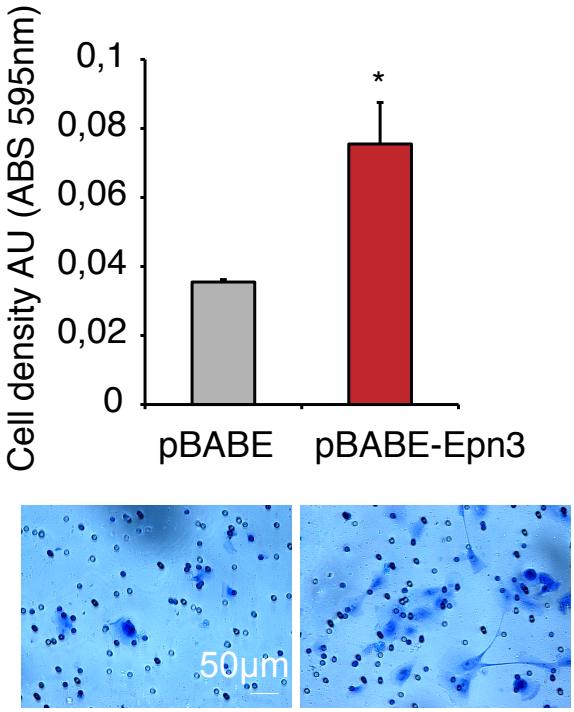


Figure 17 | Epn3 overexpression increases invasiveness of normal mammary epithelial MCF10A cells

Matrigel invasion assay of pBABE and pBABE-Epn3 MCF10A cells was performed by plating 40,000 cells in transwell chambers on a layer of growth factor reduced Matrigel. After 24 h of incubation cells that had invaded the Matrigel layer were stained with crystal violet. Membranes stained with crystal violet were then dissolved in acetic acid and absorbance was measured at 595nm. Bar graph reports mean absorbance at 595nm ± SD of an experiment performed in quadruplicate and is representative of two independent experiments. Asterisk indicates a p-value < 0.05. Representative images are shown below the graph. Magnification 20x. Scale bar 50 µm.

2.2.2 Epn3 overexpression causes an expansion of the stem cell compartment in normal breast epithelial MCF10A cells

Over the past few years, it has been demonstrated that cells that undergo EMT also acquire stem cell-like properties [148]. We decided to investigate this possibility in MCF10A Epn3-overexpressing cells. To do so we took advantage of a protocol that has been recently developed in our lab to isolate and identify mammary stem cells within a bulk population [136].

The approach we have used relies on two functional properties of mammary stem cells: 1) their relative quiescence and slow proliferation rate; 2) their capacity to survive and proliferate in an anchorage-independent manner, giving rise to clonal spheroids, named “mammospheres” [149-151]. The stem cells present in a bulk population are identified based on their ability to retain the lipophilic fluorescent dye PKH26, which is instead progressively lost by dilution in the highly proliferating differentiated/precursor cells [151]. In this experimental procedure, epithelial cells are labeled with PKH26 and plated in suspension culture to allow stem cells present in the cellular population to give rise to a clonal mammosphere. In most cases, each mammosphere is characterized by the presence of one PKH26-labeled stem cell located at the center of the mammosphere, while the existence of more than one PKH26-labeled cell suggests the presence of a cellular aggregate rather than a functional structure.

Taking advantage of this approach, we performed a mammosphere formation assay using MCF10A pBABE, pBABE-Epn3, pBABE-Epn1 and pBABE-mTwist cells. We observed that pBABE-Epn3 cells show 2-fold increase in the number of mammospheres formed *in vitro* than pBABE and pBABE-Epn1 cells (Figure 18A,B). Similar results were also obtained upon Twist overexpression that induces more than 2-fold increase of number of mammospheres (Figure 18). This

result agrees with previous studies in the literature demonstrating that Twist overexpression increases the mammosphere-forming efficiency of mammary epithelial cell lines [148].

Based on these data, we conclude that Epn3 overexpression is able to induce some stem cell-like properties like the capacity to form mammospheres *in vitro*. Together these results indicate that Epn3 is capable of inducing EMT and the EMT-associated stem-like behavior in normal epithelial cells. Epn3-induced EMT might therefore contribute to the observed increase in invasiveness and tumor aggressiveness in breast cancer through the expansion of the stem cell compartment.

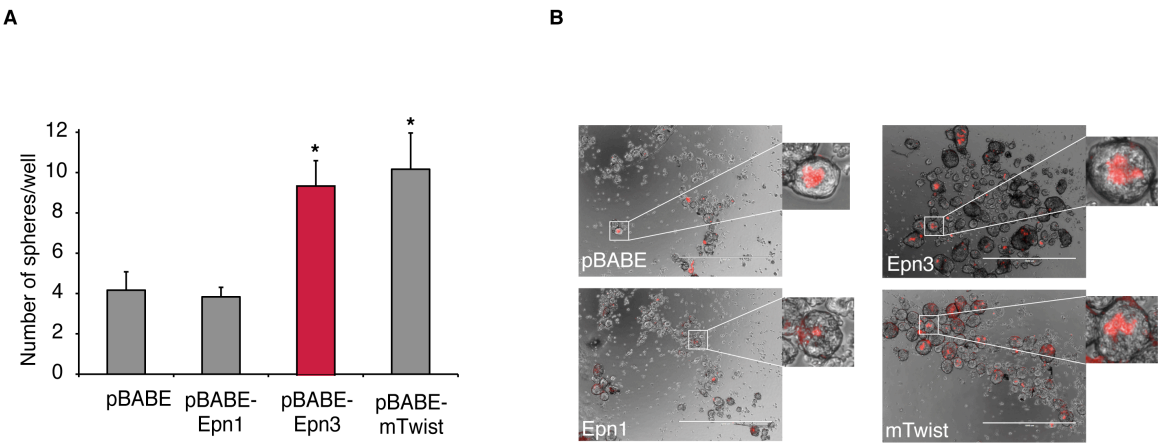


Figure 18 | Mammosphere formation assay of normal epithelial MCF10A cells

A) Mammosphere formation assay was performed with MCF10A pBABE, pBABE-Epn1, pBABE-Epn3 and pBABE-mTwist cells by trypsinizing cells and labeling them with PKH26 dye. Cells (4,000) were plated in suspension in polyHEMA-coated 6-multiwell plates in quadruplicates. Results are shown as the mean number of mammospheres per well ± SD of two independent experiments performed in quadruplicate. Asterisks indicate a p-value < 0.05. B) Representative bright-field microscopy images of mammospheres formed by MCF10A expressing either empty vector (pBABE), Epn1 Epn3 or mTwist. The boxed areas are magnified on the right and show a mammosphere derived from one PKH26-positive stem cell. Magnification 4x. Scale bar 100μm.

2.2.3 Epn3 overexpression induces EMT and increases the *in vivo* tumorigenic potential of breast tumor HCC1569 cells

We have found that Epn3 expression is associated with poor prognosis human breast tumors (Section 1.1 and section 1.2) and that Epn3 ablation in breast tumor BT474 cells, carrying *EPN3* gene amplification, is able to impair tumorigenic potential both *in vitro* and *in vivo* (Section 2.1.1 and 2.1.2).

To investigate if the overexpression of Epn3 in normal mammary epithelial MCF10A cells could induce a tumorigenic phenotype, we took advantage of the engineered MCF10A cells overexpressing Epn3 (pBABE-Epn3). We tested the capacity of these cells to grow in an anchorage-independent manner - a hallmark of cell transformation - by plating cells in a semi-solid medium. Our results showed that none of the cell lines tested (pBABE, pBABE-Epn1, pBABE-Epn3, pBABE-mTwist) formed colonies (Figure 19). On the contrary, the cell line used as positive control of the assay, MCF10A overexpressing the activated form of HER2 (Neu-T), was able to form colonies in semi-solid medium as described in the literature [152] (Figure 19).

Given this result, we decided to move towards a tumorigenic breast cell line that could be more prone to transformation. We chose as a model system the breast tumor cell line HCC1569 since this cell line shows low basal levels of Epn3, similar to the ones observed in MCF10A cells (Figure 20A). HCC1569 cells were infected with a lentiviral vector expressing Epn3 (pLVX-Epn3) and reached levels of Epn3 overexpression comparable to those of MCF10A pBABE-Epn3 cells (Figure 20A). HCC1569 cells infected with empty vector (pLVX) were used as control.

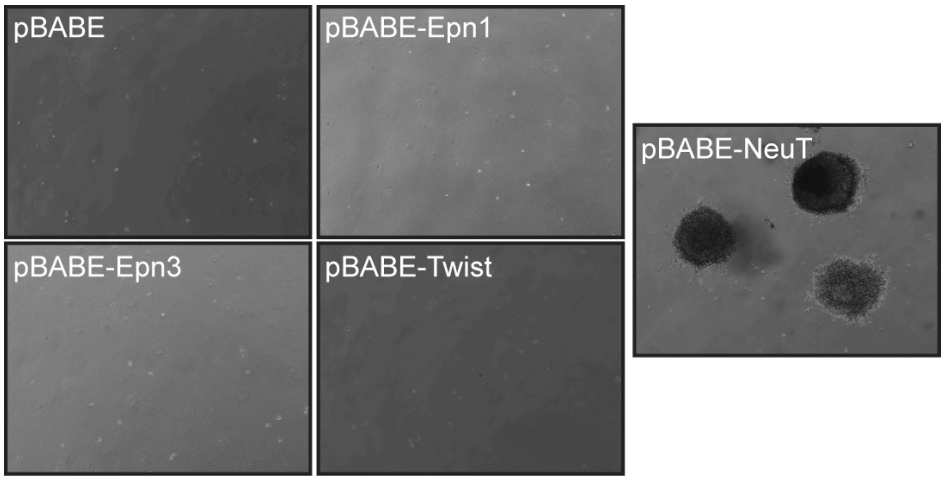


Figure 19 | Effect of Epn3 overexpression on anchorage-independent growth of normal mammary MCF10A cells

The anchorage-independent growth of MCF10A-infected cells was assessed using the soft agar assay. Cells (20,000) were plated in 6-well culture plates in a medium containing 0.3% agar. After 3 weeks in culture, the formation of colonies was evaluated by bright-field microscopy. Representative bright-field microscopy images are shown. Magnification 10x. MCF10A cells overexpressing Neu-T were used as positive control.

By performing WB analysis, we confirmed that also in the case of HCC1569 cells, Epn3 overexpression caused upregulation of the mesenchymal markers N-cadherin and vimentin (Figure 20B). Besides changes in mesenchymal markers expression, HCC1569 cells underwent a dramatic change in cell morphology upon Epn3 overexpression: while control cells formed very compact colonies, Epn3-overexpressing cells became more elongated and formed very disorganized colonies (Figure 20C), thus recapitulating the EMT phenotype already observed in MCF10A overexpressing Epn3.

Based on these *in vitro* observations, we investigated the tumorigenic potential of HCC1569 cells *in vivo*. $1*10^6$ HCC1569 pLVX-Epn3 and control cells were orthotopically injected into the inguinal mammary fat pad of NOD/SCID female immunocompromised mice. As a control, cells (HCC1569 pLVX) were injected into the contralateral mammary gland. Remarkably, Epn3-overexpressing cells formed tumor masses with increased growth rate that were ~2-fold bigger than the ones originated by control cells (Figure 21A,B). To assess the Epn3 expression levels on the tumors extracted from the mammary gland, we dissociated them into cells and performed WB analysis. We observed that the Epn3 overexpression levels similar to those displayed by the parental HCC1569 cells *in vitro* (Figure 21C).

Thus, we conclude that Epn3 overexpression in breast tumor cells HCC1569 is able to induce EMT and to increase their tumorigenic potential *in vivo*.

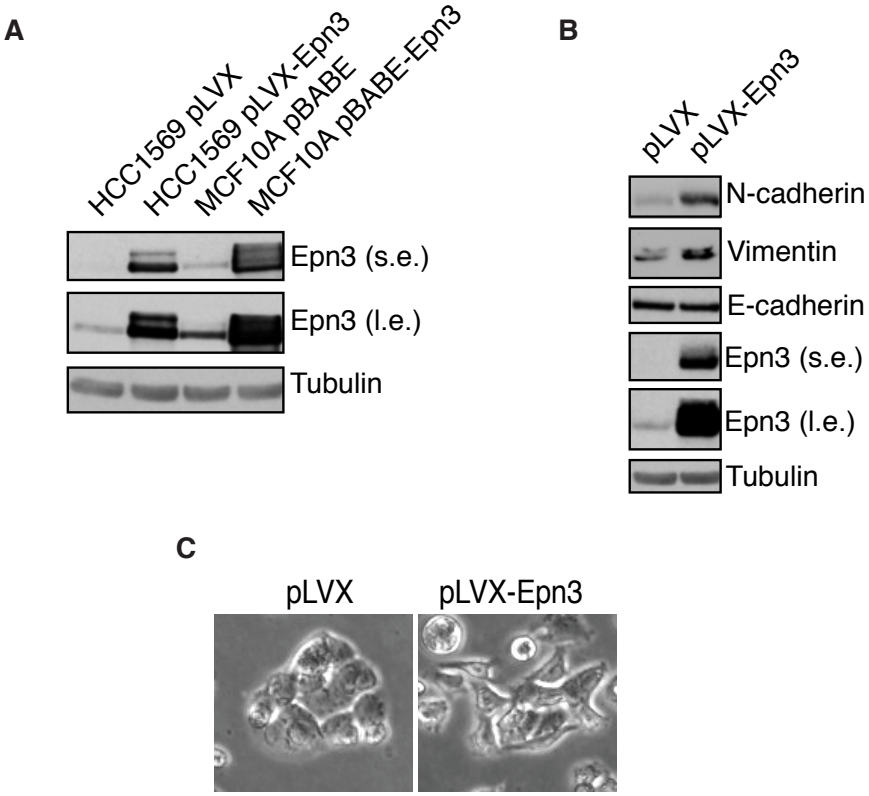


Figure 20 | *In vitro* functional characterization of Epn3 overexpression in HCC1569 cells

HCC1569 cells were infected with lentiviral vector expressing Epn3 (pLVX-Epn3) and empty vector as control (pLVX). MCF10A cells were infected with retroviral vector pBABE (empty vector) and pBABE-Epn3. A) Epn3 expression was evaluated by WB analysis in HCC1569 after lentiviral infection and was compared with the expression of Epn3 in MCF10A pBABE-Epn3. Epn3 was detected using VI31 antibody. Both short and long exposure blots are shown (s.e. and l.e.,respectively). Tubulin was used as loading control. B) The expression of N-cadherin, Vimentin, E-cadherin and Epn3 in HCC1569-infected cells was assessed by WB analysis. C) The morphology of HCC1569 cells expressing control vector pLVX or pLVX-Epn3 was assessed by phase contrast microscopy. Representative images are shown. Magnification 20x.

