



PhD degree in Systems Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples “Federico II”

Settore disciplinare: BIO/10

**Characterization of Epsin3 function in the
acquisition of a partial EMT state in breast cancer
through E-Cadherin endocytosis**

Giangreco Giovanni

Matricola n. R11142

IFOM – IEO Campus, Milan

Supervisor: Prof. Di Fiore Pier Paolo,
IEO and University of Milan

Added Supervisor: Dr. Sigismund Sara,
IEO and University of Milan

Anno accademico 2017-2018

a mia Madre e mio Padre
a mia Nonna
a mio Fratello
a Carmen

a mio Nonno

ABSTRACT

Epsin3 (EPN3) belongs to the Epsin family of endocytic adaptors, which in humans comprises 3 members: Epsin1 (EPN1), Epsin2 (EPN2) and EPN3. Differently from the other members, EPN3 is expressed at very low levels in almost all tissues and its function is unknown. Previous data have shown that EPN3 is overexpressed in about 47% of breast cancers, correlating with an aggressive tumor phenotype and increased risk of distant metastasis. At the cellular level, in vitro studies performed on MCF10A, a non-tumoral immortalized breast epithelial cell line, showed that overexpression of EPN3 (but not of EPN1) induces the appearance of fibroblast-like properties. Here, we found that EPN3 overexpression is able to induce the so-called partial epithelial-to-mesenchymal transition, pEMT, a condition characterized by a mix of epithelial and mesenchymal features that makes cells capable of easily converting between an epithelial and a mesenchymal state. Mechanistically, we showed that EPN3 interacts with the endocytic machinery (clathrin heavy chain and AP2) and with components of adherens junctions (ρ 120, β -catenin, α -catenin), and increases E-Cadherin endocytosis, suggesting an alteration of E-Cadherin turnover, known to be necessary to allow EMT. Indeed, the increased E-Cadherin endocytosis activates the pEMT through the β -catenin/TCF4 pathway. Moreover, EPN3 overexpression in MCF10A cells upregulates the TGF β pathway and synergizes with it, both in TGF β signaling and in induction of E-Cadherin endocytosis. Interestingly, EPN3-induced pEMT generates a TGF β -dependent loop that can be reverted by TCF4 or TGF β R1 KD, but it is not addicted to EPN3 expression once established.

Interestingly, endogenous EPN3 expression is also implicated in the regulation of the TGF β pathway and E-Cadherin endocytosis, indicating that EPN3 function upon overexpression is an exaggeration of its physiological function. In vivo, we generated an EPN3 inducible knock-in mouse, which overexpresses the transgene specifically in epithelial cells. High EPN3 levels did not cause significant differences during mammary gland development in puberty, but instead increased secondary and tertiary branching morphogenesis in adult virgin mice, increasing N-Cadherin expression in the epithelial ducts, supporting in vitro findings in MCF10A cells.

Together, these data show that EPN3 is a novel prognostic factor for breast cancer metastasis acquisition, acting directly on E-Cadherin endocytosis and turn-over, resulting in a pEMT highly plastic and dynamic and that could recapitulate also branching alterations in mouse mammary gland. Moreover, preliminary evidences suggest that the upstream endocytic event is able to induce a transcriptional activation that becomes independent of the first genetic lesion, a possibility that would have a profound impact in breast cancer treatment.

TABLE OF CONTENT

ABSTRACT	5
TABLE OF CONTENT	7
TABLE OF FIGURES	10
ABBREVIATIONS	12
1. INTRODUCTION	17
1.1. MECHANISMS OF ENDOCYTOSIS	18
1.1.1. CLATHRIN-MEDIATED ENDOCYTOSIS: THE CLATHRIN-COATED VESICLE FORMATION CYCLE	18
1.1.2. ENDOCYTOSIS IN THE REGULATION OF SIGNALING AND CELL RESPONSE	21
1.1.3. ROLE OF ENDOCYTOSIS IN PHYSIOLOGY	23
1.1.4. ROLE OF ENDOCYTOSIS IN CANCER	24
1.2. THE EPSIN FAMILY OF CLATHRIN ENDOCYTOIC ADAPTORS	26
1.2.1. EPSIN PROTEIN STRUCTURE	27
1.2.2. EPSIN FUNCTIONS	29
1.2.3. EPSINS IN PHYSIOLOGY	33
1.2.4. EPSINS IN PATHOLOGY	34
1.2.5. EPSIN3	35
1.3. EPITHELIAL CELL-CELL JUNCTIONS: ENDOCYTOSIS AS A MECHANISM TO REGULATE JUNCTION STABILITY AND CELL MIGRATION	37
1.3.1. CELL-CELL AND CELL- EXTRACELLULAR MATRIX JUNCTIONS	37
1.3.2. E-CADHERIN	41
1.3.3. E-CADHERIN FUNCTIONAL REGULATORS: THE CATENIN PROTEINS	42
1.3.4. AJ REGULATION: AN ENDOCYTOIC PERSPECTIVE	48
1.4. EPITHELIAL-MESENCHYMAL TRANSITION: OVERVIEW IN MAMMARY GLAND PHYSIOLOGY AND BREAST CANCER	50
1.4.1. EMT: BASIC MECHANISMS AND MOLECULAR PLAYERS	51
1.4.1.1. THE TGFB PATHWAY	54
1.4.2. EMT IN PHYSIOLOGY	58
1.4.2.1. BRANCHING MORPHOGENESIS IN MAMMARY GLAND: THE ROLE OF EMT	59
1.4.3. EMT IN PATHOLOGY	62
1.4.3.1. EMT IN BREAST CANCER: FROM DCIS TO INVASIVE CARCINOMA	65
1.5. IDENTIFICATION OF EPN3 AS A NOVEL PROGNOSTIC FACTOR IN BREAST CANCER INVASIVENESS AND METASTASIS	68
2. RATIONALE AND AIMS	75
3. RESULTS	79
3.1. EPN3 OVEREXPRESSION INDUCES A PARTIAL EMT PROGRAM IN MAMMARY EPITHELIAL MCF10A CELLS	79
3.2. REGULATION OF EPN3-INDUCED PARTIAL EMT THROUGH E-CADHERIN ENDOCYTOSIS AND ADHERENS JUNCTION STABILITY	82
3.2.1. SETTING UP OF AN EPN3-FLAG IMMUNOPRECIPITATION TO PERFORM MASS-SPECTROMETRY ANALYSES	82

TABLE OF CONTENT

3.2.2. EPN3-SPECIFIC INTERACTOME REVEALS AN ENRICHMENT OF COMPONENTS OF THE ENDOCYTIC MACHINERY AND ADHERENS JUNCTIONS	85
3.2.3. SETTING UP OF CO-IMMUNOPRECIPITATION STRATEGY TO VALIDATE MASS-SPECTROMETRY RESULTS	89
3.2.4. EPN3 CO-IMMUNOPRECIPITATES WITH CLATHRIN, E-CADHERIN AND AJ COMPONENTS	92
3.2.5. EPN3 OVEREXPRESSION ALTERS E-CADHERIN STABILITY AT THE PLASMA MEMBRANE	93
3.2.6. EPN3 OVEREXPRESSION INCREASES E-CADHERIN ENDOCYTOSIS	96
3.2.7. EPN3 OVEREXPRESSION INDUCES THE ACTIVATION OF B-CATENIN/TCF4 PATHWAY	97
3.3. EPN3-INDUCED PARTIAL EMT GENERATE A TGFB-DEPENDENT POSITIVE FEEDBACK LOOP	100
3.3.1. EPN3-INDUCED PARTIAL EMT GENERATE A TGFB-DEPENDENT POSITIVE FEEDBACK LOOP	100
3.3.2. EPN3 OVEREXPRESSION SYNERGIZES WITH TGFB SIGNALING	101
3.3.3. EPN3 OVEREXPRESSION INCREASES E-CADHERIN ENDOCYTOSIS UPON TGFB STIMULATION	104
3.3.4. EPN3-INDUCED PARTIAL EMT GENERATES A TGFB-DEPENDENT POSITIVE FEEDBACK LOOP INDEPENDENT OF EPN3 ITSELS	106
3.3.5. PHYSIOLOGICAL EPN3 REGULATES TGFB-DEPENDENT TRANSCRIPTIONAL ACTIVATION OF EMT GENES	108
3.3.6. PHYSIOLOGICAL EPN3 REGULATES E-CADHERIN ENDOCYTOSIS	110
3.4. CHARACTERIZATION OF EPN3 FUNCTION IN MOUSE MAMMARY GLAND BIOLOGY	112
3.4.1. EPN3 EXPRESSION IN THE MAMMARY GLAND OF KNOCK-IN MICE AND KNOCK-OUT MICE	114
3.4.2. EPN3 HAS A ROLE IN MOUSE MAMMARY DUCT BRANCHING	116
3.4.3. EPN3 KI MICE SHOW AN INCREASED N-CADHERIN AND KI67 EXPRESSION IN THE BRANCHING DUCTS	122
4. DISCUSSION	125
4.1. EPN3 OVEREXPRESSION INDUCES A PARTIAL EMT PHENOTYPE IN MCF10A CELLS	125
4.2. EPN3 OVEREXPRESSION INDUCES A TGFB-DEPENDENT SELF-SUSTAINING POSITIVE-FEEDBACK LOOP THAT BECOMES INDEPENDENT OF THE EPN3 INITIAL ALTERATION	127
4.3. EPN3 INTERACTS WITH THE ADHERENS JUNCTIONS AND THE CLATHRIN MACHINERY	128
4.4. EPN3 PHYSIOLOGICAL FUNCTION IN THE REGULATION OF E-CADHERIN ENDOCYTOSIS	134
4.5. EPN3 ALTERS MOUSE MAMMARY GLAND BRANCHING MORPHOGENESIS	134
4.6. FUTURE PERSPECTIVES	136
4.7. CONCLUSIONS AND PROPOSED MODEL	139
5. MATERIALS AND METHODS	141
5.1. SOLUTIONS AND BUFFERS	141
5.2. CELL CULTURE	145
5.2.1. MCF10A CELL CULTURE	145
5.2.1.1. TREATMENTS: TGFB1 AND FA 1%	146
5.2.1.2. RNA INTERFERENCE (RNAi)	146
5.2.2. MOUSE EPITHELIAL CELL CULTURE	147
5.3. CELL-BASED ASSAYS	149
5.3.1. FLUORESCENCE ACTIVATED CELL SORTING (FACS)	149
5.3.1.1. PLASMA MEMBRANE E-CADHERIN INTENSITY	150
5.3.1.2. E-CADHERIN INTERNALIZATION ASSAY	151
5.3.2. IMMUNOFLUORESCENCE (IF)	152
5.3.2.1. PLASMA MEMBRANE E-CADHERIN INTENSITY	152
5.3.2.2. E-CADHERIN INTERNALIZATION ASSAY	152
5.3.2.3. STAINING IN GROWING CONDITIONS	152

TABLE OF CONTENT

5.4. RNA-BASED ASSAYS	153
5.5. PROTEIN-BASED ASSAYS	154
5.5.1. CELL LYSIS AND WESTERN BLOT (WB)	154
5.5.2. IMMUNOPRECIPITATION (IP) AND CO-IP	155
5.5.3. MASS SPECTROMETRY ANALYSES	157
5.6. MICE HANDLING	159
5.7. MICE-BASED ASSAYS	159
5.7.1. HISTOLOGICAL PROCEDURES	159
5.7.2. IMMUNOHISTOCHEMISTRY (IHC)	160
5.8. STATISTICAL ANALYSES	161
5.9. SUPPLEMENTARY TABLES	162
REFERENCES	165
ACKNOWLEDGEMENTS	195

TABLE OF FIGURES

Figure 1: Assembly and disassembly of an endocytic clathrin-coated pit (CCP)	21
Figure 2: Structure of different members of the Epsin family	28
Figure 3: Epsin-mediated endocytosis of ubiquitinated cell surface receptors	30
Figure 4: Cell junctions typical of normal epithelial cells	40
Figure 5: E-cadherin structure and modes of interaction	42
Figure 6: β -catenin/TCF signaling pathway	44
Figure 7: Isoform-specific role of p120	47
Figure 8: p120 regulated cadherin endocytosis	50
Figure 9: EMT and epithelial plasticity	52
Figure 10: TGF β receptor signaling and the EMT program	57
Figure 11: Mouse postnatal stages in mammary gland development	61
Figure 12: The metastasis cascade	64
Figure 13: Pathobiological events associated with ductal carcinoma in situ	67
Figure 14: EPN3 is overexpressed in breast cancer patients	70
Figure 15: EPN3 overexpression in MCF10A cells induces EMT phenotype	73
Figure 16: Transcriptional expression of EMT markers in MCF10A cells following EPN3 and TWIST expression	81
Figure 17: Optimization of the anti-FLAG immunoprecipitation reaction in MCF10A-EPN3 cells	84
Figure 18: EPN3 interacts with members of endocytic machinery and AJs as determined by mass spectrometry	88
Figure 19: Optimization of EPN3 immunoprecipitation following crosslinking of protein lysates	91
Figure 20: EPN3 co-immunoprecipitates with components of endocytic machinery and adherens junctions	93
Figure 21: EPN3 destabilizes plasma membrane E-Cadherin and adherens junctions	95
Figure 22: E-Cadherin endocytosis is accelerated upon EPN3 overexpression	97
Figure 23: EPN3-induced pEMT is dependent on the β -catenin/TCF4 pathway	99
Figure 24: The effect of EPN3 upregulation on the transcriptional expression of TGF β ligands and receptors	101
Figure 25: EPN3-overexpression sensitizes MCF10A cells to TGF β -induced EMT	103
Figure 26: EPN3 overexpression cooperates with TGF β in inducing E-Cadherin endocytosis	105
Figure 27: EPN3-induced pEMT generates a TGF β -dependent positive-feedback loop not addicted to EPN3 itself	107
Figure 28: Endogenous EPN3 expression is implicated in activation of the TGF β pathway	109
Figure 29: Endogenous EPN3 mediates E-Cadherin endocytosis induced by TGF β stimulation	111

Figure 30: Representation of the strategy used to obtain FVB/EPN3 Knock-In (KI) and Knock-Out (KO) mice	113
Figure 31: Evaluation of Epn3 expression and localization in mammary gland of EPN3 KO, WT and KI mice	115
Figure 32: EPN3 KO mice show a delay in epithelial duct spreading into the fat pad at 6 weeks of age	118
Figure 33: EPN3 KI mice show a sporadic alteration of secondary branching at 12 weeks of age	119
Figure 34: EPN3 KI mice display an increased secondary duct branching at 18 weeks of age	121
Figure 35: EPN3 KI mice show an increased number of positive cells for N-Cadherin and Ki67 at 18 weeks of age	123
Figure 36: Alignment of EPN3 and EPN1 protein sequences	132
Figure 37: Working model of EPN3-dependent induction of pEMT in breast cancer	140

ABBREVIATIONS

AJ	Adherens junction
AP2	Adaptor protein 2
APC	Adenomatous polyposis coli
ATCC	American type culture collection
AW	Acid wash
BABB	Benzylalcohol : Benzylbenzoate
BC	Breast cancer
BMP	Bone morphogenetic proteins
CBD	Catenin binding domain
CCP	Clathrin-coated pit
CCV	Clathrin-coated vesicles
Cdc42	Cell division control protein 42 homolog
CDH1	E-Cadherin gene
CDH2	N-Cadherin gene
CHC	Clathrin heavy chain
CK	Cytokeratin
CK1 α	casein kinase 1 alpha
CLC	Clathrin light chain
CME	Clathrin-mediated endocytosis
CSC	Cancer stem cell
DAPI	Diazabicyclo-(2.2.2)octane antifade
DCIS	Ductal in situ carcinoma
Dll4	Delta like protein 4
DPW	Aspartate-Proline-Tryptophan
EC	Extracellular cadherin
ECM	Extracellular matrix
EEC	Endo/exocytosis cycle
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal transition
EndMT	Endothelial-to-mesenchymal transition
ENTH	Epsin-N-terminal homology domain
EpCAM	Epithelial cell adhesion molecule

EPLIN	Epithelial protein lost in neoplasm
EPN1	Eps-15 interacting protein1
EPN2	Eps-15 interacting protein2
EPN3	Eps-15 interacting protein3
ER	Estrogen
EV	Empty vector
E-Cadherin	Epithelial-cadherin
FA	Formaldehyde
FACS	Fluorescence activated cell sorting
FDR	False discovery rate
FFPE	Formalin-fixed paraffin-embedded
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FISH	Fluorescent in situ hybridization
FSC	Forward scatter
GAP	GTPase-activating proteins
GDF	Growth differentiation factors
GSK3- β	Glycogen synthase 3 beta
HER2	Epidermal growth factor 2
HGF	Hepatocyte growth factor
HGFRMet	Hepatocyte growth factor receptor
IBD	Invasive breast cancer
IF	Immunofluorescence
IGF1	Insulin-like growth factor 1
IHC	Immunohistochemistry
IP	Immunoprecipitation
JMD	Juxtamembrane domain
KD	Knock down
KEGG	Kyoto encyclopedia of genes and genomes
KI	Knock in
KO	Knock out
LqF	Liquid facets
MAGUK	Membrane-associated guanylate kinases
MEC	Mouse epithelial cell
MET	Mesenchymal-to-epithelial transition
mRNA	messenger RNA
miRNA	micro RNA

ABBREVIATIONS

siRNA	small interfering RNA
MDCK	Madin-darby canine kidney
MEF	Mouse embryonic fibroblast
MMP	Matrix metalloproteinases
MVB	Multivescicular bodies
NCE	Non-clathrin-mediated endocytosis
NECD	Notch extracellular domain
NICD	Notch intracellular domain
NPF	Asparagine-Proline-Phenylalanine
NSG mice	NOD (non-obese diabetic) - SCID (severe combined immunodeficiency) IL2RG null mice
N-Cadherin	Neural cadherin
ON	overnight
PBS	Phosphate-buffered saline
PDAC	Pancreatic ductal adenocarcinoma
PFA	Paraformaldehyde
pEMT	partial EMT
PM	Plasma membrane
PR	Progesterone
PtdIns(4,5)P ₂	Phosphatidylinositol-4,5-bisphosphate
RalBP1	Ral-binding partner 1
RIPA	Radioimmunoprecipitation assay
RNAi	RNA interference
ROS	Reactive oxygen species
RT	Room temperature
RTK	Receptor tyrosine kinase
RT-qPCR	Reverse transcriptase-quantitative polymerase chain reaction
SC	Stem cell
SDS-PAGE	Sodium dodecyl sulfate – polyacrylamide gel electrophoresis
SMAD	Small mother against decapentaplegic
SSC	Side scatter
TACE	Tumor necrosis factor- α -converting enzyme
TAK1	TGF β -activated kinase 1
TBS	Tris-buffered saline
TCF/LEF	T cell factor / lymphoid enhancer-binding factor
TCGA	The cancer genome atlas
TEB	Terminal end bud

Tf	Transferrin
TfR	Transferrin receptor
TGF β	Transforming growth factor β
TGF β R	Transforming growth factor β receptor
TJ	Tight junction
TNBC	Triple-negative breast cancer
TRAF6	Tumor necrosis factor receptor-associated factor 6
UIM	Ubiquitin interacting motifs
VEGF	Vascular endothelium growth factor
VEGFR	Vascular endothelium growth factor receptor
VHS	VPS-27 Hrs STAM
Vps27	Vacuolar protein sorting-associated protein 27
WB	Western blot
ZEB	Zinc-finger E-box-binding
ZO-1	Zonula adherens 1
ZO-2	Zonula adherens 2
ZO-3	Zonula adherens 3

1. INTRODUCTION

During my PhD, I worked on a project focused on the characterization of the function of a novel putative pro oncogenic protein, Epsin3 (EPN3), in breast cancer (BC) development and progression. This introduction aims to summarize the most relevant literature related to the different topics connected to my thesis, ranging from endocytosis to cell-cell junctions, and epithelial-to-mesenchymal transition (EMT) to mammary gland development and morphology.

One of the turning points in the (r)evolution of life was the acquisition of the ability of different cells to cooperate with each other and gain specific functions. In this context, the possibility of cells to physically interact with each other, create junctions and sense the microenvironment marks a major biological breakthrough. Nowadays, scientists are trying to understand how cells are connected to each other and why dysregulation of the processes that allow cell interaction can contribute to diseases, included cancer.

Adherens junctions (AJs) are the most studied cell-cell junctions in epithelial sheets, responsible for a wide range of functions. AJ turnover and plasma membrane (PM) stability are strongly controlled and many indications link the dysregulation of this control to cancer, in particular, to invasiveness and metastasis. Besides transcriptional regulation, protein synthesis and degradation, one of the main and more dynamic mechanisms of regulation of PM composition and of AJ turnover is exerted by the endocytic process. Thus, unraveling the mechanisms of AJ endocytosis could contribute to our understanding of how cells communicate and interact with each other, and how dysregulation of this process could contribute to cancer.

1.1. MECHANISMS OF ENDOCYTOSIS

Endocytosis is the process of intracellular uptake of molecules through the PM¹. Many routes are known to allow this process, which are characterized by the dimension and the type of the cargo, the membrane tension, and the type of endocytic machinery required. For example, phagocytosis is necessary to allow internalization of cargos bigger than 500 nm, such as bacteria, and does not require a membrane coat². Macro-pinocytosis, on the other hand, is implicated in uptake of fluidic molecules from the extracellular space and again does not require a coat³. Clathrin-mediated endocytosis (CME) is defined as the main endocytic route of PM receptors, involving the formation of small membrane invaginations (150-200 nm) coated with clathrin, called clathrin-coated pits (CCPs)⁴⁻⁷. Although CME is the best characterized mechanism of receptor endocytosis, there is increasing evidence of alternative clathrin-independent routes, i.e., caveolae-mediated^{8,9}, endoplasmic reticulum-mediated¹⁰, actin-mediated¹¹, flotillin-mediated^{12,13}, and so on (for a detailed review see¹⁴).

1.1.1. Clathrin-mediated endocytosis: the clathrin-coated vesicle formation cycle

More than 50 molecular players are implicated in the CME pathway, the best characterized being adaptor protein 2 (AP2)¹⁵, which directly links PM receptors to clathrin. Other endocytic proteins participate in the various phases of CCP formation at the PM and vesicle release in the cytosol, including Eps15¹⁶, Epsins¹⁷, FCHO1,2¹⁸, Dynamin^{19,20}, and Auxilin²¹. So far, the evidence

acquired suggests that CME can be divided into 5 steps and that the huge number of players could act as modules of protein complexes^{22,23} (Figure 1):

- i. Initiation. This is the first step in the formation of the CCP. At this step, clathrin and AP2 are concomitantly recruited to the PM, inducing an initial bending of PM. Experimental evidence suggests that the nucleation site corresponds to places with high concentrations of phosphatidylinositol-4,5-bisphosphate (PtdIns(4,5)P₂), a particular type of PM phospholipid that is enriched at the PM and in endocytic vesicles²⁴.
- ii. Growth. This step is highly connected to the functions of clathrin. The clathrin molecule is composed of 2 chains, one heavy (CHC) and one light (CLC). Normally 3 CHCs and 3 CLCs assemble to form the so-called triskelion, a trimeric complex of heavy and light chains²⁵. During CCP formation, the triskelia assemble onto the membrane producing hexagons and pentagons, creating the clathrin lattice. The function of the coat is to maintain the bending of the membrane and allow the correct formation of the pit²⁶.
- iii. Stabilization and cargo selection. The third step in CCP formation confers cargo specificity to the nascent CCP. Here, the most important component is AP2¹⁵, which presents as a heterotetramer able to bind directly PtdIns(4,5)P₂ but also to recognize cargos, both directly through specific motifs in the receptor tail, and indirectly through other adaptor proteins^{27,28}. Indeed, several adaptor proteins have a fundamental role at this stage of CME, participating in membrane bending and conferring specificity to cargo selection. For example, Dishvelled recruits Frizzled after Wnt ligand stimulation⁵, Numb

recruits Notch ²⁹, and Epsins recruit monoubiquitinated receptors, such as epidermal growth factor receptor (EGFR) ^{30,31}. It is also known that the presence of the cargo itself is required to complete CCP formation, since if the amount of cargo included in the nascent CCP does not reach a certain threshold, CCPs disassemble and are not released in the cytosol (the so-called 'abortive CCPs') ³², or their formation is delayed ³³.

- iv. Budding and scission. The critical protein in this step is dynamin, a GTPase enzyme able to perform membrane fission, assembling at the neck of the CCP (for a review on the mechanisms of dynamin action, see ¹⁹).
- v. Uncoating. Once the pit is formed, the coating is no longer required and has to be disassembled in order to be ready for the formation of other CCPs. The most relevant molecular players here are synaptojanin and auxilin. It is known that – at least in neurons – one of the first steps in uncoating is the dephosphorylation of PtdIns(4,5)P₂ into PtdIns(4)P₁ by the phosphatase synaptojanin ³⁴. After this dephosphorylation, auxilin binds to clathrin and recruits Hsc70, an ATPase that detaches the triskelia ³⁵.

These steps lead to the correct formation and internalization of the CCP, which can enter into the endosomal compartment, where the cargos can be recycled back to the PM, degraded through the lysosome compartment, or retro-transported to the Golgi compartment.

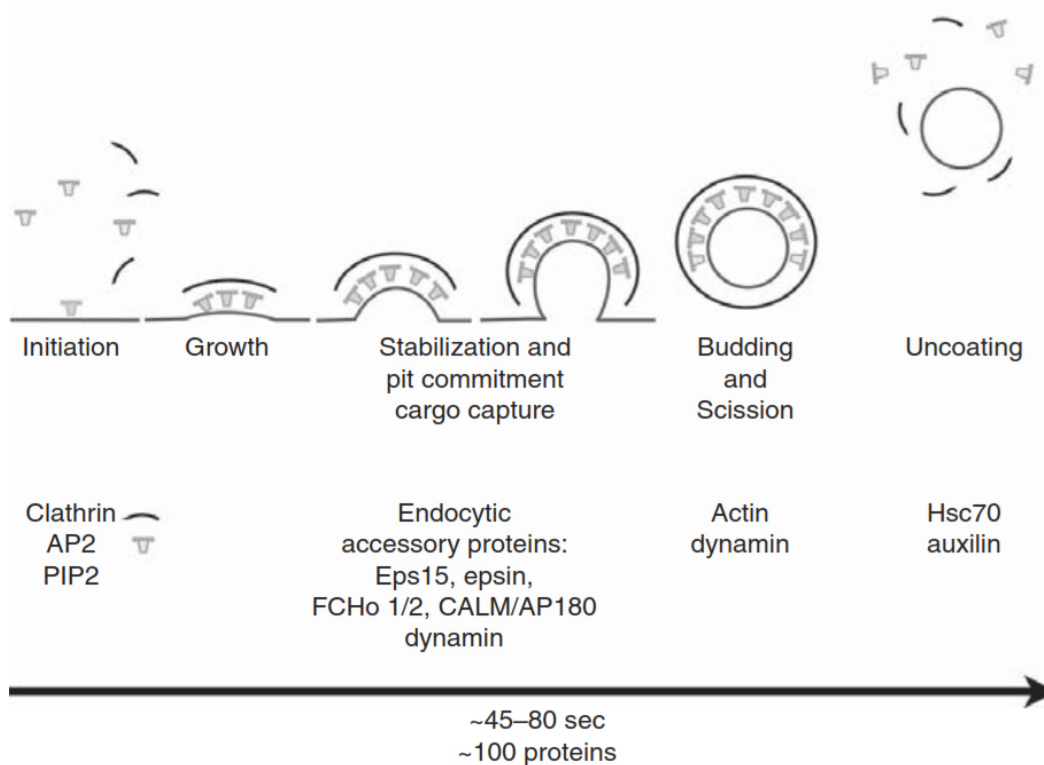


Figure 1: Assembly and disassembly of an endocytic clathrin-coated pit (CCP)

Initiation of the CCP occurs with recruitment of AP2 adaptor complexes, associated at the plasma membrane with PtdIns(4,5)P₂, that recruits clathrin triskelions. Stable growth of the CCP requires endocytic accessory proteins (Eps15, epsin, FCHo1/2, intersectin, CALM/AP180). These adaptors confer also cargo specificity to the nascent CCP. Dynamin, assisted by actin polymerization when the membrane is under tension, drives membrane scission and coated-vesicle release. Hsc70, recruited by auxilin, mediates clathrin uncoating and release of a free vesicle. Text beneath the diagram indicates the overall timescale and the stages at which various components appear to function. Figure taken from ⁶.

1.1.2. Endocytosis in the regulation of signaling and cell response

Although endocytosis has been historically ascribed as a mere mechanism necessary for cells to “eat and feed”, over the last decades this classical concept has widely changed. Endocytosis is the connection between extracellular and intracellular processes and it is critical in the regulation of receptor-mediated signaling, acting at multiple levels:

- At the PM. For example, Receptor Tyrosine Kinases (RTKs), such as EGFR, are able to induce a signaling cascade upon ligand recognition that can be regulated through receptor

availability at the membrane. Indeed, internalization of EGRF through CME sustains recycling at the PM and signaling ³⁶, while endocytosis through other clathrin-independent pathways, direct the receptor to degradation ¹⁰. Similar regulation has been observed also for transforming growth factor beta receptors (TGFβRs), a family of serine/threonine receptor kinases, where the internalization through CME sustains pathway activation, while caveolae-dependent endocytosis promotes receptor degradation ³⁷.

- Endosomal and intracellular trafficking. Endosomes are often used as signaling stations after internalization. For example, activation of MAP kinase by the EGFR is known to be sustained and to become productive in the endosomes after receptor internalization ^{38,39}. Similarly, the TGFβR continues to activate its signaling pathway even after endocytosis through CME: trafficking to the SARA-endosomes allows TGFβR to phosphorylate small mother against decapentaplegic 2 (SMAD2), which then forms a complex with SMAD4 and translocates to the nucleus where it regulates transcription of target genes (for detailed information of TGFβ / SMAD pathway see section 1.4.1.1) ⁴⁰.
- Motility and migration. Cell migration requires continual rearrangement of membrane composition and plasticity, so it is not surprising that endocytosis, particularly CME, is fundamental to this process. Specifically, migration requires rapid and constant membrane remodeling, which is achieved by highly regulated endo/exocytosis cycles (EECs) of key cargoes ⁴¹. Indeed, continuous EECs of members of the integrin family of

transmembrane adhesion receptors, such as integrin $\alpha 5\beta 1$ ⁴² and integrin $\alpha 5\beta 3$ ⁴³, allows cell movement in the surrounding microenvironment.

- Mechanical stress. Over the last decade, huge efforts have been directed to understanding how cells respond to mechanical stress. The regulation of membrane composition and turnover through endocytosis plays a critical role in this process. In 2011, the Lamaze group showed that caveolae are critical mechanical sensors that are able to respond to mechanical stress by rapid assembly/disassembly, therefore, providing the reservoir of membrane needed under mechanical perturbations⁴⁴. Interestingly, the regulation of CME is highly dependent on membrane tension: the stronger the tension, the more difficult it is for clathrin and its adaptors to bend the membrane, and – as a consequence – the action of the actin cytoskeleton becomes critical to have productive endocytosis^{11,45}. Importantly, many aspects of the interplay between mechanical stress and endocytosis are still unclear; for example, if, when, and how endocytosis is able to regulate signaling pathways in response to physical stress.

1.1.3. Role of endocytosis in physiology

The fact that endocytosis is involved in so many different cellular functions, justifies its critical role at the organismal level. Indeed, all loss-of-function mutations of clathrin and/or the major clathrin adaptor proteins results in embryo lethality: knockdown (KD) or knockout (KO) of CHC in *Drosophila melanogaster*⁴⁶ and *Caenorhabditis elegans*⁴⁷ is embryonic lethal, while AP2 is essential for mouse early embryo development⁴⁸. Moreover, endocytic homeostasis is important

in adulthood. Indeed, the uptake of albumin⁴⁹ or immunoglobulins⁵⁰ across the epithelium of lactating mouse mammary glands is tightly regulated through CME. Moreover, PICALM critically regulates transferrin (Tf) uptake in erythroid cells because it functions as a cell type-specific regulator of endocytosis of the transferrin receptor (TfR)⁵¹. Furthermore, the vast majority of receptors and adhesion molecules are internalized through CME and are responsible for many cellular and tissue phenotypes, such as TfR in iron homeostasis⁵² and LDLR in regulation of cholesterol-rich LDL particles⁵³, but also EGFR⁵⁴, Frizzled⁵, TGFβR³⁷, Cadherins^{55,56} in embryo development and tissue homeostasis.

1.1.4. Role of endocytosis in cancer

Aberrations of the endocytic program can lead to almost all types of diseases, from genetic to neurological, autoimmune to metabolic, infectious to cancer. This paragraph aims to briefly summarize and discuss alterations of internalization and trafficking programs in cancer. Given the critical role of endocytosis in cellular homeostasis, it is not surprising that alterations of endocytic factors involved in all the different steps of the endocytic process, from CCP formation to endosomal sorting, have been linked to many cancers⁵⁷. Aberrant endocytosis could impinge on many cancer-related phenotypes, in particular:

- Proliferative signaling. Alterations of endocytic players implicated in the regulation of receptor signaling are often observed in cancers. For example, the E3 ubiquitin ligase, Cbl, which is frequently mutated in human tumors, regulates ubiquitination, endocytosis and lysosomal degradation of RTKs, including hepatocyte growth factor receptor

(HGFR/Met) and EGFR⁵⁸. Cbl-mediated ubiquitination is critical to achieve long-term attenuation of signaling and its alterations cause enhanced and prolonged proliferative response. A similar phenotype is observed with mutant receptors that are inefficiently internalized, ubiquitinated and/or sorted to lysosomal degradation, resulting in aberrant signaling, as in the case of EGFRvIII, a mutant form of EGFR often found in glioblastoma that is less degraded meaning that it can sustain pathway activation⁵⁹.

- Asymmetric/symmetric division and stemness. The balance between asymmetric and symmetric division is tightly regulated by endocytosis and alterations of the machinery controlling this balance can lead to increased stem cell (SC) content, a property intimately linked to tumorigenesis^{60,61}. This is the case of Numb, an endocytic protein that regulates asymmetric cell division of mammary SCs and homeostasis of the SC compartment by acting mainly through the tumor suppressor p53 and the oncogene Notch^{60,62,63}. In BC, Numb loss leads to the concomitant loss of p53 activity and Notch activation, which culminates in expansion of the SC compartment, the acquisition of cancer stem cell (CSC) traits, and resistance to classical chemotherapy^{60,61}.
- Migration and motility. Alterations in migration and adhesion processes are fundamental for cancer invasiveness and metastasis acquisition. Cell adhesion and migration are finely regulated at many different levels; the most relevant players in these process are the junctions, both cell-cell (AJs and tight junctions) and cell-extracellular matrix (ECM; focal adhesions). The turnover and specificity of the junctions defines how a cell will behave in the space. As expected, endocytosis and its dysfunction have profound impacts on these

processes. For example, alterations in integrin composition at the PM and turnover, via altered EECs, have been observed in cancer ^{57,64}. In agreement Rab25, a GTPase that regulates integrin endosomal trafficking, has been found to be overexpressed in hepatocellular cancer, leading to anchorage-independent proliferation ^{65,66}. Cadherins also play a prominent role in cancer and we will dedicate an entire section to this topic (Section 1.3).

1.2. The Epsin family of clathrin endocytic adaptors

Epsins are a family of endocytic adaptor proteins. The first description of an epsin protein was made in 1998, when it was described as an Eps15 interactive protein, binding to the EH domain of Eps15, an already known endocytic adaptor protein ⁶⁷. Additional studies revealed that the epsin family is composed of 3 paralogs: epsin1 (EPN1) , epsin2 (EPN2) ⁶⁸ and EPN3 ⁶⁹. After 20 years, almost 300 scientific publications (PubMed data) have focused their attention on the Epsin family, unveiling a strong interest in understanding the function of this protein family. In this section, the most important features and discoveries regarding the functions of Epsin family members, in physiology and in disease, are described.

1.2.1. Epsin protein structure

Epsins share a high degree of structural similarity. They are characterized by the presence of an Eps15 N-terminal Homology (ENTH) domain highly conserved at the N-terminus of the protein, followed by 2 ubiquitin-interactive motif (UIM) motifs, 1-2 clathrin-binding motifs, and 3-8 Asp-Pro-Trp (whose aminoacid abbreviation is: DPW) motifs and 3 Asn-Pro-Phe (NPF) motifs at the C-terminus of the protein (Figure 2)⁷⁰⁻⁷². In detail:

- ENTH domain. This domain is a very highly conserved domain, first identified in EPN1⁶⁷. Its main role is to bind the phospholipid PtdIns(4,5)P₂ that is enriched at the PM and in endocytic vesicles⁷⁰. In 2000, the crystal structure of the EPN1 ENTH domain was resolved, showing that it is composed of 7 well-structured α -helices and another α -helix that is induced by epsin binding to the phospholipids⁷³. The most similar domain to the ENTH is the AP180 N-terminal Homology ANTH domain found in adaptor protein 180 AP180, belonging to another class of endocytic adaptor proteins¹⁷; however, there are similarities also with other well characterized domains, such as the VPS-27 Hrs STAM (VHS) domain of vacuolar protein sorting-associated protein 27 (Vps27) protein⁷⁴ or *armadillo* repeat of β -catenin^{70,75}.
- UIM. This motif is present in two or three copies in the various epsins and it is responsible for the binding between epsins and ubiquitinated cargos.
- Clathrin-binding motifs. These motifs are responsible for the recruitment and binding with clathrin triskelia and are flanked by AP2-binding motifs.

- DPW motifs. These motifs are responsible for the binding with AP2, the major clathrin binding partner.
- NPF motifs. These motifs are responsible for the binding with the EH domain of Eps15, Intersectin, and POB1.

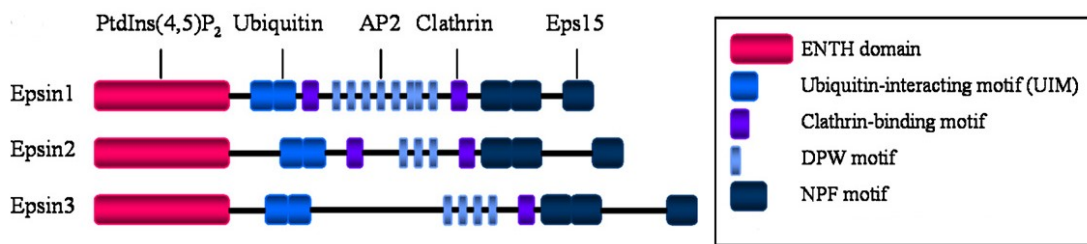


Figure 2: Structure of different members of the Epsin family

Schematic representation of epsins 1, 2 and 3 proteins. They share similar protein structure, bearing an almost identical N-terminal ENTH domain. Immediately after this domain, several ubiquitin-interacting motifs (EEElqLqAlamSkE) are situated, responsible for interaction with ubiquitinated proteins. The central region of epsin has different DPW motifs that mediate binding to AP2. These motifs are flanked by clathrin-binding motifs (LLDLD/LIELE) involved in clathrin binding. The C-terminal part of these proteins comprises NPF repeats that interact with EH domain-bearing proteins, such as Eps15 and intersectin. Adapted from ⁷¹.

The protein sequence of the epsins is highly conserved among species, with many homologues discovered along the phylogenetic tree. In *Saccharomyces cerevisiae*, two epsins, named Ent1 and Ent2, have been found as homologs of the human Epsin family; they have been characterized for their role in CME and actin regulation, and the absence of both the genes is lethal ⁷⁶. In *Xenopus laevis*, one protein closely related to epsins was identified, MP90 ⁷⁷. In *Drosophila melanogaster*, the protein liquid facets (LqF) was identified as an orthologue of epsins ⁷⁸.

1.2.2. Epsin functions

Many functions of the Epsin family have been discovered. Almost everything discovered about this family came from studies on EPN1, since it was the first epsin to be discovered and it is ubiquitously expressed. In addition, EPN1 and EPN2 seem to share redundant roles in endocytosis and development ⁷⁹. Thus, the functions that I describe here refer to EPN1, if not differently specified.