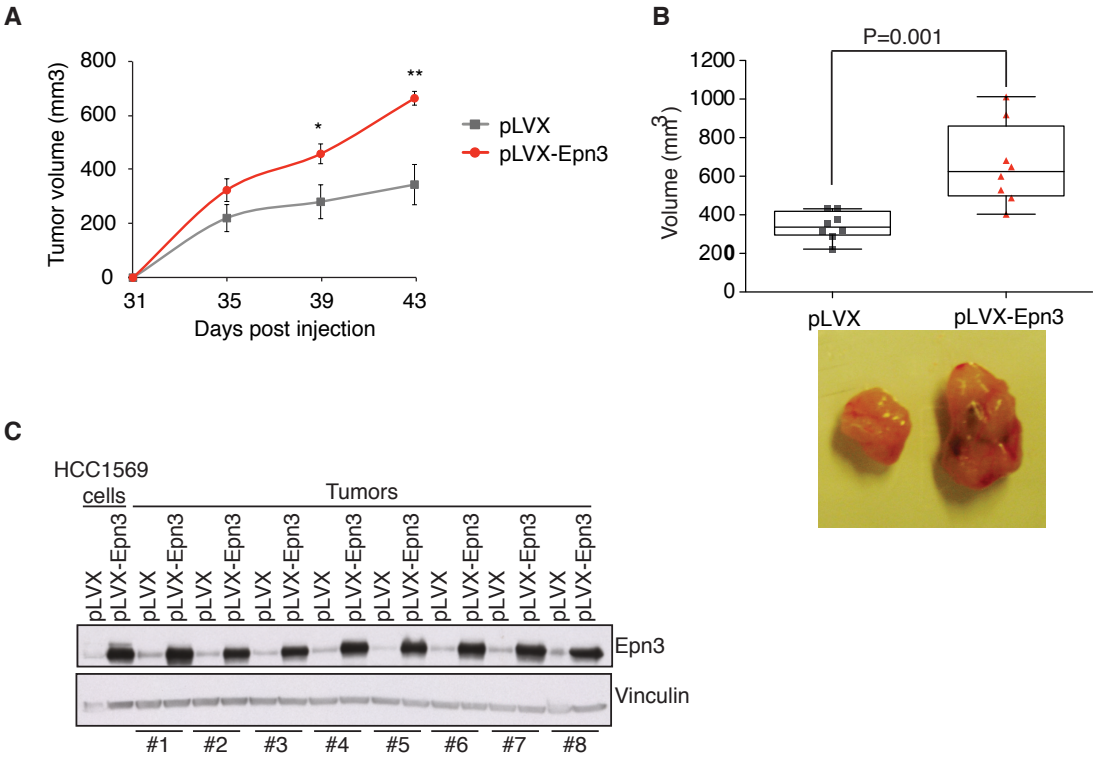


Figure 21 | Effect of Epn3 overexpression on *in vivo* tumor growth *of* HCC1569 cells

Mammary tumors were generated by injecting HCC1569 Epn3-overexpressing (pLVX-Epn3) cells and control (pLVX) cells ($1*10^6$) into the inguinal mammary fat pads of female NOD/SCID IL2 gamma-chain null mice. Tumors were grown for 43 days before being explanted. A) Kinetic evaluation of HCC1569 pLVX-Epn3 and pLVX tumor growth. Tumor volume was assessed by i*n vivo* caliper measurements at the indicated time points after injection. Asterisks indicate mean p-value < 0.05 (*) and < 0.01 (**) compared to pLVX. B) The distribution of tumor volume of HCC1569 pLVX-Epn3 and pLVX is reported in a box-plot. Representative images of tumor outgrowths are reported below the chart. C) The expression of Epn3 in mammary tumors was assessed by WB analysis using an anti-Epn3 VI31 antibody. Vinculin was used as loading control. Results are representative of 2 experiments (n=8) and represent the mean± SEM.

3 Dissection of the molecular mechanism responsible for Epn3-induced EMT

3.1 Epn3 overexpression increases basal and TGFβ-induced E-cadherin internalization in MCF10A normal breast epithelial cells

We have found that Epn3 overexpression in MCF10A is able to induce an EMT phenotype characterized by a decrease in E-cadherin protein levels and the rearrangement of cell-cell junctions (see section 2.1).

E-cadherin is the prototypical member of the classical cadherin superfamily and it represents the major component of the adherent junctions (AJs), providing cell-cell adhesion through Ca^{2+}-dependent, homophilic binding between molecules on adjacent epithelial cells (reviewed in [153]). Loss of E-cadherin expression is considered a prototypical marker of EMT and a strong rearrangement of E-cadherin-mediated cellular junctions occurs during tumor invasion and metastasis [154]. Multiple mechanisms have been described to underlie the loss of E-cadherin function and they comprise both transcriptional and post-translational regulation of E-cadherin. Besides E-cadherin regulation at the genomic level (reviewed in [155]), another mechanism known to disrupt E-cadherin's adhesive function and promote tumor aggressiveness is the endocytic trafficking with consequent downmodulation of surface E-cadherin (reviewed in [156]).

Given the established role of the epsin protein family in endocytosis, we tested the hypothesis that Epn3 could have a role in E-cadherin internalization, thus altering its turnover from the cell surface. To this aim, we set up an E-cadherin internalization assay *in vivo* using an antibody recognizing the extracellular domain of E-cadherin (see Material and Methods for details). Briefly, the antibody is allowed to bind to cells for 1h at 4°C, a temperature at which

endocytosis is inhibited. When cells are transferred at 37°C, E-cadherin internalization can occur and the effects of Epn3 overexpression on E-cadherin endocytosis assessed.

Many different pathways have been demonstrated to regulate E-cadherin internalization during EMT. One of those, the TGFβ pathway, appears to be a key regulator of cadherin turnover during EMT and to induce E-cadherin endocytosis [157]. To test whether Epn3 could potentiate TGFβ-dependent E-cadherin endocytosis, we performed the E-cadherin internalization assay in MCF10A pBABE and pBABE-Epn3 cells stimulated with TGFβ for 90 and 180 minutes. In control cells, we noticed that basal E-cadherin turn-over (in absence of any stimuli) was very slow and that no significant downregulation from the cell surface occurred (Figure 22A, left panel). On the contrary, in non-stimulated Epn3-overexpressing cells, E-cadherin was partially downregulated at time zero, an observation compatible with the anti-E-cadherin staining previously observed at the cell-cell junctions (Figure 16C), suggesting that basal E-cadherin turnover is accelerated upon Epn3 overexpression (Figure 22A, right panel). After the temperature shift from 4°C to 37°C, control cells stimulated with TGFβ for 90 minutes showed E-cadherin staining at the PM (Figure 22A, left panel, middle lane); after 180 minutes of TGFβ stimulation, we observed a partial internalization of E-cadherin in control cells (Figure 22A, left panel, bottom lane). In Epn3-overexpressing cells, E-cadherin was efficiently internalized already at 90 minutes of TGFβ stimulation and, after 180 minutes, all cells accumulated E-cadherin in the intracellular compartments without any residual PM staining (Figure 22A, right panel). Thus, our data suggest that Epn3 overexpression increases TGFβ-induced E-cadherin internalization in MCF10A cells.

We also performed co-localization studies of E-cadherin and Epn3 in MCF10A cells and the results are reported in Figure 22B. In a steady state condition (time 0, non-stimulated), both E-cadherin and Epn3 were localized at the cell surface and even though they did not co-localize entirely, Epn3 seems to be juxtaposed to E-cadherin. This observation is compatible with the fact that Epn3 has a cytoplasmic epitope while E-cadherin presents an extracellular one (Figure 22B, first lane). Importantly, after 90 and 180 minutes of TGFβ stimulation, both Epn3 and E-cadherin were delocalized from the PM to intracellular compartments with many sites of co-localization (Figure 22B, see yellow dots).

In conclusion, our results show that overexpression of Epn3 positively regulates TGFβ-dependent E-cadherin internalization and that Epn3 co-traffics with E-cadherin along the endocytic route. These observations suggest that Epn3 promotes EMT and tumor aggressiveness via its role as an endocytic protein on E-cadherin trafficking.

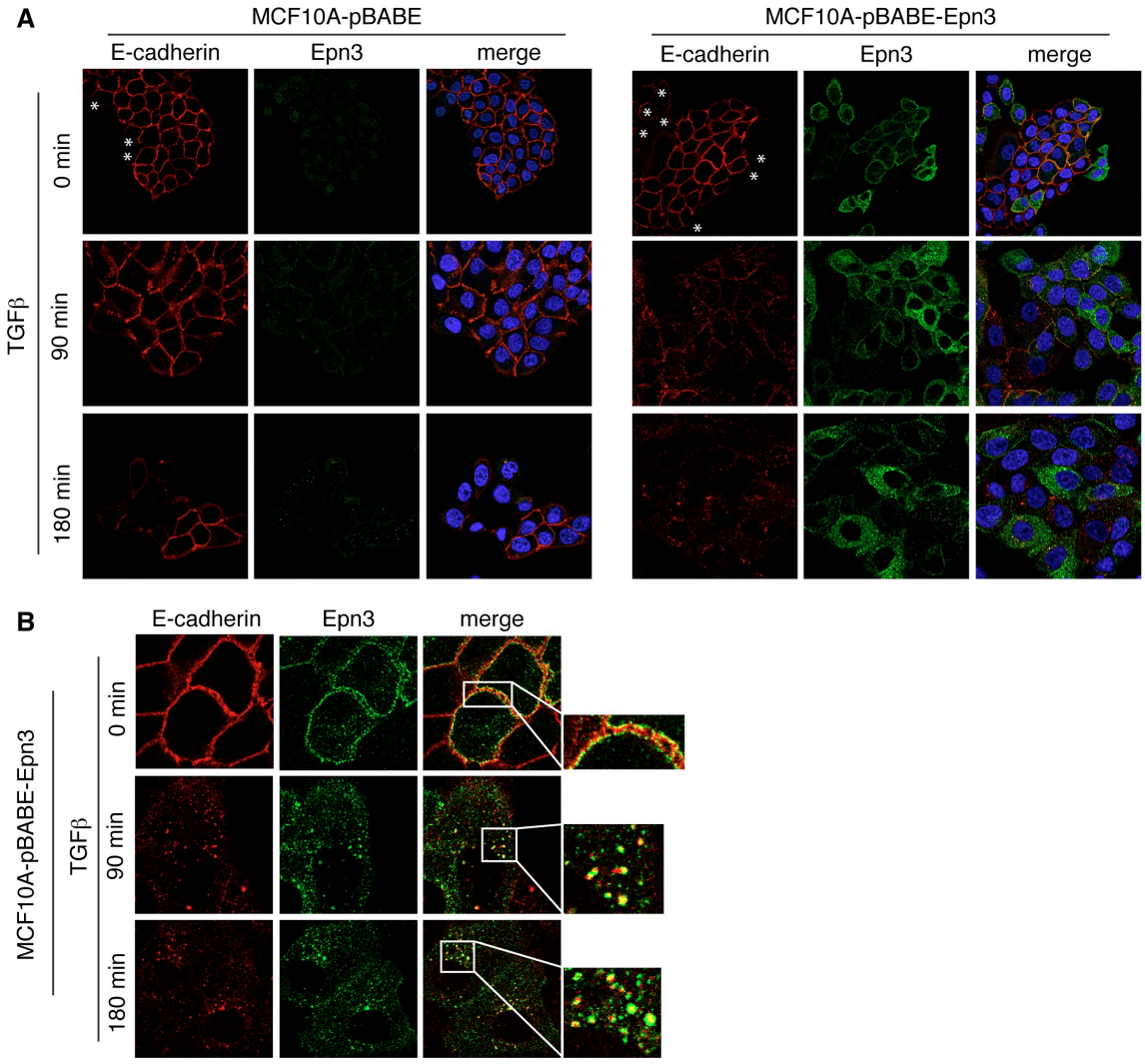


Figure 22 | Effect of Epn3 overexpression on TGFβ-induced E-cadherin internalization in normal mammary epithelial MCF10A cells

For the E-cadherin internalization assay, cells were incubated with anti-E-cadherin antibody (HECD-1, Abcam) at 4°C in PBS for 1h and then fixed in 4% PFA (0 min). For other samples, immediately after incubation at 4°C, medium containing serum and TGFβ (5ng/mL) was added to the cells at 37°C for 90 and 180 minutes (90 min and 180 min). After incubation with TGFβ, cells were fixed in 4% PFA. A) Immunofluorescence analysis of E-cadherin (red) and Epn3 expression (green). Nuclei are counterstained with DAPI (blue). White asterisks indicate cells lacking E-cadherin staining.

B) Co-localization of Epn3 and E-cadherin was evaluated by immunofluorescence analysis during E-cadherin internalization assay described above. The boxed areas are magnified on the right to better observe co-localization sites (yellow dots).

3.2 Epn3 overexpression increases TGFβ signaling and response in normal breast epithelial MCF10A cells

The observation that Epn3 potentiates TGFβ-dependent trafficking of E-cadherin, made us wonder if Epn3 overexpression could also directly regulate TGFβ signaling pathway and/or cooperate with TGFβ signaling in the establishment of EMT. We reasoned that, since TGFβ pathway is frequently subverted in cancer and is known to promote tumor progression by inducing EMT in mammary epithelial cells, this pathway could be involved in Epn3-dependent EMT (reviewed in [158]). In agreement with this possibility, we found that Epn3-overexpressing cells display an increase in the transcript levels of TGFβ receptors (TGFBR1 and TGFBR2) and ligands (TGFB1 and TGFB2) are significantly increased compared to control cells, as assessed by Q-PCR analysis (Figure 23) Moreover, TGFB2 mRNA levels further increase after stimulation with TGFβ (Figure 23). These observations unveil the existence of a transcriptional positive feedback loop required to further sustain the EMT phenotype of Epn3-overexpressing cells.

We thus investigated a possible cooperation between Epn3 and the TGFβ pathway. The TGFβ signaling pathway is well characterized and involves the recruitment and phosphorylation of serine-threonine kinase receptors, which in turn induce the phosphorylation of Smad proteins (Smad2 and Smad3). Once activated, Smad proteins are able to form multiprotein complexes and induce the transcription of target genes that are usually considered as EMT markers, including the transcription factor Snail and N-cadherin (reviewed in Massague 2005). To test our hypothesis, we stimulated Epn3-overexpressing and control MCF10A cells with TGFβ for extended times (30'-90'-6h-12h-24h) in order to follow activation of "early" signaling targets (i.e. phosphorylation of Smads) as well as "late" transcriptional response (i.e. Snail, N-cadherin), by WB and Q-PCR

analysis (Figure 24A,B). The results indicated that in Epn3-overexpressing cells there was a sustained activation of TGFβ signaling, as evidenced by the presence of a second peak phosphorylation of Smad proteins (Figure 24A) and a more sustained activation of the downstream target Snail (Figure 24A). Furthermore, pBABE-Epn3 cells, which express high basal levels of N-cadherin, showed a significantly higher activation of N-cadherin after TGFβ stimulation (Figure 24A-B). In contrast, in control cells the activation of N-cadherin expression is only slightly detectable at very late time points after TGFβ stimulation (Figure 24A,B). These data indicate that Epn3 is able to sustain TGFβ signaling and increase TGFβ-dependent transcriptional programs responsible for the establishment of EMT. Indeed, the combination of TGFβ stimulation and Epn3 overexpression further enhanced the invasion potential of MCF10A cells in matrigel invasion assay (Figure 25).

Altogether these data suggest that Epn3 is able to increase the expression of TGFβ signaling components that could act by enhancing and sustaining a mesenchymal state and invasiveness.

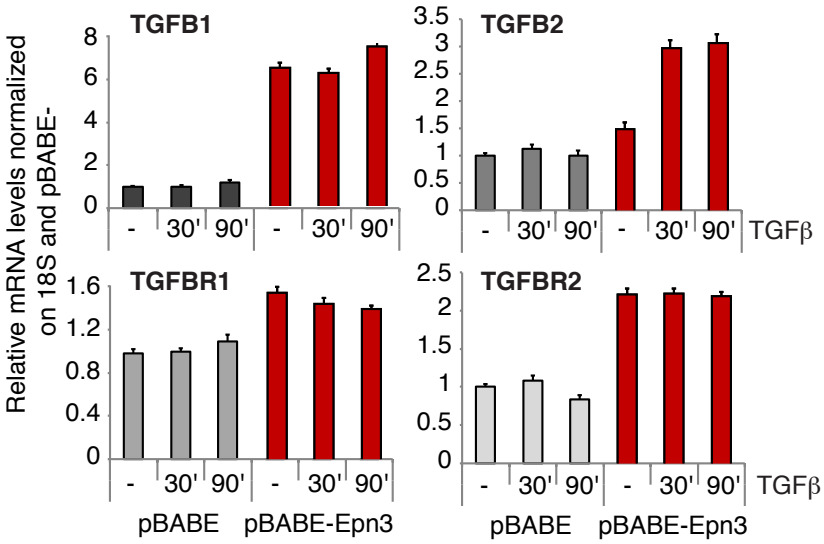


Figure 23 | Analysis of TGFβ receptor and ligands mRNA levels in MCF10A cells

MCF10A pBABE and pBABE-Epn3 cells were stimulated with TGFβ (0.3 ng/mL) for 30 and 90 minutes or cultured without stimuli (-). Transcript levels of TGFβ receptor type I (TGFBR1), type II (TGFBR2) and TGFβ ligands (TGFB1 and TGFB2) were evaluated by Q-PCR analysis. Data are normalized on the housekeeping gene 18S and are expressed as fold-change relative to the non-stimulated MCF10A pBABE cells (-). Results represent the mean ± SD of two experiments performed in triplicate.

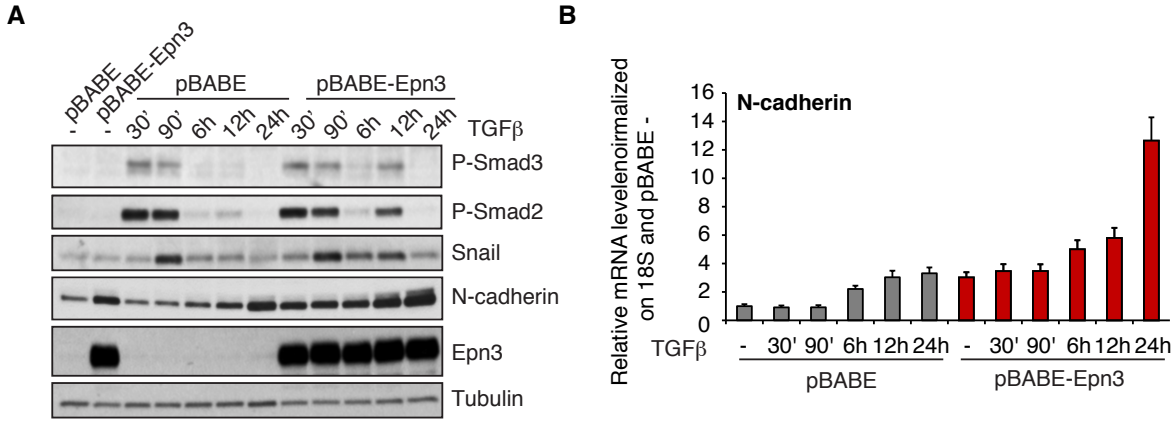


Figure 24 | Analysis of the effect of Epn3 overexpression on the TGFβ pathway

A) MCF10A pBABE and pBABE-Epn3 cells were stimulated with TGFβ (0.3 ng/mL) for the indicated time points. Expression of phosphorylated Smad2 (P-Smad2), phosphorylated Smad3 (P-Smad3), Snail, N-cadherin and Epn3 was evaluated by WB analysis. Tubulin was used as a loading control. B) N-cadherin

transcript levels in MCF10A pBABE and pBABE-Epn3 cells after stimulation with TGFβ (0.3ng/mL) were measured by Q-PCR. Data were normalized to the housekeeping gene 18S and are expressed as fold-change relative to the pBABE control cells non-stimulated with TGFβ (–). Results represent the mean ± SD of two experiments performed in triplicate.

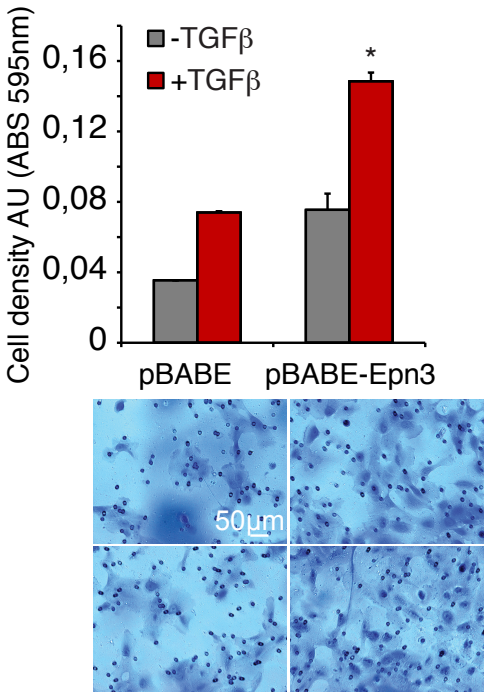


Figure 25 | Epn3 overexpression increases TGFβ–induced invasion potential of normal mammary epithelial MCF10A cells

Matrigel invasion assay was performed with pBABE and pBABE-Epn3 cells by plating 4,000 cells in the top chamber of a transwell on a layer of growth-factor reduced Matrigel. The bottom chamber contained culture medium supplemented with TGFβ (0.3ng/mL). After 24 h of incubation cells that had invaded into the Matrigel layer were stained with crystal violet. Membranes stained with crystal violet were then dissolved in acetic acid and absorbance was measured at 595nm. Bar graph reports mean absorbance at 595nm ± SD of 2 experiments performed in triplicate. Asterisk indicates a p-value < 0.05 (*). Representative images are shown below the graph. Magnification 20x.

3.3 Blocking of TGFβ signaling reverts Epn3-induced EMT phenotype in normal mammary epithelial MCF10A cells

Given the observations above described, we tested whether the Epn3-induced EMT phenotype is indeed dependent of TGFβ signaling. To this aim, we used shRNA oligos to transiently silence both TGFβ receptors (TGFβ type I and TGFβ type II receptors) in MCF10A pBABE-Epn3 and control cells. A scramble shRNA oligo was used as negative control. As assessed by Q-PCR analysis, ablation of these two receptors was efficient when KD was done singularly (TGFBR1 KD and TGFBR2 KD) and when it was performed in combination (TGFBR1+2 KD; Figure 26A). After ablation of TGFβ receptors, we observed an almost complete reversion of EMT phenotype in MCF10A Epn3-overexpressing cells. Indeed, N-cadherin and vimentin expression was reduced upon TGFβ receptors silencing, as assessed by WB and Q-PCR analysis (Figure 26B,C). Moreover, ablation of TGFβ receptors also caused a complete reversion of the morphology of Epn3-overexpressing cells from a fibroblast-like shape into an epithelial and well-rounded one (Figure 26D).

Nevertheless, to exclude potential off-target effects of transient silencing experiments, we decided to confirm these results by using a chemical inhibitor of TGFβ signaling. We used the inhibitor LY2109761 that selectively blocks the kinase activity of TGFβ receptor type I and type II, thus inhibiting Smad proteins phosphorylation [159, 160].

Our results show that treatment of MCF10A Epn3-overexpressing cells with LY2109761 for 72h was able to decrease the expression of N-cadherin and vimentin at basal levels in comparison with control cells (Figure 27A,C). We also stimulated cells with TGFβ for 90 minutes after treating cells with the inhibitor and then evaluated the activation of Snail protein. We confirmed that the inhibition

induced by LY2109761 was effective since we did not observe any activation in Snail expression after TGFβ stimulation (Figure 27B).

In conclusion, through both a genetic and a chemical approach, we have demonstrated that the blockage of TGFβ signaling reverts the EMT phenotype induced by Epn3 in MCF10A cells.

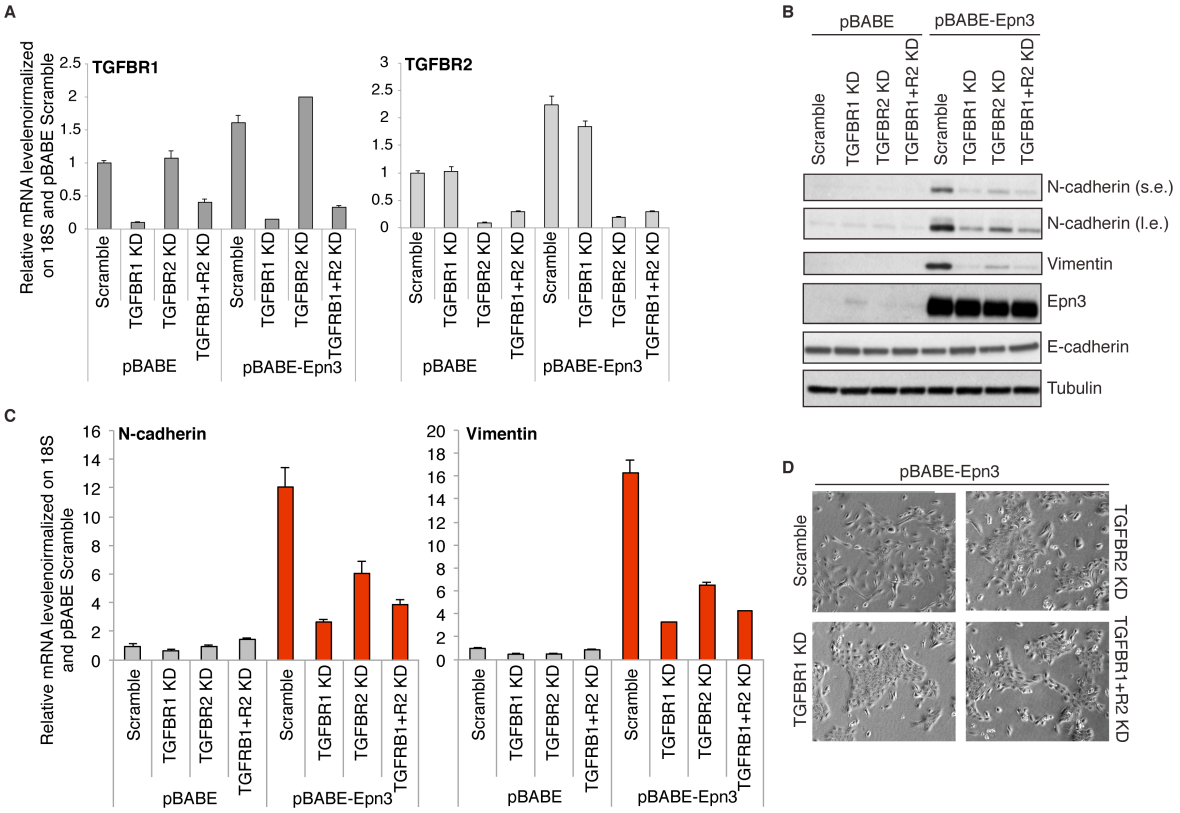


Figure 26 | Analysis of the effect of TGFβ receptor silencing in MCF10A Epn3-ovexpressing cells

MCF10A pBABE and pBABE-Epn3 cells were transiently transfected with shRNA targeting TGFβ receptor type I (TGFBR1 KD) or type II (TGFBR2 KD). Transient transfection of both shRNA was also performed (TGFBR1+R2 KD). A shRNA with a scramble oligo was used as negative control (Scramble). A) The efficiency of the silencing of TGFβ receptor type I and II was assessed by Q-PCR analysis. Data are normalized on the housekeeping gene 18S and are expressed as fold-change relative to the pBABE cells transfected with the scramble oligo . Results represent the mean ± SD of two experiments performed in triplicate. B) The expression of N-cadherin, vimentin, E-cadherin and Epn3 was evaluated by WB analysis using the following antibodies. Tubulin was used as loading control.

C) N-cadherin and vimentin transcript levels were evaluated by Q-PCR analysis. Data are normalized on the housekeeping gene 18S and are expressed as fold-change relative to the pBABE scramble cells. Results represent the mean ± SD of two experiments performed in triplicate. D) The morphology of MCF10A pBABE-Epn3 cells transfected with scramble oligo or oligos targeting TGFβ receptor type I (TGFBR1 KD) or type II (TGFBR2 KD), or both TGFβ receptor type I and type II (TGFBR1+R2 KD) was assessed by phase contrast microscopy. Representative images are shown. Magnification 4x.