It is now widely accepted that epsins have a role in the first steps of endocytosis acting as an adaptor protein. Interestingly, researchers have been able to link epsin function to specific pathways depending on the cellular context, although the molecular mechanisms governing this pathway-specific function are still to be elucidated ⁷².

- Epsin role in CME. In the first decade after the discovery of epsins, much effort was placed on unraveling their mechanism of action. An important advancement came with the discovery of the EPN1 ENTH structure ⁷³ and with the bioinformatics characterization of the domains and motifs found within the epsin sequence. Indeed, the function of epsins is to connect the 3 key components of the endocytic process: the PM bilayer via the binding of the ENTH domain to PtdIns(4,5)P₂, ubiquitinated cargo via UIMs, and the clathrin machinery, including Clathrin, AP2 and Eps15 via the clathrin-binding motifs, DPW, and NPF, respectively (Figure 3) ^{70,71}. In particular, it has been proposed that epsins could act after the nucleation step (Figure 1), contributing to cargo selection and coat assembly ^{22,23,70,71}.

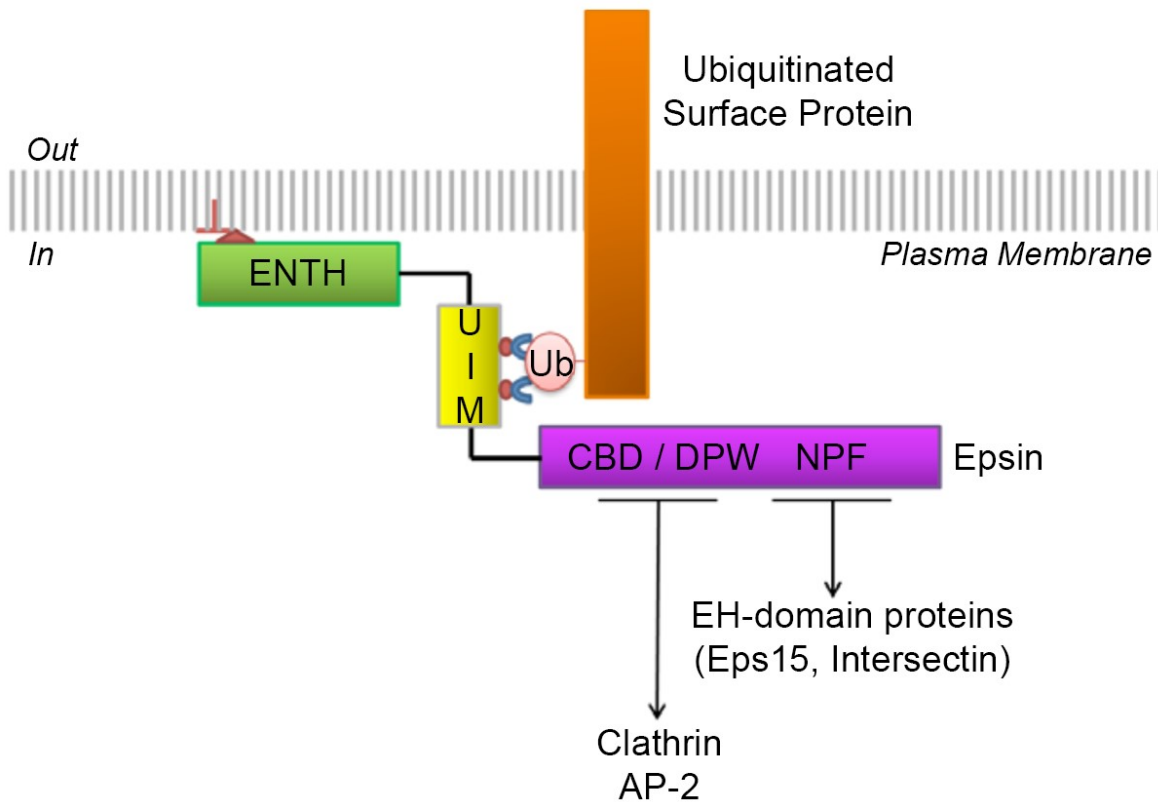


Figure 3: Epsin-mediated endocytosis of ubiquitinated cell surface receptors

ENTH, epsin N-terminal homology domain; UIM, ubiquitin interacting motif; CBD and DPW central region housing clathrin- anAP-2-binding domains, respectively; NPF, asparagine-proline-phenylalanine containing C-terminal region housing Eps-15 homology (EH) protein-binding domains; Ub, ubiquitin. Adapted from ⁸⁰

- The Notch pathway. One of the firsts pathways shown to be regulated by epsins was the Notch pathway. The Notch signaling pathway is a highly conserved pathway among species, fundamental for embryo development and cell fate ^{81,82}. The Notch receptor is activated through cleavage and the internalization of its Notch intracellular domain (NICD) in the signal-receiving cell; this event occurs upon binding of the Notch receptor extracellular domain (NECD) with Notch ligand, a transmembrane protein on the signal-sending cell, one of the most studied being delta like protein 4 (Dll4) ⁸³. When ligand and receptors on opposing cells are bound, proteolytic cleavage of the NECD and subsequent

internalization of Dll4 and NECD in the signal-sending cell, allows the endocytosis of NICD in the signal-receiving cell and activation of the pathway. It is yet not clear the exact role of epsins in this process, although it is likely that its main function is recognition of ubiquitinated Dll4 ligand in the signal-sending cell, allowing internalization of the Dll4:NECD complex⁸⁴⁻⁸⁶.

- The EGFR pathway. EGFR belongs to the ErbB family and it is one of the best characterized members because of its relevance in cancer. Indeed, EGFR is frequently overexpressed and/or mutated in cancer, leading to aberrant signaling and tumorigenesis⁸⁷. Moreover, the EGFR pathway offers opportunities for targeted anti-cancer therapies⁸⁸. EGFR is an RTK that homodimerizes or heterodimerizes with other other ErbB family members. Upon ligand binding, transphosphorylation occurs leading to receptor activation and triggering of the downstream signaling cascade⁸⁹. Ligand binding induces also EGFR endocytosis, which is dependent on the E3 ubiquitin ligase, Cbl⁹⁰⁻⁹². Many studies have shown that epsins are involved in this process: indeed EPN1 recognizes ubiquitinated EGFR in CCPs, resulting in a subsequent ubiquitination of Epsin itself by NEDD4, process called “coupled monoubiquitination”³¹. Moreover, silencing EPN1 inhibits EGFR endocytosis⁹³⁻⁹⁵. It is still not clear whether this process is dependent on CME or not, although cell context and ligand concentration are likely to influence the internalization route⁷².
- The RhoGTPase pathway. Cell division control protein 42 homolog (Cdc42) is a member of the RhoGTPase family, implicated in cell polarity, regulation of the actin cytoskeleton

and cell migration ⁹⁶. Interestingly, it has been shown that, in yeast, Ent2 is able to sequester Cdc42 GTPase-activating proteins (GAPs), allowing the activation of Cdc42 ⁹⁷. This alternative function of epsins has been ascribed to the ENTH domain, because deletion of the 2 proteins and expression of only the ENTH domain rescues the yeast defects ⁹⁷. Moreover, EPN2 was shown to regulate meiosis during mouse oocyte maturation, promoting polarity establishment through Cdc42 ⁹⁸. Interestingly, in mammals it was shown that ENTH domain interacts with Ral-binding protein 1 (RalBP1) to regulate cell migration and basement membrane invasion ⁹⁹. Indeed, the epsin-RalBP1 complex activates Rac1- and Arf6-dependent pathways to induce migration and the knock-down of epsins leads to cell migration and actin remodeling defects through Rac1 function ⁹⁹.

- The actin cytoskeleton. In addition to the already described function of epsin proteins in regulation of migration and polarity through interaction with RhoGTPases, it was shown that epsin is able to regulate CME recruiting actin cytoskeleton. Indeed, in *Dictyostelium* cells, epsin is necessary to allow the recruitment and phosphorylation of actin- and clathrin-binding protein Hip1r, in order to allow the correct internalization of clathrin coated pits ¹⁰⁰. Interestingly, the depletion of either epsin or Hip1r show deficiencies in timing and organization of actin cytoskeleton at the nascent pits ¹⁰⁰.
- The vascular endothelial growth factor (VEGF) pathway. The VEGF pathway regulates endothelial cell proliferation and migration to induce angiogenesis ^{101,102}. VEGF binds its receptor, VEGFR2, induces its ubiquitination, subsequent internalization and

downmodulation, and epsins are implicated in this internalization step¹⁰³. The ablation of epsins causes sustained VEGF/VEGFR pathway activation, causing alteration of angiogenesis^{103–107}.

- The Wnt pathway. The Wnt pathway is fundamental in the regulation of proliferation, cell migration, stemness, and cell fate determination^{108,109}. Wnt ligand binds the receptors Frizzled and LRPs, recruiting Dishvelled that in turn activates the pathway leading to β -catenin activation and translocation to the nucleus, where it regulates transcription of target genes¹¹⁰. Epsins bind Dishvelled and prevent its polyubiquitination and degradation, thereby sustaining pathway activation¹¹¹.

1.2.3. Epsins in physiology

Considering the different functions of epsins described in the previous paragraph, it is not surprising that many biological processes are sensitive to epsin expression and function. Indeed, EPN1 and EPN2 are widely expressed in almost all tissues, suggesting a role in processes common to all cells⁷⁹. In contrast, EPN3 expression is restricted mainly to gastric parietal cells¹¹² and migrating keratinocytes¹¹³, suggesting an involvement in more specialized cell-specific processes. In this paragraph, we will focus on the functions of EPN1 and EPN2, while the functions of EPN3 will be discussed in detail below.

Insights into the physiological roles of EPN1 and EPN2 have been gained through studies on transgenic KO mouse models. While the KO of either *EPN1* or *EPN2* in mice does not result in an overt phenotype, the KO of both *EPN1* and *EPN2* is embryonic lethal between stage E9.5 and

E10⁷⁹. Interestingly, *EPN1/EPN2* double KO specifically in the endothelial cellular compartment is not embryonic lethal, suggesting that the essential function of these proteins during embryonic development is connected to their roles in other cell types¹⁰³. Moreover, it was reported that the lethality of *EPN1/EPN2* double KO is not related to a generic role of epsins in endocytosis, but is instead due to a specific endocytic function of EPN1 and 2 in the Notch pathway during vasculature formation⁷⁹, however it is still unclear in which cell type/types, epsins are involved. These data are in agreement with findings obtained in *Drosophila melanogaster*, where the deletion of the only epsin member (LqF) causes embryo lethality, phenocopying the Notch deletion mutant^{85,114}. Finally, these results are in agreement also with studies in *Saccharomyces cerevisiae*, where the deletion of Ent1 and Ent2 causes lethality⁷⁶.

Together, these results suggest that, even if epsins are considered as general endocytic adaptors in CME, they probably exert a more specific function than a broad role in CME, being critical for the endocytosis of specific receptors.

1.2.4. Epsins in pathology

Alterations of epsin expression has been linked to many cancer types. Indeed, higher levels of EPN1 and EPN3 have been described in lung^{115,116}, breast^{116,117}, skin¹¹³, and colon¹¹¹. Generally, evidence points to a pro-tumorigenic role of epsins¹¹⁸ by promoting:

- Tumor angiogenesis. Given the characterized functions of epsins in the VEGF pathway, the absence of EPN1 and EPN2 in mice tumors (mouse Lewis lung carcinoma subcutaneously implanted or melanoma cells from the mouse skin cancer cell line,

B16F10) decreased tumor growth¹⁰³. The reason is because VEGF binding to VEGFR2 induces its ubiquitination, subsequent recognition by epsins, internalization and degradation; when epsins are absent, VEGFR2 activation is unregulated, resulting in vascular defects and aberrant tumor growth¹⁰³.

- Tumor proliferation. Since epsins regulate internalization and signaling of many receptors (EGFR, Notch, Wnt, etc...), it is not surprising that many researchers have postulated that epsin alterations can mimic/contribute to alterations of some receptor signaling pathways, ultimately leading to tumor growth^{72,118,119}. However, so far there is a lack of clear evidence of this contribution. An interesting study showed that epsin deficiency protects against colon cancer growth, probably through an unconventional role, destabilizing Dishevelled2 and blocking Wnt pathway activation¹¹¹.
- Tumor migration. The characterization of epsin functions in Cdc42-dependent cell migration led to the discovery that EPN1 binds to and inhibits Ral-binding partner 1 (RalBP1), a GTPase-activating protein that regulates cell migration. This inhibition increases invasion of fibrosarcoma cells, dependent on Rac1 and Arf6⁹⁹.

1.2.5. Epsin3

The *EPN3* gene, first described by Dr. Pilcher's Lab in 2000⁶⁹, is located on the human Ch.17q21.33. As already described, EPN3 protein structure is very similar to EPN1 and EPN2^{71,113}. However, despite a high grade of homology with other epsins that are ubiquitously expressed, EPN3 is expressed at very low levels in almost all tissues except for gastric parietal cells

¹¹² and migrating keratinocytes ¹¹³. The *EPN3* gene has 9 coding exons; with the central region being the most different compared to *EPN1*. Despite this difference, the *EPN3* ENTH domain shares 80% and 82% sequence identity with that of *EPN1* and *EPN2*, respectively, and the 3 NPF repeats are 100% conserved in these 3 proteins ¹¹³. Moreover, also DPW motifs and Clathrin-binding motifs are highly conserved among the 3 paralogs ^{70,71,113}.

The *Epn3* gene of *Mus musculus* produces two different transcripts, one with 96% of sequence identity to the human *EPN3* gene and the other with 85% sequence identity, due to the lack of a specific exon in the middle of the gene. Physiologically, the KO of murine *EPN3* alone results in viable offspring without major defects. However, *Epn1/Epn2/Epn3* triple KO mouse embryonic fibroblasts (MEFs) displayed a dramatic cell division defect compared with WT MEFs, impairing CME due to dysfunction of the actin cytoskeleton role in the CCP formation ¹⁰⁶; unexpectedly, *Epn1/Epn2* double KO – that is embryonic lethal in mice – show no obvious defects in MEFs, an effect that authors suggested to be linked to long-term adaptation of cells and upregulation of *EPN3* ¹⁰⁶.

As already mentioned, in humans very little is known of *EPN3* function: the fact that it is upregulated in migrating keratinocytes, in chronic wounds and uncreative colitis led to hypothesize that it could be implicated in cell migration in response to specific factors and pathological conditions ¹¹³. Consistently, in madin-darby canine kidney (MDCK) cells (epithelial cells from canine kidney) *EPN3* was found to be upregulated at the leading edge of migratory cells ¹¹⁹. Moreover, in cancer, *EPN3* has been linked with migration and invasion. *EPN3* was found to be upregulated in lung ^{115,116}, breast ^{116,117} and skin cancer ¹¹³; in breast and lung cancer this

upregulation correlates with metastasis acquisition and a worsening of the clinical outcome ¹¹⁶. EPN3 has also been shown to induce EMT in glioblastoma patients, since higher levels correlated with migratory abilities and upregulation of mesenchymal markers ¹²⁰. In contrast, *EPN3* mRNA expression was found to be downregulated in human gastric cancer tissues in the The Cancer Genome Atlas (TCGA), and, thus, it was speculated that EPN3 could act as a p53-dependent pro-apoptotic anti-tumoral factor in primary gastric cancers ¹²¹.

In conclusion, these few evidences suggest that EPN3, even if it has a very similar protein structure to the other epsin family members, could have a specific function in migration and invasion, causing a worsening of the phenotype in cancer progression, while it is still unclear the role of EPN3 in primary tumor formation.

1.3. Epithelial cell-cell junctions: endocytosis as a mechanism to regulate junction stability and cell migration

1.3.1. Cell-cell and cell- extracellular matrix junctions

One of the most important achievements during the evolution of life, was the ability to create multicellular organisms. This conquest would have never been possible without the appearance of specialized groups of molecules that attach and sense the microenvironment, connecting each cell with other cells and with the ECM: i.e., the cell junction proteins. Solid tissues need to be strong and efficient to work properly, this is achieved by the cell junctions that confer structural stability and allow communication between neighboring cells. Over the last decades, increasing evidence

has revealed that cell junctions are much more than simple pillars for the tissue stability, they are signaling stations able to regulate nuclear transcription¹²², mechanostability^{123,124}, migration¹²⁵, and proliferation¹²⁶.

There are many different cell-cell and cell-ECM junctions (Figure 4), some of which have widespread expression, while others are tissue- or cell- specific¹²⁷. The most important are:

- Tight junctions. Also known as *zonula occludens*, they are cell-cell fibril-like proteinaceous structures inserted into the PM¹²⁸. They have 2 main functions: i) separate the membrane composition of the apical and baso-lateral portions of the PM; ii) control the exchange of ions and solutes between cells¹²⁹. The main transmembrane components are occludins and claudins, regulated by the membrane-associated guanylate kinases (MAGUK) members (ZO-1, ZO-2 and ZO-3).
- AJs. Also known as *zonula adherents*, AJs are homophilic calcium-dependent cell-cell junctions that establish and maintain the epithelial sheet¹²⁹. The main component of AJs in epithelial sheets is E-Cadherin, a transmembrane protein that, via its extracellular domain, establishes junctions with other E-Cadherin molecules on neighboring cells, while its intracellular domain binds: 1) p120 catenin to stabilize the junction and inhibit E-Cadherin endocytosis^{130,131}; 2) α -catenin to connect AJs with the actin cytoskeleton¹³²; 3) β -catenin to stabilize the junction and regulate its nuclear function^{126,133,134}.
- Desmosomes. Also known as *macula adherentes*, they are the most abundant cell-cell junction in stratified epithelia¹³⁵. These junctions have a very precise structure, fundamental to maintain high adhesion between interconnecting cells. The main

extracellular components are the nonclassical cadherins, desmocollins and desmogleins, which heterophilically bind to each other in the extracellular space. These proteins are intracellularly connected, by plakophilins, plakoglobins (γ -catenin), and desmoplakins, with the cytoskeleton ¹³⁶.

- Gap junctions. These junctions allow the exchange of ions and solutes between neighboring cells. Each junction is composed of one connexon on one cell that interacts with another connexon on a neighboring cell; each connexon is composed of 6 connexin proteins ¹³⁷.
- Focal Adhesions. These cell-ECM connectors share many components and structural similarities with the hemidesmosomes. They are more broadly expressed and regulate cell migration. Integrins represent the main transmembrane protein components of focal adhesions, connected with the actin cytoskeleton through vinculin and talin proteins ^{138,139}.
- Hemidesmosomes. These junctions are found specifically in keratinocytes and are responsible for the connection between cells and the ECM. Their main components are the integrins, transmembrane proteins that connect basal membrane components with keratin filaments, through adaptor proteins called plectins ¹⁴⁰.

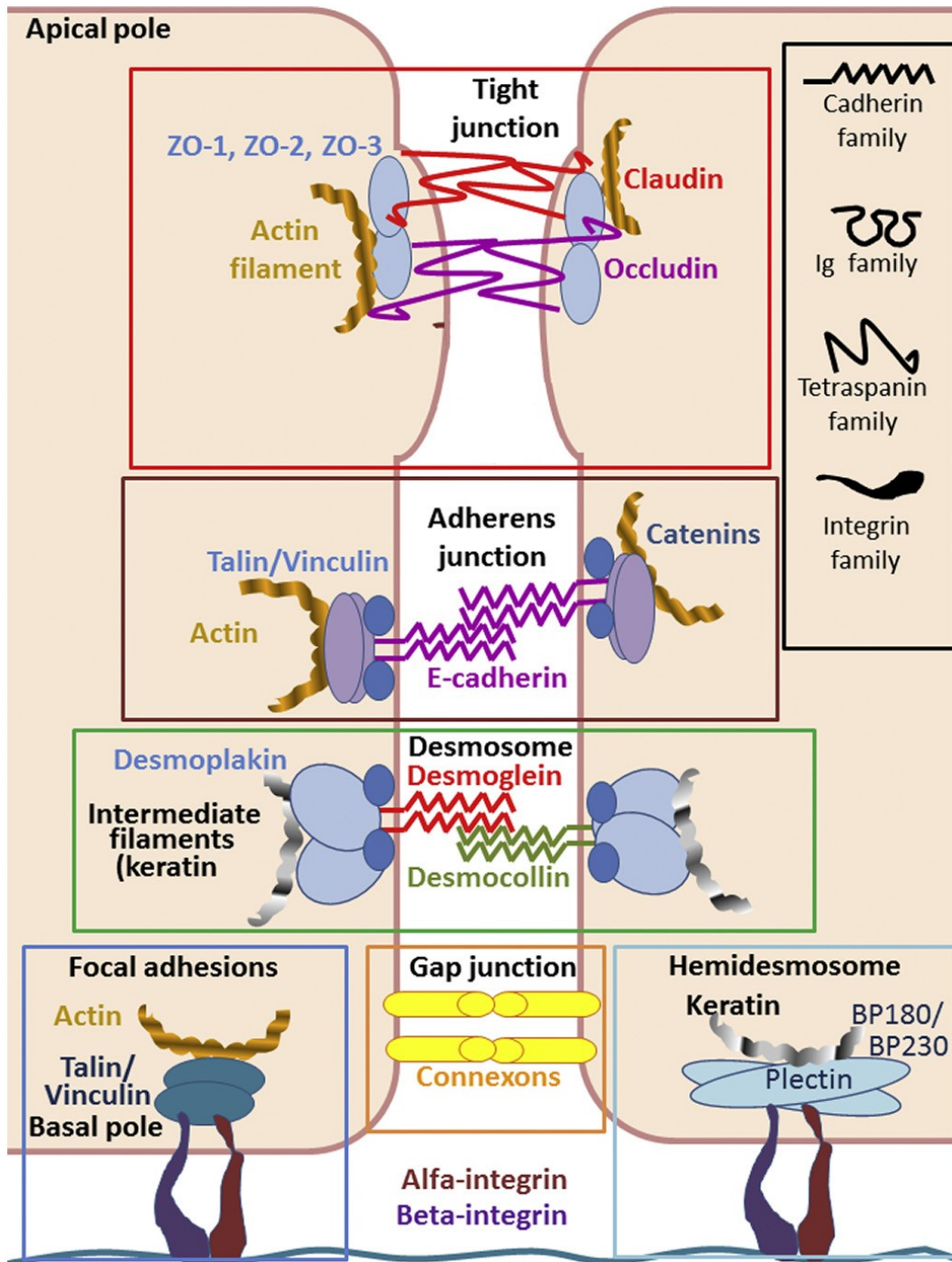


Figure 4: Cell junctions typical of normal epithelial cells

The major cell-cell junctions are: Tight Junctions, composed of the extracellular proteins Claudin and Occludin and regulated by the MAGUK members (ZO1, 2 and 3) and actin cytoskeleton; Adherens Junctions, whose main component is E-Cadherin, stabilized and regulated by the catenin family, actin cytoskeleton and actin remodeling proteins, such as talin/vinculin; Desmosomes, composed of desmoglein and desmocollin, which are intracellularly connected with intermediate filaments through desmoplakin; Gap Junctions, mainly composed of connexin proteins, that assembly themselves to form connexons. The main cell-ECM junctions are: Hemidesmosomes, specifically found in keratinocytes, whose principal transmembrane proteins are the integrins, connected to the keratin fibers through the plectins; Focal Adhesions, more generally expressed and composed of integrins connected to the actin cytoskeleton through talin/vinculin. Adapted from ¹⁴¹.

1.3.2. E-Cadherin

E-Cadherin (human gene name: *CDH1*) is a transmembrane single-pass glycoprotein that belongs to the classical Cadherin family, which establishes homophilic calcium-dependent cell-cell junctions¹²⁹. This family comprises many cadherins that usually are distinguished according to their cell/tissue expression¹⁴². E-Cadherin is the mainly expressed in epithelial sheets and it is the main component of epithelial AJs¹⁴².

The E-Cadherin protein is composed of: 5 extracellular cadherin (EC) repeat domains, one transmembrane domain, one intracellular catenin binding domain (CBD) and one juxtamembrane domain (JMD) to interact with armadillo proteins (Figure 5A). Regarding the ECs, each of these domains is able to bind the EC domains of other E-Cadherins expressed on the surface of neighboring cells. Every domain can have 2 conformations: in presence of calcium, ECs are “*on*” and able to bind other ECs; instead, without calcium, they are “*off*” and unable to form AJs¹⁴³. Studies have suggested that EC1 (the first EC domain) is fundamental to start the binding of E-Cadherins¹⁴⁴. Others have instead showed that blocking the EC4 and EC5 domains prevents the correct formation of AJs, suggesting that also these domain are responsible of the formation of the junctions^{145,146}. No matter which is the exact mechanism of trans-dimerization, cis-dimerization is necessary to create strong AJs (Figure 5B).

The JMD domain is necessary to bind the p120-catenin, an armadillo protein that confers stability to cadherins inhibiting its endocytosis and connecting with microtubules¹³³. The CBD domain connects E-Cadherin with β -catenin, which is fundamental for AJ stabilization through

α -catenin, vinculin and actin¹³⁴, and for the signaling, since β -catenin can translocate into the nucleus and activate transcription upon Wnt pathway activation and during EMT^{147,148}.

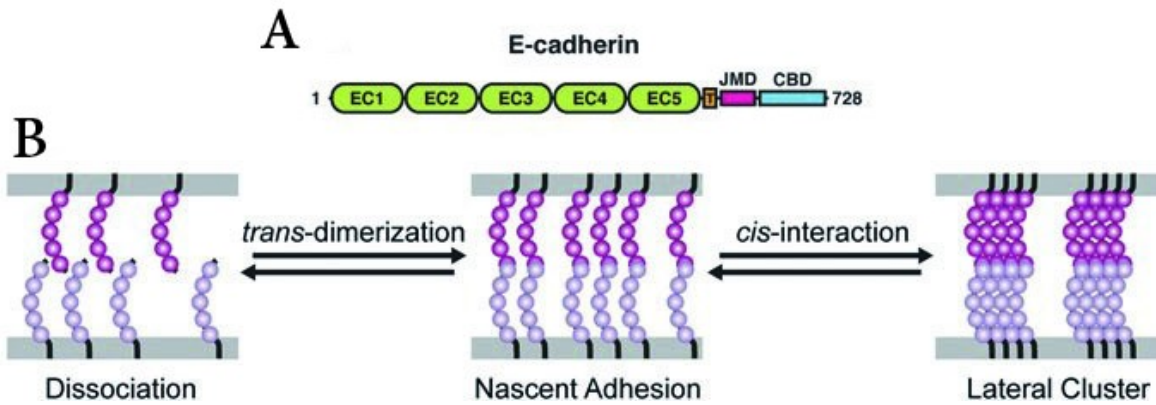


Figure 5: E-cadherin structure and modes of interaction

(A) Scheme of E-cadherin structure. E-cadherin consists of the extracellular cadherin domains 1–5 (EC1–5), the transmembrane region (T) and the cytoplasmic tail, which contains the juxtamembrane domain (JMD) and the catenin-binding domain (CBD). Amino acid residue numbers of cadherin are based on the mature form. (B) Model of cadherin *trans*-dimerization and *cis*-interaction of cadherins during cell-cell adhesion. Adapted from¹⁴⁹.

1.3.3. E-Cadherin functional regulators: the catenin proteins

The main regulators of AJs and intracellular interactors of E-Cadherin are the catenin proteins.

These proteins were first identified as being responsible for the binding and connection of cadherin with the cytoskeleton. The most important members of the Catenin family are: β -catenin, p120 and α -catenin. All but α -catenin are characterized by the presence of armadillo repeats, a sequence of 42 amino acids that compose a triple α -helix and allows the interaction with cadherins^{75,145,150}. Over the last 3 decades, many functions of this family have been proposed, which has increased dramatically the evidence that catenins are important in the regulation of many different biological processes^{129,147,151}.

β -catenin. This protein contains 13 armadillo repeats, resulting in a stoichiometric binding with E-Cadherin through its CBD, while at the C-terminus there is a transcriptional activator domain^{75,145}. In vertebrates, it was originally discovered as a regulator and partner of AJ stability, while in *Drosophila melanogaster* it was initially identified as a regulator of gene expression¹⁵². Indeed, β -catenin plays a dual function. It is widely accepted that in epithelial cells, β -catenin localizes at the AJs, linking E-Cadherin with α -catenin and actin cytoskeleton; while in response to specific growth factors or in cells more prone to do migration/invasion, β -catenin localizes mainly in the nucleus, where it works as co-transcription factor with members of the T cell factor / lymphoid enhancer-binding factor (TCF/LEF) family¹⁴⁷. Physiologically, when β -catenin is not bound to AJs, it is constitutively phosphorylated by the so-called destruction complex and degraded (Figure 6). The destruction complex is formed of 2 kinases, glycogen synthase kinase 3 beta (GSK3- β) and casein kinase 1 alpha (CK1 α) (that phosphorylate β -catenin in order to be ubiquitinated and degraded), Axin (a scaffolding protein) and adenomatous polyposis coli (APC) (that antagonizes β -catenin de-phosphorylation by phosphatases¹⁴⁷). The binding of Wnt, a growth factor important for embryonic development and tissue homeostasis¹⁵³, with LRP5/6 and Frizzled receptors induces the inhibition of the destruction complex, and stabilization and accumulation of β -catenin in the nucleus (Figure 6). In the nucleus, β -catenin binds to and works as co-transcription factor with the TCF/LEF transcription factors, activating target genes connected to EMT, cell cycle and migration¹⁴⁷. This implies that β -catenin is necessary for E-Cadherin and AJ stability, but also that E-Cadherin is a sequestering effector for β -catenin. It is widely accepted that there exists

2 different β -catenin pools, one very stable pool at the junctions and one cytosolic pool that is highly regulated and degraded in epithelial cells ¹⁵⁴. Nevertheless, to date, the existence of these two pools and the interplay between them has still not been clearly demonstrated.

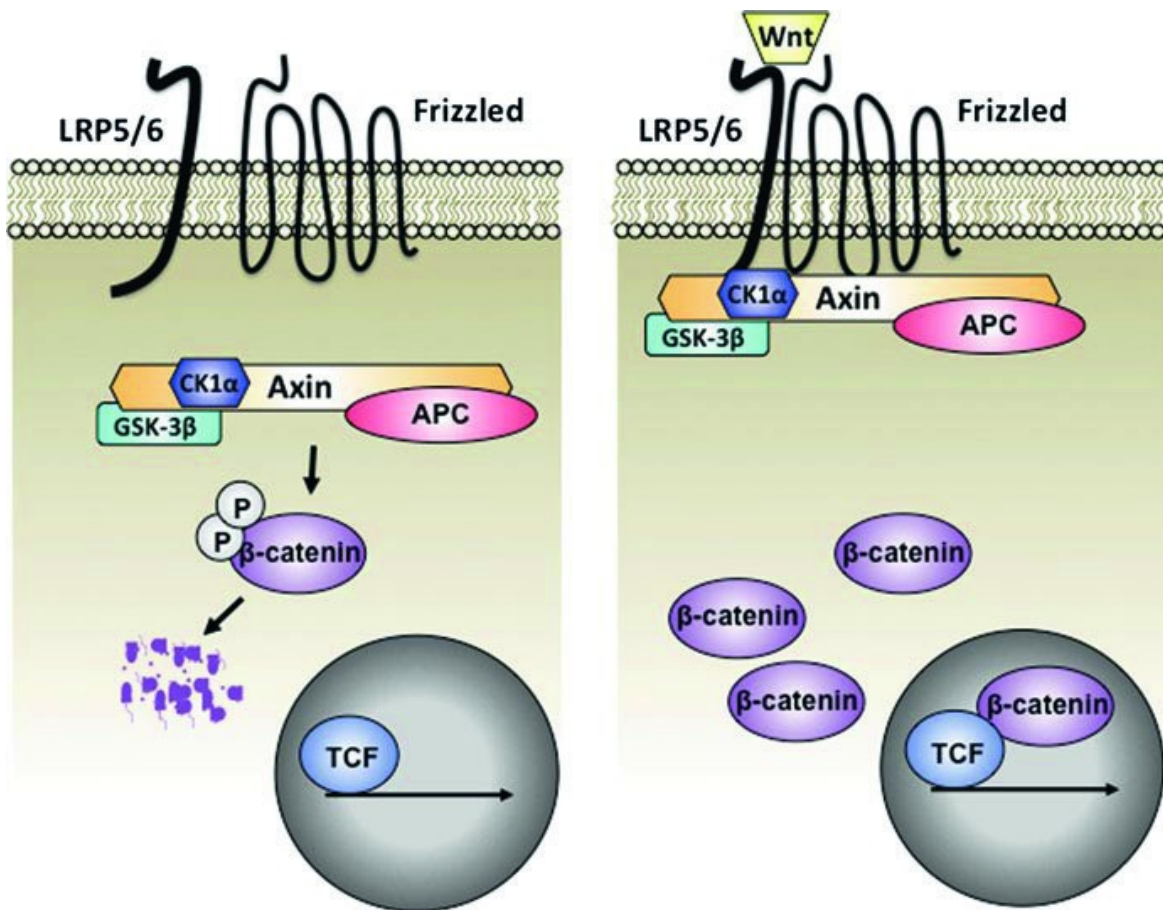


Figure 6: β -catenin/TCF signaling pathway

In the absence of Wnt (left), β -catenin that is not bound to the AJs (not shown) is continually phosphorylated by casein kinase 1 α (CK1) and glycogen synthase kinase 3 β (GSK3 β) within an Axin1 scaffold degradation complex. This phosphorylation allows β -catenin to be recognized by a specific E3 ligase (β TrCP, not shown), which catalyzes the ubiquitination and degradation of β -catenin. The adenomatous polyposis coli (APC) tumor suppressor protein participates in the phospho-destruction of β -catenin by antagonizing β -catenin de-phosphorylation by phosphatases. During Wnt activation (right), GSK3 β activity is inhibited directly by Lrp5/6, which allows β -catenin to accumulate in the cytoplasm, enter the nucleus, interact with LEF/TCF family members and promote transcription. Taken from

¹⁵⁴

p120. This catenin possesses 9 armadillo repeats that bind E-Cadherin through its JMD. p120 was initially identified as a Src substrate ¹⁵⁵ and, a few years later, as an interactor of E-Cadherin ^{156,157}. It exists as many different isoforms, of which 4 have been mainly characterized, and is subjected to different post-translational modifications, giving rise to a large number of possible combinations of p120 versions, depending on the cell type ¹⁵⁸. Isoform 1 is the longest and contains a small box at the N-terminus highly similar to the destruction box of β -catenin and highly regulated by kinases and phosphatases that are important for AJ formation and stabilization, while isoforms 2, 3 and 4 lack this box ¹⁵⁹. Of the 4 different p120 proteins, isoform 1 and 3 are the most well studied and characterized: isoform 1, thanks to the specific N terminus box, is connected to less strong junctions formations, while isoform 3 – unable to be regulated by kinases and phosphatases – strongly binds E-Cadherin and sustains epithelial junctions ^{131,158}. Interestingly, it has been proposed that, during EMT, p120 switches from isoform 3 to 1, decreasing its potential binding to E-Cadherin and resulting in increased migratory and invasive properties ^{131,160}.

The known p120 functions are connected to (Figure 7):

i) Cell motility, through RhoGTPases. Isoform 1 regulates cell motility more efficiently than other p120 isoforms, through the inhibition of RhoA and activation of Rac1. Indeed, RhoA binds preferentially at the N-terminus of the protein, where isoform 1 is phosphorylated ^{161,162}. It has been shown that phosphorylation by Fyn prevents p120/RhoA binding, while phosphorylation by Src has the opposite effect, suggesting that the N-terminus has an important regulatory role in RhoA binding and possibly activation. Finally, p120 isoform 1 – but not isoform 3 - is also known to activate Rac1 and this event leads to β -catenin nuclear translocation ¹⁶³.

ii) Activation of the Wnt pathway. Isoform 1 has a destruction box at its N-terminus similar to β -catenin, enabling the destruction complex to regulate also p120 phosphorylation status. Indeed, upon Wnt stimulation, p120 can be stabilized into the cytoplasm and shuttle into the nucleus to regulate the transcriptional repressor, Kaiso¹⁶⁴. Both isoforms 1 and 3 are known to bind Kaiso and to inhibit its repressor activity¹⁶⁴, necessary to activate target genes that are in common with β -catenin¹⁶⁵.

iii) Connection of AJs with microtubules. p120 is known to be able to link E-Cadherin with microtubules, necessary to correctly localize p120 and create AJs^{166,167}.

v) Regulation of E-Cadherin endocytosis. This topic will be discussed in detail the next paragraph (1.3.4).

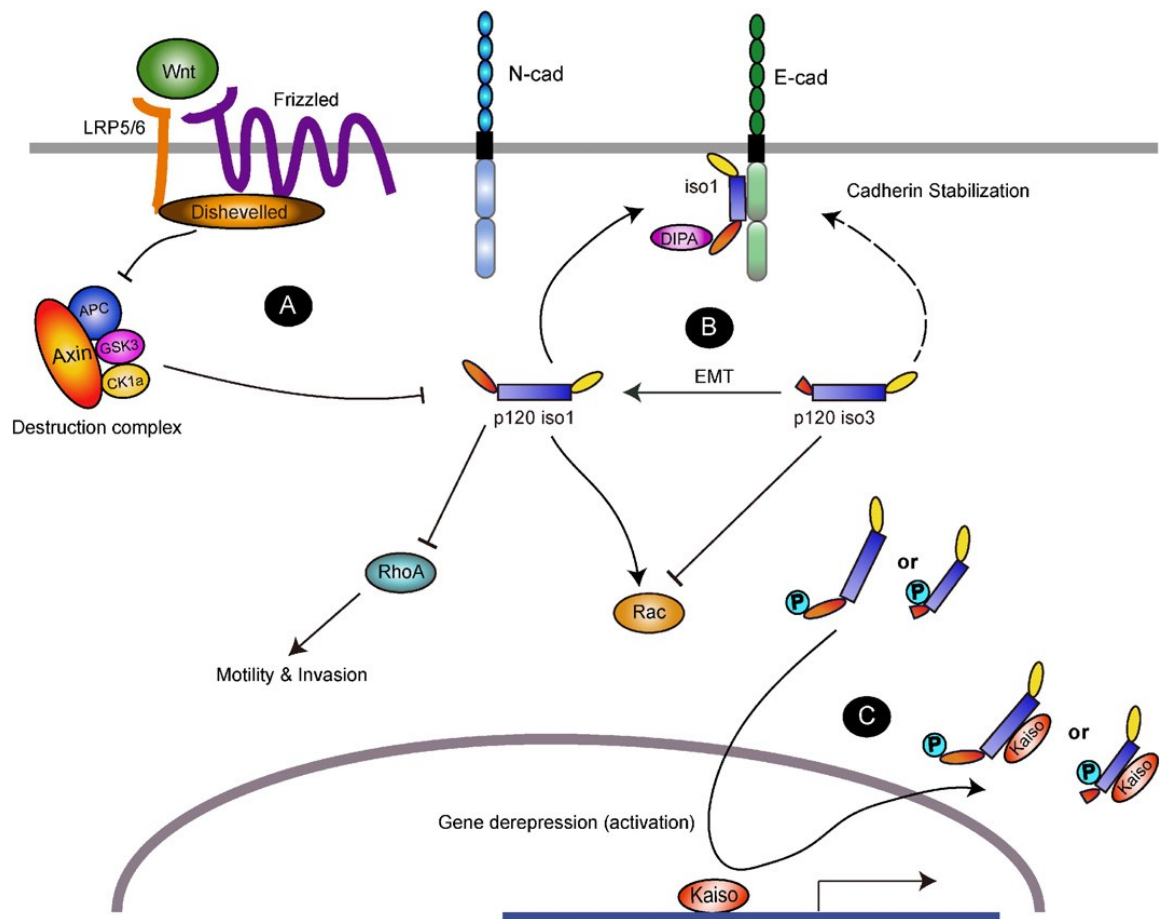


Figure 7: Isoform-specific role of p120

(A) Canonical Wnt signaling: Wnt-ligand binds to the LRP5/6 and Frizzled receptors, which in turn block the destruction complex (containing components such as Axin, APC, GSK3 β and CK1 α) from recognizing the p120-isoform1, resulting in a more stable and active isoform 1, like in the de-repression of the transcriptional repressor Kaiso depicted in (C). Canonical Wnt-signaling is less effective in protecting shorter p120-isoforms 2-4, since they lack the destruction box found only in isoform 1. Consequently, in the absence of canonical Wnt signals, isoform 1 is acted upon by the destruction complex, and delivered to the proteasome for degradation. (B) p120 isoform switching has been observed upon ectopic expression of Snail, suggesting the potential role of p120-isoforms during EMT. In addition, epithelial cell type specific regulators, ESRP1 and ESRP2, modulate p120 splicing events. Since p120-isoform 1 inhibits more effectively RhoA than isoform 3, this latter isoform induces motility or invasion. (C) p120 binds and displaces the Kaiso transcriptional repressor from its consensus binding sites, leading to gene activation. Taken from ¹⁵⁸.