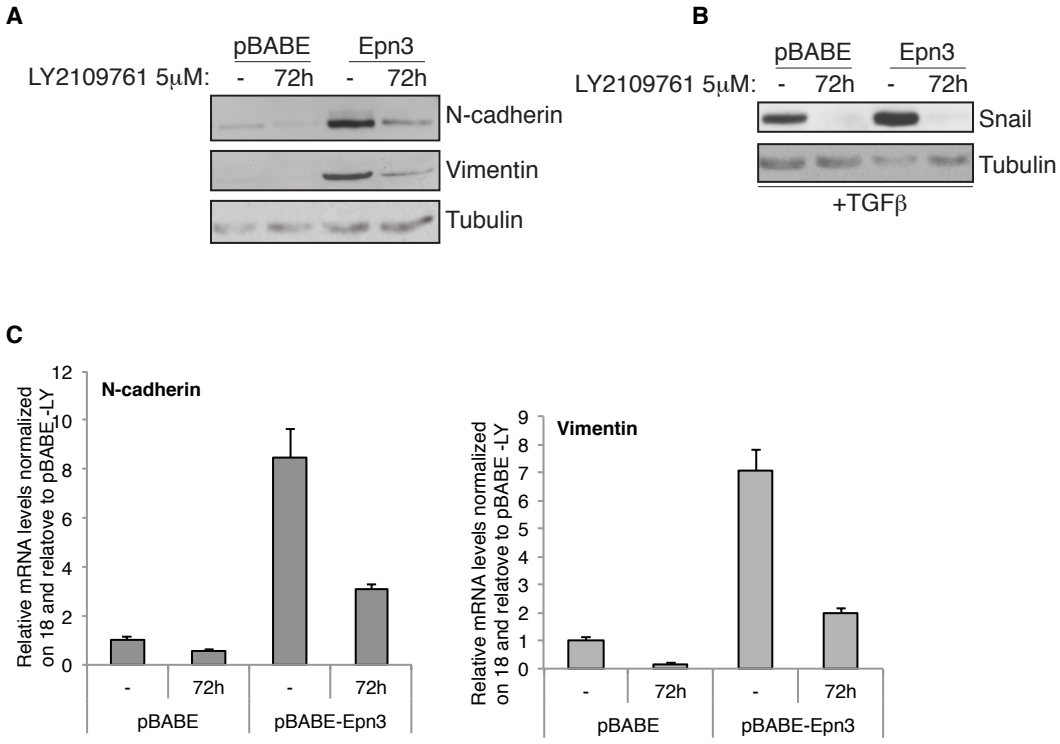


Figure 27 | Analysis of the effect of TGFβ receptors inhibitors in MCF10A Epn3-ovexpressing cells

MCF10A pBABE and pBABE-Epn3 cells were treated with TGFβ receptor kinase inhibitor LY2109761 (5μM) for 72h. DMSO treatment was used as negative control (-). A) The expression of N-cadherin and vimentin was assessed by WB analysis. Tubulin was used as loading control. B) The effect of LY2109761 was evaluated by stimulating MCF10A pBABE and pBABE-Epn3 cells with TGFβ (0,3ng/mL) for 90 minutes. The expression of Snail protein was assessed by WB analysis. Tubulin was used as loading control. C) N-cadherin and vimentin transcript levels were evaluated by Q-PCR analysis. Data are normalized on the housekeeping gene 18S and are expressed as fold-change relative to the non-treated pBABE cells (-). Results represent the mean ± SD of two experiments performed in triplicate.

DISCUSSION

1 Epn3 as novel prognostic marker in breast cancer

1.1 Epn3 expression is associated with poor prognosis in breast cancer

The present clinical management of breast cancer patients relies on the availability of robust prognostic and predictive factors, which are used to stratify patients and ensure they receive optimal treatment. In the last decade, breast cancer research was dominated by the introduction of gene expression profiling techniques aimed at the identification of specific molecular signatures that can predict tumor behavior and response to therapy (reviewed in [122].

However, many prognostic and predictive gene expression profiles that have been developed (see Introduction, section 3.3) have found poor clinical application because of difficulties in the standardization of procedures, high costs and little improvement over the traditional breast cancer markers. One other factor contributing to the slow introduction of these signatures in the clinic might also be the insufficient knowledge on the biological function of some of the genes included in these profiles. This lack of knowledge makes unclear which clinical characteristics are these signatures measuring. A good option to overcome this limitation would be, thus, to select few novel biomarkers within a molecular signature that retain the prognostic power of the signature and to assess the expression of these biomarkers in large patient cohorts using rapid, standardized and low cost tests, such as IHC assays, which are routinely used in the clinical setting to assess the expression of the molecular markers currently applied to the stratification of cancer patients (e.g., HER2, ER and PgR).

In a preliminary unpublished study performed in our lab, a gene expression profiling analysis of node-negative human breast tumor samples originated a gene signature associated with metastatic relapse in breast cancer patients, which included the endocytic protein Epn3. This first observation suggested that Epn3 transcript levels could be considered a prognostic factor associated with high risk of disease recurrence and metastasis.

To clearly establish Epn3 as prognostic biomarker in breast cancer, we analyzed the expression of Epn3 by IHC on large cohorts of breast cancer patients. We found that high Epn3 expression is significantly associated with traditional markers of aggressive breast cancer. Furthermore, regression analysis revealed that Epn3 expression could predict the risk of disease recurrence and death up to 10 years after surgery. Interestingly, these results were confirmed also in the subpopulation of HER2-negative patients, indicating that the role of Epn3 in breast cancer might be independent of HER2.

The *HER2* gene is amplified and leads to protein overexpression in approximately 18-30% of human breast tumors [138, 161]. Many studies have identified *HER2* amplification and protein overexpression as a poor clinical outcome biomarker in breast cancer patients [124, 138]. Targeted therapies have been developed to specifically treat breast cancer patients carrying *HER2* amplification. Trastuzumab (Herceptin) is a humanized recombinant antibody against HER2 and it represents the first choice therapy in HER2-positive breast cancer cases [162]. However, a large portion of metastatic HER2-positive breast cancer patients has acquired resistance to trastuzumab [163]. Since we have found that Epn3 overexpression can correlate also with HER2-positive status, Epn3 might represent an additional target to overcome the resistance phenomenon in these patients carrying both Epn3 and HER2 overexpression. On

the other hand, our findings showing that Epn3 can retain its prognostic value independently of HER2 suggest Epn3 as a possible marker to be targeted in HER2-negative patients.

Additionally, we have found that Epn3 expression is significantly associated with poor prognosis Luminal B and triple-negative tumors. Both these molecular tumor types are of particular interest to the breast cancer research community since no systematic targeted therapies have been yet developed for these cancer subtypes. In the case of triple-negative breast cancer, poor prognosis results from the absence of expression of all three key receptors deregulated in breast cancer - ER, PgR and HER2 - and, thus, from the impossibility to develop targeted therapies [111]. Luminal B cancers are ER-positive tumors with aggressive behavior, high proliferative index and poor clinical outcome [111]. Despite the expression of estrogen receptors, the driver pathway of malignancy in the Luminal B subtype does not seem to be the estrogen pathway. Indeed, recent clinical trial studies have considered targeting alternative pathways, including EGFR-dependent pathway with the drug gefitinib [164], and PI3K/Akt/mTOR with the drug everolimus [165]. Nevertheless, many efforts are needed to find new biomarkers to define new-targeted therapies for luminal B breast cancer patients.

In conclusion, we have clearly established Epn3 as a novel biomarker in breast cancer, able to stratify breast cancer patients with poor clinical outcome. Since we have set up the experimental conditions to exploit the monoclonal anti-Epn3 antibody, previously produced in our lab, to detect Epn3 expression by IHC, one possibility could be to develop clinical test to assess Epn3 expression in breast cancer patients.

1.2 *EPN3* gene is amplified in human breast tumors: "it can live with or without *HER2*"

Gene amplification is one of the most frequent mechanisms of oncogene activation in breast cancer. Of note, *EPN3* gene is located on chromosome 17 in the 17q21 region, relatively close to the *HER2* gene that is located in the 17q12 region. It is known that the amplification of the 17q12-q21 chromosomal region is the most common mechanism for HER2 activation in breast cancer and that it leads to the simultaneous activation of several other genes besides HER2 [140, 141]. Since co-amplified and co-activated genes may have an impact on disease progression and clinical behavior of HER2-positive tumors, several studies have sequenced the commonly amplified region surrounding HER2 gene. These studies delineated a minimal common region of about 300kb, the HER2 amplicon 'core', that includes a set of co-amplified and overexpressed transcripts from the 17q12-21 region [140, 141]. Some of these genes have been investigated for their oncogenic potential in breast cancer. For instance, knockdown of two genes included in the amplicon 'core', GRB7 and STARD3, decreased cell proliferation and cell-cycle progression of breast tumor cell lines harboring *HER2* amplification, suggesting that their amplification and overexpression may contribute to tumorigenic phenotype of these cells [166].

To this regard, it is worth mentioning that the *EPN3* gene is located about 12Mb away from the minimal chromosomal region surrounding *HER2* and it is not among the genes of the 17q12-21 amplicon frequently amplified with *HER2*. Of note, *Curtis et al* recently analyzed copy number, sequence changes and gene-transcription rates in approximately 2,000 breast cancers encompassing all known types. They defined some regions of sequence amplification or deletion that deregulate genes that are likely to be involved in the pathophysiology of breast

cancer [142]. Interestingly, *EPN3* is among those genes away from *HER2* amplicon that are amplified independently from HER2 and that could represent novel 'driver' cancer genes.

We have found that the *EPN3* gene is amplified in breast tumors. By performing FISH analysis on 281 human breast tumor samples, we found that *EPN3* is amplified in about 18% of breast tumors analyzed. Importantly, even though *EPN3* gene amplification can correlate with *HER2* gene amplification, about 50% of the *EPN3* amplified cases did not display *HER2* co-amplification.

Our results are in line with *EPN3* amplification data made available by the Tumor Cancer Genome Atlas (TCGA) consortium that confirmed *EPN3* gene amplification in almost 7% of all breast cancer patients, with 50% of the cases showing co-amplification of *HER2* gene (Comprehensive molecular portraits of human breast tumors, Nature 2012) (Figure 28). The slight difference between TCGA data on *EPN3* amplification and our results (7% in TGCA *vs.* 18% in our experiments) might be due to the different techniques used to detect gene amplification (Comparative Genome Hybridization-CGH in the TGCA *vs.* FISH in our experiments). Nevertheless, in both studies *EPN3* is co-amplified with *HER2* in approximately 50% of the cases. Thus, we can envision that Epn3 could confer an additive advantage to the tumorigenic potential of HER2 in those cases and might represent a possible target to improve the treatment response in HER2-positive breast cancer patients.

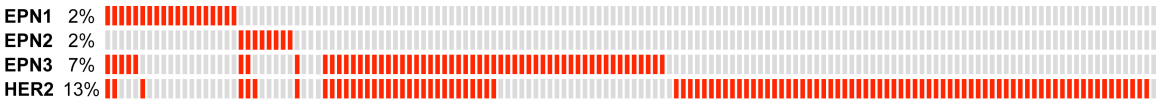


Figure 28 | Amplification of the Epsin family and HER2 in breast cancer.

The data were made available by the Tumor Cancer Genome Atlas study (Nature 2012). 778 breast invasive carcinoma cases are analyzed. Red squares represent amplified cases. Grey squares represent not amplified cases. In this image only a certain number of all the analyzed cases are represented due to space limitation, but all the altered cases are presented.

Of note, *EPN3* gene amplification seems to be a genetic phenomenon specific for breast cancer. Indeed, by looking at gene amplification in the most common cancers, *EPN3* is significantly amplified only in breast carcinomas (Comprehensive molecular portraits of human breast tumors, Nature 2012) (Figure 29). In addition, Epn3 seems to be the only member of Epsin family to be genetically altered in breast cancer, since neither *EPN1* nor *EPN2* genes are significantly amplified in breast cancer (Figure 28). These observations further confirm all the data present in literature suggesting that, differentially from Epn1 and Epn2, the tight regulation of Epn3 expression in the tissues is peculiar phenomenon of this member of the Epsin family [85] (see also Discussion section 4).

Altogether these findings strongly argue that *EPN3* amplification could represent a driver lesion in breast carcinogenesis that, importantly, can be independent from HER2 status in breast cancer patients.

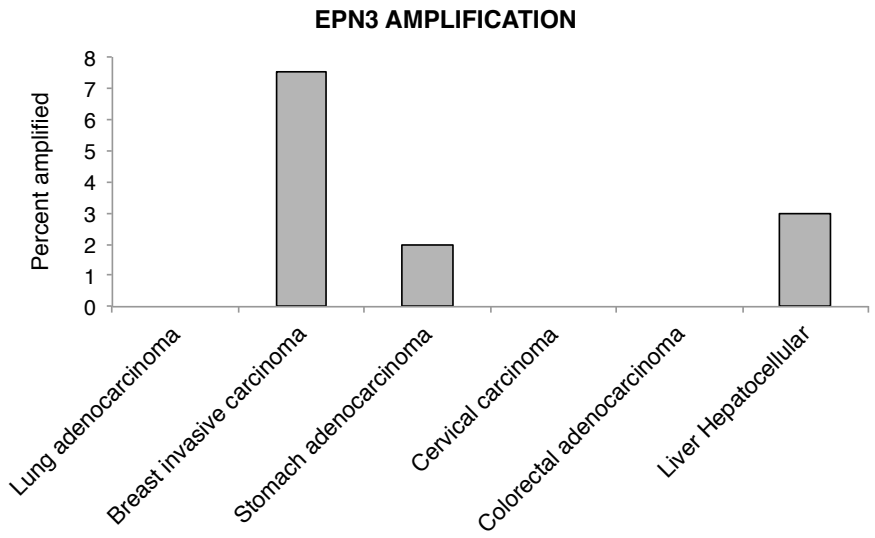


Figure 29 | *EPN3* amplification in the most common cancer types.

The data were obtained from cBioPortal for Cancer Genomics. The data are represented as percentage of amplification in each cancer type.

1.3 Epn3 is tumorigenic *in vitro* and *in vivo*

The expression of several endocytic proteins has been shown to be altered in human cancers and in some cases these alterations have been shown to be causative of a tumorigenic phenotype [69, 70]. Our findings that Epn3 expression correlates with aggressive disease and that its overexpression is driven by gene amplification, prompted us to demonstrate that Epn3 amplification is a driver lesion in breast cancer.

We demonstrated that Epn3 ablation in BT474 cells, which carry *EPN3* gene amplification and protein overexpression, results in a significant decrease of the anchorage-independent growth *in vitro* and of the tumorigenic potential *in vivo* of these cells. Conversely, downregulation of Epn3 expression in MCF7 cells, which display low levels of Epn3 expression without any genetic alteration of *EPN3* locus, does not have the same effect on tumorigenic growth *in vitro*. To

further confirm these results, we plan to evaluate the tumorigenic potential *in vivo* of MCF7 cells upon Epn3 ablation.

We also show that enforced overexpression of Epn3 in breast tumor cells, HCC1569, which express low levels of Epn3, results in increase of their tumorigenic potential *in vivo*. The same effect was not scored in normal mammary cells. Indeed, MCF10A Epn3-overexpressing cells were not able to form colonies in semi-solid medium - an hallmark of the tumorigenic phenotype. Of note, overexpression of Twist, a well-known EMT inducer gene [144], did not induce tumorigenic transformation of these cells. These results could be due either to a more permissive tumoral background of HCC1569 compared to MCF10A or to the fact that HCC1569 cells harbor HER2 amplification and overexpression.

Indeed, we have demonstrated that the amplification of *EPN3* and *HER2* significantly correlates in breast cancer, although the two events can occur independently. Importantly, both BT474 and HCC1569 breast cancer cells display amplification of the *HER2* gene. These cellular model systems are, therefore, useful to investigate a possible cooperation of Epn3 and HER2 in breast cancer progression. We will therefore investigate in future studies if inhibition of HER2 can influence the tumorigenic potential of these cells, which would suggest a cooperative effect of Epn3 and HER2 in promoting tumorigenesis. Another possibility could be to evaluate the cellular response to HER2 inhibitor treatments to understand if Epn3 could have a role in the chemoresistance induced by HER2 overexpression (reviewed in [163]).

To further investigate this aspect, it would be interesting to study the tumorigenic potential of Epn3 *in vivo* using mouse models. To this aim, we have available in our lab two Epn3 mouse models: i) Epn3-KO mouse model, generated by De Camilli's group [86], that does not show any defects in development,

behavior and breeding and that we have backcrossed in an FVB background; ii) an inducible Epn3-KI mouse model generated in house. These models could be crossed with HER2-driven breast cancer models, such as MMTV-HER2 mice [167], and will allow us to find whether Epn3 plays a role in HER2-driven tumorigenesis.

2 Epn3 overexpression induces EMT phenotype in mammary epithelial cells

2.1 EMT-like phenotype in Epn3 overexpressing cells as a mechanism for cancer progression

Tumorigenesis is a multistep process in which the accumulation of genetic alterations drives the transformation of normal cells that progressively acquire self-sufficiency in growth signals, unlimited replicative potential, resistance to cell death signals, leading ultimately to the activation of invasive and metastatic programs (reviewed in [49]).

The work described in this thesis has demonstrated that, by manipulating Epn3 expression in breast tumor cell lines, we can influence their tumorigenic growth both *in vitro* and *in vivo*. In addition, we have found that enforced Epn3 overexpression in normal MCF10A cells induces an EMT phenotype characterized by decreased expression of the epithelial marker E-cadherin and rearrangement of adherent cellular junctions. Moreover, we observed the upregulation of mesenchymal markers, N-cadherin and vimentin, and an increase in cell invasiveness. Importantly, this effect was reiterated in breast tumor HCC1569 cells, thus confirming that Epn3 is able to induce an EMT phenotype even in a tumoral background. We demonstrate, therefore, that Epn3 might exert its tumorigenic potential by reprogramming cells towards an invasive EMT phenotype.

During EMT, epithelial polarized cells undergo several biochemical changes and become fibroblastoid, mesenchymal and highly motile. This process is characterized by: i) loss of intercellular adhesion (E-cadherin); ii) upregulation of mesenchymal markers (vimentin, N-cadherin); iii) acquisition of spindle morphology; iv) increase of motility, invasiveness and metastatic capabilities

(reviewed in [143, 168, 169]). EMT has emerged as a key process during embryonic development and chronic inflammation and fibrosis, as well as cancer progression (reviewed in [143]). Indeed, during carcinoma progression, epithelial cells detach from the primary tumor and disseminate to distant tissues leading to secondary tumors.

However, it took a long time to recognize EMT as a potential mechanism for tumor progression. The major reason for this late recognition is that the involvement of EMT in cancer progression has been inferred predominantly from *in vitro* studies. EMT is a challenging process to follow *in vivo* in human tumors mainly because this phenomenon does not occur homogenously across the tumor. Indeed, in some cases, EMT takes a long time to occur and the expression of EMT markers appears mainly at the invasive front of the tumoral mass (reviewed in [170, 171]). Besides, *in vitro* models are somewhat limited: few carcinoma cells showing a well-defined epithelial phenotype can complete EMT *in vitro (*reviewed in [172]). For these reasons, an “EMT-like” term has been proposed to refer to epithelial cells that undergo a partial or “metastable” EMT program, which means that cells presenting this phenotype exhibit both epithelial and mesenchymal features (reviewed in [172, 173]). An “EMT-like” phenotype is consistent with what we have observed in the case of Epn3 overexpression in mammary cells and could explain why the cells did not completely lose E-cadherin expression while concomitantly expressing mesenchymal markers.

Breast cancer is considered a heterogeneous disease and gene expression studies have identified different molecular subtypes [111, 137]. Of note, it has been shown that the EMT signature is observed in the basal-like subgroup of breast cancers characterized by unfavorable prognosis, high rate of disease relapse and distant metastasis [174]. In agreement, not only we have observed an

EMT-like phenotype upon Epn3 overexpression in breast cell lines, but we have also found a strong correlation between Epn3 expression and poor prognosis breast cancer subtypes, including basal-like tumors. Therefore, it would be interesting to perform a meta-analysis on breast cancer expression data to uncover a possible association between Epn3 expression and an EMT signature in invasive breast carcinomas.

2.2 Epn3: a novel regulator of the network of EMT, stemness and endocytosis?

Despite recent medical advances, including early detection, adjuvant and targeted therapies, tumor relapse and resistance to therapy remain the principal causes of death in breast cancer patients.

An emerging hypothesis is that tumor progression is driven by a subpopulation of cells, termed cancer stem cells (CSCs), that have the ability to self-renew, form tumor spheres *in vitro* and regenerate tumors in xenotransplant systems [175]. Interestingly, several studies have established a crucial link between the activation of an EMT program and the acquisition of molecular and functional properties of stem cells [148, 176]. Indeed, *Mani* and colleagues have reported that the induction of EMT in immortalized human mammary epithelial cells, by ectopic expression of either transcription factors Twist or Snail, results in mesenchymal cells with increased ability to form mammospheres in suspension culture and increased expression of CSCs markers, such as $CD44^{+}/CD24^{low}$ [148].

Since we have found that Epn3 overexpression induces an EMT phenotype in normal human mammary epithelial MCF10A cells, we tested the possibility that these cells had also acquired some stem cell-like properties. Mammosphere formation assays revealed that Epn3 overexpression increases the capacity of

these cells to form mammospheres *in vitro*. Although these experiments were performed only in immortalized cells without allowing propagation of mammospheres, they suggest a role for Epn3 in the maintenance of the stem cell compartment. However, to further confirm these results, we now plan to perform mammosphere serial propagation assays in primary mammary stem cells overexpressing or not Epn3. In addition, we will also perform experiments in normal mammary stem cells to assess whether overexpression of Epn3 can skew the mode of stem cell division from asymmetric (typical of a normal stem cell) towards a symmetric one (typical of a putative cancer stem cells) [149, 151]. An extreme scenario would be one in which Epn3 overexpression reverts progenitor cells to a stem cell-like state, as shown in the case of Sox9 and Slug in the work from Weinberg's group [177]. In breast cancer, an increased proportion of CSCs within the tumor accompanies malignant progression: poorly differentiated, more aggressive breast carcinomas display a higher CSCs number than well differentiated, less aggressive tumors [151]. Under these conditions, tumorigenesis might depend on the expansion of the CSC compartment resulting from the CSCs mode of division being skewed towards a symmetric mode [149, 151]. Endocytosis is known to play a crucial role in the regulation of the asymmetrical cell division and the partitioning of cell fate determinants (reviewed in [1, 2]). Therefore, it is possible that alteration of endocytic pathways can lead to the deregulation of the asymmetric division program and induce changes in the content of CSCs.

A prominent example is provided by the endocytic protein Numb, which is asymmetrical partitioned during the genesis of the sensory organ of *Drosophila* and confers different fates to the two daughter cells originated from the asymmetric division of the precursor cell (SOP). It has been proposed that Numb

is able to counteract the activity of Notch receptor by inducing its internalization and degradation [73, 178]. In breast tumors, Numb expression is frequently lost and such loss correlates with poor prognosis and expression of CSCs markers ([75, 77, 179], which in the light of the findings that poorly differentiated breast tumors harbor high CSCs-content means that Numb-negative tumors are less differentiated and more aggressive [151]. Numb represents a leading example of the strong connection between endocytosis, homeostasis of the stem cell compartment and malignancies.

We now provide data arguing for another endocytic protein, Epn3, being involved in changes in the endocytic matrix that can affect the development of naturally occurring malignancies. We can envision a scenario in which Epn3 might induce alterations in the CSC homeostasis through EMT, which lead to an expansion of the CSC compartment, acquisition of a motile/invasive phenotype and malignant transformation.

3 A novel endocytic function of Epn3 regulates TGFβ pathway and response

3.1 Epn3 functions as an endocytic protein in E-cadherin internalization

We have already discussed that the loss of E-cadherin expression is the prototypical marker of EMT and it is considered a crucial step in the transition from papilloma to carcinoma [180]. Multiple mechanisms have been described to underlie the loss of E-cadherin function and they comprise transcriptional and post-translational regulation of E-cadherin.

Transcriptional repression of the E-cadherin gene is mediated by several transcription factors, including Snail, Slug and Twist [144, 181, 182]. All these transcriptional factors are upregulated and activated during EMT processes (reviewed in [155]).

Besides E-cadherin genomic regulation, the endocytic trafficking of surface E-cadherin is also required to disrupt its adhesive function and to promote tumor aggressiveness (reviewed in [156]. E-cadherin cell entry routes comprise clathrin-mediated endocytosis [183], caveolae-dependent endocytosis [184] and also macropinocytosis [47]. Stimulation with growth factors or induction of EMT can regulate E-cadherin endocytosis that results either in recycling or degradation.

TGFβ appears to be a key regulator of cadherin turnover during EMT, as it has been demonstrated that TGFβ and Ras-Raf signaling cooperate in inducing E-cadherin internalization through CME [157].

We have shown here that Epn3 overexpression in normal mammary MCF10A cells results in a reduction of E-cadherin protein levels and rearrangement of cellular junctions. We have also demonstrated that Epn3-overexpressing cells display a reduction of E-cadherin staining at the PM that

suggests an increase of basal E-cadherin turnover. Importantly, Epn3-overexpressing cells significantly accelerated TGFβ-induced E-cadherin internalization and downregulation from the cell surface. These results suggest a cooperative effect of Epn3 and TGFβ in inducing E-cadherin removal from the PM and its internalization.

Of note, E-cadherin forms a complex with catenins localized at the PM, including β-catenin. Once E-cadherin is internalized during EMT, β-catenin translocates into the nucleus and acts as transcriptional activator together with T cell factor (TCF/LEF) complex and regulates the expression of genes considered markers of EMT, including Snail family transcription factors [185, 186]. Thus, one possibility would be that Epn3 induces a β-catenin-dependent transcriptional program. To test this hypothesis, in future studies we plan to score if, upon Epn3overexpression, β-catenin is translocated into the nucleus and shows increased transcriptional activity. On the other hand, we will investigate whether Epn3 interacts directly or in a complex with E-cadherin at the cell surface. One possibility might be that Epn3 interacts with E-cadherin in an ubiquitin-dependent manner, since E-cadherin can undergo ubiquitination [187] and Epn3 contains UIMs. However, testing this hypothesis will require further investigation by performing structural function analysis of Epn3.

In conclusion, our observations suggest that Epn3 is able to induce EMT by acting as endocytic adaptor on E-cadherin trafficking.

3.2 Epn3-induced EMT phenotype is dependent on TGFβ signaling

Epn3 overexpression in MCF10A cells causes EMT and these effects are mediated to some degree by Epn3-induced upregulation of the expression of TGFβ ligands and receptors. Such upregulation sensitizes cells to TGFβ signaling, as we detected an increased TGFβ response (Smad-2 and Smad-3 phosphorylation, upregulation of Snail and of N-cadherin, and invasiveness) in Epn3-overexpressing cells.

TGFβ signaling strongly regulates tumor initiation, progression and metastasis. In breast cancer, TGFβ signaling plays a dual role in tumorigenesis. In early lesions, TGFβ acts as tumor suppressor through its growth inhibitory effects [188]. In contrast, TGFβ promotes metastatic spread in later stages of tumorigenesis. In numerous models of breast cancer associated with invasion and metastasis, activated TGFβ signaling induces increased aggressiveness. For example, in mice overexpressing the Neu oncogene, the concomitant overexpression of TGFβ protein increased the number of circulating cancer cells and lung metastases [189]. Likewise, ablation of TGFβ signaling in the same model decreases the number of lung metastases while also decreasing the latency of primary tumor growth, again emphasizing the dual functions of TGFβ in tumorigenesis [190].

Additionally, clinical evidence suggests that TGFβ ligands are overexpressed either at the mRNA or protein level in breast tumors. Intense TGFβ immunostaining correlates with disease progression to metastasis in breast carcinomas [191, 192]. Moreover, high TGFβ ligands levels correlates with poor outcome in breast cancer patients [193].

Because TGFβ signaling is a key inducer of the EMT program in a tumorigenic context, it is likely that the ability of TGFβ to promote metastasis is at least in part due to its ability to induce EMT (reviewed in [194-196]). However, a recent paper from Polyak' s group suggests an alternative mechanism. Expression profiling of a population of breast epithelial stem cells showed enrichment in transcripts associated with cell motility, angiogenesis and TGFβ pathway, thus suggesting a possible role for TGFβ in mammary stem cell maintenance [197]. Interestingly, Weinberg' s group had already demonstrated that mesenchymal cells are able to maintain their residence in the mesenchymal–stem cell state by ongoing autocrine signaling. The work established that the network of reinforcing autocrine loops engages also TGFβ pathway [198].

Based on the existing wide knowledge on the circuitries between TGFβ, EMT and breast cancer, we can speculate that our results suggest a possible new role for the endocytic protein Epn3. According to our proposed model, Epn3 acts as endocytic protein on E-cadherin trafficking. Downregulation of E-cadherin induces an EMT program that involves upregulation of mesenchymal proteins that control cell motility and invasion. Among all the activated transcripts, there are components of TGFβ signaling, such as TGFβ proteins and receptors that act by maintaining and eventually further increasing the EMT state upon stimulation.

The dependence of the Epn3-induced EMT phenotype on TGFβ in our mammary cell system was confirmed by an almost complete reversion of the Epn3-induced EMT after treatment with a TGFβR inhibitor or after silencing TGFβR. Given these *in vitro* results, we are now interested in assessing whether the ability of Epn3 to increase the tumorigenic potential of breast tumor cells *in vivo* is also dependent on TGFβ. To this aim, we are planning to evaluate whether

treatment with TGFβ inhibitor is able to reduce the tumorigenic capacity of breast tumor cells upon Epn3 expression in xenografts experiments.