α -catenin. This is the only catenin without armadillo repeats, implying that it is not able to bind directly E-Cadherin. Indeed, α -catenin connects the actin cytoskeleton and E-Cadherin through β -catenin binding. α -catenin/ β -catenin complex binds to actin through the interplay of a specific adaptor, epithelial protein lost in neoplasm (EPLIN) ¹⁶⁸.

1.3.4. AJ regulation: an endocytic perspective

Regulation of E-Cadherin levels and life-time at the PM is fundamental to control the cellular adhesive/migratory status. For this reason, E-Cadherin is highly regulated at multiple levels and a lot of effort has been employed to unravel these regulatory mechanisms¹⁶⁹. One of the best characterized mechanisms is the transcriptional regulation and epigenetic control of the E-Cadherin gene (*CDH1*). *CDH1* is transcriptionally repressed by transcription factors implicated in EMT, such as Snail and Twist that recruit histone deacetylases to the *CDH1* promoter^{170,171}.

Transcriptional regulation is the most effective strategy to obtain long-term (hours to days) stable control of adhesion and migration, but to rapidly and dynamically rearrange junctional stability and to maintain plasticity, regulation of AJs through endocytosis and trafficking is the main mechanism^{169,172,173}. E-Cadherin endocytosis can occur through CME, caveolin-mediated endocytosis or micropinocytosis. While very little is known about these latter two pathways of internalization, many studies have been focused on E-Cadherin CME¹⁷³. The master regulator of E-Cadherin PM stability and endocytosis is p120¹⁷³ (Figure 8). Although the exact mechanisms involved are still not completely clear, over the last decade it has been clarified that the p120 isoform 1 is able to bind to and stabilize the JMD of E-Cadherin through different motifs¹⁷⁴. This domain probably has a dual function, since there are both putative endocytic binding sequences for clathrin adaptors and motifs for p120 binding¹⁷⁵. In epithelial cells, p120 binds to the JMD of E-Cadherin, masking the DEE, VFEEE and LL motifs present in the E-Cadherin intracellular tail that are important also for CME, since they are recognized by clathrin adaptors, such as AP2^{174,176}. In this way, p120, blocks the accessibility of these adaptors to E-

Cadherin, stabilizing the AJs. Indeed, E-Cadherin internalization increases when p120 is mutated and unable to bind to E-Cadherin ^{176,177}.

Post-translational modifications play a fundamental role in E-Cadherin stability. Interestingly, in 2002 for the first time an E-Cadherin E3-ubiquitin ligase, Hakai, was identified ¹⁷⁸. It was shown that overexpression of Hakai in MDCK cells induces E-Cadherin internalization and degradation ¹⁷⁸. Interestingly, Src-dependent phosphorylation of two tyrosine residues in the JMD of E-Cadherin is necessary for Hakai binding to E-Cadherin itself ¹⁷⁸. Thus, p120 seems to unmask both endocytic sites for AP2, as well as ubiquitination sites for Hakai; when Src kinase phosphorylates the JMD, p120 detaches and Hakai ubiquitinates E-Cadherin, recruiting the clathrin endocytic machinery and inducing E-Cadherin internalization and degradation ¹⁷³.

Since the tyrosine residues of the JMD are not highly conserved among species and among other cadherins, this process is most likely highly specific and dependent on the cell type and on the type of cadherin involved. Indeed, other E3-ubiquitin ligases have been discovered, such as K5 for VE-cadherin ¹⁷⁹.

Importantly, E-Cadherin internalization is highly regulated also by growth factors, such as EGF ¹⁸⁰, HGF ¹⁸¹, FGF ¹⁸², and TGF β ¹⁸³. Indeed, TGF β cooperates with Raf-1 to induce E-Cadherin endocytosis and EMT in mammary epithelial cells ¹⁸³.

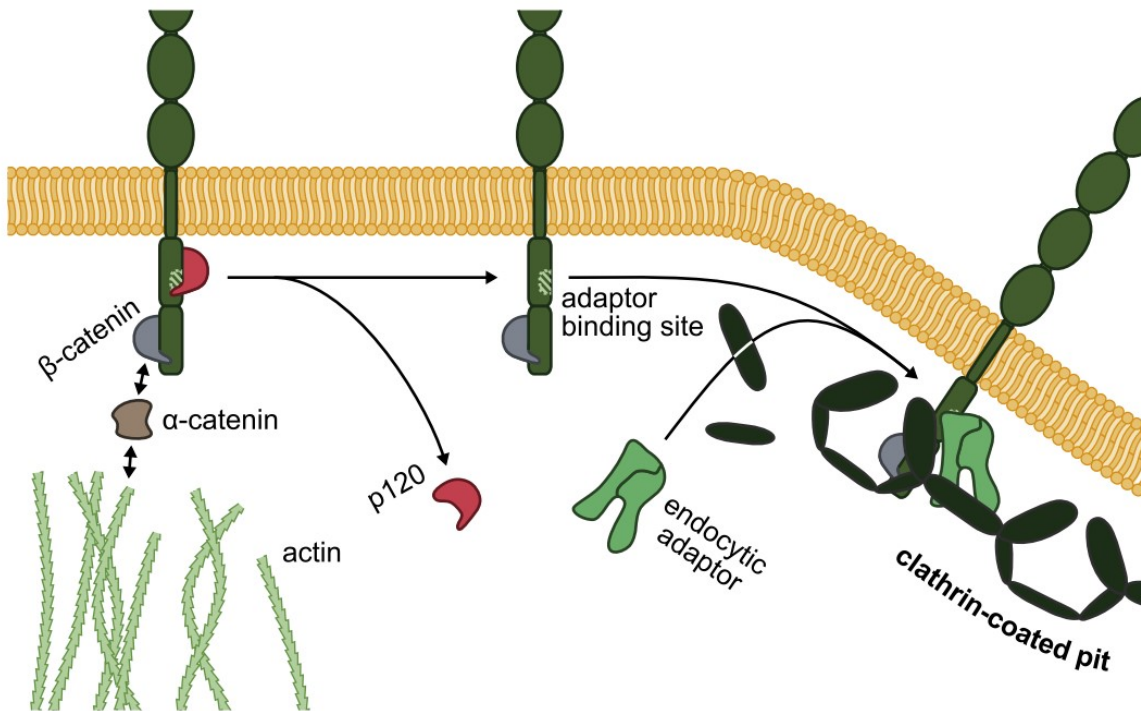


Figure 8: p120 regulated cadherin endocytosis

p120 binds to cadherins and masks an endocytic adaptor binding site. When p120 dissociates from the cadherin, the adaptor binding site is exposed, allowing the endocytic adaptor to bind to the cadherin, triggering cadherin endocytosis through a clathrin-coated pit. Taken from ¹⁷².

1.4. Epithelial-mesenchymal transition: overview in mammary gland physiology and breast cancer

EMT is a plastic, highly dynamic process characterized by the conversion of cells with epithelial properties into cells with mesenchymal-like properties ¹⁸⁴. During this transition, epithelial cells lose their adherent features and acquire migratory behavior ¹⁸⁴⁻¹⁸⁷. EMT is only half of the transition process, indeed once the migration has occurred, it is often necessary to revert back again through the so-called Mesenchymal-to-Epithelial Transition (MET) ¹⁸⁶. Moreover, it is now widely accepted that EMT is not a binary process, but there are different intermediate states,

named partial EMT (pEMT), characterized by the coexistence of epithelial and mesenchymal markers,^{184,188,189}

Physiologically, EMT-MET are important processes that occur during both embryonic development and in adulthood. Dysregulation of EMT-MET has been proposed for a long time to be important in the acquisition of metastatic potential, causing a more aggressive phenotype, cancer recurrence, drug resistance and CSC expansion¹⁹⁰.

This chapter aims to describe the most relevant findings in the EMT field, focusing on the molecular mechanisms at the basis of this transition and how EMT dysregulation leads to BC metastasis.

1.4.1. EMT: basic mechanisms and molecular players

In physiology, EMT is a highly regulated process that allows a controlled transition of epithelial cells into mesenchymal-like cells. The process is composed of different steps, many of which have been well studied and understood; however, the exact molecular players can be very different depending by the cell contexts¹⁹¹.

In general, EMT is a transcriptional program that requires an ordered rearrangement of the adhesive and migratory features of a cell (Figure 9). One of the first steps is the removal/rearrangement of all the epithelial properties, in particular AJs, tight junctions, desmosomal junctions, focal adhesions, substituting cadherins and integrins important for cell stability with other molecules able to create more transient bindings¹⁹². The consequence of this

phase is a change in cellular shape that reflects specific molecular rearrangements in the cell-cell adhesion compartment.

After these changes, or in concert with them, the upregulation of mesenchymal markers and transcription factors occurs¹⁹³. This step is necessary to acquire a full EMT, characterized by highly upregulated mesenchymal markers (also at the transcriptional level) and a typical fibroblast-like shape. Of note, both in physiology and in pathology, this multistep process has different metastable intermediate states that can then lead to the acquisition of a full EMT, or can be the endpoint of the process^{189,194}. In these metastable pEMT states, the coexistence of both epithelial and mesenchymal markers confers a more plastic behavior, which is more prone to rapid changes in cell morphology¹⁸⁴.

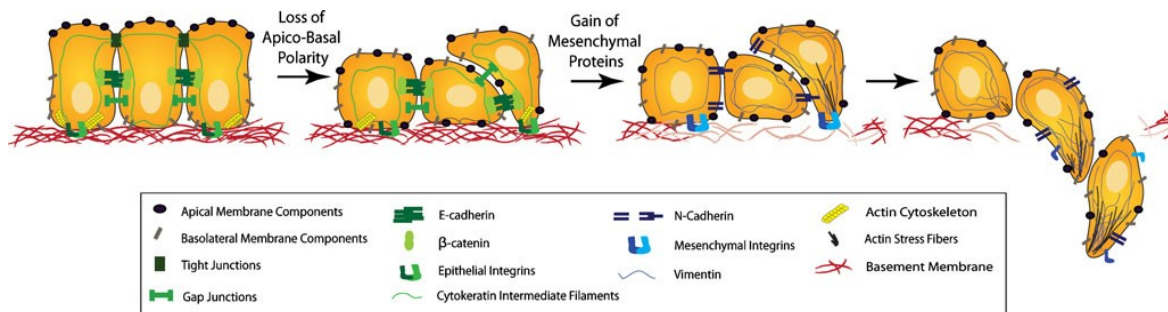


Figure 9: EMT and epithelial plasticity

During EMT, one of the first events is the loss of apico-basal polarity from epithelial cells. Tight junctions, which typically maintain apico-basal polarity, are disrupted allowing the mixing of apical and basolateral membrane proteins. AJs and gap junctions are disassembled and cell surface proteins, such as E-cadherin and epithelial-specific integrins (green), are replaced by N-cadherin and integrins specific to extracellular components (blue). The actin cytoskeleton appears as stress fibers and accumulates at areas of cell protrusions, while cytokeratins are substituted by vimentin. Moreover, the underlying basement membrane is degraded to allow cell invasion into the surrounding stroma. Taken from¹⁹¹.

From the molecular point of view, full EMT is a transcriptional program and transcription factors are fundamental players; however, over the years, other alternative/complementary

regulatory layers have been discovered, including: post-translational modifications¹⁷⁸, proteasomal degradation¹⁹⁵, miRNA regulation¹⁹⁶ and endocytosis⁵⁶. One possible explanation is that the transcriptional/epigenetic reprogramming determines a more effective and long-term phenotype, but often the necessity to locally and rapidly change cell shape and properties (e.g., adhesion, migration) can be more easily attained through post-transcriptional regulations^{169,172,173} and these mechanisms have been linked to the pEMT state^{184,188,197}.

The key steps of EMT are:

- Downregulation of epithelial properties. The downregulation of AJs, tight junctions and other cell-cell junctions is necessary to start the EMT process. Indeed, E-Cadherins, claudins, occludins and epithelial keratins are all downregulated^{184,191}. In particular, E-Cadherin is the key determinant of the epithelial state and it is tightly controlled. As already described in section 1.3.4, E-Cadherin endocytosis and turnover are thought to play a critical role in the initial disruption of AJs, followed by transcriptional regulation of the E-cadherin gene. Indeed, mesenchymal transcription factors, such as Snail, Slug, Zeb, and Twist, are able to silence the E-Cadherin transcript, further decreasing cell-cell adhesion and leading to a more advanced EMT state. The downregulation of epithelial markers causes also a shift in the cytoskeleton composition and localization, replacing peripheral actin cortex with actin stress fibers, more prone to sense the environment and modulate cell shape¹⁷⁰.
- Upregulation of mesenchymal markers. The loss of E-Cadherin (or PM E-Cadherin) causes the upregulation of mesenchymal markers, such as N-Cadherin, Vimentin, and

Fibronectin-1^{184,191}. These components are able to create more transient and dynamic interactions between cells and the ECM. Also in this case, the main players are the transcription factors Snail, Slug, Zeb, and Twist, which are able to induce transcription of the mesenchymal markers.

- Growth factor signaling. Both under physiological and pathological conditions, the signaling event that often induces the cascade is growth factor stimulation. Signaling cascades downstream of these growth factors are able to regulate both transcriptionally and post-translationally the critical components needed to induce EMT. The most well-known growth factors able to induce EMT are Wnt¹⁹⁸, HGF¹⁹⁹, FGF²⁰⁰, and TGF β ²⁰¹.

1.4.1.1. The TGF β pathway

The TGF β superfamily of ligands includes at least 30 genes in mammals, of which the most important are: 3 TGF β genes, 4 activin β -chains, 10 bone morphogenetic proteins (BMPs), nodal, and 11 growth differentiation factors (GDFs)²⁰². This superfamily is implicated in almost all biological processes: cell growth, adhesion, migration, differentiation and apoptosis. The TGF β superfamily regulates development, tissue regeneration, cancer, genetic disorders, and fibrosis.²⁰². Among the TGF β superfamily, TGF β proteins are the best characterized and they have been highly connected to EMT.

Regarding the TGF β signaling pathway, its activation converges on different downstream effectors (Figure 10), the most important of which are the SMAD proteins, but TGF β ligands have also been linked to nuclear regulation through other molecular players:

- SMADs. The canonical TGF β pathway is based on the binding of the ligands to the TGF β Rs. These receptors are the only known serine/threonine kinase receptors. There are 2 types: receptors I and II ²⁰³. Receptors normally are monomeric and inactive; upon ligand stimulation, receptors I and II heterotetramerize ²⁰⁴, then TGF β R2 phosphorylates TGF β R1, which recruits the regulatory Smads: R-SMADs: SMAD1, SMAD2, SMAD3, SMAD5 and SMAD8 ²⁰³. TGF β R1 phosphorylates R-SMAD, activating them. At this point, cofactor-SMADs (SMAD4) bind activated R-SMAD and allow nuclear translocation, where SMADs work as transcription factors upregulating mesenchymal markers (i.e., Snail, Slug, Zeb, Twist) and downregulating epithelial markers (i.e., E-Cadherin) ^{170,205}.

Considering the multistep activation of the EMT, it is not surprising that also the TGF β pathway is a multistep process. Among the early target genes sensitive to the TGF β pathway, there are the Snail and Slug transcription factors, leading to reversible pEMT ^{199,206}. Among the late target genes is the hub transcription factor, Zeb1: once it is activated, EMT becomes independent of TGF β pathway ¹³⁰.

- Tumor necrosis factor receptor-associated factor 6 (TRAF6). Activation of TGF β Rs leads to polyubiquitination and activation of TRAF6 and TGF β -activated kinase 1 (TAK1),

which phosphorylate and activate p38 MAPK, which, in turn, activates the MAPK pathway that converges on the SMADs²⁰⁷.

- Tumor necrosis factor- α -converting enzyme (TACE). Activation of TRAF6 by TGF β R1 induces TACE proteases that cleave the intracellular domain of the TGF β R1, which then translocates to the nucleus, where it regulates chromatin status and gene expression²⁰⁸.
- Par complex. The TGF β pathway induces loss of epithelial polarity by phosphorylation of Par6 via TGF β R2, leading to recruitment of the E3 ubiquitin ligase, Smurf1, that ubiquitinates RhoA, targeting it to degradation. In this way, loss of polarity and regulation of cytoskeleton can contribute to EMT acquisition²⁰⁹.
- Src. One last mechanisms of EMT activation to be proposed is the activation of Src kinase by TGF β ²¹⁰. Although the exact mechanism is still unclear, it seems that TGF β could induce the activation of Src, which causes Src-dependent E-Cadherin phosphorylation and ubiquitination, leading to EMT^{210,211}.

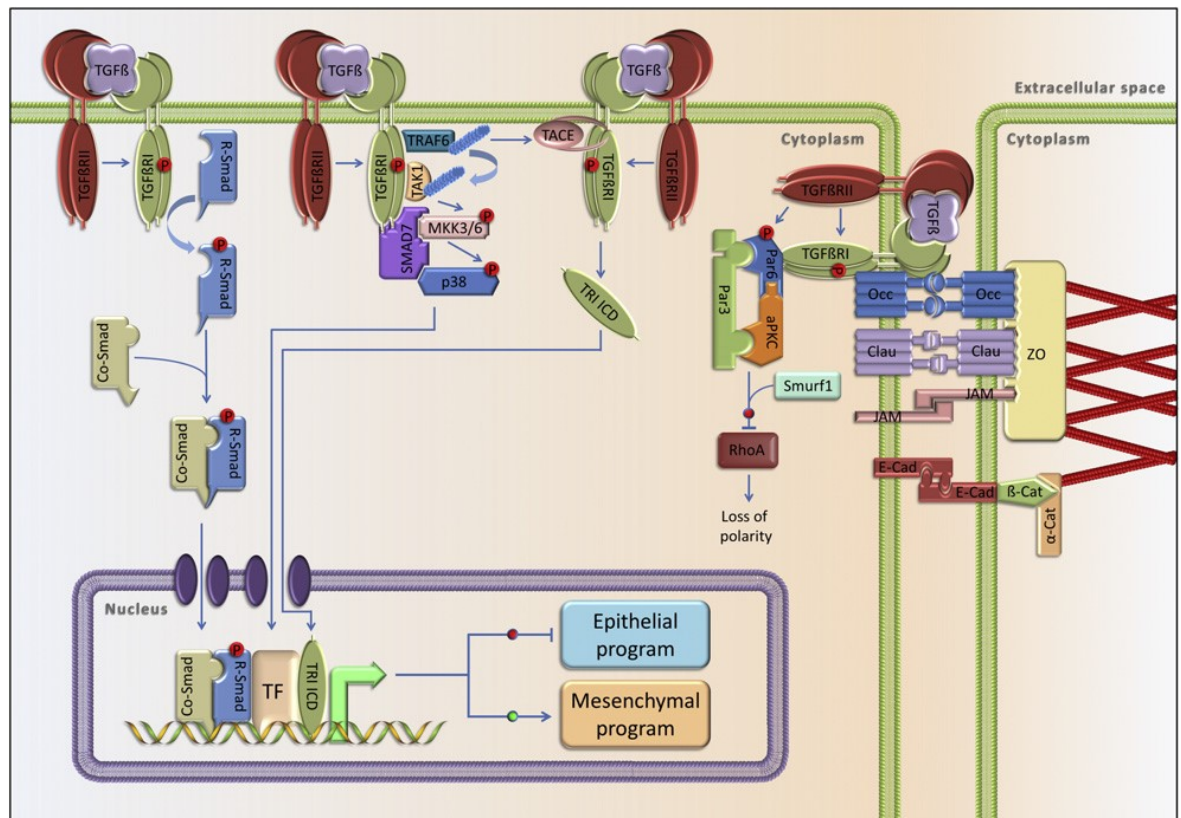


Figure 10: TGF β receptor signaling and the EMT program

Two adjacent epithelial cells are shown, interacting through tight junctions and adherens junctions. Four different scenarios of TGF β receptor tetrameric complexes are shown providing different signaling outputs (from left to right): (i) TGF β receptor signaling to Smads, with R-Smad phosphorylation and association with the Co-Smads, nuclear translocation and binding to DNA together with transcription factors (TF). (ii) TGF β receptor signaling to p38 MAPK via polyubiquitination (blue circles) of TRAF6 and TAK1, and phosphorylation of MKK3/6 and p38. Smad7 works as scaffold for the kinases. (iii) The TGF β receptor complex with TRAF6 activates the protease TACE that cleaves the type I receptor and releases its intracellular domain (ICD) that moves to the nucleus and binds to chromatin. These signaling pathways regulate transcription of genes leading to EMT. (iv) TGF β receptor signaling to the polarity complex (Par6, Par3, aPKC) via receptors localized at tight junctions, with the type II receptor phosphorylating Par6 and recruiting the ubiquitin ligase Smurf1 that degrades RhoA. Taken from ²⁰¹.

In cancer, TGF β is peculiar because it has a dual function: it works as a strong tumor suppressor in the primary tumor mass, since it inhibits proliferation and induces apoptosis; but it is also pro-tumorigenic promoting cancer invasiveness and metastasis through EMT activation ²⁰¹. Indeed, mice overexpressing the oncogene ErbB2/Neu and a constitutively active form of TGF β R1 in mammary gland, present a delay in tumor onset, but show increased lung metastasis when the tumor reaches a substantial size ²¹².

1.4.2. EMT in physiology

Physiologically, EMT is an important process that occurs during both embryonic development and in adulthood. In particular, the best characterized stages of embryonic development that involve EMT are:

- Gastrulation. The first evidence of the involvement of EMT in gastrulation was obtained in *Drosophila melanogaster*. During this process, the basement membrane that limits the epiblast is disrupted, and EMT of the epithelial cells is then required for them to enter the primitive streak. This process is highly regulated through the activation of FGFR²¹³, upregulation of Snail and Twist, and EMT switch of E-Cadherin to N-Cadherin²¹⁴.
- Neuroectoderm formation. The formation of the neural crest cells, during the neurula stage, starts with EMT of cells coming from the embryonic ectoderm and neuroectoderm. Neural crest cells migrate in order to form facial and cervical structures, glial cells, melanocytes, and the peripheral nervous system^{184,191}. Also in this case, EMT occurs through the downregulation of E-Cadherin via Snail²¹⁵ and production of growth factors, such as BMPs and FGFs²¹⁶.
- Heart development. Heart development happens during gastrulation, when the myocardium activates endocardial cells to undergo EMT and create a mesenchymal cardiac cushion needed for cardiac heart valves. In this context, among other regulatory factors, TGFβ regulates and induces mesenchymal transition of endothelial cells (EndMT)

²¹⁷.

In adulthood, the situation is less clear, the reason being because – physiologically – cells are more stable in the adult organism than during embryonic development. One strong exception in adulthood is represented by the mammary gland, an atypical gland characterized by the continuous proliferation, invasion and apoptosis of epithelial cells during branching morphogenesis. Moreover, EMT is important in tissue repair and wound healing, for example after injuries.

1.4.2.1. Branching morphogenesis in mammary gland: the role of EMT

The mammary gland, differently from almost all the other tissues, develops mainly postnatally. Indeed, this gland matures through 4 main steps: embryonic, pubertal, virginity and reproductive^{218–220} (Figure 11):

- Embryonic. The mammary gland is composed of 2 types of compartments, the epithelial and the stromal, deriving from the ectoderm and mesoderm, respectively. In mice, during E10.5, the milk line is derived from the ectoderm, giving rise to the first mammary identification²²¹. This mammary epithelium becomes a ball that infiltrates the mesenchyme at E15.5, where it becomes surrounded by the mammary mesenchyme²¹⁸. This primordial mammary bud gives rise to a rudimental ductal tree in the newborn, composed of primary, not well defined, epithelial ducts. At this stage the mammary gland stops growing, until the puberty begins.

- Pubertal. Epithelial ducts are composed of multiple layers of luminal epithelial cells, directly exposed to the lumen of the duct; luminal cells are surrounded by a single layer of myoepithelial cells, which contain SCs, and are in direct contact with stromal cells (fibroblasts, immune cells, adipocytes) and stromal ECM. At this stage, primary ducts terminate with the so-called Terminal End Buds (TEBs): these are club-shaped structures that penetrate the fat pad through proliferation of the first layer of cells, called cap cells²²². Puberty in mice starts at 3-4 weeks of age and it ends at 10-12 weeks of age. After a strong release of estrogens from the ovaries and of growth factors, such as IGF1 from the liver, TEBs infiltrate the fat pad with the primary epithelial ducts^{218,223}. Sometimes secondary branching from the primary ducts can occur, creating a tree-like structure that, at the end of the puberty, occupies about around 60% of the fat pad.
- Virgin adulthood. The mammary gland of adult epithelial ducts in virgin mice does not display TEBs, and the epithelial tree is composed of primary ducts, with sporadic secondary and tertiary branching. The gland shuffles between low and high branching status, according to the estrous cycle. At this stage, the main regulator of mammary gland morphology is progesterone, whose levels rise more than 4-fold in the diestrus phase causing an increase of duct branching²²⁴.
- Reproduction. Progesterone and estrogen hormones are necessary to prepare a lactation-competent gland in case of pregnancy. During pregnancy, progesterone and prolactin are necessary to create alveoli structures at the edge of each epithelial duct^{225,226}. These alveoli

are the ones responsible for milk production and secretion during lactation; at the end of this stage they tend to disappear almost completely through apoptosis²¹⁸.

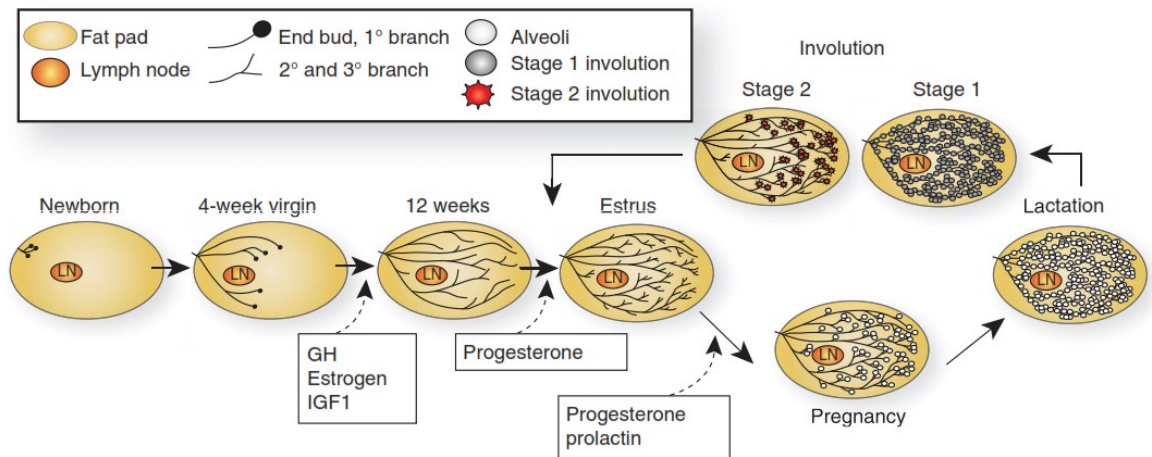


Figure 11: Mouse postnatal stages in mammary gland development

At birth, the mammary epithelium is composed of a rudimentary tree, with only few small ducts until puberty (4 weeks in mice). In puberty, specific hormones and growth factors – estrogen, insulin-like growth factor-1 (IGF-1), growth hormone (GH) – induce a process called ductal morphogenesis that fills the fat pad with the epithelial mammary tree. In the mature virgin, secondary and tertiary branching occurs according to progesterone levels during the estrus cycle, but alveologensis only occurs upon pregnancy with the induction of prolactin (PRL), which together with progesterone, fuels the growth of alveolar cells. During lactogenesis, PRL stimulation induces milk production and secretion, until a lack of demand at weaning. Involution signals remodel back the mammary gland to its original adult state. Taken from²¹⁸.

Since the mammary gland develops mainly postnatally, it has been deeply studied over the decades. One of the most interesting and difficult processes to be characterized is branching morphogenesis that comprises the expansion of the epithelial ducts through TEBs during puberty, and continues in post-puberty according to the estrous cycle²²⁰. In these process, proliferation is probably the most important event, in particular, cap cells (the external layer of TEBs, composed of SCs) allow the infiltration into the fat pad through proliferation. Nevertheless, loss of cell polarity and migration are fundamental processes during branching morphogenesis, and they are

intimately linked to EMT²¹⁹. Indeed, TEBs and epithelial adult ducts bifurcate through a process involving pEMT¹⁹¹. Cap cells and myoepithelial cells reveal patterns of epithelial plasticity, such as alteration of β -catenin localization²²⁷ and secretion of matrix metalloproteinases (MMPs) to modulate the basement membrane^{228,229}. Moreover, activation of Cripto-1 pathway has been shown to induce pEMT²³⁰. Of note, none of these factors induce a full EMT program, but – more appropriately – they confer mesenchymal ability to cells confined in the epithelium^{184,219}.

Also TGF β is important for branching morphogenesis, but its main function has been elucidated as a negative regulator of epithelial bifurcation^{218–220}: overexpression of TGF β 1 in the mammary gland results in suppression of ductal outgrowth²³¹, while TGF β 1-deficient mice show increased ductal development²³². However, it is still not clear how and where TGF β ligands should work: they are produced by epithelial cells²³³ and transplantation of TGF β 1-/- epithelium into the fat pad of WT mice causes increased ductal expansion²³⁴. On the other hand, loss of TGF β R2 in the stroma accelerates epithelial branching^{235,236}. Therefore, the exact mechanism and site of action of TGF β in suppressing branching morphogenesis is still unclear.

1.4.3. EMT in pathology

EMT, during adulthood, is fundamental to maintain epithelial integrity, as EMT occurs during wound healing, tissue repair and fibrosis^{184,237–239}. Indeed, in order to repair a wound, epithelial cells need to undergo pEMT to recreate the correct epithelial sheet. This pEMT is driven by Slug under the control of EGFR²⁴⁰, since Slug-deficient mice fail to repair wounds²⁴¹.

Another important case is fibrosis, a severe disorder caused by the accumulation of ECM secreted by myofibroblasts. The presence of a chronic inflammation status has been shown to represent the cause of fibrosis in a specific organ or tissue, that can lead also to organ failure¹⁸⁴. The mechanisms are still not well understood, but one of the possible causes is aberrant EMT. In particular, it is known that TGF β and the downstream effectors SMAD proteins are activated during tubular interstitial fibrosis²⁴². Moreover, reactivation of Snail in renal epithelial cells can induce renal fibrosis and failure²⁴³. Despite these and other findings, to date clear evidence is still lacking that, *in vivo*, migration of epithelial cells through EMT occurs in fibrosis¹⁸⁴.

A different case is cancer. The concept is that cancer cells from the primary tumor undergo different EMT states in order to escape and disseminate into a secondary site, giving rise to a metastasis mass (Figure 12)^{184,191}. To obtain a metastasis mass, it is fundamental that cells, once they have invaded and reached the secondary site, return to an epithelial-like morphology, reverting EMT through the MET. Thus, dysregulation of the EMT-MET program in cancer can lead to cancer recurrence, drug resistance and expansion of CSCs¹⁹⁰.

In agreement with this hypothesis, many different EMT markers correlate with more aggressive cancer phenotypes and metastasis acquisition in numerous solid carcinomas. Indeed, mesenchymal markers, such as vimentin, N-Cadherin, α -SMA, and desmin, are often found upregulated in solid secondary tumors²³⁹. Conversely, E-Cadherin is downregulated to allow EMT, promoting metastatic behavior and increasing tumor invasiveness^{244,245}.

However, there are still unresolved questions regarding the connection between EMT and cancer: first of all, still very few EMT-based therapeutic strategies are effective, indicating that

tumor heterogeneity in patients is much bigger than the variability observed in cell cultures or mice models¹⁸⁴; second, in many cases, invasiveness and metastasis do not present clear evidence of an EMT signature¹⁸⁴. Thus, there is still the need to better characterize the mechanisms *in vivo*.

In the last years, increasing evidence has accumulated that pEMT states could explain the high heterogeneity observed in real tumors. Indeed, it was shown – both in cell lines and in mouse models – the existence of many intermediate EMT states that occur during tumor invasiveness in skin, lung, liver and mammary primary tumors^{246,247}. These intermediate states confer different cellular plasticity and metastatic potential. Moreover, in pancreatic, breast and colorectal carcinomas, the existence of pEMT states has been causatively linked to the re-localization of epithelial markers, such as E-Cadherin, and not to the direct transcriptional downregulation, showing for the first time that, *in vivo*, alterations of epithelial marker trafficking and turnover could explain a more aggressive tumor phenotype^{197,247}.

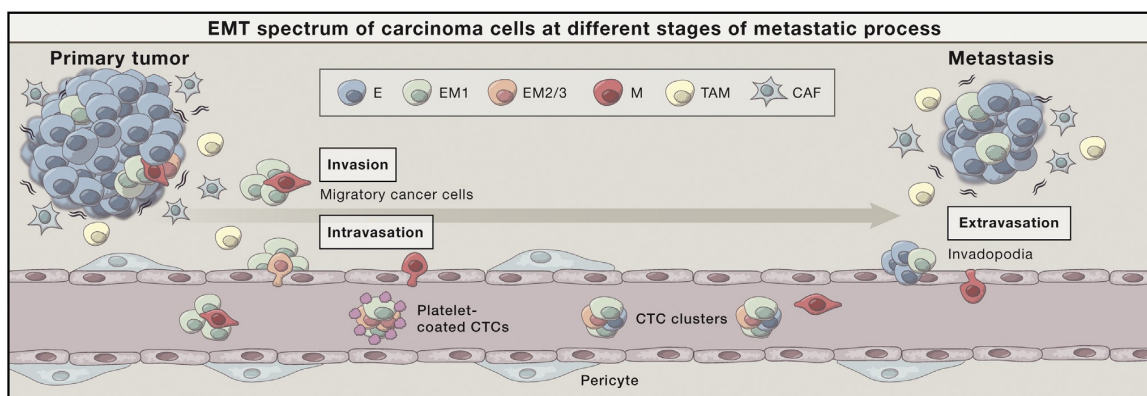


Figure 12: The metastasis cascade

EMT is one of the few events in the primary tumor that could occur upon the interaction between cancer cells and its microenvironment, in which cancer-associated fibroblasts (CAFs) and immune cells can be found, such as tumor-associated macrophages (TAMs). Some cells are possibly recruited prior to the establishment of primary and metastatic malignancies. The majority of circulating tumor cells (CTCs) exhibit a pEMT phenotype. Cells that undergo EMT in the primary tumor can also help epithelial cancer cells to invade. (*continued on next page*)

CTCs can be derived from carcinoma cells that undergo EMT in situ in the primary tumor or they can acquire pEMT in circulation, particularly, when in clusters exposed to high concentrations of TGF β originating from associated platelets. Cancer cells in metastases are epithelial-like, suggesting that mesenchymal cells have to revert to the epithelial phenotype to give rise to a secondary tumor. Taken from¹⁸⁴.

Among the different solid cancers, BC is one of the best characterized. This is because it is the most common female cancer. Indeed, according to World Health Organization data, in 2018, it is estimated that 627,000 women died from BC worldwide, representing 15% of all cancer deaths among women. Furthermore, BC is, experimentally, one of the easiest to study, due to the fact that the mammary gland is easily accessible compared to other organs and tissues.

1.4.3.1. EMT in breast cancer: from DCIS to invasive carcinoma

BCs are classified according to the expression of specific membrane receptors by immunohistochemistry analyses²⁴⁸: i.e., estrogen receptor (ER), progesterone receptor (PR) and epidermal growth factor receptor 2 (HER2/NeuN). Moreover, the proliferation status, through the Ki67 marker, is always considered in the final classification of the tumor. There are 5 BC subtypes^{248,249}:

- Luminal A. ER+, PR+, HER2- and Ki67-. These tumors are hormone-responsive and tend to grow slowly, having the best prognosis.
- Luminal B. ER+, PR+, HER2- and Ki67+. These tumors are hormone-responsive and tend to grow faster than Luminal A, having a slightly worse prognosis.
- Normal-like. ER+, PR+, HER2- and Ki67-. These tumors are very similar to Luminal A tumors, but with a slightly worse prognosis.

- HER2-enriched. ER-, PR- and HER2+. These tumors are hormone-receptor negative and have a worse prognosis than all the above subtypes, but they are effectively treated with the targeted therapy: HER2-specific antibodies, e.g., Trastuzumab.
- Triple-negative/basal-like. ER-, PR-, HER2- and basal marker+. These tumors are hormone-receptor negative and represent the subtype with the worst overall prognosis.

Independently of their molecular properties, BC can arise from different sites within the mammary gland. The most common type of BC is ductal carcinoma *in situ* (DCIS)^{250,251} that arises from the epithelial milk ducts. Many studies have tried to unveil the different steps of BC appearance and evolution. In general, epithelial cells from a normal ductal lumen can give rise to benign proliferative changes that lead to atypical hyperplasia, a pre-neoplastic histological feature. DCIS represents the primary tumor mass that can give rise to invasive breast cancer (IBC) and subsequent metastasis formation (Figure 13)²⁵².

While it is still not clear how exactly cancer cells appear and create a tumor mass, it is widely accepted that invasive carcinoma appears almost every time from a DCIS²⁵². DCIS is a layer of myoepithelial cells surrounded by basement membrane^{250,253}. They are mainly characterized by genome instability, loss of the tumor suppressor p53 and overexpression of oncogenes, such as HER2/NeuN, ER and PR²⁵². IBC, instead, shows cancer myoepithelial cells not confined in space through a basement membrane²⁵⁰.

The fact that IBC derives from DCIS is corroborated by numerous observations. Firstly, they are very often found at the same anatomical sites^{250,254}. Moreover, the genomic instability is

rarely discordant between DCIS and IBC of the same sites, corroborating the idea that IBC is evolutionary linked to DCIS^{250,255,256}.

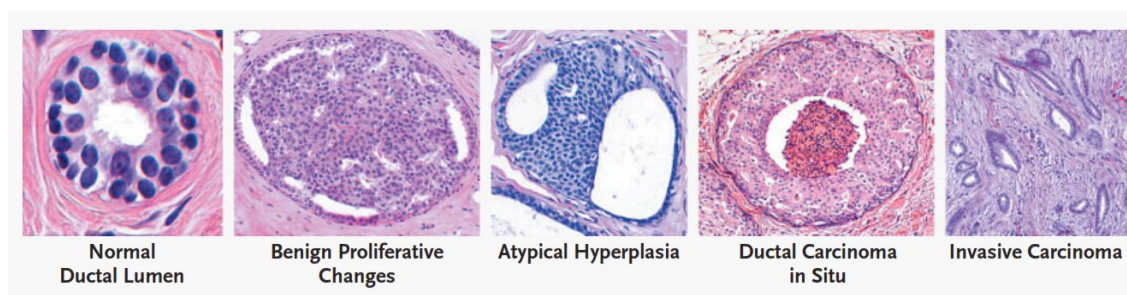


Figure 13: Pathobiological events associated with ductal carcinoma in situ

The transition from healthy tissue to preinvasive lesions, such as ductal carcinoma in situ, to breast cancer is shown. The majority of the changes should be evident at the DCIS stage. These changes include loss of physiological cell-cycle regulation, genetic and epigenetic rearrangements, which give rise to cancer. The transition to invasive carcinoma includes other events, such as changes in the surrounding stroma. Taken from²⁵².

In BC, poor prognosis correlates almost always with IBC and metastasis acquisition²⁵⁷.

The mechanisms of transition from DCIS to IBC and then to metastatic disease are far from being elucidated. However, EMT/pEMT is thought to be a key process in this transition. Indeed, in IBC, E-Cadherin is lost, while fibronectin and collagen I/III are highly upregulated^{191,258}. Moreover, TGF β , Wnt and MMP pathways have been shown to be relevant for BC aggressiveness and metastasis formation^{184,191,259}.

In general, as already discussed in other paragraphs, EMT occurs locally and often with co-existence of epithelial and mesenchymal markers (pEMT), explaining why it is so difficult to unveil the exact mechanisms of EMT in the transition from DCIS to IBC and also to prevent it²⁵⁰. Importantly, in 2012, it was shown for the first time that EMT markers correlate with IBC more than with DCIS through a very elegant approach that allowed the gene expression profile analyses

of the different cellular compartments in BC ²⁵¹, increasing the evidence of a role for EMT/pEMT in metastasis acquisition.

1.5. Identification of EPN3 as a novel prognostic factor in breast cancer invasiveness and metastasis

Previous screening studies in our lab focused on identification of endocytic proteins that are aberrantly expressed in cancer, revealed that the *EPN3* gene was amplified in ~9.6% of BC cases, as determined by fluorescence *in situ* hybridisation (FISH) analysis of a cohort of 219 BC patients who underwent treatment at the European Institute of Oncology (IEO) in Milan (Figure 14A). Considering that the *EPN3* gene is located on the human chromosome 17q21.33, 10.8 Mbps away from the frequently amplified BC oncogene *ERBB2* (also known as HER2/NeuN), we also investigated whether *EPN3* and *ERBB2* were co-amplified in the IEO BC cohort. Among the *EPN3* amplified cases, approximately half (52%) of them displayed co-amplification of *ERBB2* (Figure 14A). Extending these findings to a larger cohort, through a survey of the TCGA Breast Invasive Carcinoma cohort (N=1080), the *EPN3* gene was found to be amplified in ~8.0% of BCs, and co-amplified with *ERBB2* in ~57% of cases (Figure 14A), confirming the data from the IEO cohort.

To investigate whether alterations in *EPN3* transcript levels were reflected at the protein level, we analyzed the IEO cohort of BC patients by immunohistochemistry (IHC) to evaluate *EPN3* protein expression. The results confirmed that in almost all *EPN3* amplified cases (19/21), protein overexpression also occurred (Figure 14B). Notably, in 47% of all BC cases, *EPN3* protein was found to be overexpressed compared to very low levels expressed in normal breast tissue,

suggesting that EPN3 can be upregulated through gene amplification or other mechanisms (Figure 14B).

To establish the clinical relevance of these results, EPN3 expression was analyzed by IHC and RT-qPCR in a large consecutive cohort of BC patients (N=2453) who underwent treatment at IEO between 1997-2000 and for whom complete clinicopathological follow-up was available, as well as formalin-fixed paraffin-embedded (FFPE) blocks stored in the IEO Biobank. In this analysis, we found that EPN3 was overexpressed at the mRNA and protein levels in 50% and 37,6% respectively, confirming results from the initial cohort. Moreover, in a multivariable analysis, we found that EPN3 overexpression in IHC or RT-qPCR analysis, correlated with an increased risk of distant metastasis, but not with loco-regional relapse (Figure 14C). Importantly, higher levels of EPN3 correlated with poor prognosis also in HER2-negative patients (Figure 14D). Together, these data indicate that EPN3 status is an independent prognostic predictor in BC and that EPN3 overexpression could contribute to BC progression independently of the oncogene HER2.

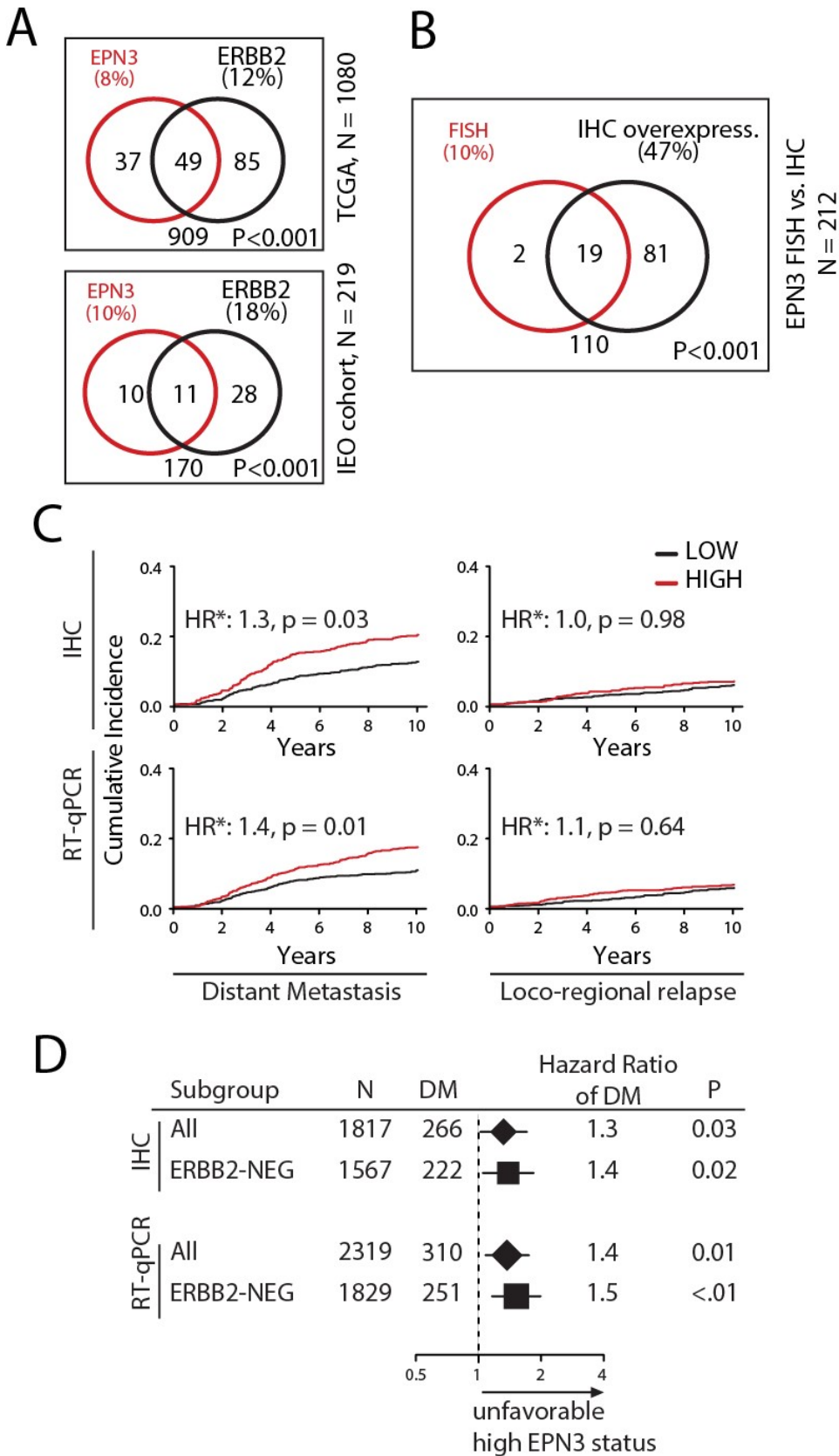


Figure 14: EPN3 is overexpressed in breast cancer patients

(A) Venn diagram representation of EPN3 and ERBB2 amplification. (continued on next page)

Percentages (%) of amplification are in parenthesis. Top, TCGA Breast Invasive Carcinoma cohort (N = 1080, TCGA Research Network). Bottom, IEO cohort (N = 219); EPN3 and ERBB2 were considered amplified when EPN3/CEP17 ratio >2.5, and ERBB2/CEP17 ratio \geq 2.0, respectively. CEP17 is the housekeeping reference gene. P is the p-value of the association between EPN3 and ERBB2 amplification. (B) Venn diagram representation of EPN3 amplification (FISH) and EPN3 protein overexpression (IHC; score >1.0) in the IEO cohort (N = 212). P is the p-value of overlap. (C) Cumulative incidence of distant metastasis or loco-regional relapse in BC patients of the IEO consecutive cohort (N = 2453), stratified according to EPN3 protein levels (top, IHC) or mRNA (bottom, RT-qPCR) levels. HR* is the multivariable hazard ratio. (D) Forest plot of the multivariable hazard ratios of distant metastasis and 95% Wald confidence intervals (whiskers) in the entire cohort of patients (All) and in the ERBB2-negative (ERBB2-NEG) or lymph node-negative (pN0) subgroups of patients stratified by EPN3 protein (IHC) or mRNA (RT-qPCR) expression levels. The size of the solid squares and diamonds is proportional to the number of distant metastases. The number (N) of patients and distant metastases (DM) in each group is indicated. Hazard ratios were estimated with a Cox proportional hazards multivariable model, adjusted for Grade, Ki-67, ERBB2 status, estrogen/progesterone receptor status, tumor size (pT), number of positive lymph nodes (pN) and age at surgery (as appropriate). P is the Wald test P-value.

In order to understand the biological role of EPN3 in BC, EPN3 was overexpressed in MCF10A, a non-tumoral, immortalized, breast epithelial cell line that does not show EPN3 alterations. We confirmed that the level of EPN3 overexpression achieved in the MCF10A cell model was comparable with the levels observed in BC cells, by comparing with BT747, a tumoral, immortalized, breast epithelial cell line that displays amplification of the *EPN3* genetic locus (Figure 15A). We also overexpressed the related protein, EPN1, as a negative control, since it did not show altered expression in BC in our initial screening of endocytic proteins (Figure 15A). Importantly, the overexpression of EPN3 – but not of EPN1 – induced a morphological change in MCF10A cells, compatible with EMT (Figure 15B). Indeed, MCF10A-EPN3 cells displayed upregulation of N-Cadherin and Vimentin at the protein and mRNA levels, while they showed downregulation of E-Cadherin mainly at the protein level, and to a lesser extent at the transcriptional level (Figure 15C-D). Importantly, EPN1 overexpression showed no significant morphological changes and only a minor increase of the EMT markers, suggesting a specific function for EPN3. (Figure 15C-D). As a positive control, we also overexpressed the well-known

INTRODUCTION

EMT inducer, Twist, in MCF10A cells, which causes a more advanced EMT phenotype as compared to EPN3 with a stronger morphological change and with major effects on EMT markers.

In conclusion, these data indicate that EPN3 is a novel independent prognostic factor in BC that predicts the risk of distant metastasis, and that could contribute to BC progression through a mechanism involving the activation of EMT.

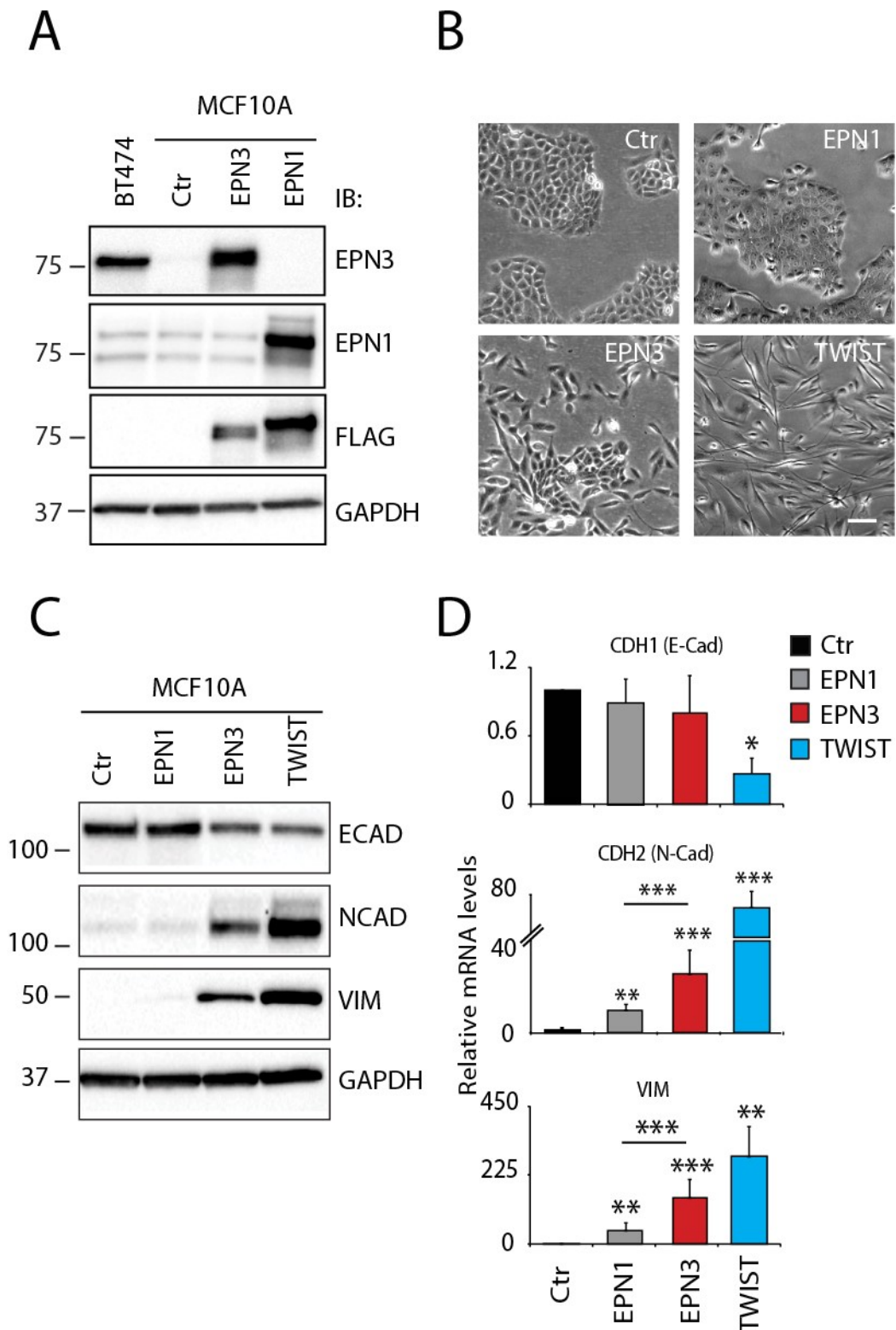


Figure 15: EPN3 overexpression in MCF10A cells induces EMT phenotype

(A) MCF10A cells were infected with vectors expressing FLAG-EPN3 (EPN3) or FLAG-EPN1 (EPN1) or empty vector (Ctr), and immunoblotted (IB) as shown. BT474 cells were used as positive control for EPN3 overexpression in BC. GAPDH, loading control. Molecular weight (MW) markers are shown on the left. (B) Phase contrast microscopy of cells as in (A), compared to MCF10A-TWIST. Bar, 50 μ m. (continued on next page)

INTRODUCTION

(C) IB analysis of EMT markers in the indicated cell lines. GAPDH, loading control. Molecular weight (MW) markers are shown on the left. (D) RT-qPCR analyses of EMT markers in the indicated cell lines. Data are normalized to the 18S gene and expressed relative to Ctr cells (mean \pm SD of three independent experiments for EPN3 and EPN1, two independent experiments for TWIST). p-values vs. Ctr; *p < 0,05; **p < 0,01; ***p < 0,001 .

2. RATIONALE AND AIMS

EPN3 is the least characterized member of the Epsin family of endocytic adaptors, which in *Homo sapiens* comprises of 3 members⁷⁰⁻⁷². These proteins have been studied for their role in CME, in particular, in the regulation of CCP maturation^{23,70,71,73}. Since their discovery 20 years ago⁶⁷, epsins have been connected with the endocytic regulation of specific pathways, such as VEGF¹⁰³⁻¹⁰⁷, Notch⁸⁴⁻⁸⁶, EGFR⁹³⁻⁹⁵, Wnt¹¹¹, and RhoGTPase^{97,98}, both in immortalized cell lines and *in vivo* in mouse models. These data have led to the hypothesis that the epsin family is, most likely, implicated in regulation of specific signaling pathways in a cell-context dependent manner, rather than being involved in a general endocytic mechanism of CCP formation⁷².

The peculiarity of EPN3 is that, even if very similar to EPN1 and EPN2 at the structural level, it is expressed at very low levels in almost all tissues, except for gastric parietal cells¹¹² and migrating keratinocytes¹¹³, suggesting a more specialized function of EPN3 compared with other epsins. Our preliminary data indicate that EPN3 is overexpressed in 40% of all BCs, and correlates with more a aggressive cancer phenotype and increased risk of distant metastasis. Moreover, overexpression of EPN3 in MCF10A cells, a non-tumoral immortalized breast epithelial cell line, induces the acquisition of a partial mesenchymal-like morphology. Based on these observations, we hypothesized that higher levels of EPN3 in breast cancer could induce invasive and migratory properties through EMT acquisition, ultimately leading to higher risk of metastasis formation.

The overall aims of this PhD project were to:

- Understand the molecular mechanisms governing EPN3-induced pEMT in MCF10A cells.

We wanted also to investigate whether the role of EPN3 in cancer represents a gain-of-function role linked to its overexpression, or instead an exaggeration of its physiological function.

- Characterize the role of EPN3 in mouse mammary gland biology and tumorigenesis. To this aim, we have used: i) a conditional EPN3 knock-in (KI) mouse that overexpresses the transgene in the epithelial compartment, to mimic the tumoral context, to study the possible appearance and development of mammary tumors and/or metastasis ; ii) an EPN3 KO mouse to characterize the endogenous role of EPN3 in mouse mammary gland physiology and development.

Towards understanding the molecular mechanisms governing EPN3-induced pEMT in MCF10A cells, we first attempted to gain an in-depth understanding of the effects of EPN3 overexpression in MCF10A cells by comparing the expression of different EMT markers in MCF10A-EPN3 cells with those expressed in MCF10A-TWIST cells (Results, Section 1). Next, we tried to unravel the molecular mechanisms of EPN3 function using an unbiased approach based on the characterization of EPN3-specific interactors in MCF10A-EPN3 cells by mass spectrometry (Results, Section 2). This analysis led to a series of investigations focused on the role of EPN3 in E-Cadherin endocytosis and AJ stability and the crosstalk with the TGF β signaling pathway under pathological and physiological conditions (Results, Sections 2 and 3). Finally, using *in vivo* EPN3

KI and KO mouse models, we investigated the role of EPN3 in the regulation of mouse mammary gland morphogenesis (Results, Section 4).

3. RESULTS

3.1. Epn3 overexpression induces a partial EMT program in mammary epithelial MCF10A cells

As discussed in the preliminary data paragraph (see Introduction, Section 1.5), the overexpression of EPN3 in MCF10A cells (MCF10A-EPN3) induces a change in cellular morphology, with the appearance of a mesenchymal-like phenotype compared to MCF10A control cells (MCF10A-Ctr), and causes the upregulation of mesenchymal markers, such as, N-Cadherin and Vimentin and the slight downregulation of the epithelial marker, E-Cadherin, at the protein level.

To gain insights into the morphological and molecular changes observed in MCF10A-EPN3 cells, we further characterized the mRNA expression levels of different known EMT factors (Figure 16). In particular, we tested the transcriptional status of EMT markers – E-Cadherin, N-Cadherin and Vimentin – and EMT transcription factors – Snail and Zeb1 –, which are considered as master regulators of the EMT transcriptional program^{186,239,260,261}. We compared MCF10A-EPN3 cells with the epithelial-like control MCF10A-Ctr cells, and with MCF10A cells overexpressing the EMT inducer TWIST (MCF10A-TWIST), which display a strong mesenchymal phenotype and upregulation of EMT markers. While both mesenchymal markers (N-Cadherin and Vimentin) and transcription factors (Snail and Zeb1) are upregulated upon Epn3 or Twist overexpression, the epithelial marker, E-Cadherin, is downregulated, at the mRNA level, only in MCF10A-TWIST cells and not in MCF10A-EPN3 cells (Figure 16).

Taking into account these results, the overexpression of Epn3 in mammary epithelial cells appears to induce the so-called pEMT state^{184,188,189}), characterized by the co-occurrence of epithelial (E-Cadherin protein is only partially downregulated) and mesenchymal properties (e.g., expression of N-Cadherin and Vimentin). In addition, the fact that E-Cadherin transcript levels are unchanged, while the protein levels are decreased, suggests post-transcriptional regulation of E-Cadherin by EPN3.

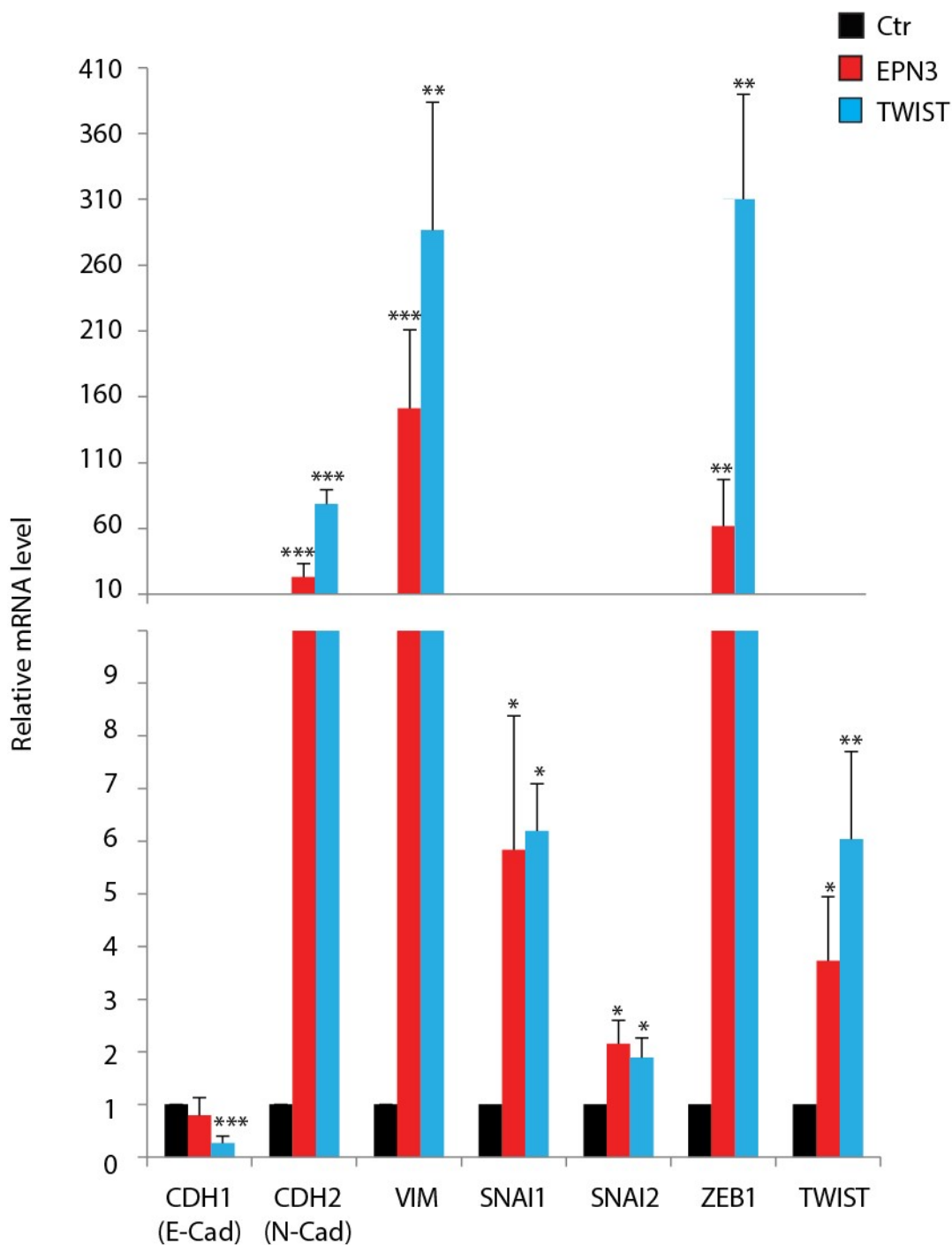


Figure 16: Transcriptional expression of EMT markers in MCF10A cells following EPN3 and TWIST expression

Relative mRNA expression levels of the indicated EMT markers in MCF10A-control (Ctr), -EPN3-overexpressing and -TWIST overexpressing cells were evaluated by RT-qPCR analysis. Data were normalized on 3 housekeeping genes (18S, GAPDH, ACTB) and expressed as relative to levels in the MCF10A-Ctr sample (mean \pm SD of at least three independent experiments). p-values vs. Ctr; * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$.

3.2. Regulation of EPN3-induced partial EMT through E-Cadherin endocytosis and adherens junction stability

In order to unravel the molecular mechanisms of EPN3-induced pEMT in MCF10A cells, we decided to adopt an unbiased approach involving the characterization of all proteins that interact physically with EPN3, i.e., the “EPN3 interactome”. To this aim, we coupled immunoprecipitation (IP) with mass spectrometry to identify proteins that bind, directly or indirectly, with EPN3. In the following sections, I have described the set-up and optimization of the experiment and the subsequent validation. After that, I have further characterized the mechanism of action, showing that the binding between EPN3, adherens junction components and clathrin machinery is functional and instrumental for EPN3-induced pEMT in MCF10A.

3.2.1. Setting up of an EPN3-FLAG immunoprecipitation to perform mass-spectrometry analyses

For the EPN3 IP step, we took advantage of the fact that the ectopically expressed EPN3 in MCF10A cells is FLAG-tagged at the N-terminus of the protein. We first analyzed by Western Blot (WB) the anti-FLAG (α -FLAG) IP efficiency in MCF10A-EPN3 cells, and observed that almost all FLAG-Epn3 was immunoprecipitated (Figure 17A). Then, we compared by Coomassie gel the α -FLAG co-immunoprecipitating proteins from MCF10A-Ctr and MCF10A-EPN3 cells. The analyses showed many bands that co-immunoprecipitated with the α -FLAG antibody in both

the control and the FLAG-EPN3 overexpressing cells (Figure 17B), suggesting that, under the conditions used, the α -FLAG antibody was very unspecific.

For this reason, we set up a more stringent and specific co-immunoprecipitation (Co-IP) protocol. In particular, we:

- incubated the lysates with agarose-sepharose beads before performing the IP, in order to reduce the non-specific bindings to the beads;
- decreased the antibody concentration from 75 μ g to 25 μ g per mg of lysate;
- increased salt concentration during the IP washes, from NaCl 150 mM to 250mM;
- increased number of washes, from 4 to 5.

By WB, we verified that the IP efficiency was maintained under these more stringent experimental conditions (Figure 17C). We then repeated the Coomassie gel analysis, which revealed that the number of unspecific bands in the MCF10A-Ctr sample was markedly decreased (Figure 17D). In the MCF10A-EPN3 sample, we observed a major specific band corresponding to EPN3 bait (Figure 17D). Thus, we used this settings to perform IP α -FLAG (from MCF10A-Ctr and MCF10A-EPN3 cells) and mass spectrometry analyses to identify Epn3 interactors.

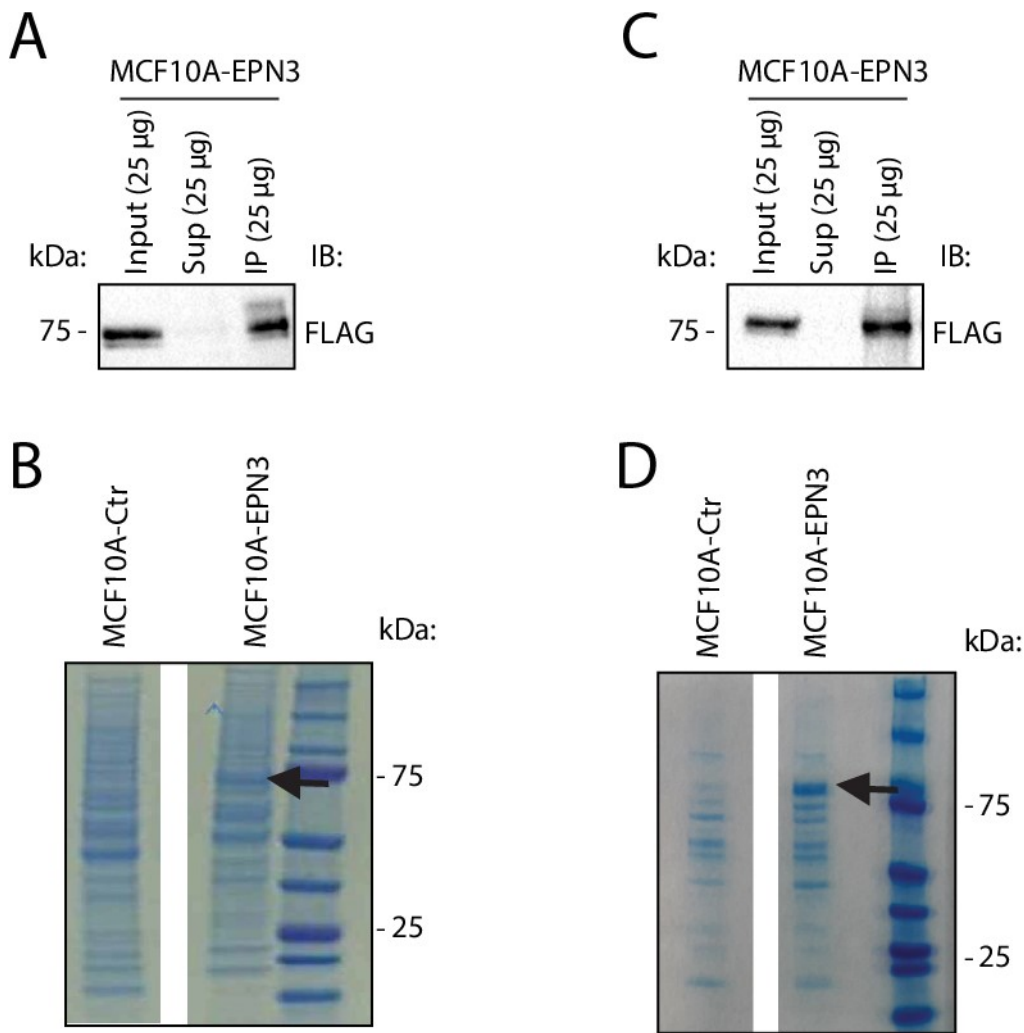


Figure 17: Optimization of the anti-FLAG immunoprecipitation reaction in MCF10A-EPN3 cells

A) Western Blot (WB) analysis of lysates from MCF10A-EPN3 cells after anti-FLAG immunoprecipitation (IP) with the α -FLAG antibody. Input: 25 μ g of totally lysate; Sup: 25 μ g of supernatant; IP: 25 μ g of total IP. MW markers are shown on the left. B) Coomassie gel of lysates from MCF10A-Ctr and -EPN3 cells after anti-FLAG immunoprecipitation with α -FLAG antibody. Black arrow indicates the EPN3 bait in the MCF10A-EPN3 sample. MW markers are shown on the right. C) WB analyses of lysates from MCF10A-EPN3 cells after anti-FLAG immunoprecipitation with α -FLAG antibody upon setting of new IP conditions. Input: 25 μ g of totally lysate; Sup: 25 μ g of supernatant; IP: 25 μ g of total IP. MW markers are shown on the left. D) Coomassie gel of lysates from MCF10A-Ctr and -EPN3 cells after anti-FLAG immunoprecipitation with α -FLAG antibody under new more stringent IP conditions. Black arrow indicates the EPN3 bait in the MCF10A-EPN3 sample. MW markers are shown on the right.

3.2.2. Epn3-specific interactome reveals an enrichment of components of the endocytic machinery and adherens junctions

In order to identify the EPN3 interactome, after the Coomassie gel, we digested and processed the samples to analyze the peptides through the use of a mass spectrometer. For each protein obtained, we considered them as identified only when at least 2 different peptides were recognized and at least one of them had to be unique, not redundant with other proteins. Moreover, to identify only EPN3-specific interactors and to be as much stringent as possible, we selected in the mass spectrometry analysis only those proteins that were found in the IP of MCF10A-EPN3, but not of MCF10A-Ctr, cells, exclude: I called these interactors “On/Off interactors”, to reflect their specific presence only in the MCF10A-EPN3 sample. We obtained a list of 430 EPN3 On/Off interactors from a total of 1314 protein identified (Figure 18A). Of note, despite the optimization of the α -FLAG IP, we still observed that the majority of the identified proteins were common to both EPN3 and Ctr samples (818 proteins, Figure 18A), suggesting that they are non-specific interactors.

Nevertheless, we compared the list of EPN3-specific interactors (430) with the list of non-specific interactors (66+818) and submitted them to Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway database through EnrichR, a specific enrichment analyses software²⁶² (Figure 18B). We found – among the 10 most statistically significant enriched pathways – a strong presence of proteins implicated in ribosome biogenesis and spliceosome both in MCF10A-EPN3-derived samples and in -Ctr samples, suggesting that proteins belonging to these complexes were not specifically bound. Interestingly, we found a statistically significant enrichment of proteins

related to the endocytic pathway only in EPN3 sample. Moreover, in this sample, we also observed a specific enrichment for a network of 4 pathways that are highly interconnected: AJs, TJs, cytoskeleton and focal adhesion (Figure 18B).

I concentrated the attention on the most relevant components of the clathrin machinery and the AJs (Figure 18C). Indeed, these results showed that EPN3, whose role was almost unknown, is able to bind many different components of the clathrin machinery: both the CHC and the CLC, but also different adaptor proteins, like the main components of the AP-2 complex, Cbl (Figure 18C, left). These reinforced the idea that EPN3 role is as an endocytic protein. Moreover, the presence of proteins related to components of membrane stability, cell migration and identity (such as AJs, TJs, cytoskeleton and focal adhesion) suggested me that EPN3 could conduct its function on PM components implicated in identification and acquisition of EMT phenotype. The previous findings on the characterization of the pEMT showed that E-Cadherin (the main component of the AJs), was slightly downregulated at the protein level but not paired by a transcriptional repression, pointing to a putative question mark on the regulation of AJs in the EPN3-induced pEMT. For these reasons the attention was focused on the AJ members, strongly regulated during EMT. Interestingly, except for E-Cadherin, EPN3 was found to be linked with different components of the junctions (p120 , β -catenin, α -catenin) but also with regulators and interactors of AJs (ZO1, PARD3, Actinin4) (Figure 18C, right).

Interestingly, destabilization of AJs and downregulation of E-Cadherin are required for the establishment of the full EMT phenotype and the acquisition of invasive traits; indeed, it has been shown that decreased E-Cadherin levels promotes metastatic behavior and increases tumor

cell invasion ¹⁶⁰. More interestingly, dysregulated E-cadherin trafficking leading to its downregulation from the PM and relocalization into the cytosol (in the absence of complete loss of E-Cadherin levels) has been linked to the acquisition of the pEMT state ^{183,184,189,197}. Thus, a possible physical and functional interaction of Epn3 with components of AJs and/or E-Cadherin itself might directly explain the EPN3-induced pEMT phenotype.

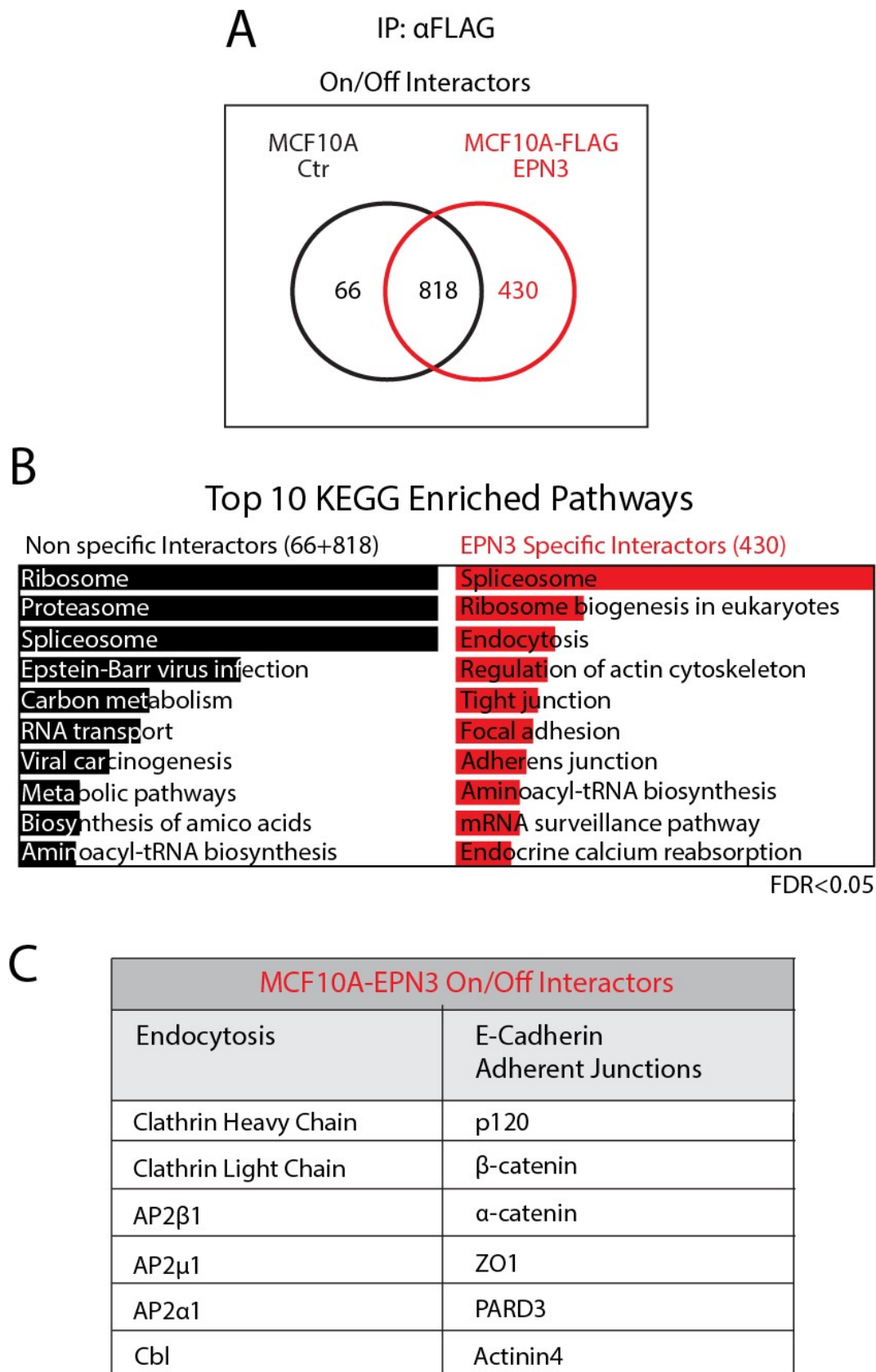


Figure 18: EPN3 interacts with members of endocytic machinery and AJs as determined by mass spectrometry

(continued on next page)

A) Top: VENN diagram of 1314 proteins identified by mass spectrometry analysis of proteins immunoprecipitating from lysates, incubated with the α -FLAG antibody, from MCF10A-Ctr and MCF10A-EPN3 cells. The proteins have been assigned to either MCF10A-Ctr (66) or -EPN3 (430) when present only in that sample, the others (818) have been linked to the MCF10A-Ctr specific proteins and considered as contaminants (66+818). The 430 proteins identified only in MCF10A-EPN3 cells are called On/Off Interactors. Analyses performed with the help of Dr. Confalonieri Stefano (IFOM-IEO Campus). B) Table of the top 10 statistically significant enriched pathways (FDR <0.05) in EPN3 On/Off interactors versus non-specific interactors determined using EnrichR software and the KEGG Pathway database. The length of each pathway represents the statistical significant enrichment, computed by multiplying the p-value of the Fisher's exact test with the z-score of the deviation from the expected rank. C) Table of main interactors of EPN3 in the endocytic pathway and in the E-Cadherin / AJ pathway.