These observations might have a strong clinic impact. Of note, clinical trials with anti-TGFβ agents are ongoing (reviewed in [196]). Since TGFβ has a strong impact also in cell physiology, due to its function as cytostatic and homeostatic function in the cells, it is possible that treatment with anti-TGFβ drugs might have to be limited: for example, it might be used against metastases only for certain tumors, maybe in limited doses, and in combination with other inhibitors of the EMT program.

In conclusion, if we score some effects on Epn3-driven tumorigenesis by treating the mice with TGFβ inhibitors, then we can think to translate these findings to the clinical setting by performing some correlations between Epn3 and TGFβ signatures in human breast tumor samples. In this scenario, Epn3 could be used as marker to stratify breast cancer patients that could benefit from a TGFβ inhibiting treatment.

4 The physiological role of Epn3 as the missing piece of the puzzle

Growing evidence suggests that Epn3 is the only epsin family member aberrantly expressed under pathological conditions [85], supporting the possibility that it might be involved in critical cellular functions.

The present study has demonstrated that Epn3 is overexpressed in breast cancer and that Epn3 overexpression correlates with disease progression. Importantly, this pathological role in breast progression seems to be specific for Epn3 and not shared by the other members of the epsin family. Indeed, as we have already discussed, neither *EPN1* nor *EPN2* genes are significantly amplified in breast cancer (see Discussion, section 1.2 and Figure 28). Additionally, by overexpressing Epn1 in breast cell lines we did not observe the EMT-like phenotype scored upon Epn3 overexpression (see Results, section 2.2.1), a feature that we believe is crucial for cancer progression.

These observations prompted us to hypothesize that Epn3 might have evolved an additional/specialized function, closely linked to molecular pathways differentially activated in cancer cells. Amino acid sequence alignment of the Epsin family members reveals that all three proteins show high similarity and conserved domain organization (about 80% of identity) (Figure 30). Future proteomic and structural function studies will help to unveil differences between the epsin family members that do not emerge simply from sequence and structural analysis of the proteins. On the one hand, we will engineer constructs carrying deletions or point mutations in the different functional domains of Epn3 (ENTH, UIM, DPW, Clathrin box, NPF) (see also Introduction, section 2.1.1). The resulting mutant proteins will be expressed in normal and tumor breast cells and these cells will be tested for their ability to undergo EMT as well as in other biological and tumorigenic assays. Such assays will probably allow the identification of the domain responsible for the

peculiar oncogenic properties of Epn3. On the other hand, we plan to generate Epn1-Epn3 chimeric proteins. This approach has the advantage of not being biased towards known endocytic functions of epsins.

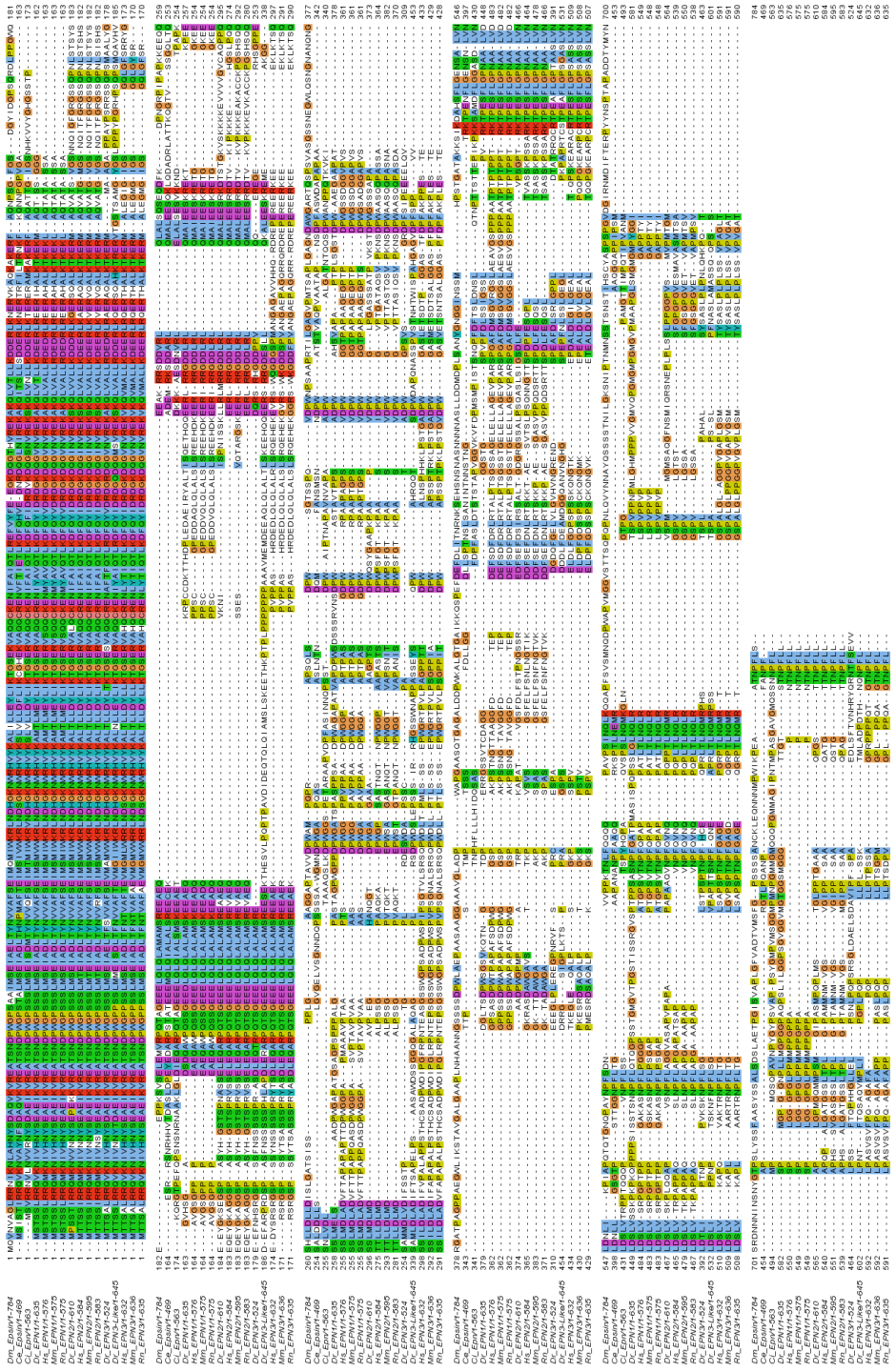


Figure 30 | Amino acid sequence alignment of Epn1, Epn2 and Epn3.

The indicated protein sequences were taken from NCBI-Protein database and were aligned using CLUSTALW.

Based on its highly conserved domain structure in comparison to other epsins, Epn3 is predicted to be an endocytic protein. This assumption is further strengthened by data demonstrating that Epn3 is localized in the endocytic compartment and that it interacts with known endocytic/trafficking proteins, such as AP2 and clathrin ([86] and *our unpublished data*). Given the important crosstalk that is emerging between endocytosis, signaling and cancer, it is possible that Epn3 exerts an oncogenic activity through an endocytosis-based mechanism. We have found that Epn3 overexpression in normal breast cells is able to increase E-cadherin internalization thus inducing EMT-like phenotype and contributing to cancer progression. This effect on E-cadherin trafficking seems to be specific for this cargo, since our previous data in the lab demonstrated that Epn3 overexpression does not influence the trafficking of other cargos, such as EGFR, transferrin receptor (TfR) and TGFβ receptor.

One important unanswered question remains about the physiological role of Epn3. Indeed, it would be very interesting to understand whether the function of Epn3 as adaptor protein on E-cadherin internalization is a specific oncogenic function that cancer cells acquire upon *EPN3* gene amplification and protein overexpression, or whether Epn3 still retains this endocytic function in cell physiology. To address this, we can take advantage of the Epn3-KO mouse model previously described by De Camilli and colleagues. Since Epn3-KO mouse does not show any particular phenotype, one can think of some overlapping function of the other two epsins. However, a more accurate description of this mouse model is still missing. It is now clear that Epn3 exerts a role in the migration and invasion of breast cell lines and this is in complete agreement with the work that first identified Epn3 as a novel transcript upregulated by migrating keratinocytes in cutaneous wounds [85]. Therefore, we can use cells derived from Epn3-KO mice (i.e. mouse

fibroblast or breast epithelial cells) to perform endocytic assays to confirm the role of Epn3 on E-cadherin trafficking also in physiological conditions. Moreover, we can also perform biological assays, such as wound healing *in vivo* assays, to understand whether Epn3 is necessary for the migration and reorganization of tissues after injury.

Altogether, these approaches will allow an even more accurate dissection of Epn3 role in physiology and cancer. The main possible scenarios are: i) Epn3 exerts a certain function in physiology that becomes aberrant upon the genetic liaison that occurs in breast cancer leading to protein overexpression; ii) Epn3 protein, when overexpressed due to gene amplification, in a tumoral context acquires a novel aberrant function that contributes to tumorigenesis. Both these hypotheses are compatible with the fact that Epn3 is expressed at very low levels in all normal homeostatic tissues, while it is upregulated almost exclusively in pathological conditions, where it can exert its function contributing to malignancy.

BIBLIOGRAPHY

1. Polo, S. and P.P. Di Fiore, *Endocytosis conducts the cell signaling orchestra.* Cell, 2006. **124**(5): p. 897-900.
2. Scita, G. and P.P. Di Fiore, *The endocytic matrix.* Nature, 2010. **463**(7280): p. 464-73.
3. Lanzetti, L. and P.P. Di Fiore, *Endocytosis and cancer: an 'insider' network with dangerous liaisons.* Traffic, 2008. **9**(12): p. 2011-21.
4. Sigismund, S., et al., *Endocytosis and signaling: cell logistics shape the eukaryotic cell plan.* Physiological reviews, 2012. **92**(1): p. 273-366.
5. Swanson, J.A., *Shaping cups into phagosomes and macropinosomes.* Nature reviews. Molecular cell biology, 2008. **9**(8): p. 639-49.
6. Kerr, M.C. and R.D. Teasdale, *Defining macropinocytosis.* Traffic, 2009. **10**(4): p. 364-71.
7. Insall, R.H. and L.M. Machesky, *Actin dynamics at the leading edge: from simple machinery to complex networks.* Developmental cell, 2009. **17**(3): p. 310-22.
8. Mayor, S. and R.E. Pagano, *Pathways of clathrin-independent endocytosis.* Nature reviews. Molecular cell biology, 2007. **8**(8): p. 603-12.
9. McMahon, H.T. and E. Boucrot, *Molecular mechanism and physiological functions of clathrin-mediated endocytosis.* Nature reviews. Molecular cell biology, 2011. **12**(8): p. 517-33.
10. Yao, D., et al., *Transforming growth factor-beta receptors interact with AP2 by direct binding to beta2 subunit.* Molecular biology of the cell, 2002. **13**(11): p. 4001-12.
11. Sorkin, A., *Cargo recognition during clathrin-mediated endocytosis: a team effort.* Current opinion in cell biology, 2004. **16**(4): p. 392-9.
12. Traub, L.M., *Tickets to ride: selecting cargo for clathrin-regulated internalization.* Nature reviews. Molecular cell biology, 2009. **10**(9): p. 583-96.
13. Sigismund, S., et al., *Clathrin-independent endocytosis of ubiquitinated cargos.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(8): p. 2760-5.
14. Johannessen, L.E., et al., *Activation of the epidermal growth factor (EGF) receptor induces formation of EGF receptor- and Grb2-containing clathrin-coated pits.* Molecular and cellular biology, 2006. **26**(2): p. 389-401.
15. Shenoy, S.K. and R.J. Lefkowitz, *Multifaceted roles of beta-arrestins in the regulation of seven-membrane-spanning receptor trafficking and signalling.* The Biochemical journal, 2003. **375**(Pt 3): p. 503-15.
16. Howes, M.T., et al., *Clathrin-independent carriers form a high capacity endocytic sorting system at the leading edge of migrating cells.* The Journal of cell biology, 2010. **190**(4): p. 675-91.
17. Naslavsky, N., R. Weigert, and J.G. Donaldson, *Characterization of a nonclathrin endocytic pathway: membrane cargo and lipid requirements.* Molecular biology of the cell, 2004. **15**(8): p. 3542-52.
18. Lamaze, C., et al., *Interleukin 2 receptors and detergent-resistant membrane domains define a clathrin-independent endocytic pathway.* Molecular cell, 2001. **7**(3): p. 661-71.

19. Sabharanjak, S., et al., *GPI-anchored proteins are delivered to recycling endosomes via a distinct cdc42-regulated, clathrin-independent pinocytic pathway.* Developmental cell, 2002. **2**(4): p. 411-23.
20. Haugh, J.M. and T. Meyer, *Active EGF receptors have limited access to PtdIns(4,5)P(2) in endosomes: implications for phospholipase C and PI 3-kinase signaling.* Journal of cell science, 2002. **115**(Pt 2): p. 303-10.
21. McLaughlin, S. and D. Murray, *Plasma membrane phosphoinositide organization by protein electrostatics.* Nature, 2005. **438**(7068): p. 605-11.
22. Posor, Y., et al., *Spatiotemporal control of endocytosis by phosphatidylinositol-3,4-bisphosphate.* Nature, 2013. **499**(7457): p. 233-7.
23. Puthenveedu, M.A. and M. von Zastrow, *Cargo regulates clathrin-coated pit dynamics.* Cell, 2006. **127**(1): p. 113-24.
24. Fortini, M.E. and D. Bilder, *Endocytic regulation of Notch signaling.* Current opinion in genetics & development, 2009. **19**(4): p. 323-8.
25. Windler, S.L. and D. Bilder, *Endocytic internalization routes required for delta/notch signaling.* Current biology : CB, 2010. **20**(6): p. 538-43.
26. Vaccari, T., et al., *Endosomal entry regulates Notch receptor activation in Drosophila melanogaster.* The Journal of cell biology, 2008. **180**(4): p. 755-62.
27. Meloty-Kapella, L., et al., *Notch ligand endocytosis generates mechanical pulling force dependent on dynamin, epsins, and actin.* Developmental cell, 2012. **22**(6): p. 1299-312.
28. Rajan, A., et al., *The Arp2/3 complex and WASp are required for apical trafficking of Delta into microvilli during cell fate specification of sensory organ precursors.* Nature cell biology, 2009. **11**(7): p. 815-24.
29. Di Guglielmo, G.M., et al., *Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover.* Nat Cell Biol, 2003. **5**(5): p. 410-21.
30. Sigismund, S., et al., *Threshold-controlled ubiquitination of the EGFR directs receptor fate.* The EMBO journal, 2013. **32**(15): p. 2140-57.
31. Sigismund, S., et al., *Clathrin-mediated internalization is essential for sustained EGFR signaling but dispensable for degradation.* Developmental cell, 2008. **15**(2): p. 209-19.
32. Stenmark, H., *Rab GTPases as coordinators of vesicle traffic.* Nature reviews. Molecular cell biology, 2009. **10**(8): p. 513-25.
33. Sorkin, A. and M. von Zastrow, *Endocytosis and signalling: intertwining molecular networks.* Nature reviews. Molecular cell biology, 2009. **10**(9): p. 609-22.
34. Pol, A., M. Calvo, and C. Enrich, *Isolated endosomes from quiescent rat liver contain the signal transduction machinery. Differential distribution of activated Raf-1 and Mek in the endocytic compartment.* FEBS letters, 1998. **441**(1): p. 34-8.
35. Wunderlich, W., et al., *A novel 14-kilodalton protein interacts with the mitogen-activated protein kinase scaffold mp1 on a late endosomal/lysosomal compartment.* The Journal of cell biology, 2001. **152**(4): p. 765-76.
36. Teis, D., et al., *p14-MP1-MEK1 signaling regulates endosomal traffic and cellular proliferation during tissue homeostasis.* The Journal of cell biology, 2006. **175**(6): p. 861-8.
37. Nada, S., et al., *The novel lipid raft adaptor p18 controls endosome dynamics by anchoring the MEK-ERK pathway to late endosomes.* The EMBO journal, 2009. **28**(5): p. 477-89.