3.2.3. Setting up of co-immunoprecipitation strategy to validate mass-spectrometry results

Given that EPN3 belongs to a family of endocytic adaptors and given the well-established link between deregulated E-Cadherin endocytosis and EMT¹⁸³, we concentrated our attention on the endocytic machinery (CHC) and the AJ components (p120, β -catenin) identified in the EPN3 interactome, to evaluate whether EPN3 could link AJs to the clathrin machinery.

Since in the mass spectrometry experiment we detected many non-specific proteins, although we could not detect E-Cadherin itself co-immunoprecipitating with EPN3, we decided to optimize our co-IP experiments. First, we used a specific α -EPN3 polyclonal antibody generated in our laboratory instead of the α -FLAG antibody, which revealed to be very unspecific in IP experiments. The use of α -EPN3 antibody also allowed us to IP the endogenous EPN3 from MCF10A cells, and to assess whether also the endogenous protein is able to interact with AJs. Second, in order to increase the co-IP efficiency, we treated cells - prior to lysis - with a cell-permeable crosslinking agent, 1% formaldehyde (FA), in order to stabilize the protein complexes formed within the cells²⁶³. This treatment is routinely used to stabilize weak and/or dynamic

protein-protein interactions. Moreover, FA is the crosslinker with the shortest arm (1 nM), able to generate covalent bindings between proteins that are in close proximity to each other, while minimizing the risk of artefacts²⁶³.

For the crosslinking treatment of cells, we followed the protocol suggested by²⁶³. Briefly, we treated MCF10A-Ctr and -EPN3 cells with 1% FA or with PBS for 10 minutes at room temperature (RT). This time was optimized in order to allow the action of the crosslinker and to avoid too much exposure of cells to the crosslinker that might cause artefacts and aggregates. The 1% FA treatment ensured stable crosslinking of proteins until samples are heated for at least 30 minutes at 95°C. After crosslinking, cells were then lysed in RIPA lysis buffer (see Methods). Lysates were diluted in sample buffer, treated for 5 minutes at 65°C or 30 minutes at 95°C and loaded on SDS-page gels to check for the crosslinking efficiency. As positive controls for the efficacy of the crosslinking, we tested the homodimerization of Integrin- β 1 by WB. We observed that Integrin- β 1 homodimers were stabilized upon 1% FA treatment, as evidenced by a slower migrating band, compared to the non-crosslinked lysate, at a MW corresponding to Integrin- β 1 homodimers. The treatment of samples at 95°C for 30 minutes was able to revert the crosslinking showing only Integrin- β 1 monomers (Figure 19A).

We then set the co-IP with the α -EPN3-specific antibody on the cross-linked lysate. Indeed, we tested the IP efficiency upon crosslinking in MCF10A-Ctr and -EPN3-overexpressing cells; with this antibody, we could IP both the endogenous and the overexpressed EPN3 with an efficiency of about 50% (Figure 19B). We therefore used this experimental setting to validate the

mass spectrometry results, since it allows the stabilization of labile/low stoichiometry interactions and the comparison of endogenous and overexpressed EPN3.

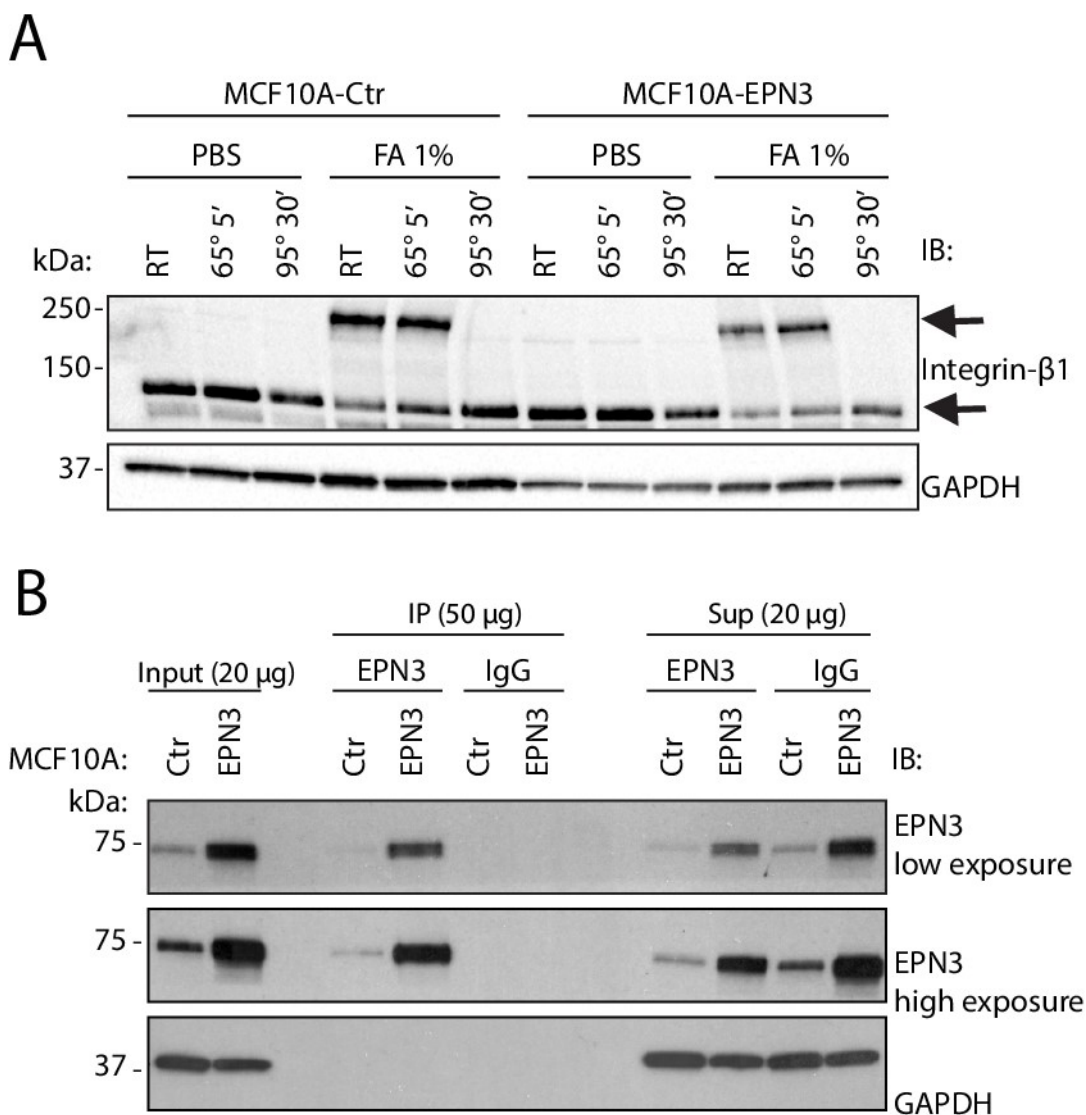


Figure 19: Optimization of EPN3 immunoprecipitation following crosslinking of protein lysates

A) WB analysis of Integrin- β 1 homodimerization in MCF10A-Ctr and MCF10A-EPN3 samples treated or not with 1% formaldehyde and treated or not at 65°C or 95°C. Black arrows indicate the homodimer (250 kDa) or the monomer (125 kDa) of Integrin- β 1. GAPDH was used as a loading control. MW markers are shown on the left. B) WB analyses of IP experiment with α -EPN3 antibody in -Ctr and -EPN3 cells. Input: 20 μ g of total lysate; Sup: 20 μ g of supernatant; IP: 50 μ g of total IP. GAPDH was used as a loading control. MW markers are shown on the left.

3.2.4. Epn3 co-immunoprecipitates with Clathrin, E-Cadherin and AJ components

Using the above described settings, we confirmed the interaction of EPN3 with CHC, unveiling also that it is a strong interactor (Figure 20, top). Moreover, we confirmed the interaction with p120 and β -catenin, components of the AJs. Interestingly, we were also able to score an interaction with E-Cadherin (Figure 20, bottom).

These interactors are recovered in the co-IP both with the endogenous and the overexpressed EPN3 protein. Of note, however, while EPN3 overexpression is about 10-15 fold in MCF10A-EPN3 compared to MCF10A-Ctr, none of the interactors showed such an increase in the co-IP from EPN3-overexpressing cells compared to control cells; instead, we observed an enrichment of only 2-4 folds. One possible explanation could be that EPN3 levels reached saturation, arguing that EPN3 is not the limiting factor in the interaction; alternatively, in MCF10A-EPN3 cells the interaction are increased but they are more dynamic because it allows a faster turnover of E-Cadherin at the PM, as it is shown in the next paragraphs.

These results, in summary, show that both endogenous and overexpressed EPN3 are able to bind the clathrin machinery and the AJ components. EPN3 overexpression increases these interactions.

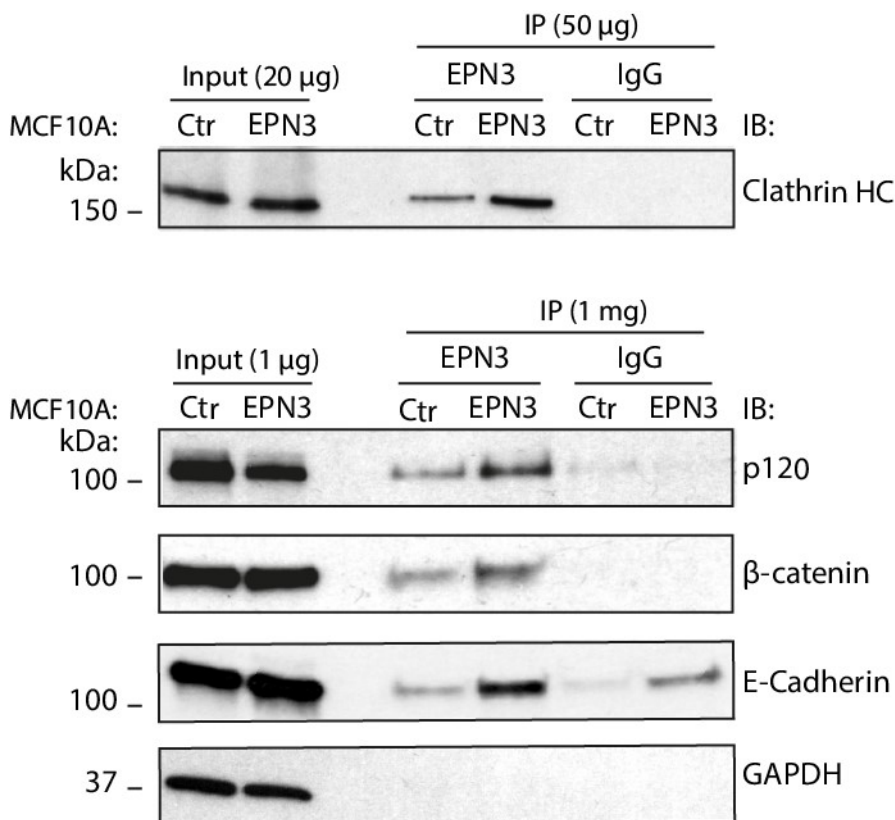


Figure 20: EPN3 co-immunoprecipitates with components of endocytic machinery and adherens junctions

(TOP) WB analysis of Clathrin heavy chain (HC) in the IP experiment with α -EPN3 antibody in MCF10A-Ctr and -EPN3 cells. Input: 20 μ g of totally lysate; IP: 50 μ g of total IP. MW markers are shown on the left. (BOTTOM) WB analysis of p120, β -catenin and E-Cadherin in the IP experiment with α -EPN3 antibody in MCF10A-Ctr and -EPN3 cells. Input: 1 μ g of totally lysate; IP: 1 mg of total IP. GAPDH was used as loading control. MW markers are shown on the left.

3.2.5. EPN3 overexpression alters E-Cadherin stability at the Plasma Membrane

Our results show that EPN3 is able to bind AJs components and clathrin machinery suggesting a putative alteration of E-Cadherin endocytosis at the PM. Interestingly, EPN3 overexpression also causes a partial downregulation of the E-Cadherin protein levels, but not of its transcriptional levels, allowing us to suppose that alteration of PM E-Cadherin stability could be sufficient to give rise to a pEMT program.

For these reasons, we decided to check the localization of E-Cadherin in MCF10A-Ctr and MCF10A-EPN3 cells (Figure 21A). Indeed, control cells grow as epithelial clusters and E-Cadherin appears to be mainly localized at the PM both, in low and high density fields of view. Instead, MCF10A-EPN3 cells represent a mixture of epithelial and mesenchymal properties, showing almost no E-Cadherin staining in scattered single cells, instead in epithelial clusters E-Cadherin appears to be both at the PM and in the cytosol, with the formation of destabilized AJs (Figure 21A).

To test the possibility that EPN3 is altering E-Cadherin stability at the PM, we performed an *in vivo* staining for PM E-Cadherin: MCF10A-Ctr and -EPN3 cells in suspension were labelled at +4°C with a specific antibody recognizing the extracellular domain of E-Cadherin, followed by staining with fluorescently-labelled secondary antibody. We subsequently analyzed by fluorescence-activated cell sorting (FACS) the percentage of E-Cadherin-positive cells. While the epithelial-like MCF10A-Ctr cells were all positive for PM E-Cadherin (98.9%), 34.3% of MCF10A-EPN3 cells had already lost PM E-Cadherin at steady state, confirming the hypothesis that Epn3 influences the stability of AJs (Figure 21B).

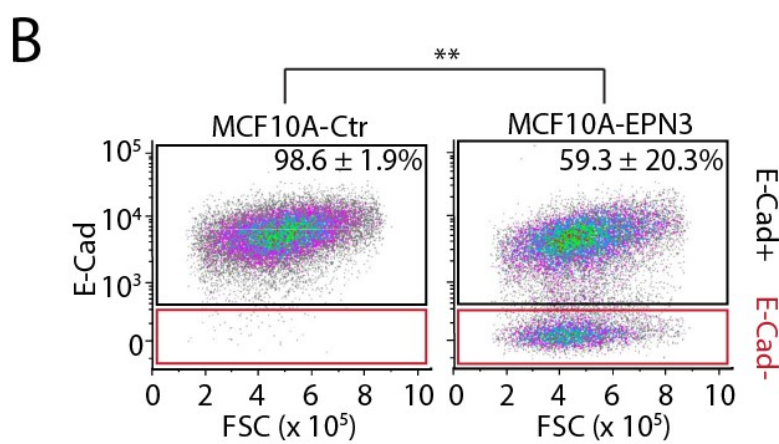
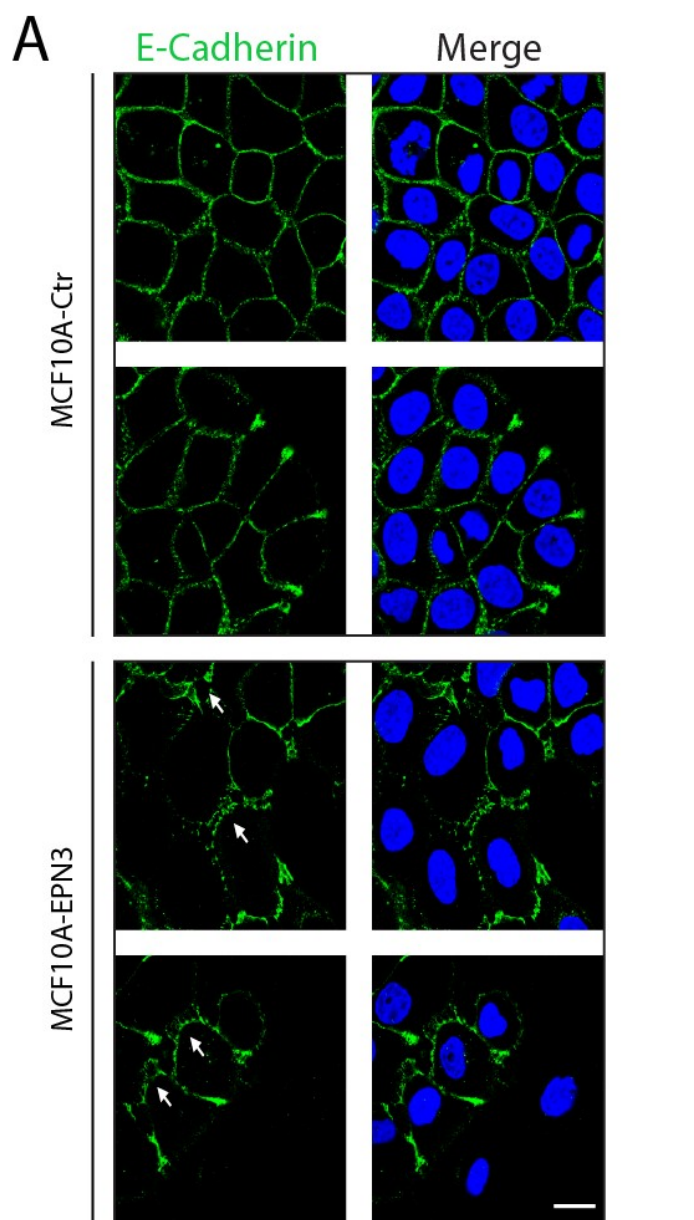


Figure 21: EPN3 destabilizes plasma membrane E-Cadherin and adherens junctions
(continued on next page)

A) Confocal images of MCF10A-Ctr and MCF10A-EPN3 cells stained for E-Cadherin (green) in fixed and permeabilized samples. White arrows indicate areas of plasma membrane (PM) E-Cadherin with destabilized junctions. Nuclei were stained with DAPI (blue). Bar scale, 30 μ m. B) Representative FACS analyses of MCF10A-Ctr and MCF10A-EPN3 cells stained *in vivo* for E-Cadherin. Black and red gates represent E-Cad positive and E-Cad negative cells, respectively, and have been set according to the not stained control samples. MCF10A-Ctr display 99.6% of cells positive to E-Cad, while the portion of MCF10A-EPN3 positive is 59.3%. FSC, forward scatter. On the top percentage of E-Cad positive cells \pm SD in eighteen independent experiments. P-value was calculated with Student's t test, two-tailed. **, <0.01.

3.2.6. EPN3 overexpression increases E-Cadherin endocytosis

Overall our results suggest that EPN3 could act as a classical endocytic protein and its overexpression could cause an altered trafficking and turnover of E-Cadherin. To directly test this hypothesis, we performed an E-Cadherin *in vivo* internalization assay. We took advantage of the antibody previously described to stain the extracellular domain of E-Cadherin *in vivo*, then cells were put back at 37°C or not, for the indicated time points (10, 30, 90 and 180 minutes), allowing the internalization of the antibody-E-Cadherin complex. We removed the antibody-E-Cadherin complexes remaining at the PM at each time point through an acid wash (pH 2.5), and checked by FACS analyses the amount of internalized E-Cadherin. The results showed that in MCF10A-Ctr cells, the internalized E-Cadherin increased linearly overtime; while in MCF10A-EPN3 cells the kinetics of E-Cadherin internalization was significantly accelerated compared to control cells (Figure 22), arguing that Epn3 overexpression increases the rate of E-cadherin endocytosis.

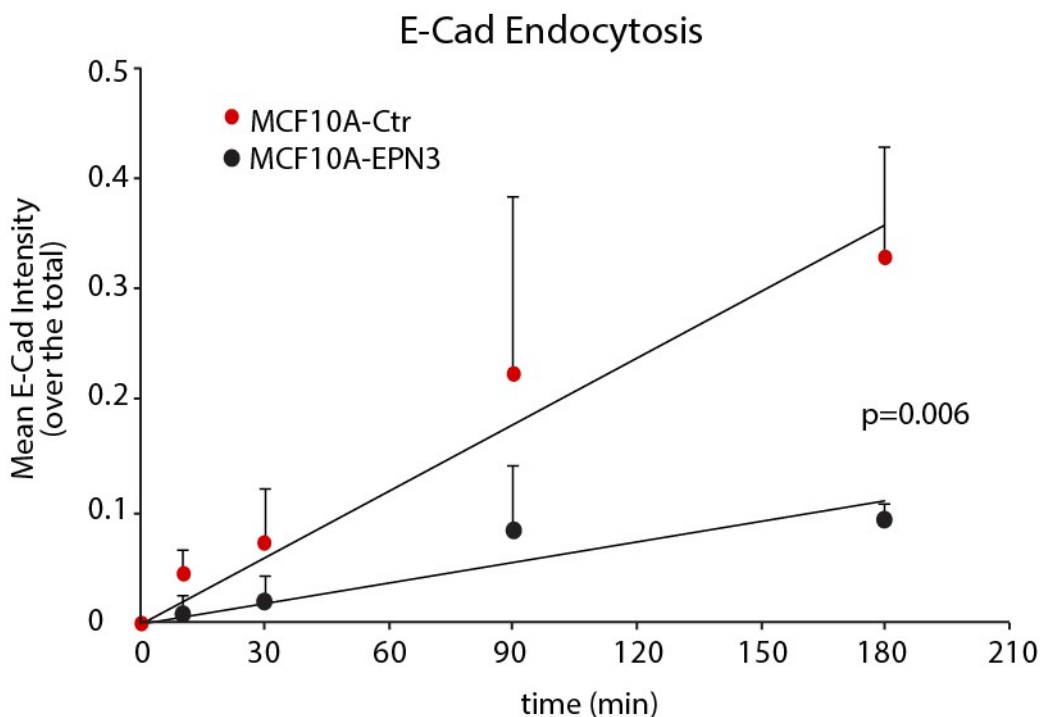


Figure 22: E-Cadherin endocytosis is accelerated upon EPN3 overexpression

Time-course analysis of E-Cadherin in vivo internalization assay in MCF10A-Ctr and MCF10A-EPN3 cells measured through FACS analysis. Cells were incubated in vivo with a specific E-Cadherin antibody that recognizes the extracellular domain at 4°C. After the secondary antibody incubation, cells were fixed at 0 minutes or incubated again at 37°C in order to evaluate E-Cadherin internalization. To remove the excess of PM E-Cadherin, acid wash treatment was applied. On the y-axis, data are reported as fraction of mean fluorescence intensity of internalized E-Cadherin signal over the total \pm SD of at least seven independent experiment. P-value was calculated with Student's t test, two-tailed.

3.2.7. EPN3 overexpression induces the activation of β -Catenin/TCF4 pathway

The destabilization/downregulation of AJs is known to be necessary to induce acquisition of mesenchymal properties. One of the most characterized pathways downstream of E-Cadherin downregulation is the activation and translocation of β -catenin into the nucleus¹⁴⁷. Active β -catenin (dephosphorylated) acts as a co-transcription factor with the TCF/LEF family to transcriptionally activate target genes linked to EMT acquisition¹⁴⁷. For this reason, we tested the localization of β -catenin in MCF10A-Ctr and MCF10A-EPN3 cells by immunofluorescence. In the epithelial-like MCF10A-Ctr cells, active β -catenin was localized both at the PM and in the

nucleus, while in MCF10A-EPN3 cells β -catenin showed an increased nuclear staining. We quantified the amount of total and nuclear active β -catenin per cell based on staining intensity, and found that in MCF10A-EPN3 cells there was a statistically significant enrichment of the ratio of nuclear/total β -catenin compared to control cells (Figure 23A).

Moreover, upon transient knock-down (KD) of one of the best characterized partners of β -catenin in the nucleus, TCF4, a member of the TCF/LEF family of transcription factors, we were able to completely revert the acquisition of a pEMT state in MCF10A cells, as assessed by reversion to an epithelial morphology and downregulation of mesenchymal markers, N-Cadherin and Vimentin (Figure 23B).

These results suggest that EPN3 overexpression induces E-Cadherin internalization and destabilization of cell-cell junctions, which, in turn, triggers β -catenin nuclear translocation and TCF4-dependent upregulation of EMT target genes.

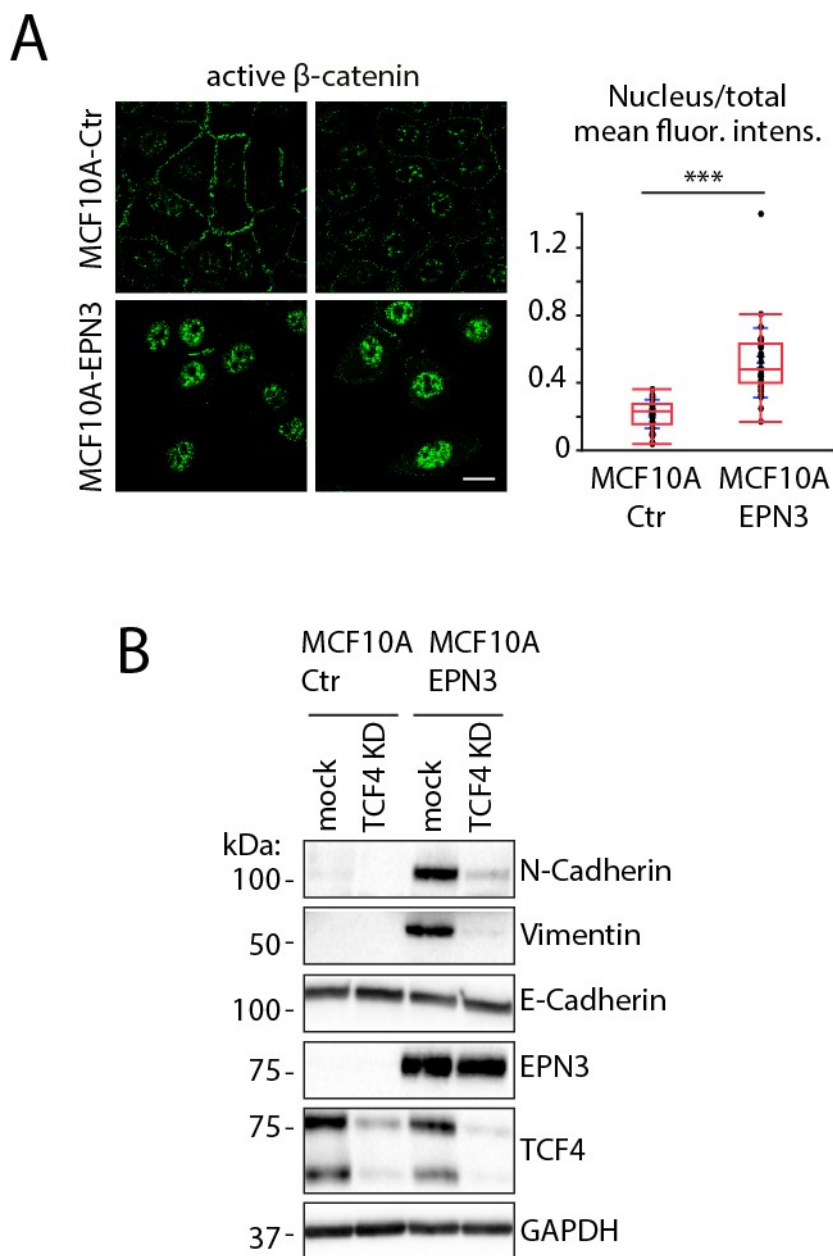


Figure 23: EPN3-induced pEMT is dependent on the β -catenin/TCF4 pathway

A) (LEFT) Confocal images of MCF10A-Ctr and MCF10A-EPN3 cells stained for active β -catenin. Bar scale, 30 μ m. (RIGHT) Box plot of the ratio between nuclear and total fluorescence intensity per cell in MCF10A-Ctr (N=34) and -EPN3 (N=36) cells in one representative experiment out of three. The lines of the boxes represent – from the highest to the lowest – the 95th, 75th, 50th, 25th, 5th percentile. -value was calculated with Student's t test, two-tailed. ***, <0.001. Analyses performed with the help of Dr. Freddi Stefano (IFOM-IEO Campus). B) WB analysis of EMT markers in lysates from MCF10A-Ctr and -EPN3 mock transfected or TCF4 KD cells. GAPDH was used as a loading control. Experiment performed by Schiano Lomoriello Irene (IFOM-IEO Campus).

3.3. EPN3-induced partial EMT generates a TGF β -dependent positive feedback loop

3.3.1. EPN3-induced partial EMT generates a TGF β -dependent positive feedback loop

The TGF β pathway is one of the most relevant and best characterized pathways involved in the activation of EMT²⁰¹. Given the strong connection between the TGF β circuitry and mesenchymal acquisition, we decided to investigate the transcriptional status of the TGF β pathway in EPN3 overexpressing cells, investigating the expression of both TGF β ligands and receptors. In humans, there are 3 TGF β ligands and 2 TGF β receptors. Interestingly, we found that 2 out of 3 TGF β ligands are upregulated upon Epn3 overexpression and that the both receptors are increased compared to Ctr samples, even if not in a statistically significant manner. Moreover, as for EMT markers (see Figure 16), MCF10A-EPN3 cells confirm to be less strong in upregulation of mesenchymal markers, suggesting that the pEMT acquisition could be linked to the TGF β pathway (Figure 24).

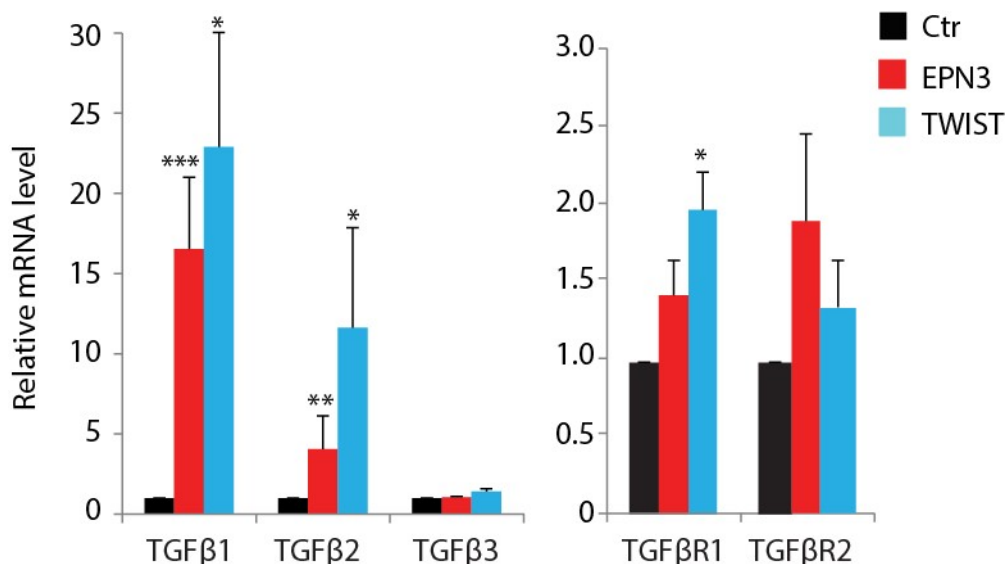


Figure 24: The effect of EPN3 upregulation on the transcriptional expression of TGFβ ligands and receptors

Relative mRNA expression levels of the indicated genes in MCF10A-control (Ctr), -EPN3-overexpressing and -TWIST overexpressing cells were evaluated by RT-qPCR analysis. Data were normalized on 3 housekeeping genes (18S, GAPDH, ACTB) and expressed as relative to levels in the MCF10A-Ctr sample (mean ± SD of at least three independent experiments). p-values vs. Ctr; *p < 0.05; **p < 0.01; ***p < 0.001.

3.3.2. EPN3 overexpression synergizes with TGFβ signaling

Given that, upon EPN3 overexpression, TGFβ ligands and receptors are upregulated, we performed an in-depth characterization of the role of the TGFβ pathway in EPN3-induced pEMT. We decided to test the response of MCF10A-Ctr cells and MCF10A-EPN3 cells to TGFβ stimulation.

TGFβ receptors, after ligand recognition, are able to dimerize (both homo and heterodimerization), autophosphorylate through their receptor kinase activity and activate a Smad-dependent signaling cascade. The Smad protein family is composed of co-transcription factors that, upon phosphorylation, translocate into the nucleus and activate the transcription of early (i.e., Snail) and late (i.e., Zeb1) target genes able to induce EMT.

We thus checked the activation of the canonical TGF β pathway in MCF10A-Ctr and MCF10A-EPN3 cells, in order to investigate a possible synergy of EPN3 with TGF β stimulation. To this aim, we stimulated cells with two different concentrations of TGF β – a suboptimal dose, 0.75 ng/ml, and a saturating dose, 5 ng/ml, – and we analyzed Snail, as an early target of TGF β stimulation, and Zeb1, as a late target (Figure 25). As expected, TGF β stimulation in MCF10A-Ctr cells activated Snail at day1 and day2, with a decrease at later time-points; moreover, its activation was enhanced at increasing ligand concentration (Figure 25). In contrast, the late mesenchymal markers (N-Cadherin, Vimentin) were upregulated only after prolonged treatment with TGF β (day7 and 14) at the highest dose (5 ng/ml, Figure 25). In contrast, Zeb1 is not upregulated in control cells, at both doses, even after 14 days of treatment with TGF β , and E-Cadherin protein levels were not downregulated in control cells in any of the conditions tested. Thus, in control cells and in the time-frame analyzed, chronic stimulation with saturating TGF β concentrations for two weeks is able to induce a partial mesenchymal switch.

In MCF10A-EPN3 cells, TGF β ligand stimulation showed a sustainment of Snail levels compared to control cells, which remained upregulated at day2 and day7, in particular at 5 ng/ml of TGF β . Moreover, Zeb1 levels, which were not increased in MCF10A-Ctr cells by TGF β treatment at any dose/time, were upregulated in MCF10A-EPN3 starting from d7, suggesting that the EMT program was enhanced compared to control cells. Indeed, late mesenchymal markers (N-Cadherin, Vimentin) were more highly upregulated in EPN3-overexpressing cells compared to control cells already at limiting TGF β doses, and E-Cadherin was downregulated after 14 days of treatment at saturating ligand concentrations (Figure 25).

These results confirm that Epn3 is able to induce pEMT, where cells show a mix of epithelial and mesenchymal features, with the upregulation of mesenchymal markers while maintaining E-Cadherin expression. The pEMT mediated by Epn3 is more prone to convert to an advanced EMT state upon treatment with TGF β , compared to cells in which EPN3 is not overexpressed.

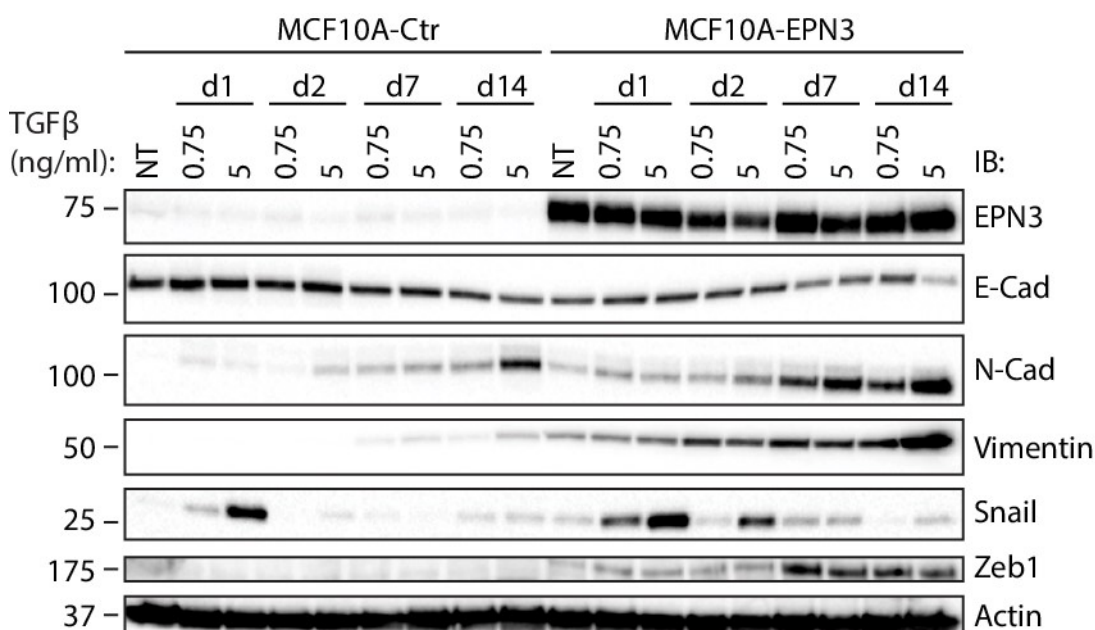


Figure 25: EPN3-overexpression sensitizes MCF10A cells to TGF β -induced EMT

WB analysis of MCF10A-Ctr and MCF10A-EPN3 cells treated with TGF β at the indicated concentrations and time-points to check activation of the TGF β pathway and expression of EMT markers. Actin was used as a loading control. MW markers are shown on the left. Experiment performed by Schiano Lomoriello Irene (IFOM-IEO Campus).

3.3.3. EPN3 overexpression increases E-Cadherin endocytosis upon TGF β stimulation

It is known that E-Cadherin stability and endocytosis are regulated by different growth factors, among others also TGF β is able to accelerate AJs turnover¹⁸³. Given our results on the cooperation between EPN3 and TGF β , we decided to test whether also E-Cadherin endocytosis is accelerated upon TGF β stimulation. We performed the E-Cadherin *in vivo* internalization assay, as described in Section 3.2.6, in MCF10A-Ctr or -EPN3 cells stimulated or not with TGF β for 30 minutes. As expected, TGF β treatment accelerated E-Cadherin endocytosis in control cells. E-Cadherin endocytosis was further increased in MCF10A-EPN3 cells stimulated TGF β compared to control cells and to unstimulated MCF10A-EPN3 cells, confirming that Epn3-overexpressing cells are more responsive and prone to TGF β ligand stimulation (Figure 26).