38. Irannejad, R., et al., *Conformational biosensors reveal GPCR signalling from endosomes.* Nature, 2013. **495**(7442): p. 534-8.
39. Tsukazaki, T., et al., *SARA, a FYVE domain protein that recruits Smad2 to the TGFbeta receptor.* Cell, 1998. **95**(6): p. 779-91.
40. Hayes, S., A. Chawla, and S. Corvera, *TGF beta receptor internalization into EEA1-enriched early endosomes: role in signaling to Smad2.* The Journal of cell biology, 2002. **158**(7): p. 1239-49.
41. Disanza, A., et al., *Endocytosis and spatial restriction of cell signaling.* Molecular oncology, 2009. **3**(4): p. 280-96.
42. Bailly, M., et al., *Epidermal growth factor receptor distribution during chemotactic responses.* Molecular biology of the cell, 2000. **11**(11): p. 3873-83.
43. Dormann, D. and C.J. Weijer, *Chemotactic cell movement during development.* Current opinion in genetics & development, 2003. **13**(4): p. 358-64.
44. Palamidessi, A., et al., *Endocytic trafficking of Rac is required for the spatial restriction of signaling in cell migration.* Cell, 2008. **134**(1): p. 135-47.
45. Takeichi, M., *Morphogenetic roles of classic cadherins.* Current opinion in cell biology, 1995. **7**(5): p. 619-27.
46. Bryant, D.M. and J.L. Stow, *The ins and outs of E-cadherin trafficking.* Trends in cell biology, 2004. **14**(8): p. 427-34.
47. Bryant, D.M., et al., *EGF induces macropinocytosis and SNX1-modulated recycling of E-cadherin.* Journal of cell science, 2007. **120**(Pt 10): p. 1818-28.
48. Parachoniak, C.A. and M. Park, *Dynamics of receptor trafficking in tumorigenicity.* Trends in cell biology, 2012. **22**(5): p. 231-40.
49. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
50. Mosesson, Y., G.B. Mills, and Y. Yarden, *Derailed endocytosis: an emerging feature of cancer.* Nature reviews. Cancer, 2008. **8**(11): p. 835-50.
51. Lynch, T.J., et al., *Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib.* The New England journal of medicine, 2004. **350**(21): p. 2129-39.
52. Mulloy, R., et al., *Epidermal growth factor receptor mutants from human lung cancers exhibit enhanced catalytic activity and increased sensitivity to gefitinib.* Cancer research, 2007. **67**(5): p. 2325-30.
53. Chung, B.M., et al., *Aberrant trafficking of NSCLC-associated EGFR mutants through the endocytic recycling pathway promotes interaction with Src.* BMC cell biology, 2009. **10**: p. 84.
54. Moscatello, D.K., et al., *Frequent expression of a mutant epidermal growth factor receptor in multiple human tumors.* Cancer research, 1995. **55**(23): p. 5536-9.
55. Batra, S.K., et al., *Epidermal growth factor ligand-independent, unregulated, cell-transforming potential of a naturally occurring human mutant EGFRvIII gene.* Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research, 1995. **6**(10): p. 1251-9.
56. Tang, C.K., et al., *Epidermal growth factor receptor vIII enhances tumorigenicity in human breast cancer.* Cancer research, 2000. **60**(11): p. 3081-7.

57. Grandal, M.V., et al., *EGFRvIII escapes down-regulation due to impaired internalization and sorting to lysosomes.* Carcinogenesis, 2007. **28**(7): p. 1408-17.
58. Christensen, J.G., J. Burrows, and R. Salgia, *c-Met as a target for human cancer and characterization of inhibitors for therapeutic intervention.* Cancer letters, 2005. **225**(1): p. 1-26.
59. Jeffers, M., et al., *Activating mutations for the met tyrosine kinase receptor in human cancer.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(21): p. 11445-50.
60. Jeffers, M.F. and G.F. Vande Woude, *Activating mutations in the Met receptor overcome the requirement for autophosphorylation of tyrosines crucial for wild type signaling.* Oncogene, 1999. **18**(36): p. 5120-5.
61. Joffre, C., et al., *A direct role for Met endocytosis in tumorigenesis.* Nature cell biology, 2011. **13**(7): p. 827-37.
62. Thien, C.B. and W.Y. Langdon, *Cbl: many adaptations to regulate protein tyrosine kinases.* Nature reviews. Molecular cell biology, 2001. **2**(4): p. 294-307.
63. Sargin, B., et al., *Flt3-dependent transformation by inactivating c-Cbl mutations in AML.* Blood, 2007. **110**(3): p. 1004-12.
64. Sanada, M., et al., *Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms.* Nature, 2009. **460**(7257): p. 904-8.
65. Metzler, M., et al., *HIP1 functions in clathrin-mediated endocytosis through binding to clathrin and adaptor protein 2.* The Journal of biological chemistry, 2001. **276**(42): p. 39271-6.
66. Rao, D.S., et al., *Altered receptor trafficking in Huntingtin Interacting Protein 1-transformed cells.* Cancer cell, 2003. **3**(5): p. 471-82.
67. Bradley, S.V., et al., *Aberrant Huntingtin interacting protein 1 in lymphoid malignancies.* Cancer research, 2007. **67**(18): p. 8923-31.
68. Bradley, S.V., et al., *Huntingtin interacting protein 1 is a novel brain tumor marker that associates with epidermal growth factor receptor.* Cancer research, 2007. **67**(8): p. 3609-15.
69. Rao, D.S., et al., *Huntingtin-interacting protein 1 is overexpressed in prostate and colon cancer and is critical for cellular survival.* The Journal of clinical investigation, 2002. **110**(3): p. 351-60.
70. Cheng, K.W., et al., *The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers.* Nature medicine, 2004. **10**(11): p. 1251-6.
71. Caswell, P.T., et al., *Rab25 associates with alpha5beta1 integrin to promote invasive migration in 3D microenvironments.* Developmental cell, 2007. **13**(4): p. 496-510.
72. Dozynkiewicz, M.A., et al., *Rab25 and CLIC3 collaborate to promote integrin recycling from late endosomes/lysosomes and drive cancer progression.* Developmental cell, 2012. **22**(1): p. 131-45.
73. Rhyu, M.S., L.Y. Jan, and Y.N. Jan, *Asymmetric distribution of numb protein during division of the sensory organ precursor cell confers distinct fates to daughter cells.* Cell, 1994. **76**(3): p. 477-91.
74. Santolini, E., et al., *Numb is an endocytic protein.* The Journal of cell biology, 2000. **151**(6): p. 1345-52.
75. Pece, S., et al., *Loss of negative regulation by Numb over Notch is relevant to human breast carcinogenesis.* The Journal of cell biology, 2004. **167**(2): p. 215-21.

76. Westhoff, B., et al., *Alterations of the Notch pathway in lung cancer.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(52): p. 22293-8.
77. Colaluca, I.N., et al., *NUMB controls p53 tumour suppressor activity.* Nature, 2008. **451**(7174): p. 76-80.
78. Harris, A.L., *Hypoxia--a key regulatory factor in tumour growth.* Nature reviews. Cancer, 2002. **2**(1): p. 38-47.
79. Wang, Y., et al., *Regulation of endocytosis via the oxygen-sensing pathway.* Nature medicine, 2009. **15**(3): p. 319-24.
80. Zhong, H., et al., *Overexpression of hypoxia-inducible factor 1alpha in common human cancers and their metastases.* Cancer research, 1999. **59**(22): p. 5830-5.
81. Chen, H., et al., *Epsin is an EH-domain-binding protein implicated in clathrin-mediated endocytosis.* Nature, 1998. **394**(6695): p. 793-7.
82. Yamabhai, M., et al., *Intersectin, a novel adaptor protein with two Eps15 homology and five Src homology 3 domains.* The Journal of biological chemistry, 1998. **273**(47): p. 31401-7.
83. Morinaka, K., et al., *Epsin binds to the EH domain of POB1 and regulates receptor-mediated endocytosis.* Oncogene, 1999. **18**(43): p. 5915-22.
84. Rosenthal, J.A., et al., *The epsins define a family of proteins that interact with components of the clathrin coat and contain a new protein module.* The Journal of biological chemistry, 1999. **274**(48): p. 33959-65.
85. Spradling, K.D., et al., *Epsin 3 is a novel extracellular matrix-induced transcript specific to wounded epithelia.* J Biol Chem, 2001. **276**(31): p. 29257-67.
86. Ko, G., et al., *Selective high-level expression of epsin 3 in gastric parietal cells, where it is localized at endocytic sites of apical canaliculi.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21511-6.
87. Mills, I.G., et al., *EpsinR: an AP1/clathrin interacting protein involved in vesicle trafficking.* The Journal of cell biology, 2003. **160**(2): p. 213-22.
88. Saint-Pol, A., et al., *Clathrin adaptor epsinR is required for retrograde sorting on early endosomal membranes.* Developmental cell, 2004. **6**(4): p. 525-38.
89. De Camilli, P., et al., *The ENTH domain.* FEBS letters, 2002. **513**(1): p. 11-8.
90. Kay, B.K., et al., *Identification of a novel domain shared by putative components of the endocytic and cytoskeletal machinery.* Protein science : a publication of the Protein Society, 1999. **8**(2): p. 435-8.
91. Hyman, J., et al., *Epsin 1 undergoes nucleocytosolic shuttling and its eps15 interactor NH(2)-terminal homology (ENTH) domain, structurally similar to Armadillo and HEAT repeats, interacts with the transcription factor promyelocytic leukemia Zn(2)+ finger protein (PLZF).* The Journal of cell biology, 2000. **149**(3): p. 537-46.
92. Itoh, T., et al., *Role of the ENTH domain in phosphatidylinositol-4,5-bisphosphate binding and endocytosis.* Science, 2001. **291**(5506): p. 1047-51.
93. Ford, M.G., et al., *Curvature of clathrin-coated pits driven by epsin.* Nature, 2002. **419**(6905): p. 361-6.
94. Horvath, C.A., et al., *Epsin: inducing membrane curvature.* The international journal of biochemistry & cell biology, 2007. **39**(10): p. 1765-70.
95. Vecchi, M., et al., *Nucleocytoplasmic shuttling of endocytic proteins.* The Journal of cell biology, 2001. **153**(7): p. 1511-7.

96. Polo, S., et al., *A single motif responsible for ubiquitin recognition and monoubiquitination in endocytic proteins.* Nature, 2002. **416**(6879): p. 451-5.
97. Oldham, C.E., et al., *The ubiquitin-interacting motifs target the endocytic adaptor protein epsin for ubiquitination.* Current biology : CB, 2002. **12**(13): p. 1112-6.
98. Drake, M.T., M.A. Downs, and L.M. Traub, *Epsin binds to clathrin by associating directly with the clathrin-terminal domain. Evidence for cooperative binding through two discrete sites.* The Journal of biological chemistry, 2000. **275**(9): p. 6479-89.
99. Salcini, A.E., et al., *Epidermal growth factor pathway substrate 15, Eps15.* The international journal of biochemistry & cell biology, 1999. **31**(8): p. 805-9.
100. Santolini, E., et al., *The EH network.* Experimental cell research, 1999. **253**(1): p. 186-209.
101. Wang, W. and G. Struhl, *Drosophila Epsin mediates a select endocytic pathway that DSL ligands must enter to activate Notch.* Development, 2004. **131**(21): p. 5367-80.
102. Wang, W. and G. Struhl, *Distinct roles for Mind bomb, Neuralized and Epsin in mediating DSL endocytosis and signaling in Drosophila.* Development, 2005. **132**(12): p. 2883-94.
103. Chen, H., et al., *Embryonic arrest at midgestation and disruption of Notch signaling produced by the absence of both epsin 1 and epsin 2 in mice.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(33): p. 13838-43.
104. Pasula, S., et al., *Endothelial epsin deficiency decreases tumor growth by enhancing VEGF signaling.* The Journal of clinical investigation, 2012. **122**(12): p. 4424-38.
105. Aguilar, R.C., et al., *Epsin N-terminal homology domains perform an essential function regulating Cdc42 through binding Cdc42 GTPase-activating proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(11): p. 4116-21.
106. Coon, B.G., et al., *The epsin family of endocytic adaptors promotes fibrosarcoma migration and invasion.* The Journal of biological chemistry, 2010. **285**(43): p. 33073-81.
107. Chen, H., et al., *The interaction of epsin and Eps15 with the clathrin adaptor AP-2 is inhibited by mitotic phosphorylation and enhanced by stimulation-dependent dephosphorylation in nerve terminals.* The Journal of biological chemistry, 1999. **274**(6): p. 3257-60.
108. Rosse, C., et al., *RLIP, an effector of the Ral GTPases, is a platform for Cdk1 to phosphorylate epsin during the switch off of endocytosis in mitosis.* The Journal of biological chemistry, 2003. **278**(33): p. 30597-604.
109. Liu, Z. and Y. Zheng, *A requirement for epsin in mitotic membrane and spindle organization.* The Journal of cell biology, 2009. **186**(4): p. 473-80.
110. Boucrot, E., et al., *Membrane fission is promoted by insertion of amphipathic helices and is restricted by crescent BAR domains.* Cell, 2012. **149**(1): p. 124-36.
111. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
112. Malhotra, G.K., et al., *Histological, molecular and functional subtypes of breast cancers.* Cancer biology & therapy, 2010. **10**(10): p. 955-60.