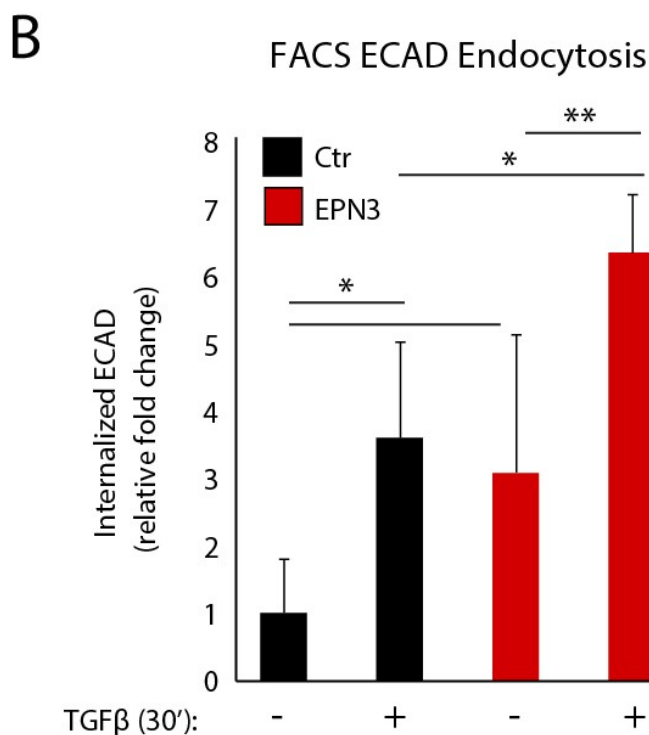
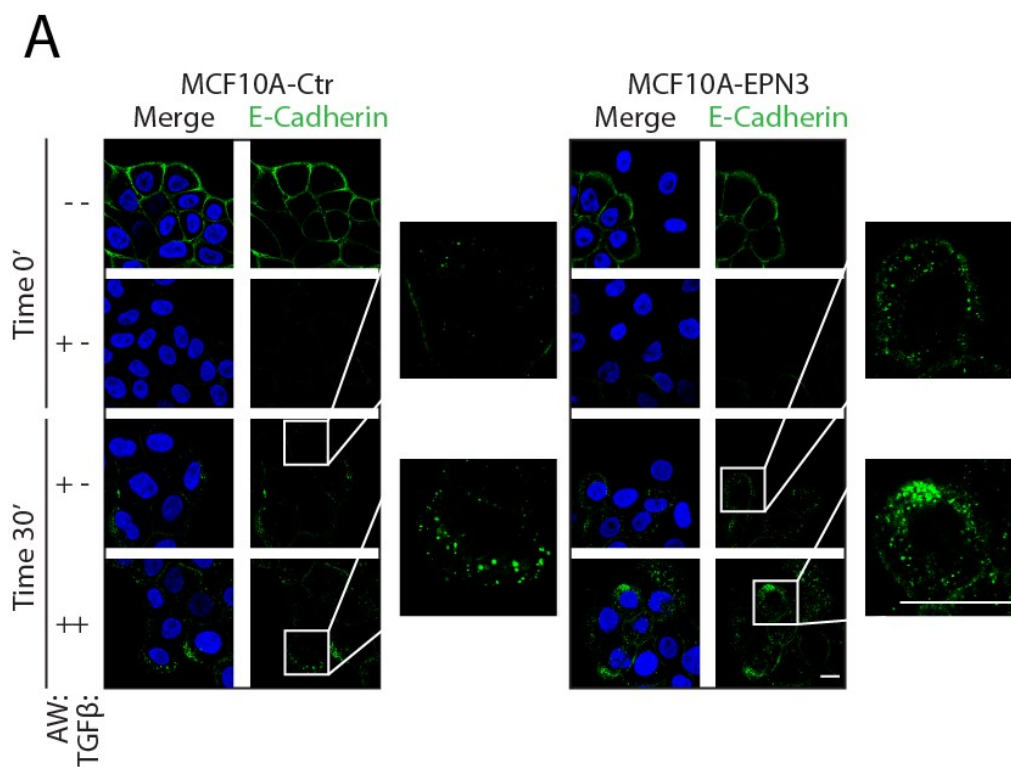


Figure 26: EPN3 overexpression cooperates with TGFβ in inducing E-Cadherin endocytosis

A) Confocal images of MCF10A-Ctr and MCF10A-EPN3 cells stained for E-Cadherin (green) after in vivo E-Cadherin internalization assay at 30 minutes with or without TGFβ stimulation. DAPI-stained nuclei (blue). Bar scale, 30 μm. B) Quantification of E-Cadherin in vivo internalization at 30 minutes with or without TGFβ stimulation measured through FACS analysis. On the y-axis, data are reported as fold-change relative to untreated MCF10A-Ctr cells of four independent experiment. P-value was calculated with Student's t test, two-tailed. p-values: *p < 0.05; **p < 0.01.

3.3.4. EPN3-induced partial EMT generates a TGF β -dependent positive feedback loop independent of EPN3 itself

Our results indicate that EPN3 overexpression upregulates the expression of TGF β ligands and receptors at the mRNA level. Moreover, TGF β stimulation cooperates with EPN3 in the activation of early and late target genes. These data suggest that EPN3-induced pEMT could generate a TGF β -dependent positive-feedback loop. To test this hypothesis, we decided to KD TGF β R1 in MCF10A-Ctr and MCF10A-EPN3 cells and assess effects on both the morphology and the different EMT markers at protein levels. The KD of TGF β R1 was able to revert the EPN3-induced pEMT, as cells reverted to an epithelial-like morphology (Figure 27A) and downregulated the EMT markers (Figure 27B).

Enhanced E-Cadherin endocytosis upon Epn3-overexpression, followed by β -catenin-dependent EMT activation, ultimately causes the upregulation of the TGF β pathway and augmented responsiveness to TGF β stimulation, generating a TGF β -dependent positive-feedback loop that maintains the pEMT state. That said, we asked whether the EPN3-induced pEMT is still addicted or not to EPN3 expression.

To test this issue, we KD EPN3 both in control and in overexpressing cells and checked the levels of expression of EMT markers. Differently from what occurred upon TGF β R1 KD, MCF10A-EPN3 cells still remained mesenchymal-like upon EPN3 KD (Figure 27A), and did not show any decrease in the expression of mesenchymal markers at the protein level (Figure 27B). These data argue that EPN3 overexpression induces a pEMT transcriptional program with the

establishment of TGF β -mediated self-sustaining positive-feedback loop, which renders the cells no longer addicted to the initial alteration, i.e., EPN3 overexpression.

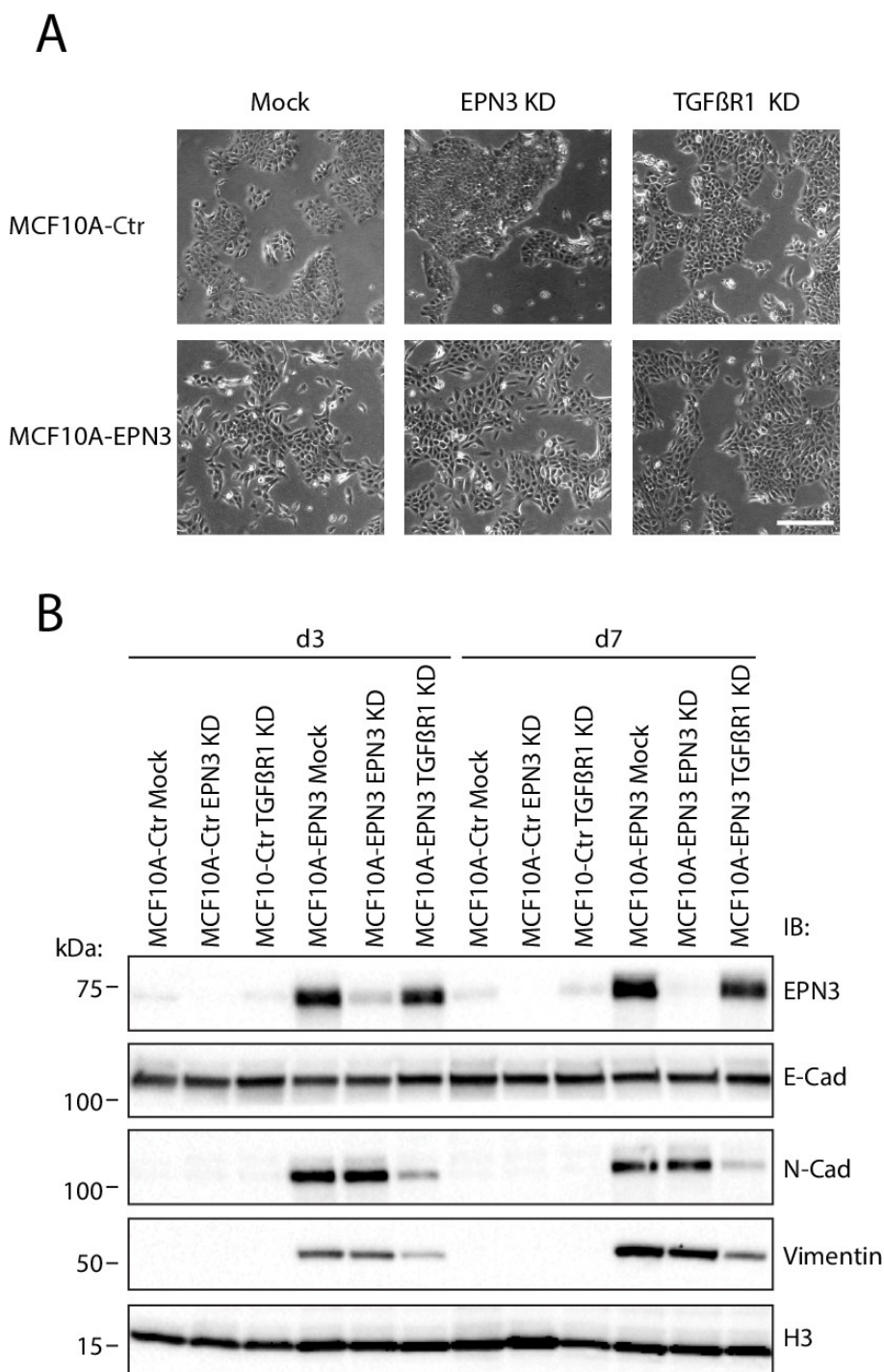


Figure 27: EPN3-induces pEMT generates a TGF β -dependent positive-feedback loop not addicted to EPN3 itself
(continued on next page)

A) Bright field representative images of MCF10A-Ctr and MCF10A-EPN3 cells treated with small-interfering RNA (siRNA) oligos against EPN3 or TGF β R1, or mock transfected, at 7 days of KD. Bar scale, 250 μ m. B) WB analysis of the cells described in A) at 3 or 7 days of KD to assess EMT markers expression. H3 was used as a loading control. MW markers are shown on the left.

3.3.5. Physiological EPN3 regulates TGF β -dependent transcriptional activation of EMT genes

One of the main questions arising from the results is whether the EPN3 function upon overexpression is an exaggeration of its physiological role or instead represents a new gain-of-function. Indeed, the Co-IP experiments showed that physiological EPN3 is able to bind – although not as strongly as the ectopically expressed EPN3 – clathrin, p120, β -catenin and E-Cadherin (see Section 3.2.4). For this reason, we decided to KD physiological EPN3 (Figure 28A) and ask whether also endogenous EPN3 influences TGF β responsiveness. We performed the KD of endogenous EPN3 in MCF10A-Ctr cells and then stimulated cells with TGF β for different time-points. We observed that the transcriptional upregulation of EMT markers, such as *CDH1* (N-Cadherin), *VIM* (Vimentin), *ZEB1* and *TGFB1* induced upon TGF β treatment, is decreased in EPN3 KD cells after 24 hours of TGF β treatment, suggesting that endogenous EPN3 contributes to the responsiveness of cells to TGF β (Figure 28B). These results confirm that the function of overexpressed EPN3 is most likely an exaggeration of its endogenous function.

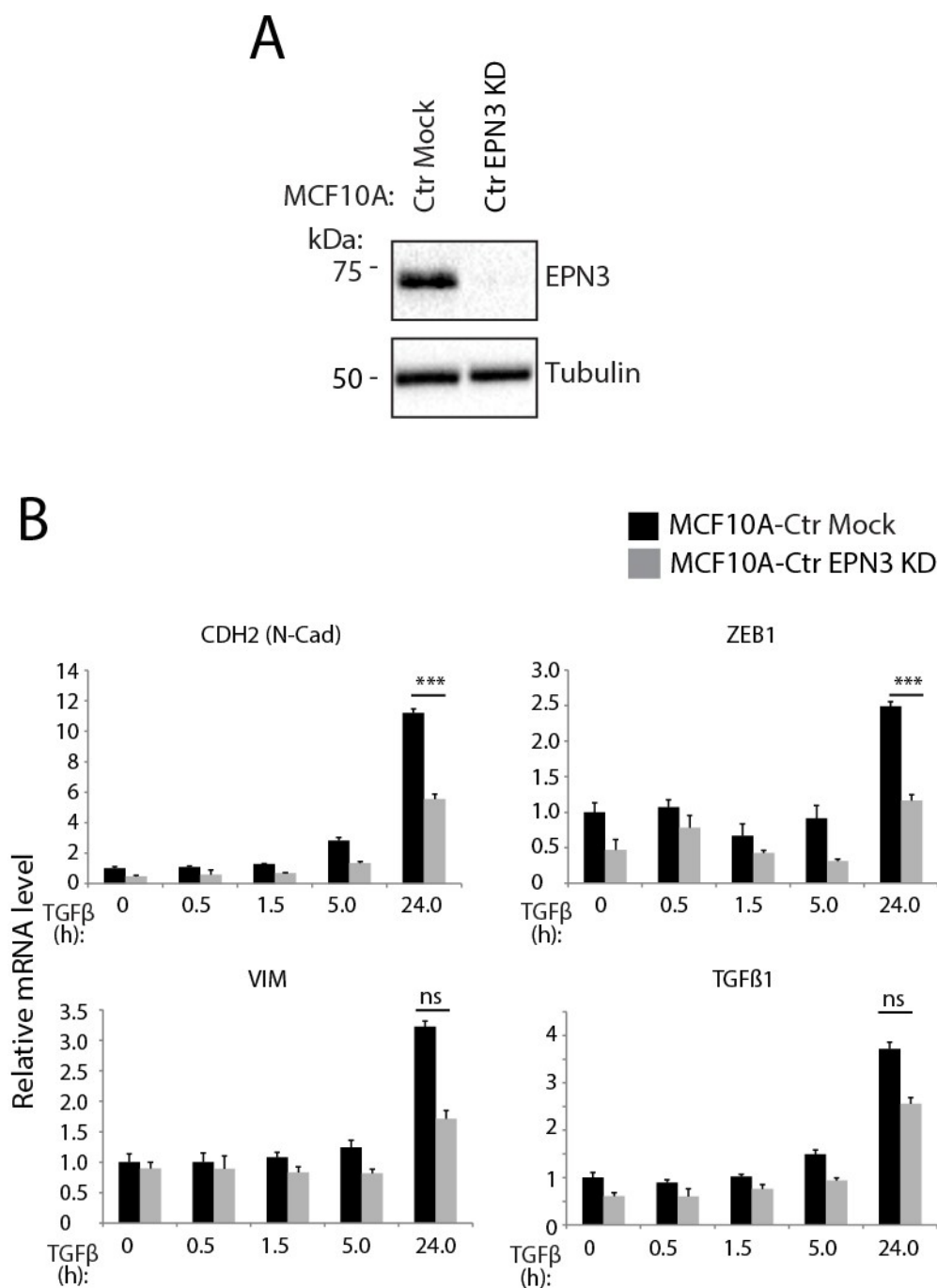


Figure 28: Endogenous EPN3 expression is implicated in activation of the TGFβ pathway

A) WB analysis of EPN3 levels in MCF10A-Ctr mock transfected or EPN3 siRNA transfected cells. Tubulin was used as a loading control. MW markers are shown on the left. B) Relative mRNA expression levels of the indicated genes were evaluated by RT-qPCR analysis. MCF10A-Ctr Mock or EPN3 KD cells, as described in A, were stimulated or not with TGFβ (5 n/ml) at the indicated time-points. Data were normalized on 3 housekeeping genes (18S, GAPDH, ACTB) and expressed as relative to the levels in the untreated MCF10A-Ctr Mock sample (mean ± SD of two independent experiments). p-values vs MCF10A-Ctr Mock treated for 24h; ***p < 0.001; ns, not significant.

3.3.6. Physiological EPN3 regulates E-Cadherin endocytosis

Our data suggest that a physiological function of EPN3 is to allow TGF β -induced EMT transcriptional activation. Since endogenous EPN3 binds to AJs, albeit to a lesser extent than ectopic EPN3 (see Section 3.2.4), we asked whether also E-Cadherin endocytosis, upon TGF β stimulation, is altered by KD of endogenous EPN3. Using the E-Cadherin *in vivo* internalization assay (see Section 3.2.6), we showed that the KD of endogenous EPN3 reduced E-Cadherin endocytosis after TGF β stimulation for 180 minutes, compared with controls (Figure 29). These data provide further support to the notion that the effects of EPN3 overexpression might be an exaggeration of its physiological function.

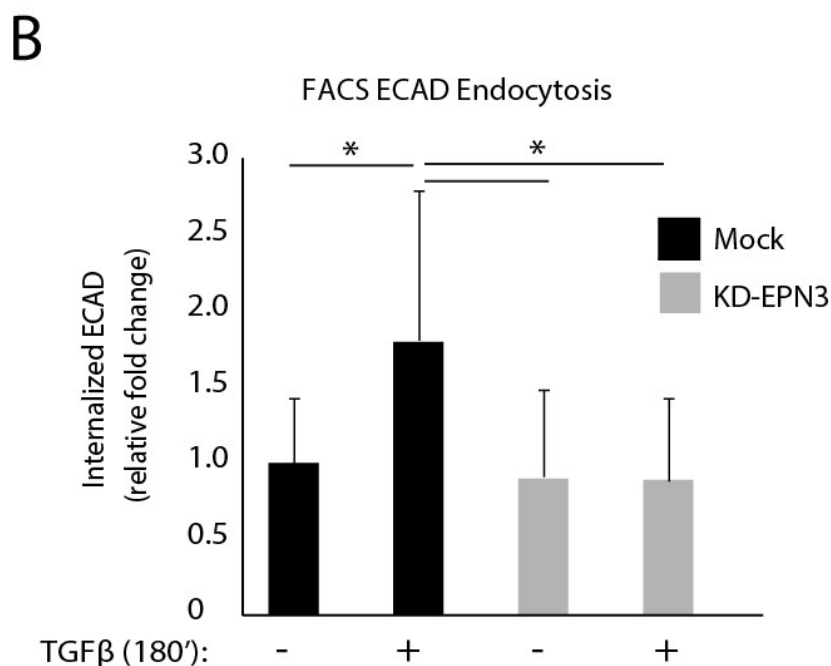
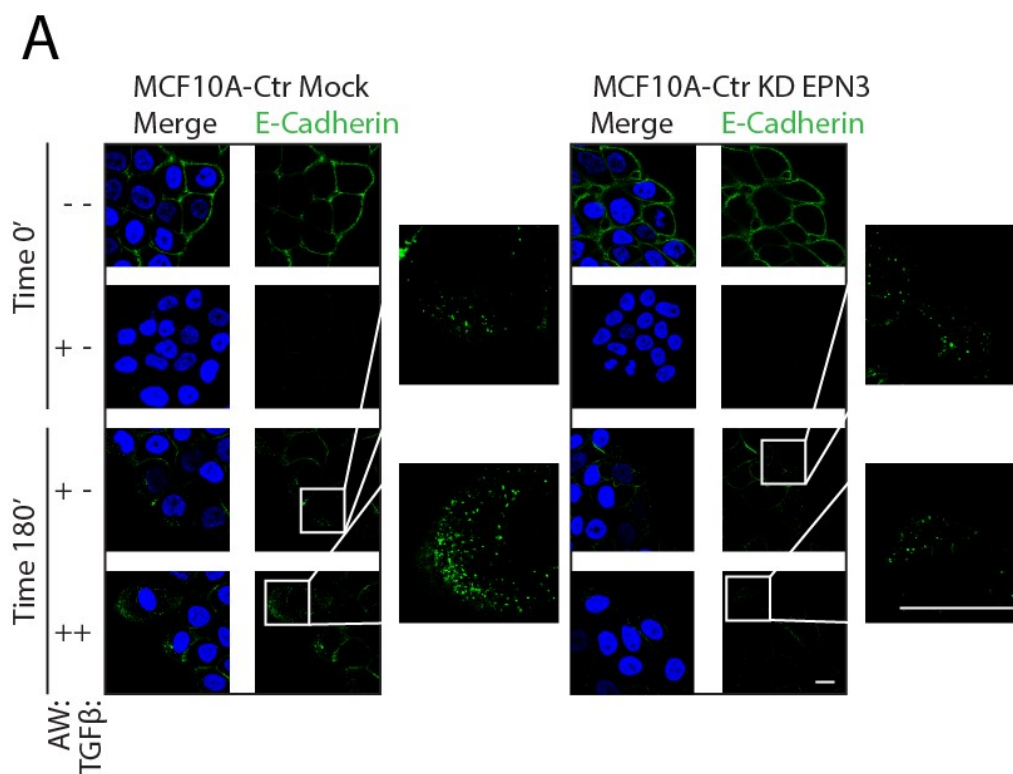


Figure 29: Endogenous EPN3 mediates E-Cadherin endocytosis induced by TGFβ stimulation

A) MCF10A-Ctr mock transfected or EPN3 siRNA transfected cells were treated of not with TGFβ (5 ng/ml) for 180 minutes and E-Cadherin endocytosis was measured by the in vivo internalization assay. Confocal images of the indicated cells stained for E-Cadherin (green) are shown. Nuclei were stained with DAPI (blue). Bar scale, 30 μm. B) Quantification of E-Cadherin endocytosis in cells treated as described in (A), measured through FACS analysis. On the y-axis, data are reported as fold-change relative to untreated MCF10A-Ctr Mock cells of four independent experiment. P-value was calculated with Student's t test, two-tailed. p-values: *p < 0.05.

3.4. Characterization of EPN3 function in mouse mammary gland biology

EPN3 function is poorly understood in humans and in others species. The finding of an EPN3 role in breast cancer and the characterization of the molecular mechanisms governing EPN3-mediated acquisition of pEMT in MCF10A cells, suggested the necessity to unravel the *in vivo* role of this protein in the context of mammary gland physiology and tumorigenesis. *Mus musculus* mammary gland development recapitulates that of *Homo sapiens*: it is characterized by a pubertal phase, dependent on the presence of different hormones and growth factors, and by an adult phase that – in virginity – shuffles between different stages of branching according to the mouse estrus cycle²²⁴. These similarities allows the use of mouse models as powerful tools to study and characterize the roles of different factors in the breast biology and tumorigenesis.

In our studies, we aimed to: i) mimic the breast tumorigenic context and ii) study the role of physiological EPN3 in the mouse mammary gland.

i) To mimic the tumoral context, we generated FVB/EPN3 KI mice that overexpress the human transgene under a strong ubiquitous promoter (UbiC) inserted in the ROSA 26 locus; a stop codon flanked by Lox-P sites is positioned between the promoter and the transgene and allows the overexpression of EPN3 only in presence of Cre recombinase. These mice were crossed with mice bearing Cre recombinase under the Keratin-5 promoter (active only in basal epithelial cells, K5-Cre)²⁶⁴, in order to allow EPN3 overexpression specifically in the epithelial compartment (Figure 30A).

ii) In order to unravel the physiological function of EPN3, we took advantage of FVB/EPN3 KO mice generated by De Camilli's lab¹¹² that ubiquitously lack the EPN3 gene (Figure 30B).

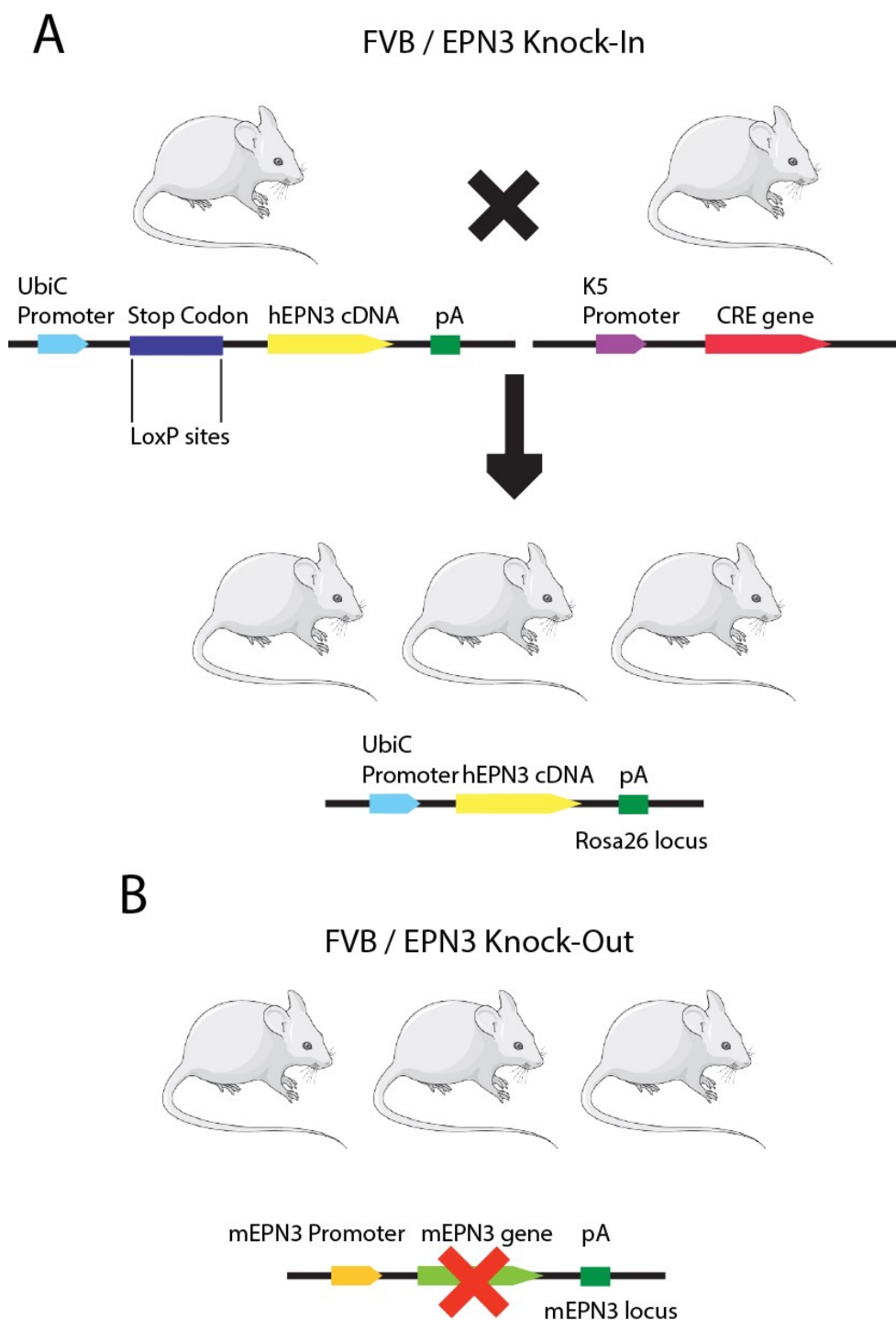


Figure 30: Representation of the strategy used to obtain FVB/EPN3 Knock-In (KI) and Knock-Out (KO) mice

A) FVB/EPN3 KI mice were obtained by crossing mice bearing the CRE recombinase gene under Keratin-5 promoter with mice carrying the human EPN3 cDNA under a strong promoter in the Rosa26 locus, but with a stop codon in between and flanked by 2 loxP sites. When these mice are crossed, the Cre enzyme cuts the loxP sites and removes the stop codon, allowing the overexpression of EPN3 in the epithelial compartment (Keratin-5 positive). B) EPN3 KO mice completely lack all the 9 coding exons of the murine gene.

3.4.1. EPN3 expression in the mammary gland of knock-in mice and knock-out mice

In order to check the level of EPN3 expression in our mouse models, we digested the mammary glands of EPN3 WT, KI and KO mice in post-puberty and purified the epithelial cells to perform WB analysis. We observed that amount of level of EPN3 overexpression in epithelial cells derived from EPN3 KI mice was comparable to that observed in BT474 cells (i.e., human breast cancer cells carrying *EPN3* gene amplification), suggesting a level of overexpression that mimics the tumoral context (Figure 31A). In WB analysis, we were not able to detect endogenous EPN3 protein expression in EPN3 WT mice, confirming that it is expressed at very low levels in the mouse mammary gland, as showed also by De Camilli's lab ¹¹² (Figure 31A). For this reason, we analyzed also the mRNA expression and ensured that EPN3 is expressed in EPN3 WT mice – although at low levels – in the normal mammary gland at this stage of development, while it was not detectable in EPN3 KO mice (Figure 31B).

Moreover, we performed IHC analysis on FFPE samples of the mammary gland of EPN3 KO, WT and KI mice to gain information about EPN3 localization. The IHC analysis showed that EPN3 overexpression in the mammary gland of KI mice occurs specifically in the epithelial compartment, but not in the surrounding cells (Figure 31C). Moreover, the IHC analysis of FFPE samples of mammary gland of EPN3 WT mice confirmed a low but specific expression of endogenous EPN3 in the epithelial cells of the mammary ducts, which disappeared in EPN3 KO mice (Figure 31C).

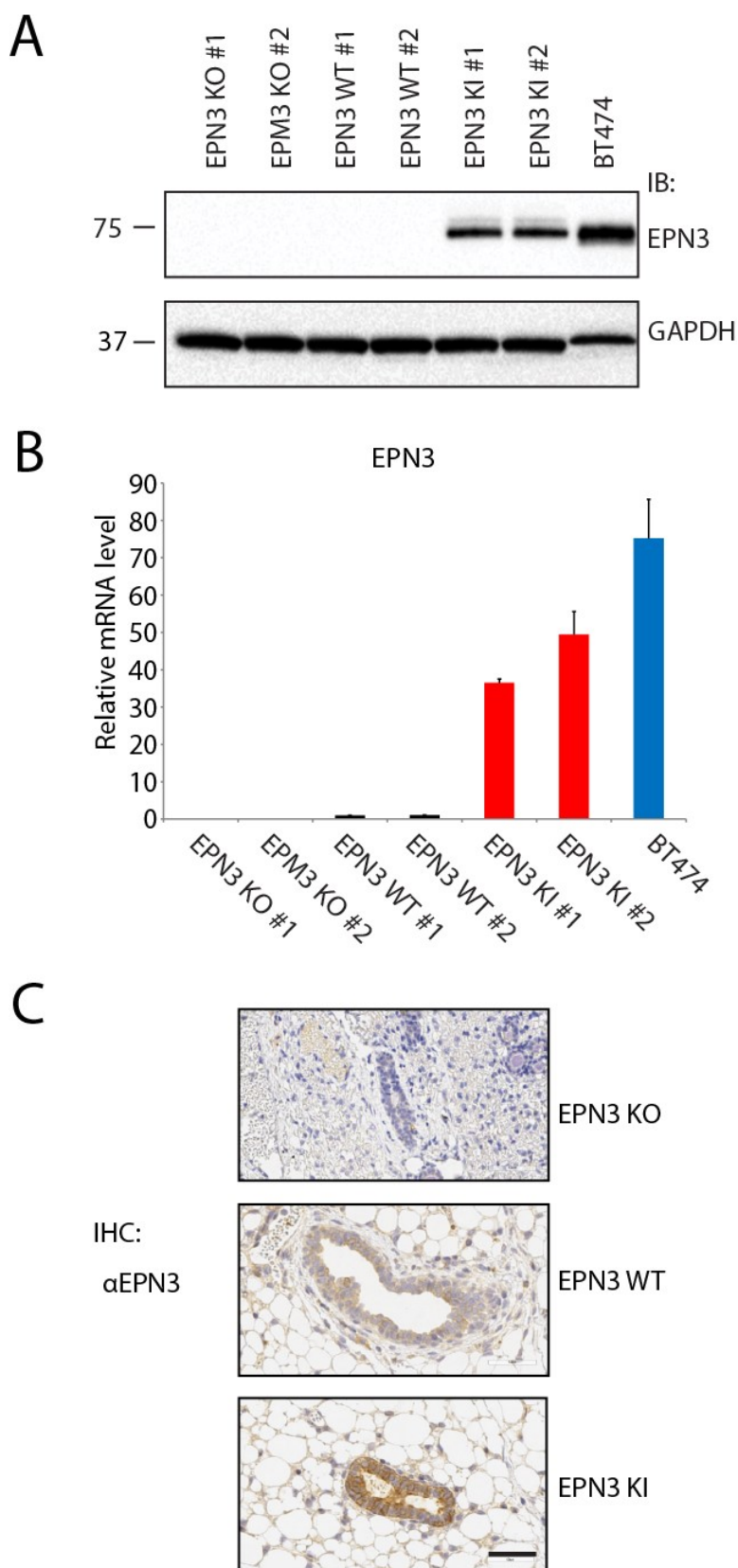


Figure 31: Evaluation of Epn3 expression and localization in mammary gland of EPN3 KO, WT and KI mice

(continued on next page)

A) WB analysis of EPN3 expression in mouse mammary epithelial cell lysates of 2 different EPN3 KO, WT and KI mice. BT474 is a human breast cancer cell line bearing EPN3 gene amplification and was used as a positive control. GAPDH was used as a loading control. B) Relative EPN3 mRNA expression levels in mouse mammary epithelial cell lysates of KO, WT and KI mice determined by RT-qPCR. Data were normalized on the 18S housekeeping gene and expressed as relative to levels in the WT #1 mouse. C) Immunohistochemistry (IHC) analysis of EPN3 expression levels in the mammary gland of EPN3 KO, WT and KI mice. Images show localization in epithelial ducts of EPN3 WT and KI mice. Bar, 50 μ m. Pictures are representative of multiple experiments from at least six different mice.

3.4.2. EPN3 has a role in mouse mammary duct branching

The role of EPN3 in the mammary gland is unknown. For this reason, we decided to analyze mammary glands of EPN3 KI and KO mice at different stages of development in virginity, from puberty to adult phase. The mouse mammary gland at birth is composed of a bud at the beginning of the fat pad²¹⁸. At puberty, about 3-4 weeks of age, the massive release of hormones and growth factors induces proliferation and invasion of epithelial ducts into the surrounding fat pad. The puberty stage ends at about 10-12 weeks of age, when the fat pad has been infiltrated by epithelial ducts. After this stage, the mammary gland of virgin mice shuffles between stages with low and high branching according to the estrus cycle phase.

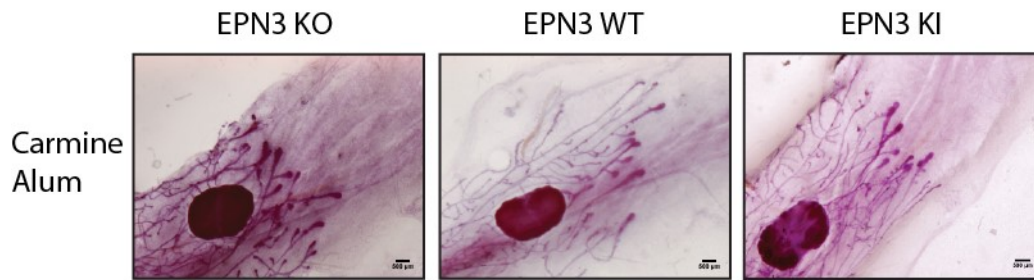
To analyze the mammary gland morphology, we performed whole mount staining, a technique that allows the visualization of the localization, morphology and amount of epithelial ducts that invade into the fat pad²⁶⁵. This analysis showed that:

1) **EPN3 KO mice display a delay in epithelial duct invasion of the fat pad at puberty stage.**

At puberty (6 weeks of age), WT mammary gland is characterized by the presence of the TEBs, bulb-shaped structures, precursors of the ducts, that direct the growth of the ducts throughout the surrounding fat pad, in order to then give rise to the fully-developed

mammary gland. EPN3 KI mice did not show any phenotypic alteration of the mammary gland at this stage, while EPN3 KO mice displayed a delay in the spreading of TEBs into the fat pad (Figure 32A). In order to perform a quantitative analysis, we decided to use as reference the lymph node, located approximately in the middle of the mouse mammary gland. We calculated the distance of the longest TEB from the lymph node and confirmed that EPN3 KO mice presented a statistically significant delay of the growth of epithelial ducts into the fat pad (Figure 32B).

A



B

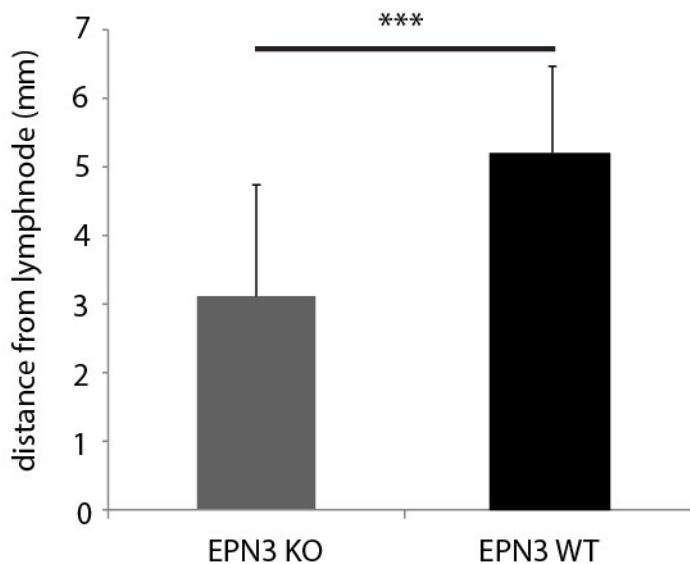


Figure 32: EPN3 KO mice show a delay in epithelial duct spreading into the fat pad at 6 weeks of age

A) Whole mount analysis through Carmine Alum staining of the mammary gland of 6-week old EPN3 KO, WT and KI mice. Bar scale, 500 μ m. At least six mice were evaluated for each sample. B) Quantification of the distance in millimeters of the furthest epithelial duct from the lymph node of the KO and WT mice of panel A).

- 2) **EPN3 KI mice show a sporadic branching alteration at the end of puberty.** At 12 weeks of age, at the end of the puberty stage, epithelial ducts have infiltrated the fat pad of the gland, and the development of the gland is complete. WT mice show the typical mammary gland morphology, with few primary epithelial branching and some sporadic

secondary branching. Compared to WT mice, a subset of EPN3 KI mice showed an increased secondary and tertiary branching (Figure 33). Interestingly, the epithelial ducts of EPN3 KO mice have rescued their delay and completely filled the fat pad, showing no phenotypic alteration (Figure 33).

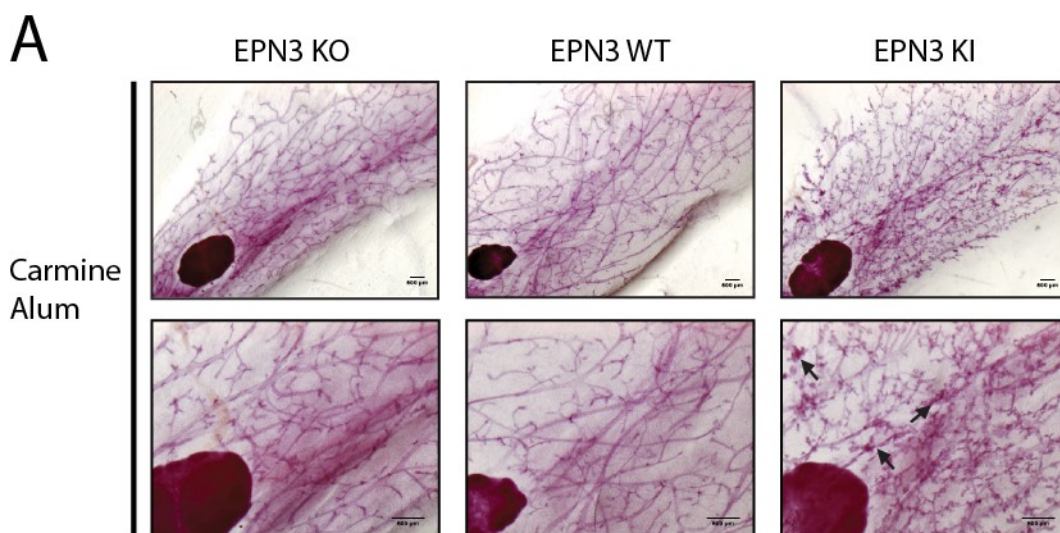


Figure 33: EPN3 KI mice show a sporadic alteration of secondary branching at 12 weeks of age
Whole mount analysis through Carmine Alum staining of the mammary glands of 12-week old EPN3 KO, WT and KI mice. Black arrows indicate increased secondary and tertiary branching; Bar scale, 500 μ m. At least seven mice were evaluated for each sample.

- 3) **EPN3 KI mice display a strong increase in duct branching post-puberty.** At 18 weeks of age, the mammary gland of a virgin mouse is completely mature and is responsive to hormonal changes, shuffling between low and high branching phases according to the estrus cycle. In WT mice, the majority of the mammary glands displayed primary branching, as expected. Very interestingly, EPN3 KI mice confirmed the phenotype sporadically observed at 12 weeks of age, showing an increase of secondary and tertiary branching compared to WT mice (Figure 34A). In contrast, EPN3 KO mice showed a

sporadic alteration of primary branching, unveiling an increased number of nascent ducts that do not appear to develop into mature epithelial ducts. Moreover, in some cases, we observed an enlargement of the diameter of the primary ducts, called duct ectasia (Figure 34A). By employing a specific ImageJ plug-in that evaluates the complexity of a grid, we evaluated the number of duct intersections (see Material and Methods). Interestingly, we observed that EPN3 KI mice show a statistically significant enrichment of epithelial connections, a proxy to measure the number of branching ducts (Figure 34B). Instead, the EPN3 KO mice did not reveal any statistically significant difference compared to WT mice.