113. Li, C.I., D.J. Uribe, and J.R. Daling, *Clinical characteristics of different histologic types of breast cancer.* British journal of cancer, 2005. **93**(9): p. 1046-52.
114. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(14): p. 8418-23.
115. Sorlie, T., et al., *Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms.* BMC genomics, 2006. **7**: p. 127.
116. Sotiriou, C. and L. Pusztai, *Gene-expression signatures in breast cancer.* The New England journal of medicine, 2009. **360**(8): p. 790-800.
117. Prat, A., et al., *Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.* Breast cancer research : BCR, 2010. **12**(5): p. R68.
118. Millikan, R.C., et al., *Epidemiology of basal-like breast cancer.* Breast cancer research and treatment, 2008. **109**(1): p. 123-39.
119. Bosch, A., et al., *Triple-negative breast cancer: molecular features, pathogenesis, treatment and current lines of research.* Cancer treatment reviews, 2010. **36**(3): p. 206-15.
120. Kamangar, F., G.M. Dores, and W.F. Anderson, *Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2006. **24**(14): p. 2137-50.
121. Abner, A.L., et al., *Correlation of tumor size and axillary lymph node involvement with prognosis in patients with T1 breast carcinoma.* Cancer, 1998. **83**(12): p. 2502-8.
122. Rakha, E.A., et al., *Breast cancer prognostic classification in the molecular era: the role of histological grade.* Breast cancer research : BCR, 2010. **12**(4): p. 207.
123. Jones, R.L., et al., *The prognostic significance of Ki67 before and after neoadjuvant chemotherapy in breast cancer.* Breast cancer research and treatment, 2009. **116**(1): p. 53-68.
124. Ross, J.S., et al., *The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy.* The oncologist, 2003. **8**(4): p. 307-25.
125. Lim, E., O. Metzger-Filho, and E.P. Winer, *The natural history of hormone receptor-positive breast cancer.* Oncology, 2012. **26**(8): p. 688-94, 696.
126. Jahanzeb, M., *Adjuvant trastuzumab therapy for HER2-positive breast cancer.* Clinical breast cancer, 2008. **8**(4): p. 324-33.
127. Nahta, R. and F.J. Esteva, *Trastuzumab: triumphs and tribulations.* Oncogene, 2007. **26**(25): p. 3637-43.
128. Rakha, E.A., et al., *Expression profiling technology: its contribution to our understanding of breast cancer.* Histopathology, 2008. **52**(1): p. 67-81.
129. van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer.* Nature, 2002. **415**(6871): p. 530-6.
130. Buyse, M., et al., *Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer.* Journal of the National Cancer Institute, 2006. **98**(17): p. 1183-92.
131. Minn, A.J., et al., *Genes that mediate breast cancer metastasis to lung.* Nature, 2005. **436**(7050): p. 518-24.

132. Wang, Y., et al., *Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer.* Lancet, 2005. **365**(9460): p. 671-9.
133. Ross, J.S., et al., *Commercialized multigene predictors of clinical outcome for breast cancer.* The oncologist, 2008. **13**(5): p. 477-93.
134. Paik, S., et al., *A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.* The New England journal of medicine, 2004. **351**(27): p. 2817-26.
135. Ventura, A., et al., *Cre-lox-regulated conditional RNA interference from transgenes.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(28): p. 10380-5.
136. Tosoni, D., P.P. Di Fiore, and S. Pece, *Functional purification of human and mouse mammary stem cells.* Methods in molecular biology, 2012. **916**: p. 59-79.
137. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(19): p. 10869-74.
138. Slamon, D.J., et al., *Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene.* Science, 1987. **235**(4785): p. 177-82.
139. Knuutila, S., et al., *DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies.* The American journal of pathology, 1998. **152**(5): p. 1107-23.
140. Luoh, S.W., *Amplification and expression of genes from the 17q11 approximately q12 amplicon in breast cancer cells.* Cancer genetics and cytogenetics, 2002. **136**(1): p. 43-7.
141. Kauraniemi, P., et al., *New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays.* Cancer research, 2001. **61**(22): p. 8235-40.
142. Curtis, C., et al., *The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.* Nature, 2012. **486**(7403): p. 346-52.
143. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* The Journal of clinical investigation, 2009. **119**(6): p. 1420-8.
144. Yang, J., et al., *Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis.* Cell, 2004. **117**(7): p. 927-39.
145. Cavallaro, U. and G. Christofori, *Cell adhesion and signalling by cadherins and Ig-CAMs in cancer.* Nature reviews. Cancer, 2004. **4**(2): p. 118-32.
146. Kim, J.B., et al., *N-Cadherin extracellular repeat 4 mediates epithelial to mesenchymal transition and increased motility.* The Journal of cell biology, 2000. **151**(6): p. 1193-206.
147. Ayollo, D.V., et al., *Rearrangements of the actin cytoskeleton and E-cadherin-based adherens junctions caused by neoplasic transformation change cell-cell interactions.* PloS one, 2009. **4**(11): p. e8027.
148. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.
149. Cicalese, A., et al., *The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells.* Cell, 2009. **138**(6): p. 1083-95.
150. Dontu, G., et al., *In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells.* Genes & development, 2003. **17**(10): p. 1253-70.

151. Pece, S., et al., *Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content.* Cell, 2010. **140**(1): p. 62-73.
152. Lucassen, E., et al., *The effects of the neuN and neuT genes on differentiation and transformation of mammary epithelial cells.* Journal of cell science, 1994. **107 (Pt 10)**: p. 2919-29.
153. Angst, B.D., C. Marcozzi, and A.I. Magee, *The cadherin superfamily.* Journal of cell science, 2001. **114**(Pt 4): p. 625-6.
154. Frixen, U.H., et al., *E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells.* The Journal of cell biology, 1991. **113**(1): p. 173-85.
155. Peinado, H., D. Olmeda, and A. Cano, *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nature reviews. Cancer, 2007. **7**(6): p. 415-28.
156. Delva, E. and A.P. Kowalczyk, *Regulation of cadherin trafficking.* Traffic, 2009. **10**(3): p. 259-67.
157. Janda, E., et al., *Raf plus TGFbeta-dependent EMT is initiated by endocytosis and lysosomal degradation of E-cadherin.* Oncogene, 2006. **25**(54): p. 7117-30.
158. Zavadil, J. and E.P. Bottinger, *TGF-beta and epithelial-to-mesenchymal transitions.* Oncogene, 2005. **24**(37): p. 5764-74.
159. Zhang, M., et al., *Blockade of TGF-beta signaling by the TGFbetaR-I kinase inhibitor LY2109761 enhances radiation response and prolongs survival in glioblastoma.* Cancer research, 2011. **71**(23): p. 7155-67.
160. Melisi, D., et al., *LY2109761, a novel transforming growth factor beta receptor type I and type II dual inhibitor, as a therapeutic approach to suppressing pancreatic cancer metastasis.* Molecular cancer therapeutics, 2008. **7**(4): p. 829-40.
161. Owens, M.A., B.C. Horten, and M.M. Da Silva, *HER2 amplification ratios by fluorescence in situ hybridization and correlation with immunohistochemistry in a cohort of 6556 breast cancer tissues.* Clinical breast cancer, 2004. **5**(1): p. 63-9.
162. Shalaby, M.R., et al., *Development of humanized bispecific antibodies reactive with cytotoxic lymphocytes and tumor cells overexpressing the HER2 protooncogene.* The Journal of experimental medicine, 1992. **175**(1): p. 217-25.
163. Fang, L., et al., *Targeted therapy in breast cancer: what's new?* Swiss medical weekly, 2011. **141**: p. w13231.
164. Osborne, C.K., et al., *Gefitinib or placebo in combination with tamoxifen in patients with hormone receptor-positive metastatic breast cancer: a randomized phase II study.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2011. **17**(5): p. 1147-59.
165. Baselga, J., et al., *Everolimus in postmenopausal hormone-receptor-positive advanced breast cancer.* The New England journal of medicine, 2012. **366**(6): p. 520-9.
166. Kao, J. and J.R. Pollack, *RNA interference-based functional dissection of the 17q12 amplicon in breast cancer reveals contribution of coamplified genes.* Genes, chromosomes & cancer, 2006. **45**(8): p. 761-9.
167. Guy, C.T., et al., *Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(22): p. 10578-82.

168. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.
169. Yilmaz, M. and G. Christofori, *EMT, the cytoskeleton, and cancer cell invasion.* Cancer metastasis reviews, 2009. **28**(1-2): p. 15-33.
170. Christofori, G., *New signals from the invasive front.* Nature, 2006. **441**(7092): p. 444-50.
171. Micalizzi, D.S., S.M. Farabaugh, and H.L. Ford, *Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression.* Journal of mammary gland biology and neoplasia, 2010. **15**(2): p. 117-34.
172. Savagner, P., *The epithelial-mesenchymal transition (EMT) phenomenon.* Annals of oncology : official journal of the European Society for Medical Oncology / ESMO, 2010. **21 Suppl 7**: p. vii89-92.
173. Klymkowsky, M.W. and P. Savagner, *Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe.* The American journal of pathology, 2009. **174**(5): p. 1588-93.
174. Sarrio, D., et al., *Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype.* Cancer research, 2008. **68**(4): p. 989-97.
175. Gupta, P.B., C.L. Chaffer, and R.A. Weinberg, *Cancer stem cells: mirage or reality?* Nature medicine, 2009. **15**(9): p. 1010-2.
176. Morel, A.P., et al., *Generation of breast cancer stem cells through epithelial-mesenchymal transition.* PloS one, 2008. **3**(8): p. e2888.
177. Guo, W., et al., *Slug and Sox9 cooperatively determine the mammary stem cell state.* Cell, 2012. **148**(5): p. 1015-28.
178. Spana, E.P. and C.Q. Doe, *Numb antagonizes Notch signaling to specify sibling neuron cell fates.* Neuron, 1996. **17**(1): p. 21-6.
179. Rennstam, K., et al., *Numb protein expression correlates with a basal-like phenotype and cancer stem cell markers in primary breast cancer.* Breast cancer research and treatment, 2010. **122**(2): p. 315-24.
180. Perl, A.K., et al., *A causal role for E-cadherin in the transition from adenoma to carcinoma.* Nature, 1998. **392**(6672): p. 190-3.
181. Cano, A., et al., *The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.* Nature cell biology, 2000. **2**(2): p. 76-83.
182. Casas, E., et al., *Snail2 is an essential mediator of Twist1-induced epithelial mesenchymal transition and metastasis.* Cancer research, 2011. **71**(1): p. 245-54.
183. Ivanov, A.I., A. Nusrat, and C.A. Parkos, *Endocytosis of epithelial apical junctional proteins by a clathrin-mediated pathway into a unique storage compartment.* Molecular biology of the cell, 2004. **15**(1): p. 176-88.
184. Lu, Z., et al., *Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion.* Cancer cell, 2003. **4**(6): p. 499-515.
185. Kim, K., Z. Lu, and E.D. Hay, *Direct evidence for a role of beta-catenin/LEF-1 signaling pathway in induction of EMT.* Cell biology international, 2002. **26**(5): p. 463-76.
186. Yook, J.I., et al., *A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells.* Nature cell biology, 2006. **8**(12): p. 1398-406.
187. Fujita, Y., et al., *Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex.* Nature cell biology, 2002. **4**(3): p. 222-31.

188. Reynisdottir, I., et al., *Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta.* Genes & development, 1995. **9**(15): p. 1831-45.
189. Muraoka, R.S., et al., *Increased malignancy of Neu-induced mammary tumors overexpressing active transforming growth factor beta1.* Molecular and cellular biology, 2003. **23**(23): p. 8691-703.
190. Siegel, P.M., et al., *Transforming growth factor beta signaling impairs Neu-induced mammary tumorigenesis while promoting pulmonary metastasis.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(14): p. 8430-5.
191. Walker, R.A. and S.J. Dearing, *Transforming growth factor beta 1 in ductal carcinoma in situ and invasive carcinomas of the breast.* European journal of cancer, 1992. **28**(2-3): p. 641-4.
192. Dalal, B.I., P.A. Keown, and A.H. Greenberg, *Immunocytochemical localization of secreted transforming growth factor-beta 1 to the advancing edges of primary tumors and to lymph node metastases of human mammary carcinoma.* The American journal of pathology, 1993. **143**(2): p. 381-9.
193. Ghellal, A., et al., *Prognostic significance of TGF beta 1 and TGF beta 3 in human breast carcinoma.* Anticancer research, 2000. **20**(6B): p. 4413-8.
194. Bierie, B. and H.L. Moses, *Gain or loss of TGFbeta signaling in mammary carcinoma cells can promote metastasis.* Cell cycle, 2009. **8**(20): p. 3319-27.
195. Massague, J., *TGFbeta in Cancer.* Cell, 2008. **134**(2): p. 215-30.
196. Heldin, C.H., M. Vanlandewijck, and A. Moustakas, *Regulation of EMT by TGFbeta in cancer.* FEBS letters, 2012. **586**(14): p. 1959-70.
197. Shipitsin, M., et al., *Molecular definition of breast tumor heterogeneity.* Cancer cell, 2007. **11**(3): p. 259-73.
198. Scheel, C., et al., *Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast.* Cell, 2011. **145**(6): p. 926-40.

ACKNOWLEDGMENTS

I would like to thank all the people that made this thesis work possible.

Above all, I thank my Ph.D. supervisor Pier Paolo Di Fiore, who gave me the opportunity to join his group four years ago and carry out this research project. He was a great mentor and he provided me with invaluable scientific knowledge, support and guidance through all these years.

A special thanks goes to Sara Sigismund who supported me day by day in the lab discussing experiments and science. She always gave me a great scientific and personal encouragement and I really appreciate everything she did for me.

I also acknowledge Manuela Vecchi who started this project some years ago and continued to be very helpful in all these years. Thanks also to Stefano Confalonieri, Giovanni Bertalot and Chiara Luise who performed IHC experiments and statistical analysis.

I am grateful to all my colleagues and labmates who made all my days in the lab extremely pleased and fun. In particular I thank Alexia Conte for sharing with me good and bad moments and Veronica Algisi for the valuable advices along all these years. A huge thanks goes to Roberta Pascolutti: we started together the PhD program and we shared all the stressful and enjoyable moments helping each other in all the possible ways.

Finally, I wish to thank Raquel Carvalhosa and Pascale Romano for critically reading my thesis.