These results showed that EPN3 overexpression is able to alter branching turnover, increasing the number of epithelial ducts and their connections in the mammary gland. In contrast, the deprivation of endogenous EPN3 causes a delay in fat pad invasion at early stage of development, while it displays a mild decrease in branching post-puberty. However, together with the increased branching observed upon EPN3 overexpression, these data suggest that the EPN3 has an impact on branching morphogenesis.

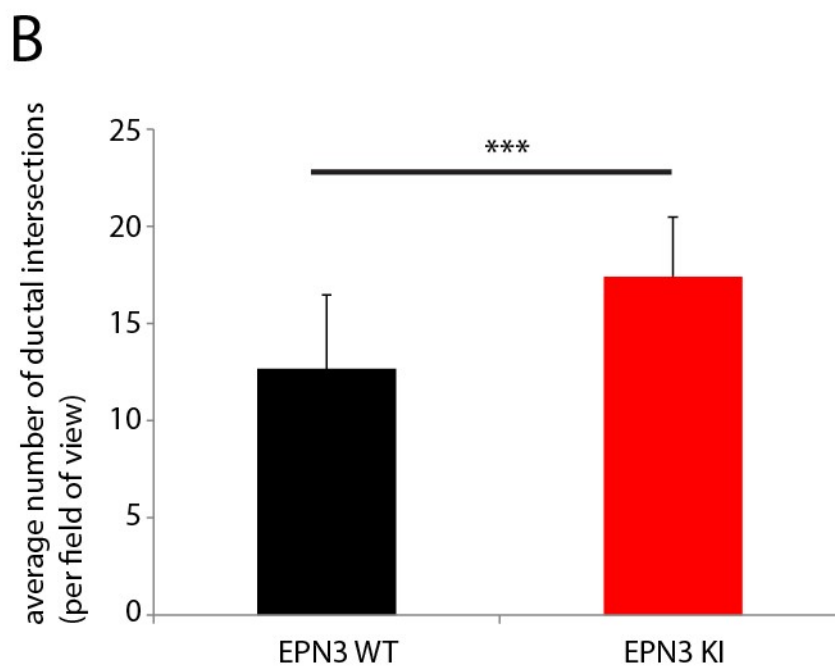
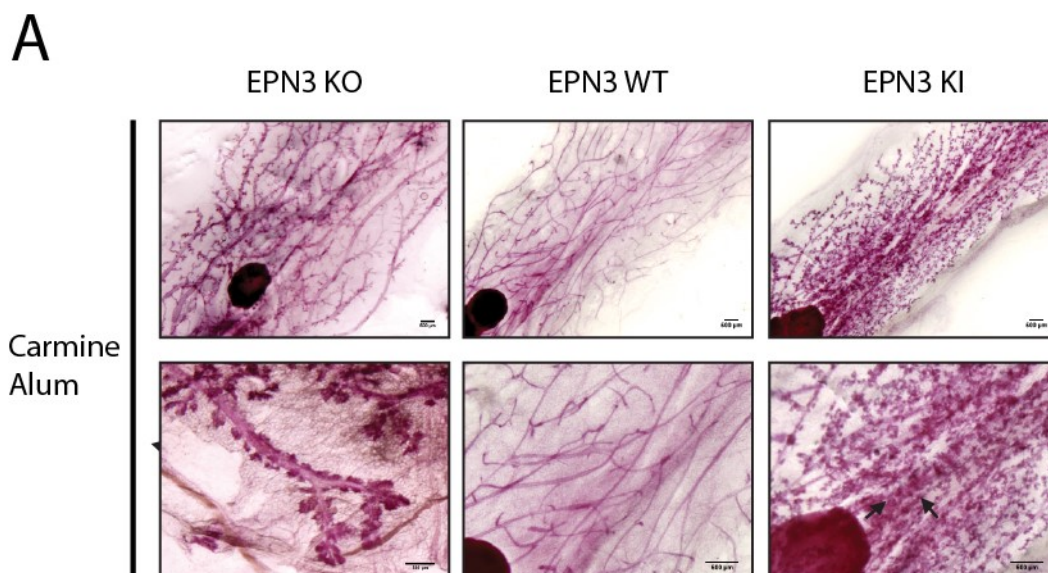


Figure 34: EPN3 KI mice display an increased secondary duct branching at 18 weeks of age

A) Whole mount analysis through Carmine Alum staining of the mammary gland of 18 week-old EPN3 KO, WT and KI mice. Black arrows indicate increased secondary and tertiary branching; blue arrow indicates alteration of primary branching; black head arrow indicates duct ectasia. Bar, 500 μ m. At least nine mice were evaluated for each sample. B) Quantification of the average number of ductal intersections per field of view in EPN3 WT and KI mice. Analyses performed with the help of Dr. Freddi Stefano (IFOM-IEO Campus).

3.4.3. EPN3 KI mice show an increased N-Cadherin and Ki67 expression in the branching ducts

Branching morphogenesis in the mouse mammary gland is regulated by a plethora of different pathways and factors, among them EMT is known to play an important role in this process²¹⁹. For this reason, we tested by IHC the staining of the EMT markers, E-Cadherin, N-Cadherin and Vimentin at 18 weeks of age. While E-Cadherin staining did not show alterations in its expression and localization (Figure 35A), EPN3 KI mice displayed an increased level of N-Cadherin positive cells in the epithelial ducts compared to EPN3 WT and KO mice (Figure 35B). These data suggest that the alterations in mammary gland morphogenesis observed in EPN3 KI mice could be due to activation of an EMT program. Unfortunately, we were not able to obtain a good staining for Vimentin with these conditions of IHC (data not shown).

The activation of EMT in mammary duct branching is necessary to allow the morphological changes required to migrate into the surrounding fat pad. It is also known that proliferation has an important role in the correct formation of the epithelial ducts²¹⁹; for this reason, we asked whether in EPN3 KI mice displayed also different levels of proliferation in epithelial cells. We used Ki67 as a proliferation marker and observed that EPN3 KI mice have an increased number of Ki67-positive cells compared to WT and KO mice (Figure 35C).

In conclusion, these data suggest that EPN3 overexpression could act on branching regulation, possibly altering both the migratory and proliferative ability of epithelial cells.

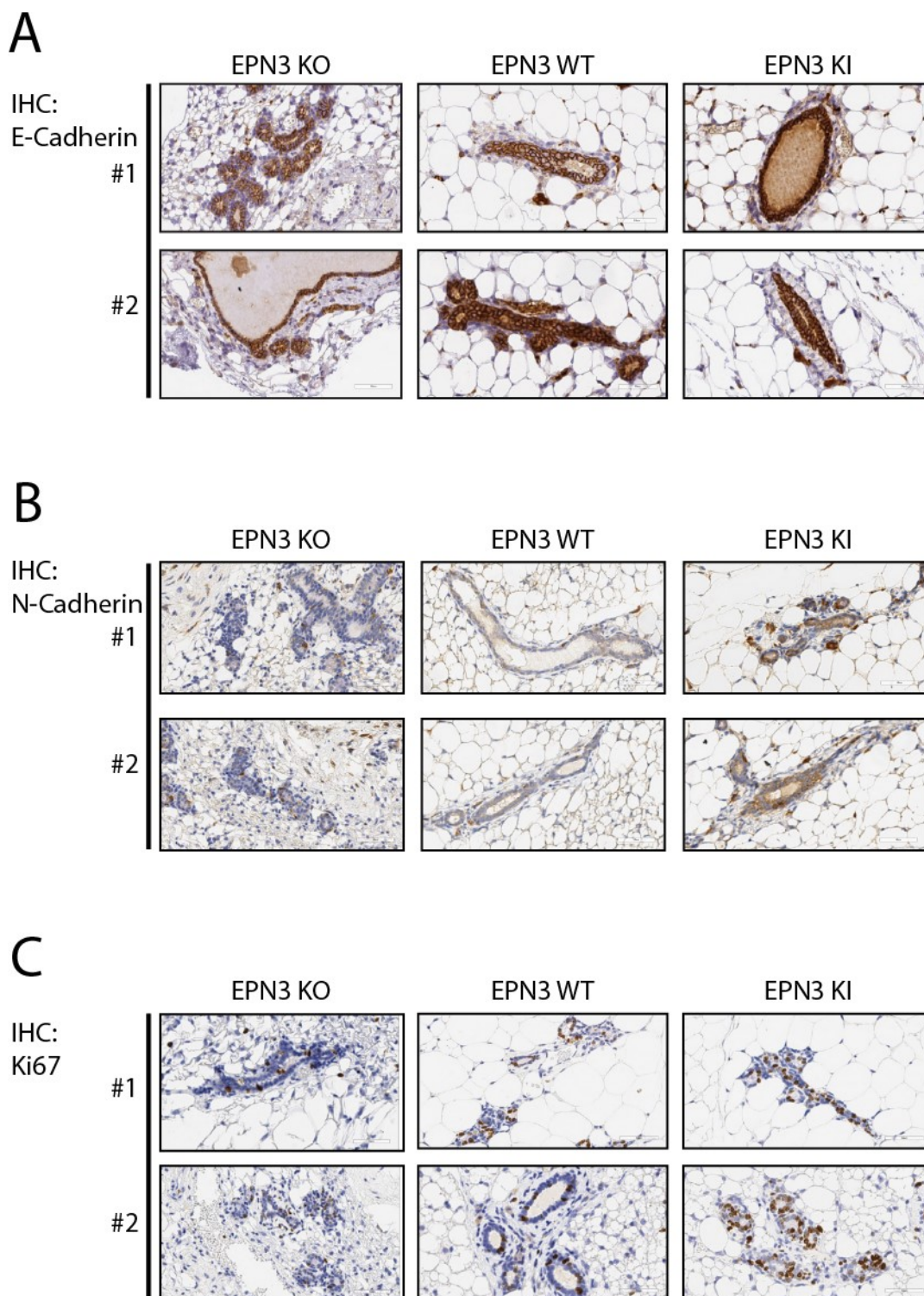


Figure 35: EPN3 KI mice show an increased number of positive cells for N-Cadherin and Ki67 at 18 weeks of age

A) Immunohistochemistry (IHC) analysis of E-Cadherin expression in the epithelial ducts of EPN3 KO, WT and KI mice. Bar, 50 μ m. At least two mice were evaluated for each sample. B) IHC analysis of N-Cadherin expression in the epithelial ducts of EPN3 KO, WT and KI mice. Bar, 50 μ m. At least two mice were evaluated for each sample. C) IHC analysis of Ki67 expression in the epithelial ducts of EPN3 KO, WT and KI mice. Bar, 50 μ m. At least two mice were evaluated for each sample.

4. DISCUSSION

The studies conducted during this PhD project have unveiled the role of a novel prognostic factor in breast cancer, EPN3. The findings obtained suggest a possible mechanism of action of EPN3 in breast cancer invasion and metastasis, based on EMT acquisition through an endocytosis-dependent circuitry. Moreover, an initial characterization of EPN3 function in mouse mammary gland biology suggests a role of EPN3 in ductal branching morphogenesis.

4.1. EPN3 overexpression induces a partial EMT phenotype in MCF10A cells

EMT is a multistep process that requires the concomitant loss of epithelial characteristics and the acquisition of mesenchymal ones ^{184,239}. In the last decade, increasing evidence supports the notion that the acquisition of a full, irreversible, EMT might not be compatible with the ability of cancer cells to metastasize to distant sites ^{184,188,189}. This is because EMT is only part of the process required to allow cells to migrate to and to invade the surrounding tissues: MET is also important once cancer cells have migrated to and reached other sites, to allow the colonization of these sites. Plasticity is therefore one of the most important features that cancer cells use to modulate their behavior in order to adapt to different microenvironments.

Starting from this consideration, many groups have shown the existence of different intermediate EMT phenotypes that are encompassed in the term of pEMT ^{184,188,189,194,246,247,266}. pEMT is characterized by the simultaneous presence of epithelial and mesenchymal markers and by high plasticity. In agreement with this, our results indicate that the overexpression of EPN3 in

the non-transformed epithelial cell line, MCF10A, is characterized by the upregulation of mesenchymal markers, such as N-Cadherin, Vimentin, TGF β ligands and receptors, and the transcription factors, Snail1 and Zeb1, concomitant with the expression of the epithelial marker E-Cadherin, which is only partially reduced. These observations suggest that EPN3 is not able by itself to activate a full EMT program, and instead keeps cells in a pEMT state. In accordance with this, Zhang *et al.*¹³⁰ showed that MCF10A is a model cell system able to acquire pEMT states. They demonstrated that a full EMT, induced by chronic stimulation with high doses of TGF β , is a two-step program: the first step is characterized by Snail upregulation and by the presence of both epithelial and mesenchymal properties, and it is reversible upon TGF β deprivation. The second step is Zeb1-dependent and gives rise to a complete mesenchymal transition that is not reversible upon TGF β deprivation and largely irreversible after treatment with TGF β inhibitors.

In our case, the EPN3-induced pEMT upregulates both Snail and Zeb1, but maintains consistent levels of E-Cadherin; moreover, TGF β R1 KD is able to cause an almost complete reversion of this phenotype. Importantly, MCF10A-EPN3 cells are still sensitive to TGF β stimulation, as treatment with TGF β , already at a limiting dose, is able to push EPN3-overexpressing cells to a more advanced EMT state, with a higher level of upregulation of all mesenchymal markers, including Snail and Zeb1. These data suggest that the upregulation of Snail and Zeb1 upon EPN3 overexpression is quantitatively not sufficient to reach a full EMT state, as described by Zhang *et al.*¹³⁰.

Different EMT intermediate states have been described^{189,194,246,266}. Importantly, many tumors are composed of a mix of epithelial and mesenchymal markers¹⁸⁹, allowing the speculation that these intermediate, plastic phenotypes lead to more aggressive, invasive cancer

cells, metastasis and drug resistance ²⁶⁷⁻²⁶⁹. Indeed, cells that are in the pEMT state have the highest plasticity and are able to rapidly undergo full EMT, allowing them invade the surrounding tissues, reach distant organs and then to perform the reverse transition, MET, to colonize the new sites and to proliferate ¹⁷¹. In agreement, EPN3 overexpression positively correlates with an aggressive cancer phenotype and increased risk of distant metastasis in breast cancer patients.

4.2. EPN3 overexpression induces a TGF β -dependent self-sustaining positive-feedback loop that becomes independent of the EPN3 initial alteration

The activation of transcription factors that induce a transcriptional feedback loop is very well characterized, especially in EMT through the TGF β pathway ^{130,170,184}. It is also known that upstream events, such as endocytosis, have profound impacts on signaling and on the regulation of transcriptional programs ^{1,270}. Our data show that EPN3 overexpression represents one of these upstream events, since it induces the upregulation of TGF β ligands and receptors that, in turn, create a positive-feedback loop that auto-sustains itself. The TGF β loop induced by EPN3 is not in saturating conditions, since MCF10A-EPN3 cells are more prone to TGF β stimulation both in the activation of the downstream signaling pathways and in the induction of E-Cadherin internalization.

Interestingly, we have preliminary evidence that EPN3-induced pEMT acquires novel functions and becomes independent of EPN3 expression itself. Thus, our results show that the original genetic lesion that started the pEMT process, once cells have been instructed through the generation of self-sustaining loops, becomes no longer necessary. It is also reasonable to speculate

that epigenetic and/or micro-environmental changes within the tumor mass might cause profound rearrangements of the biological identity of cancer cell that cannot be reverted by targeting the primary genetic lesion. Indeed, if the hypothesis that EPN3 induces a mesenchymal switch and, then, it is no longer necessary, is correct, the obvious consequence would be that the choice to design drugs against this molecule could be useless for tumor treatment, or maybe even detrimental. Indeed, inefficacy of drug treatments and the development of drug resistance is one of the main causes of poor prognosis and death in cancer; only 9.6% of newly synthesized drugs are able to complete clinical trials and enter into the drug market ²⁷¹. These preliminary considerations, if confirmed and extended to other cases, could help inform decision-making on the kind of targeted therapy to invest in.

4.3. EPN3 interacts with the adherens junctions and the clathrin machinery

The possibility that EMT/pEMT can be regulated through mechanisms that fall outside the transcriptional program has been widely investigated in the last decade and is now well accepted by the scientific community ^{184,188,189}. Given that EMT has to start from the downregulation of epithelial markers, a lot of effort has been spent to try to unveil all the possible mechanisms of E-Cadherin regulation. Many groups have supposed that regulation of E-Cadherin endocytosis and trafficking occur during tumor progression in order to allow pEMT ^{197,246,247}: pancreatic, breast, skin, liver and colorectal carcinomas have been linked to the existence of many intermediate EMT states that occur during tumor invasiveness ^{197,246,247}. For instance, in pancreatic, breast and colorectal carcinomas, the existence of pEMT states has been causally linked to the re-localization

of epithelial markers, such as E-Cadherin, and not to their direct transcriptional downregulation. Although E-Cadherin trafficking was not directly measured, these data suggested for the first time that, *in vivo*, alterations of epithelial marker trafficking and turnover could explain a more aggressive cancer phenotype^{197,247}. Our results indicate that EPN3 overexpression induces a pEMT state that is associated with enhanced E-Cadherin endocytosis, and correlates with increased risk of distant metastasis.

The Epsin family was discovered around 20 years ago and immediately recognized as an endocytic protein family⁶⁷. The majority of the studies on this family have focused on EPN1, and showed that this protein is indeed working as an endocytic adaptor for different cargoes⁷². Interestingly, epsins (in particular EPN1) have been linked with cell migration: EPN1 is also able to regulate migration through the binding with RalBP1⁹⁹, moreover it is able to recruit actin at the nascent pit through Hip1r¹⁰⁰. On the other side, EPN1 regulates the Notch pathway through the endocytosis of its ligand, Dll4⁸³, that is structurally similar to the activin type II receptors (ActRIIs), belonging to the superfamily of TGF β receptor superfamily. Interestingly, it was shown that ActRIIs internalization and signaling are regulated by activin receptor-interacting protein 2 (ARIP2) and RalBP1²⁷², putatively connecting EPN1 with ActRIIs and, more in general, with the TGF β receptor superfamily.

Concerning EPN3, this family member has been shown to co-localize with vesicles positive for clathrin and the endocytic factor, EHD2^{112,113}. Moreover, it has been suggested that EPN3 has a role in cancer cell migration and EMT since it is strongly upregulated at the leading edge of migrating MDCK cells¹¹⁹ and its overexpression in glioblastoma cells induces the appearance of mesenchymal markers¹²⁰. Apart from these specific cases, EPN3 is very rarely

expressed in normal tissues, suggesting that it could have a specialized function, in contrast to the other 2 members of the Epsin family (EPN1 and Epn2), which are widely expressed in adult tissues¹¹².

Due to the lack of knowledge on EPN3, we had no *a priori* idea of its potential function in breast cancer. To unveil the mechanism of EPN3 in pEMT induction in MCF10A cells, we applied an unbiased approach, defining through mass spectrometry analysis the “EPN3 interactome” in MCF10A-EPN3 cells. This analysis revealed a strong enrichment in the EPN3 interactome of proteins belonging to the endocytic machinery (clathrin heavy and light chain, AP2 complex, Cbl.). Moreover, the second most highly enriched category in the EPN3 interactome included proteins belonging to AJs, TJs, focal adhesions and the actin cytoskeleton. By combining this result with the knowledge regarding cell adhesion regulation and EMT, we focused our attention on AJs, since they are the most important regulator of the transition. Moreover, given the already described link between EPN1, actin cytoskeleton and TGF β receptor superfamily, we decided to exclude from further analyses the known interactors of EPN1 (linked to the actin cytoskeleton and cell migration), since in MCF10A cells we didn't observe any difference upon EPN1 overexpression, suggesting that the EPN3 mechanism could be different from the characterized EPN1 functions.

Importantly, we validated the interaction of EPN3 with AJ components, and we also showed a weak, probably indirect, but reproducible interaction, with E-Cadherin. Interestingly, in 2014, an E-Cadherin interactome identified through quantitative proteomics, revealed a putative weak interaction with EPN3 in MKN28 cells (human gastric adenocarcinoma), which, however, was not subsequently studied and characterized²⁷³. In agreement with this study, we showed for

the first time that EPN3 possibly connects the clathrin machinery to the E-Cadherin complex, thereby regulating E-Cadherin trafficking. Indeed, we revealed that MCF10A-EPN3 cells have less E-Cadherin at the PM and display increased E-Cadherin internalization, indicating that the overexpression of EPN3 could enhanced E-Cadherin turnover at the PM. Moreover, TGF β stimulation sustains E-Cadherin endocytosis in MCF10A-EPN3 cells, consistent with our results showing cooperation between EPN3 and the TGF β pathway, and the creation of a positive-feedback loop.

Our data are in agreement with the scenario that EPN3 causes pEMT through increased E-Cadherin endocytosis. One critical piece of evidence that corroborates this possibility is the fact that we scored more active β -catenin in the nucleus, suggesting that a consequence of enhanced E-Cadherin PM turnover via endocytosis is the destabilization AJs complex and the release of β -catenin, which relocates to the nucleus where it activates the transcription of EMT genes. In agreement, the ablation of the major partner of β -catenin in the nucleus, the transcription factor, TCF4, is able to revert the pEMT induced by EPN3.

Another important aspect concerns the fact that EPN1 overexpression only partially mimics the effects of EPN3 overexpression, despite sharing a very similar protein structure (amino acid sequence 47% identical) and a highly conserved predicted tertiary structure. The reason for the specificity of EPN3 in pEMT induction is still unclear; it has been speculated that the ENTH domain of epsins is very similar to the armadillo repeats of the catenin family ⁷⁰, a sequence of 42 amino acids that allows the interaction with cadherins ^{75,145,150}. Thus, one possibility could be that EPN3, but not EPN1, binds the E-Cadherin via its ENTH. However, the ENTH domains of EPN1 and EPN3 are almost identical [80-82% sequence identity], ¹¹³, thus,

making this possibility unlikely, although still to be experimentally verified. We note that the central region is the most divergent between EPN1 and EPN3 (Figure 36), even though it does not contain a known domain or motif specific for EPN3 and it is largely unstructured. Further structure-functional studies are required to resolve this issue.

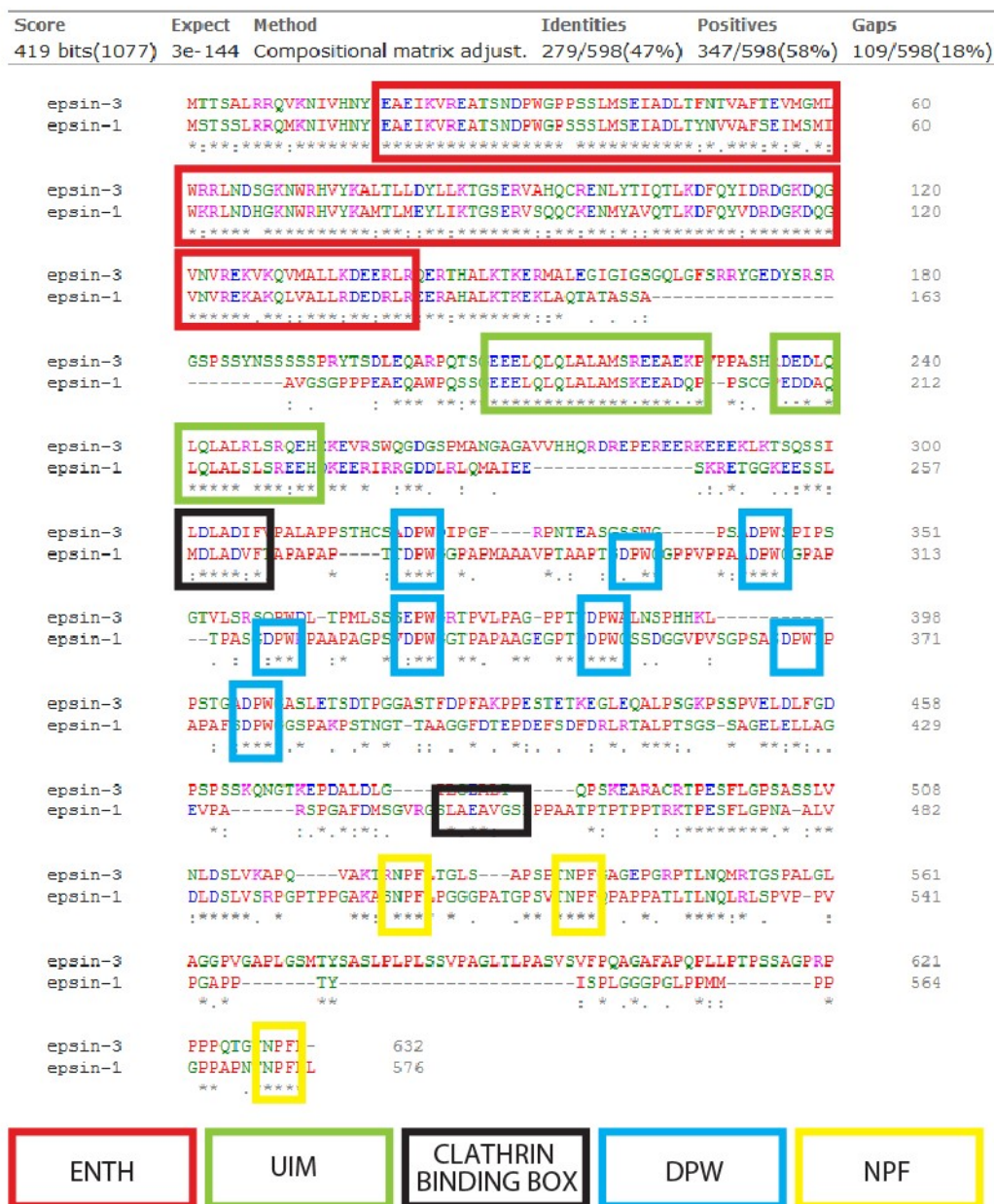


Figure 36: Alignment of EPN3 and EPN1 protein sequences

Alignment of the EPN3 and EPN1 protein sequences performed using the Clustal Omega tool available on EMBL-EBI (<https://www.ebi.ac.uk/Tools/msa/clustalo/>), together with the score attributed to the alignment and the probability of having the same alignment by chance (Expect) and the method used to align the two sequences through the

blast2seq tool available on NCBI (<https://blast.ncbi.nlm.nih.gov/>). The number and the relative percentage of identical (Identities) and similar (Positives) residues, as well as gaps inserted in the alignment, also are shown. Residues with similar chemical and physical properties are represented by the same colour (Red: hydrophobic; Blu: positively charged; Green: polar; Purple: negatively charged). Legend of colored boxes used in the alignment is shown to highlight specific domains and motifs of epsin proteins.

4.4. EPN3 physiological function in the regulation of E-Cadherin endocytosis

Having characterized EPN3 function in an overexpressing system, we concentrated our attention on the physiological role of EPN3. Interestingly, we showed that also endogenous EPN3 is able to bind AJs components, albeit to a lesser extent than overexpressed EPN3, suggesting that the overexpression of EPN3 is enhancing its physiological function. In agreement with this possibility, E-Cadherin endocytosis induced upon TGF β stimulation is impaired by EPN3 ablation. Concomitantly, also the TGF β -induced transcriptional response is reduced in EPN3-KD cells. Although we cannot exclude that EPN3 might have also other more direct functions in the activation of the TGF β pathway, our results converge on the possibility that EPN3 has a physiological role in TGF β -induced E-Cadherin endocytosis, which is a critical mediator of the TGF β -dependent transcriptional response that leads to the activation of EMT genes. These conclusion are in line with the experiments performed by Wang *et al.*¹²⁰, where physiological EPN3 is implicated in the maintenance of a mesenchymal phenotype in glioma cells.

In conclusion, our data show that physiological EPN3 is probably implicated in more specific functions than in a general formation of CCP.

4.5. EPN3 alters mouse mammary gland branching morphogenesis

The role of EPN3 in tumorigenesis is largely unknown. In 2016, it was proposed for the first time that EPN3 expression negatively correlates with metastasis-free survival both in breast cancer and non-small cell lung cancer¹¹⁶. According to our and these data, EPN3 worsens the phenotype, and

it is implicated in distant metastasis acquisition. In order to study the EPN3 tumorigenic role *in vivo*, we generated specific conditional EPN3 KI mice that overexpress the transgene just in the epithelial compartment. The levels of overexpression are comparable with those observed in BT474, a breast cancer cell line with *EPN3* gene amplification. However, as expected, we did not observe any neoplastic or pre-neoplastic lesions in these mice. The reason might be that EPN3 overexpression *per se* is not sufficient for tumor initiation and growth, but it can cooperate with other lesions, while it might be sufficient for the acquisition of invasive, metastatic traits. To evaluate this, we are currently crossing EPN3 KI mice with ErbB2/Neu-N mice, which overexpress the form of ErbB2 that causes low frequency and slow proliferative but aggressive breast tumors ^{274,275}, to score for possible cooperation in tumor appearance and metastasis formation.

Interestingly, the evaluation of the mammary glands of EPN3 KI mice shows that EPN3 overexpression alters branching morphogenesis, increasing secondary and tertiary epithelial ducts with higher expression of a proliferation (Ki67) and EMT (N-Cadherin) markers. This phenotype is intimately linked to the regulation of EMT, indeed, β -catenin ²²⁷, MMPs ^{228,229}, Cripto-1 ²³⁰ and TGF β ^{218–220} have been showed to regulate EMT during mouse mammary gland development. Of note, our results obtained *in vivo* and *in vitro* partially overlap, and EPN3-induced pEMT could in theory explain the phenotypes observed *in vivo*. However, we could not detect by IHC any defect in E-Cadherin localization or expression in EPN3 KI mice. Experiments with more appropriate techniques (e.g., immunofluorescence on FFPE samples) are required in order to better visualize E-Cadherin localization. Thus, although the EPN3-dependent mechanisms operating *in vivo*

during mammary gland development are far from being elucidated, our findings justify future investigations in this area.

Regarding the physiological role of EPN3 in mice, De Camilli's group described in 2010 that EPN3 is usually expressed at very low levels in most tissues, with the exception of gastric parietal cells in the stomach, where it cooperates in the control of acidification of the stomach lumen through endo/exocytosis of vesicles containing the H/K ATPase¹¹². De Camilli's group generated EPN3 KO mice, which showed no obvious phenotypic defects, probably due to compensatory effects of the other epsins¹¹². We took advantage of these mice to characterize EPN3 function in the mammary gland. Our results indicate that EPN3 ablation causes a delay of TEB infiltration into the fat pad during puberty, although these mice recover completely from the delay at the end of puberty. In adulthood, virgin mice show alterations in primary branching, with a low penetrance behavior that does not cause an impairment in the function of the gland during pregnancy and lactation, since the litters grow up healthy and normal. Moreover, since EPN3 KO mice ubiquitously lack the gene, we cannot exclude that the observed phenotypes in mouse mammary gland could be induced by other effects of EPN3 ablation in distinct organs or glands.

4.6. Future perspectives

i) Enhanced E-Cadherin endocytosis is the first event in the EPN3-induced pEMT

We aim to demonstrate that E-Cadherin endocytosis is the initial event through which EPN3 overexpression induces pEMT in MCF10A. To address this point, we will take advantage of an

inducible EPN3 model system. We will infect MCF10A cells with a doxycycline-inducible EPN3 expression vector and with an empty vector as control. These cells should remain epithelial since EPN3 overexpression is not induced. Upon doxycycline treatment, EPN3 overexpression should occur, causing induction of pEMT. With this system, we will investigate the timing and the steps involved in pEMT induction. Moreover, we will directly test whether increased E-Cadherin endocytosis is upstream of the activation of EMT by blocking E-Cadherin endocytosis concomitantly with EPN3 overexpression and checking whether this affects pEMT establishment.

In order to block E-Cadherin endocytosis, we will take advantage of two different experimental strategies:

- a. we will use an anti-E-Cadherin monoclonal antibody, generated by Petrova *et al.*²⁷⁶, which binds specifically to the extracellular domain of E-Cadherin, and was shown to maintain E-Cadherin at the PM, favoring AJ stabilization and clustering of the cells.
- b. we will overexpress p120, an inhibitor of E-Cadherin endocytosis^{277,278}. We will overexpress the full-length form or a truncated version, missing the first 346 amino acids, that should work as dominant negative by binding to E-Cadherin, but not undergoing post-translational modifications involved in p120 release from E-Cadherin that lead to its endocytosis^{277,278}.

ii) Characterization of the self-sustaining loop leading to pEMT

Our results suggest that once the TGF β -dependent transcriptional program has been established, EPN3 endocytic function is no longer required. To formally prove this hypothesis, we will take advantage of the EPN3-inducible expression system in MCF10A cells. We will inhibit E-Cadherin

endocytosis in these cells (as described in point 1) or we will perform EPN3 ablation, before and after the induction of EPN3 overexpression and the establishment of pEMT. If our hypothesis is true, once pEMT is established, we would expect no reversion of the mesenchymal state, both upon ablation of EPN3 or inhibition of E-Cadherin endocytosis, while a reversion of EMT is expected before pEMT is fully established.

iii) Molecular mechanism of EPN3-mediated E-cadherin endocytosis

We want to unravel the molecular mechanism governing EPN3 function at AJs. In order to further dissect the mechanism, we plan to perform structure-function studies to identify the region directly involved in E-Cadherin endocytosis and EMT induction. To this aim, we will exploit a chimeric approach taking advantage of the fact that EPN1 does not phenocopy EPN3 in inducing EMT and E-cadherin endocytosis. Chimeras, where the different domains and motifs of EPN1/EPN3 are swapped, will be expressed in MCF10A cells and tested for their ability to recruit AJ components, to enhance E-cadherin endocytosis and to induce EMT.

iv) EPN3 *in vivo* role in breast tumorigenesis

We want to understand the *in vivo* role of EPN3 in mouse mammary gland tumorigenesis.

- a. We will evaluate through immunofluorescence the localization and expression of EMT markers (E-Cadherin, N-Cadherin, Vimentin, beta-catenin etc.) in FFPE samples of mammary gland tissue from EPN3 WT, KO and KI mice.
- b. We will extract mammary epithelial cells from adult post-puberty EPN3 WT and KI mice and transplant them into the cleared fat pads of immunocompromised mice (NOD-SCID-IL2R-

gamma-null), according to previously described protocols ²⁷⁹, in order to evaluate the ability to EPN3-overexpressing mammary epithelial cells to give rise to a normal mammary gland or instead induce pre-neoplastic/neoplastic events.

c. We will test whether EPN3 overexpression cooperates with other oncogenes in breast tumorigenesis, by crossing EPN3 KI mice with ErbB2/Neu-N mice – bearing the overexpressed WT form of ErbB2 that causes tumor appearance with low frequency ^{274,275} – to score for possible acceleration in tumor appearance/development, and alteration in morphology. We will also check for increased metastasis in these mice.

4.7. Conclusions and proposed model

Our *in vitro* results led us to propose a model for EPN3 function in E-Cadherin endocytosis and in the activation of pEMT in breast epithelial cells (Figure 37): EPN3 physiologically is able to regulate TGF β -induced E-Cadherin endocytosis and turnover, finally, having an impact also on the TGF β -mediated transcriptional response. Upon overexpression, EPN3 increases E-Cadherin internalization (A) and leads to the activation of the β -catenin/TCF4 pathway (B), inducing the transcription of EMT target genes (Snail, Zeb1, N-Cadherin, Vimentin and TGF β ligands and receptors; C). The upregulation of the TGF β pathway creates a self-sustained positive-feedback loop (D), directly activating EMT genes (E) and further stimulating E-Cadherin endocytosis (F). Preliminary evidence suggests that the TGF β loop, once generated, becomes largely independent of EPN3 endocytic function and leads to the acquisition of a pEMT state.

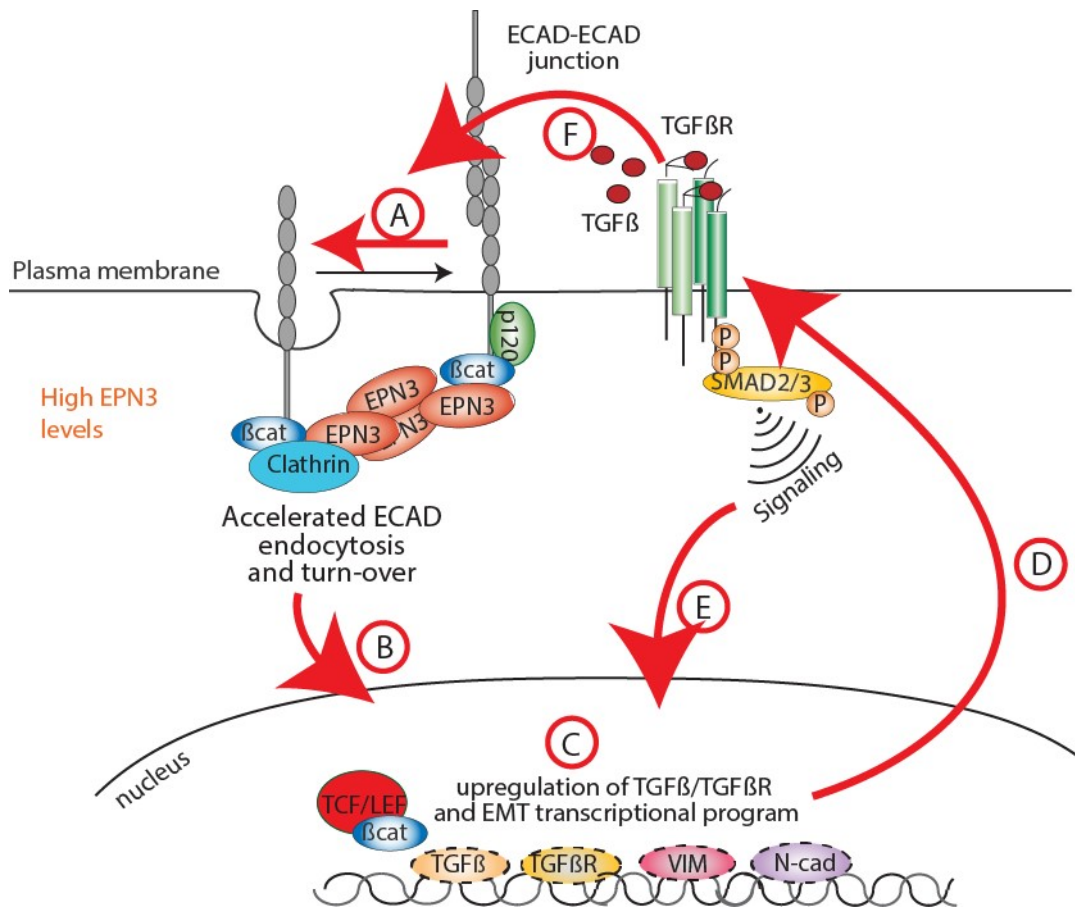


Figure 37: Working model of EPN3-dependent induction of pEMT in breast cancer
 See text for details.

5. MATERIALS AND METHODS

5.1. Solutions and buffers

Phosphate-buffered saline (PBS)

NaCl	137 mM
KCl	2.7 mM
Na ₂ HPO ₄	10 mM
KH ₂ PO ₄	2 mM

8 g of NaCl, 0.2 g of KCl, 1.44 g of Na₂HPO₄, and 0.24 g of KH₂PO₄ are solved in 800 ml of distilled H₂O (ddH₂O). The pH is adjusted to 7.4 with HCl and volume is brought to 1 liter with ddH₂O.

Tris-HCl 1M

121.1 g of Tris base are dissolved in 800 ml distilled H₂O. The pH is adjusted to 7.4, 7.6 or 8.0 (depending on the case) with HCl, and ddH₂O is added to bring volume to 1 liter.

Tris-buffered saline (TBS)

NaCl	137 mM
KCl	2.7 mM
Tris HCl pH 7.4	25 mM

8 g of NaCl, 0.2 g of KCl, and 3 g of Tris base are solved in 800 ml of distilled H₂O.

The pH is adjusted to 7.4 with HCl and ddH₂O is added to bring volume to 1 liter.

10X SDS-PAGE running buffer

Glycine	192 mM
SDS	15
Tris HCl pH 7.4	250 mM

1X JS buffer #1: for total lysis and FLAG IP

Hepes pH 7.5	0.05M
NaCl	0.15M
Glycerol	10%
TritonX 100	1%
MgCl ₂	0.0015M
EGTA	0.005M
NaPyrophosphate	0.02M
NaF	0.05M
PMSF	0.002M
Na ₃ VO ₄	0.01M
Protease Inhibitors	1:500
H ₂ O	Up to volume

200X Proteases inhibitor cocktail from Calbiochem, sodium pyrophosphate pH 7.5 20 mM, sodium fluoride 250 mM, PMSF 2 mM, and sodium orthovanadate 10 mM are added to the buffer just before use.

1X JS buffer #2: for FLAG IP with strong washes

Hepes pH 7.5	0.05M
NaCl	0.25M
Glycerol	10%
TritonX 100	1%
MgCl ₂	0.0015M
EGTA	0.005M
NaPyrophosphate	0.02M
NaF	0.05M
PMSF	0.002M
Na ₃ VO ₄	0.01M
Protease Inhibitors	1:500
H ₂ O	Up to volume

200X Proteases inhibitor cocktail from Calbiochem, sodium pyrophosphate pH 7.5 20 mM, sodium fluoride 250 mM, PMSF 2 mM, and sodium orthovanadate 10 mM are added to the buffer just before use.

1X RIPA buffer

Tris pH 7.4	50 mM
NaCl	150 mM
EDTA	1 mM
TritonX 100	1%
NaDoc	1%
SDS	0.1%
NaPyrophosphate	0.02 M
NaF	0.05 M
PMSF	0.002 M

MATERIALS AND METHODS

Na ₃ VO ₄	0.01 M
Protease Inhibitors	1:500
H ₂ O	Up to volume

200X Proteases inhibitor cocktail from Calbiochem, sodium pyrophosphate pH 7.5 20 mM, sodium fluoride 250 mM, PMSF 2 mM, and sodium orthovanadate 10 mM are added to the buffer just before use.

4X Laemmli buffer

Tris pH 6.8	200 mM
SDS	8%
Glycerol	40%
Beta Mercaptoethanol	2.4%
Bromophenol Blue	Just till it becomes blue

Diluted to 1X for protein loading into the gel.

Acid wash #1

Glycine	100 mM
ddH ₂ O	Up to volume

Once the solution was prepared, pH was controlled and adjusted to 2.2 with HCl 37%.

Acid wash #2

NaCl	500 mM
Acetic Acid Glacial	200 mM
ddH ₂ O	Up to volume

Once the solution was prepared, pH was controlled and adjusted to 2.5 with HCl 37%.

Mouse epithelial ductal staining

Carminio powder (Sigma C1022)	1 g
Aluminium Potassium Sulfate (Sigma A7167)	2.5 g
ddH ₂ O	500 ml

Mix for at least 2 h, then boil for 7' in microwave and then filter (0.45 μ m).

5.2. Cell culture**5.2.1. MCF10A cell culture**

MCF10A cells were from the American Type Culture Collection (ATCC). They were cultured in 1:1 mixture of DMEM and HAM'S-F12 (Gibco, Life Technologies), supplemented with 5% horse serum (Invitrogen), 2 mM Glutamine, 50 ng/ml cholera toxin, 10 μ g/ml insulin, 500 ng/ml hydrocortisone, 1% Pen-Strep antibiotics (Sigma) and fresh 20 ng/ml human EGF (Invitrogen). MCF10A cells were tested for STR profile and mycoplasma. TGF β 1 (Peprotech) stimulation of MCF10A cells was performed on cells, at 40% confluence, previously starved overnight in standard medium lacking EGF, followed by stimulation with sub-optimal TGF β 1 doses (0.75 ng/ml) or saturating doses (5 ng/ml) for the indicated time points. Cells were cultured at 37°C in a humidified atmosphere containing 5% CO₂.

5.2.1.1. Treatments: TGF β 1 AND FA 1%

TGF β 1 (Peprotech) stimulation of MCF10A cells was performed on cells, at 40% confluence, previously starved overnight in standard medium lacking EGF, followed by stimulation with sub-optimal TGF β 1 doses (0.75 ng/ml) or saturating doses (5 ng/ml) for the indicated time points. Cells were cultured at 37°C in a humidified atmosphere containing 5% CO₂.

To crosslink proteins *in vivo* in cells, MCF10A cells were washed twice in PBS, then Crosslinking Buffer (Formaldehyde 1% in ddH₂O) was added for 10 minutes at RT, maintaining the plates on horizontal shaker. Then, the buffer was removed and Glycine Buffer (Glycine 0.125 M in ddH₂O) was added for 5 minutes at RT shaking horizontally. Then cells were washed twice with cold PBS and lysed in RIPA by scraping directly on plates (300 μ l RIPA buffer/150 mm plate, cells at 80-90% confluent); see the dedicated paragraph 5.4.1 for detailed information. Cells were left 30 minutes on ice, then sonicated (20 cycles at 4°C: 30 s sonication and 30 s pause). Sonicated samples were centrifuged at 16,000 g 4°C for 40 minutes and the supernatant was collected for analysis. The protocol was adapted from ²⁶³.

5.2.1.2. RNA interference (RNAi)

RNAi was performed with Lipofectamine RNAimax reagent from Invitrogen, according to manufacturer's instructions. 1,000,000 cells were transfected in reverse, meaning that liposomes complex carrying small-interfering RNAs (siRNAs) were allowed to form for 20 minutes at RT

and then added to trypsinized cells in suspension. For the second cycle, cells were trypsinized and treated in the same way.

For transient KD of EPN3, cells were subjected to single reverse transfection with 10 nM ON-TARGETplus siRNA (Dharmacon) and analyzed 3, 4 or 5 days after transfection; in case of 7 days, cells were subjected to a second cycle of reverse transfection at day 3 or 4. For transient KD of TGF β R1, cells were subjected to double reverse transfection with 8 nM smart pool of RNAi oligos (Dharmacon), and analyzed 3 or 7 days after transfection. For transient KD of TCF4, cells were subjected to double transfection with 15 nM smart pool oligos (Ribocxx). As a negative control, we used cells transfected just with Lipofectamine RNAimax. See Supplementary Table 1 for the nucleotide sequence of siRNAs employed.

5.2.2. Mouse epithelial cell culture

Mouse epithelial cells (MECs) used to test EPN3 expression in FVB/EPN3 WT, KO and KI mice were extracted from mammary gland of mice in post-puberty. Briefly, after mice sacrifice, both inguinal and thoracic mammary glands were put in a culture dish and cut with scissors, usually left and right glands were kept separated. Then, the mechanically digested glands were maintained in a 1:1 mixture of DMEM and Ham's F12 medium (Gibco, Life Technologies), supplemented with HEPES 25 mM, 1% Pen-Strep (Sigma), Liberase TM 20 μ g/ml (Roche) and Collagenase Type 3 150 Units/ml (Worthington): 5 ml per 5 mouse mammary glands (usually left and right) in a 50 ml Falcon tube. The samples were left overnight (ON) in loosely capped tubes, at 37°C in a humidified atmosphere containing 5% CO₂.

MATERIALS AND METHODS

The day after, the cells were gently re-suspended without making bubbles (if the ON digestion worked, the gland should break down), then incubated for another 1 h. After that, 30 ml of PBS were added and cells centrifuged 335 g for 5 minutes at room temperature (RT). The supernatant was removed sucking the interphase between pellet and the undigested mammary glands, without removing the fat. Around 5 ml of solution were left and 5 ml of 0.25% Trypsin was added to resuspend cells, which were then maintained in the incubator for 40 minutes with loosely capped tubes, resuspending every 10 minutes.

25 ml of DMEM and Ham's F12 medium (Gibco, Life Technologies), supplemented with Hepes 25mM, Fetal Calf Serum 10% (Sigma Aldrich) and DNaseI (Sigma Aldrich), were added to resuspend the digested mammary glands and incubated for 5 minutes in the incubator. Then samples were centrifuged at 335 g for 5 minutes at RT. The supernatant was removed, including the fat layer to avoid adipose cells, muscle cells and other contaminants. 3 ml of MEBM medium were added and cells resuspended, then plated cells on collagen coated plates (ThermoFisher Scientific) in order to allow the attachment only of epithelial cells and fibroblasts: usually 2 wells for 5 mammary glands. Then cells were incubated for 24 hours cells.

The day after, supernatant was removed and cells were washed twice with PBS, before 1 ml of 0.25% Trypsin was added and cells were incubated for 30 seconds. After that, the Trypsin was aspirated to selectively reduce the fibroblasts, then 1 ml of 0.25% Trypsin was added again for 3-5 minutes. Cells were recovered in DMEM and Ham's F12 medium (Gibco, Life Technologies), supplemented with Hepes 25 mM and Fetal Calf Serum 10% (Sigma Aldrich). Cells were centrifuged 335 g for 5 minutes and resuspended in DMEM and HAM'S-F12 (Gibco, Life

Technologies), supplemented with 5% horse serum (Invitrogen), 2 mM Glutamine, 50 ng/ml cholera toxin, 10 µg/ml insulin, 500 ng/ml hydrocortisone, 1% Pen-Strep antibiotics (Sigma) and fresh 5 ng/ml human EGF (Invitrogen). Cells were plated in a 6-well plate (100,000 cells per well) for an additional 24 h, then they were lysed to evaluate RNA and protein expression.

5.3. Cell-based assays

5.3.1. Fluorescence activated cell sorting (FACS)

Flow cytometry is a widely used imaging technique to analyze cells based on morphological and protein expression features: it detects and measures fluorescence intensity. The source of fluorescence is that of a fluorescent reporter, i.e., fluorochromes conjugated to antibodies that detect antigens present in/on cells. During FACS acquisition cells must be resuspended as a single-cell solution. When cells pass through the nozzle, the fluid sheath induces them to move from a turbulent regimen to a laminar regimen where they can be scanned one by one by the laser. The first parameter which can be evaluated by FACS analysis is the morphology of the cells. When cells are hit by the laser, light scatters in all of the directions, and detectors measure the scattering. Two physical parameters can be detected by the detectors: side scatter (SSC) and forward scatter (FSC). SSC measures the granularity of the cells, while FSC corresponds to their size. Importantly, FACS also detects fluorescent markers. In this case, the laser excites the fluorophores with different wavelengths. If the wavelength is correct, the electrons of the fluorophore pass from the ground state to the excited state. The electrons then fall back and return to their ground state and

energy is transferred in a form of a photon, which has its own wavelength (emission spectrum). The wavelength produced by the emission is captured by photomultipliers, which convert the signals into digital data.

5.3.1.1. Plasma membrane E-Cadherin intensity

Plasma membrane E-Cadherin was evaluated using an anti-E-cadherin antibody (HECD-1) from Abcam, recognizing the extracellular domain of E-cadherin^{56,280}. 140,000 MCF10A-Ctr, MCF10A-Ctr Mock or MCF10A-Ctr EPN3 KD cells, or 180,000 MCF10A-EPN3 cells were plated 1.5 days before analysis on 35 mm dishes. The day of the experiment, cells were at a confluence between 60% and 80%. Cells were washed twice and left for 3 hours in incubator in medium without EGF. After that, cells were incubated with the HECD-1 E-Cadherin antibody (Abcam) – dilution 1:25 in 300 µl of medium – for 1 hour at 4°C. Then cells were incubated with a secondary antibody, Alexa-488 (dilution 1:200 in 500 µl of medium), for 30 minutes at 4°C. After that, cells were washed twice with PBS and 0.25% Trypsin was added for 15-20 minutes. Then cells were recovered with medium in a 15 ml Falcon, centrifuged 335 g for 5 minutes, resuspended with 3 ml PBS, centrifuged again 335 g for 5 minutes and resuspended in 500 µl PBS. Then 500 µl of 2% formaldehyde were added for 15 minutes on ice to fix cells. Subsequently, cells were centrifuged 335 g for 5 minutes and resuspended in PBS with EDTA 2 mM to be analyzed at FACS Celesta (BD).

As a negative control, samples stained only with secondary antibody were used. The analysis was performed through the FlowJo Software (LLC). Briefly, doublets and clumps of cells were

excluded through the SSC Area over Height parameters; then the cells with correct morphology were selected through the FSC Area over SSC Area parameters. Finally, E-Cadherin mean fluorescence intensity of the population was used to compare the different samples.

5.3.1.2. E-Cadherin internalization assay

To evaluate E-Cadherin internalization, it both FACS and IF analyses were used. The same protocol previously described (Section 5.3.1.1) was used to evaluate PM E-Cadherin. After antibody incubation, cells were washed twice with PBS and reincubated at 37°C with complete medium \pm TGF β 5 ng/ml (total volume of 1.5 ml). Then cells, after the indicated time points, were washed once with PBS and treated with Acid Wash (AW) solution. 2 different AW have been tested, showing the same efficiency: for AW1, cells were treated on ice for 3 x 45 s, then 2 PBS washes were performed; for AW2, cells were treated on ice once for 5 minutes, then washed once with Medium with HEPES 20 mM and once with PBS. In case of IF, glass coverslips were fixed for 10 minutes with 4% paraformaldehyde (PFA) at RT, followed by IF to stain nuclei and analyze by confocal microscope.

For FACS analysis, 0.25% Trypsin was added to cells for 90 s and then cells were recovered in a 15 ml Falcon tube with medium. Then cells were processed as previously described to be fixed and analyzed (see Section 5.3.1.1). To evaluate the amount of internalized E-Cadherin, the mean fluorescence intensity of each time-point was divided by the total amount of fluorescence (internalized and PM E-Cadherin), and the time 0 minutes was set as background and its value subtracted from the other time-points.

5.3.2. Immunofluorescence (IF)

5.3.2.1. Plasma membrane E-Cadherin intensity

In order to evaluate cell shape and E-Cadherin localization, the same experiment described in the Paragraph 5.3.1.1 was performed. After the acid wash treatment step, coverslips were fixed for 10 minutes with PFA 4% at RT, followed by IF to stain nuclei and analyze by confocal microscope.

5.3.2.2. E-Cadherin internalization assay

In order to evaluate E-Cadherin internalization also by IF, the same experiment described in the Paragraph 5.3.1.2 was performed. After the acid wash treatment step, coverslips were fixed for 10 minutes with PFA 4% at RT, followed by IF to stain nuclei and analyze by confocal microscope.

5.3.2.3. Staining in growing conditions

150,000 MCF10A cells per well of a 6-well plate, were plated 2 days before the fixation on glass coverslips. Cells, plated on glass coverslips, were fixed in 4% PFA for 10 minutes at RT for E-Cadherin staining, or cold methanol for 10 minutes at -20°C for active β -catenin staining, then washed with PBS and permeabilized in PBS 0.1% Triton X-100 (Sigma) for 10 minutes at RT. To prevent non-specific binding of the antibodies, cells were incubated with PBS in the presence of 2% BSA for 30 minutes. Primary antibodies were diluted in PBS 0.2% BSA: E-Cadherin antibody

(BD) and active β -catenin (Millipore) diluted 1:200. After 1 hour of incubation at RT, coverslips were washed 3 times with PBS. Cells were then incubated for 30 minutes at RT with the appropriate secondary antibody, Alexa 488-conjugated (dilution 1:400, Molecular Probes). After three washes in PBS, coverslips were mounted in a 90% glycerol solution containing diazabicyclo-(2.2.2)octane antifade (DAPI, dilution 1:5000, Sigma). Images were obtained using a Leica TCS SP8 confocal microscope equipped with a 63X oil objective and processed using ImageJ.

Quantification of active β -catenin

Images were processed with Image J software. For each image (.tiff), a DAPI mask was used to identify nuclei and then measure the mean fluorescence intensity of active β -catenin in the nuclei per each cell. The mean fluorescence intensity of total cells was instead calculated dividing the mean fluorescence intensity of the entire field of view by the number of cells in that field. Finally, the results were reported as relative fold-change compared to EV of the ratio of active β -catenin staining in nuclei/total. The analysis was performed with the help of Dr. Freddi Stefano (IFOM/IEO, Milan).

5.4. RNA-based assays

Total RNA extraction and quantitative Real Time PCR

Cells were lysed by removing cell culture medium, washing twice with PBS and adding Trizol (Life Technologies) directly to the cell plates at 60-80% of confluence and lysate was clarified by

centrifugation at 12,000 g for 15 minutes at 4°C. The aqueous phase was used to extract total RNA by using RNeasy kit (Quiagen) according to manufacturer's protocol. Single strand cDNA was obtained by SuperScript VILO reverse transcriptase (Invitrogen) according to manufacturer's protocol.

RT-qPCR experiments were performed according to Taqman chemistry (Thermo Fisher Scientific) in technical triplicates. Data were normalized on the average expression of housekeeping genes, *18S*, *ACTB* and *GAPDH*, and results were reported as relative mRNA expression compared to control by the use of the $2^{-\Delta\Delta C_t}$ method. See **Supplementary Table 2** for Taqman assays.

5.5. Protein-based assays

5.5.1. Cell lysis and Western Blot (WB)

Cells were lysed on ice by removing cell culture medium, washing twice with PBS and by adding RIPA buffer directly to the cell plates at 60-80% of confluence. Cell lysate was clarified by centrifugation at 16,000 g for 20 minutes at 4°C and supernatant was transferred to a new tube. Protein concentration was measured by Bradford Assay (Biorad). Protein lysates were prepared in Laemmli Buffer 1X (diluted from the 4X). Protein lysates were run on gradient 4-20% pre-cast Gels (Biorad) at 250 Volts and transferred by using Trans-Blot (Biorad) according to manufacturer's instructions. Filters were blocked with 5% non-fat dried milk (Applichem) diluted in TBS 0.1% Tween (TBS-T) and then incubated ON with primary antibody according to the data

sheet. Following 3 washes with TBS-T, filters were then incubated with the appropriate secondary antibody conjugated with horseradish peroxidase. After 3 more washes, the signal was detected at Chemidoc (Biorad) using Amersham ECL Western Blotting Detection Reagent (GE Healthcare), for strong signals, or with Clarity Western ECL Blotting Substrates (Biorad) or with SuperSignal West Femto (Thermo Fisher) for low signals. See **Supplementary Table 3** for the list of antibodies employed.

5.5.2. Immunoprecipitation (IP) and Co-IP

1,000,000 MCF10A-Ctr cells and 3,000,000 MCF10A-EPN3 cells were plated the first day on 150 mm plates. 2 days after cells were washed twice in PBS and EGF-starved ON. The morning after, cells, at a confluence between 80-95%, were washed twice in PBS and lysed: JS was used for α -FLAG antibody IP settings and mass spectrometry analysis; RIPA was used for α -EPN3 Rabbit antibody Co-IP.

α -FLAG IP

FLAG M2 antibody 50% slurry (Sigma) was used at a concentration of 25 μ g per mg of lysate, in a volume of 1 ml. MCF10A-Ctr and -EPN3 cells were lysed in JS buffer. Lysate was incubated with a mixture 1:1 of Protein A and Protein G for 1 hour at 4°C in rotation to reduce the non-specific binding. Lysates were washed twice with JS buffer by centrifugation (16,000 g, 15 s), and then quantified for protein amount using the Bradford Assay (Biorad). Experiments were performed

MATERIALS AND METHODS

starting from 1 mg of lysate for each condition for 2 hours at 4°C in rotation. At least 4 washes were then performed in JS buffer and elution in Laemmli Buffer 2x was performed.

In order to prepare the samples for mass spectrometry analysis, FLAG Peptide Elution was performed: after the washes, the lysates were incubated with 2 volumes of FLAG Peptide 300 ng/ml (diluted in TBS 1X) for 15 minutes at 4°C in rotation. Then lysates were centrifuged (16,000 g, 15 minutes, 4°C) and the beads reincubated with fresh FLAG Peptide again. The first and second cycles were mixed and then concentrated through an AMICON 10 kDa Eppendorf (Merck Millipore), by centrifugation for 15 minutes at 16,000 g, 4°C. Samples were resuspended in Laemmli Buffer 2x. Mass spectrometry was performed on fresh lysates.

α -EPN3 Rabbit IP

EPN3 Rabbit antibody from Pilcher's lab was used at a concentration of 5 ng per mg of lysate, in a volume of 400 μ l. MCF10A-Ctr and -EPN3 cells were lysed in RIPA buffer after crosslinking treatment. Lysate was incubated with a mixture 1:1 of Protein A and Protein G for 1 hour at 4°C in rotation to reduce the non-specific binding. Lysates were washed twice with JS buffer by centrifugation (16,000 g, 15 seconds), and then quantified for protein amount using the Bradford Assay (Biorad). Experiments were performed starting from 1 mg of lysate for each condition for 2 hours at 4°C in rotation. Protein G was added and incubated for an additional 1 h. At least 4 washes were then performed in JS buffer and elution in Laemmli Buffer 2x was performed. Co-IP experiments were performed on fresh lysates.

5.5.3. Mass spectrometry analyses

IP with α -FLAG antibody of EPN3 was performed, followed by resolution on a SDS-PAGE gel and Coomassie staining. Successively, samples were Trypsin digested and analyzed with a label-free mass spectrometry analysis. Proteins were resolved by SDS-PAGE on a gradient gel (4–12% Tris–HCl Precast Gel, Invitrogen) and stained with Colloidal Instant Blue (Expedeon). Enzymatic in-gel digestion was performed essentially as previously described in ²⁸¹. Briefly, samples were subjected to reduction in 10 mM DTT for 1 hour at 56°C, followed by alkylation with 55 mM iodoacetamide for 45 minutes at RT, in the dark. Digestion was carried out saturating the gel with 12.5 ng/ml sequencing grade modified Trypsin (Roche) in 50 mM ammonium bicarbonate, ON. Peptide mixtures were acidified with formic acid (final concentration 0.1%), extracted from gel slices with two rounds of washes (in 80% acetonitrile (ACN)/0.1% formic acid and then in 100% ACN, respectively) and concentrated to 100 μ l in a vacuum concentrator (Eppendorf). Each sample was loaded onto homemade C18-Stage Tips, for concentration and desalting prior LC-MS/MS analysis ²⁸². 3 μ l of each sample were analyzed as technical replicate on a nLC–ESI–MS–MS quadrupole Orbitrap QExactive-HF mass spectrometer (Thermo Fisher Scientific). Peptides separation was achieved on a linear gradient from 100% solvent A (2% ACN, 0.1% formic acid) to 60% solvent B (80% acetonitrile, 0.1% formic acid) over 37 minutes and from 60 to 100% solvent B in 8 minutes at a constant flow rate of 0.25 μ l/minutes on UHPLC Easy-nLC 1000 (Thermo Scientific) connected to a 25-cm fused-silica emitter of 75 μ m inner diameter (New Objective, Inc. Woburn, MA, USA), packed in-house with ReproSil-Pur C18-AQ 1.9 μ m beads (Dr Maisch GmbH, Ammerbuch, Germany) using a high-pressure bomb loader (Proxeon, Odense, Denmark).

MATERIALS AND METHODS

Mass spectrometry data were acquired using a data-dependent top 15 method for higher-energy C-trap dissociation (HCD) fragmentation. Survey full scan mass spectrometry spectra (300–1650 Th) were acquired in the Orbitrap with 60000 resolution, AGC target 3e6, IT 20 ms. For HCD spectra, resolution was set to 15000 at m/z 200, AGC target 1e5, IT 80 ms; NCE 28% and isolation width 1.2 m/z .

For quantitative proteomics, “.Raw” data were processed with MaxQuant (ver. 1.5.2.8) searching against the database “uniprot_human_2015_03” setting Trypsin specificity and up to two missed cleavages. Cysteine Carbamidomethyl was used as fixed modification, methionine oxidation and protein N-terminal acetylation as variable modifications. Mass deviation for MS-MS peaks was set at 20 ppm. The peptides and protein false discovery rates (FDR) were set to 0.01; the minimal length required for a peptide was six amino acids; a minimum of two peptides and at least one unique peptide were required for high-confidence protein identification. The lists of identified proteins were filtered to eliminate reverse hits and known contaminants.

To evaluate the specific EPN3 interactors, the On/Off interactors were selected from the LFQ intensity of each protein in the two samples, MCF10A-Ctr and -EPN3. Once the list of On/Off interactors was obtained, it was compared to all the others interactors (the ones in common between the two samples and the On/Off of MCF10A-Ctr sample). These lists were submitted into the KEGG Pathway database (v. 2016) through EnrichR, a specific enrichment analysis software (Kuleshov et al., 2016) with an FDR < 0.05. The mass spectrometry analyses was performed thanks to the IFOM Proteomic Facility, the subsequent pathway analysis was performed with the help of Dr. Confalonieri Stefano (IFOM/IEO, Milan).

5.6. Mice handling

FVB/EPN3 KO mice were a kind gift from De Camilli's lab ¹¹², and ubiquitously lack the *EPN3* gene. FVB/EPN3 KI mice were previously generated with overexpression of the human transgene under a strong ubiquitous promoter (UbiC) inserted in the ROSA 26 locus; a stop codon flanked by Lox-P sites is positioned between the promoter and the transgene, and allows overexpression of EPN3 only in presence of Cre recombinase. When these mice are crossed with mice bearing Cre recombinase under the Keratin-5 promoter (active only in basal epithelial cells, K5-Cre) ²⁶⁴, EPN3 overexpression occurs specifically in the epithelial compartment. Mice were maintained according to the Italian Law, with the permission obtained through the Research Project 27/2015.

5.7. Mice-based assays

5.7.1. Histological procedures

Mammary glands from WT, EPN3-KO and EPN3-KI FVB female mice, were processed for whole-mount analysis, as described previously ²⁶⁵. Briefly, the forth left mammary gland was removed and stretched out flat onto a glass slide, maintaining integrity and shape. Then, it was fixed in EtOH 100%/glacial acetic acid (ratio 25%:75%) at RT from ON to over weekend; subsequently, one EtOH 70% wash of 30 minutes and one ddH₂O wash of 10 minutes were performed. Then, staining in Carmine Alum solution ON at RT was performed. After that,

sequential washes in EtOH 70%, 95% and 100% for 30 minutes were performed, followed by fat clarification in BABB solution (Benzylalcohol (Sigma):Benzylbenzoate (Sigma), ratio 1:2). Microscope acquisition was performed with a stereomicroscope (Leica).

Quantification of ductal length

Images (.tiff) were analyzed to understand the distance of the last observed epithelial duct from the lymph node with ImageJ software. Briefly, a first line was designed perpendicular to the length of the mammary gland, touching the extremity of the lymph node closer to the end of the gland; a second line starting from the last epithelial duct was designed, ending perpendicularly to the first line. The length of this second line was calculated, according to the pixel size of the picture, in order to measure in millimeters the distance from the lymph node of every mammary duct analyzed.

Quantification of branching

Branching quantification was performed using the Scholl analysis, through ImageJ software^{283,284}. Briefly, images (.tiff) were subjected to a binary mask that was analyzed through the plug-in School analyses of ImageJ. The analysis was performed by Dr. Freddi Stefano (IFOM/IEO, Milan).

5.7.2. Immunohistochemistry (IHC)

Mammary glands were fixed in 1% formaldehyde for 6 hours at RT and then ON at 4°C. The day after, the glands were washed twice in PBS and then paraffin-embedded to obtain FFPE samples.

Paraffin sections were twice de-paraffinized with Bio Clear (Bio-Optica) for 15 minutes and hydrated through graded alcohol series (100%, 95%, 70% ethanol and water) for 5 minutes. Antigen unmasking was performed with 0.1 mM citrate buffer (pH=6) or EDTA (pH=8) for 50 minutes at 95°C. Slides were cooled down for 20 minutes at RT then washed in water and treated with 3% H₂O₂ for 5 minutes at RT. Then, slides were pre-incubated with an antibody mixture (2% BSA, 2% normal goat serum, 0.02% Tween20 in TBS) for 20 minutes at RT and stained with primary antibody for 1 hour at 37°C. As primary antibodies, we used: Ki67 1:200, mouse (Dako), E-cadherin 1:250 mouse (Dako), N-cadherin 1:250 mouse (Dako), EPN3 1:30.000, mouse (home-made). After two washes with TBS, slides were incubated with a secondary antibody ready to use (DAKO Envision system HRP rabbit or mouse) for 30 minutes at RT and washed twice again in TBS. The sections were incubated in peroxidase substrate solution (DAB DAKO) from 2-10 minutes (1 drop of chromogen/ml buffer) and the reaction was blocked in water. Slides were counterstained with hematoxylin for 5 s, dehydrated through graded alcohol series (water and 70%, 95%, 100% ethanol) for 5 minutes each, washed with Bio Clear (Bio-Optica) twice for 1 minutes and finally mounted with Eukitt (Kindler GmbH). IHC analyses were performed with the help of IEO Molecular Pathology Unit.

5.8. Statistical Analyses

All statistical analyses were performed using Excel and JMP software. Student T-Test and Two-way ANOVA with Bonferroni post-tests were used to analyze the experiments and the relative statistical significance of their differences. Some of the statistical analyses performed were

obtained with the help of Dr. Confalonieri Stefano (IFOM/IEO, Milan). In all cases p-value <0.05

*, p-value <0.01 **, p-value <0.001 ***.

5.9. Supplementary Tables

Supplementary Table 1: RNAi oligo sequences (3' – 5')

TGFBR1 ON-TARGETplus SMART pool	GAGAAGAACGUUCGUGGUU UGCGAGAACUAUUGUGUUA GACCACAGACAAAGUUAUA CGAGAUAGGCCGUUUGUAU
EPN3 ON-TARGETplus siRNA	GUACAAGGCUCUAACAUUG
TCF7L2 iBONI siRNA pool	AUAAUACAGAACCAACUCCCC UUCUUCCAAACUUUCCCGGCCCCC UUAAGUCUGCUGCCUACCCC

Supplementary Table 2: TaqMan assays

Gene symbol	Assay ID	Ref Seq
TGFB1	Hs00998133_m1	NM_000660.5
TGFB2	Hs00234244_m1	NM_003238.3
TGFB3	Hs01086000_m1	NM_003239.3
TGFBR1	Hs00610320_m1	NM_004612.2
TGFBR2	Hs00234253_m1	NM_003242.5
18S	Hs9999901_s1	NR_003286.2
GAPDH	Hs9999905_m1	NM_002046.5
ACTB	Hs9999903_m1	NM_001101.3
CDH1	Hs00170423_m1	NM_004360.3

CDH2	Hs00169953_m1	NM_001792.3
EPN3	mm00660955_m1	NM_027984.3
VIM	Hs00185584_m1	NM_003380.3
SNAI1	Hs00195591_m1	NM_005985.3
SNAI2	Hs00950344_m1	NM_003068.4
ZEB1	Hs00232783_m1	NM_01128128.2

Supplementary Table 3: List of antibodies

Antibody	Producer	Clone or Epitope	Catalog number	Dilution
EPN3	homemade	VI31, epitope 464-483		1:1000 WB
EPN3	homemade	monoclonal		1:30.000 IHC
EPN1/Epn2	homemade	ZZ3, epitope 249-401		1:100 WB
FLAG	Sigma	Clone M2	F1804	1:1000 WB
Twist	Santa Cruz	Twist2C1a	sc-811417	1:500 WB
GAPDH	Santa Cruz	6C5	sc-32233	1:1000 WB
E-Cadherin	BD	Clone 36	610181	1:1000 WB 1:200 IF
E-Cadherin	Abcam	HECD-1	Ab1416	1:25 Int. Assay
E-Cadherin	Dako	NCH-38	M3612	1:250 IHC
Vimentin	BD	RV202	550513	1:1000 WB
N-Cadherin	BD	Clone 32	610920	1:1000 WB
N-Cadherin	Dako	6G11	M3613	1:250 IHC
Snail	Cell Signaling	L70G2	3895	1:1000 WB
pSMAD2	Cell Signaling	138D4	3108	1:1000 WB
Zeb1	Santa Cruz	H102	sc-25388	1:500 WB
TCF4	Cell Signaling	C48H11	2569	1:1000 WB

MATERIALS AND METHODS

β -Catenin total	BD	14/ β -Catenin	610153	1:1000 WB
β -Catenin active	Millipore	8E7	05-665	1:100 WB
p120	Cell Signaling	Polyclonal	4989	1:1000 WB
Integrin- β 1	homemade	Polyclonal		1:1000 WB
Actin	homemade	Polyclonal		1:1000 WB
Clathrin HC	BD	Clone 23	610500	1:1000 WB
H3	Abcam	Polyclonal	Ab1791	1:3000 WB
Tubulin	homemade			1:1000 WB
Ki67	Dako	MIB-1	M7240	1:200 IHC

REFERENCES

1. Sigismund, S. *et al.* Endocytosis and Signaling: Cell Logistics Shape the Eukaryotic Cell Plan. *Physiol. Rev.* **92**, 273–366 (2012).
2. Swanson, J. A. Shaping cups into phagosomes and macropinosomes. *Nat. Rev. Mol. Cell Biol.* **9**, 639–649 (2008).
3. Kerr, M. C. & Teasdale, R. D. Defining Macropinocytosis. *Traffic* **10**, 364–371 (2009).
4. Doherty, G. J. & McMahon, H. T. Mechanisms of Endocytosis. *Annu. Rev. Biochem.* **78**, 857–902 (2009).
5. Yu, A. *et al.* Association of Dishevelled with the clathrin AP-2 adaptor is required for Frizzled endocytosis and planar cell polarity signaling. *Dev. Cell* **12**, 129–41 (2007).
6. Kirchhausen, T., Owen, D. & Harrison, S. C. Molecular Structure, Function, and Dynamics of Clathrin-Mediated Membrane Traffic. *Cold Spring Harb. Perspect. Biol.* **6**, a016725–a016725 (2014).
7. Robinson, M. S. Forty Years of Clathrin-coated Vesicles. *Traffic* **16**, 1210–1238 (2015).
8. Nabi, I. R. Cavin fever: regulating caveolae. *Nat. Cell Biol.* **11**, 789–791 (2009).
9. Hansen, C. G. & Nichols, B. J. Molecular mechanisms of clathrin-independent endocytosis. *J. Cell Sci.* **122**, 1713–1721 (2009).
10. Caldieri, G. *et al.* Reticulon 3-dependent ER-PM contact sites control EGFR nonclathrin endocytosis. *Science (80-.).* **356**, 617–624 (2017).
11. Ferguson, J. P. *et al.* Mechanoregulation of clathrin-mediated endocytosis. *J. Cell*

- Sci.* **130**, 3631–3636 (2017).
12. Glebov, O. O., Bright, N. A. & Nichols, B. J. Flotillin-1 defines a clathrin-independent endocytic pathway in mammalian cells. *Nat. Cell Biol.* **8**, 46–54 (2006).
 13. Frick, M. *et al.* Coassembly of flotillins induces formation of membrane microdomains, membrane curvature, and vesicle budding. *Curr. Biol.* **17**, 1151–6 (2007).
 14. Johannes, L., Parton, R. G., Bassereau, P. & Mayor, S. Building endocytic pits without clathrin. *Nat. Rev. Mol. Cell Biol.* **16**, 311–321 (2015).
 15. Pearse, B. M., Smith, C. J. & Owen, D. J. Clathrin coat construction in endocytosis. *Curr. Opin. Struct. Biol.* **10**, 220–8 (2000).
 16. Fallon, L. *et al.* A regulated interaction with the UIM protein Eps15 implicates parkin in EGF receptor trafficking and PI(3)K-Akt signalling. *Nat. Cell Biol.* **8**, 834–42 (2006).
 17. Ford, M. G. J. *et al.* Curvature of clathrin-coated pits driven by epsin. *Nature* **419**, 361–6 (2002).
 18. Henne, W. M. *et al.* FCHO proteins are nucleators of clathrin-mediated endocytosis. *Science* **328**, 1281–4 (2010).
 19. Antonny, B. *et al.* Membrane fission by dynamin: what we know and what we need to know. *EMBO J.* **35**, 2270–2284 (2016).
 20. Sweitzer, S. M. & Hinshaw, J. E. Dynamin Undergoes a GTP-Dependent Conformational Change Causing Vesiculation. *Cell* **93**, 1021–1029 (1998).
 21. Ungewickell, E. *et al.* Role of auxilin in uncoating clathrin-coated vesicles. *Nature* **378**, 632–635 (1995).

22. McMahon, H. T. & Boucrot, E. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **12**, 517–533 (2011).
23. Kaksonen, M. & Roux, A. Mechanisms of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **19**, 313–326 (2018).
24. Höning, S. *et al.* Phosphatidylinositol-(4,5)-Bisphosphate Regulates Sorting Signal Recognition by the Clathrin-Associated Adaptor Complex AP2. *Mol. Cell* **18**, 519–531 (2005).
25. Kirchhausen, T. & Harrison, S. C. Protein organization in clathrin trimers. *Cell* **23**, 755–61 (1981).
26. Dannhauser, P. N. & Ungewickell, E. J. Reconstitution of clathrin-coated bud and vesicle formation with minimal components. *Nat. Cell Biol.* **14**, 634–9 (2012).
27. Collins, B. M., McCoy, A. J., Kent, H. M., Evans, P. R. & Owen, D. J. Molecular architecture and functional model of the endocytic AP2 complex. *Cell* **109**, 523–35 (2002).
28. Kelly, B. T. *et al.* A structural explanation for the binding of endocytic dileucine motifs by the AP2 complex. *Nature* **456**, 976–979 (2008).
29. Guo, M., Jan, L. Y. & Jan, Y. N. Control of daughter cell fates during asymmetric division: interaction of Numb and Notch. *Neuron* **17**, 27–41 (1996).
30. Oldham, C. E., Mohny, R. P., Miller, S. L. H., Hanes, R. N. & O’Byrne, J. P. The ubiquitin-interacting motifs target the endocytic adaptor protein epsin for ubiquitination. *Curr. Biol.* **12**, 1112–6 (2002).
31. Woelk, T. *et al.* Molecular mechanisms of coupled monoubiquitination. *Nat. Cell Biol.* **8**, 1246–54 (2006).
32. Ehrlich, M. *et al.* Endocytosis by Random Initiation and Stabilization of Clathrin-

- Coated Pits. *Cell* **118**, 591–605 (2004).
33. Henry, A. G. *et al.* Regulation of endocytic clathrin dynamics by cargo ubiquitination. *Dev. Cell* **23**, 519–32 (2012).
34. Chang-Ileto, B. *et al.* Synaptojanin 1-Mediated PI(4,5)P₂ Hydrolysis Is Modulated by Membrane Curvature and Facilitates Membrane Fission. *Dev. Cell* **20**, 206–218 (2011).
35. Massol, R. H., Boll, W., Griffin, A. M. & Kirchhausen, T. A burst of auxilin recruitment determines the onset of clathrin-coated vesicle uncoating. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10265–10270 (2006).
36. Sigismund, S. *et al.* Threshold-controlled ubiquitination of the EGFR directs receptor fate. *EMBO J.* **32**, 2140–2157 (2013).
37. Balogh, P., Katz, S. & Kiss, A. L. The Role of Endocytic Pathways in TGF- β Signaling. *Pathol. Oncol. Res.* **19**, 141–148 (2013).
38. Miaczynska, M., Pelkmans, L. & Zerial, M. Not just a sink: endosomes in control of signal transduction. *Curr. Opin. Cell Biol.* **16**, 400–406 (2004).
39. Villaseñor, R., Nonaka, H., Del Conte-Zerial, P., Kalaidzidis, Y. & Zerial, M. Regulation of EGFR signal transduction by analogue-to-digital conversion in endosomes. *Elife* **4**, (2015).
40. Tsukazaki, T., Chiang, T. A., Davison, A. F., Attisano, L. & Wrana, J. L. SARA, a FYVE Domain Protein that Recruits Smad2 to the TGF β Receptor. *Cell* **95**, 779–791 (1998).
41. Sigismund, S. & Scita, G. The ‘endocytic matrix reloaded’ and its impact on the plasticity of migratory strategies. *Curr. Opin. Cell Biol.* **54**, 9–17 (2018).
42. Caswell, P. T. *et al.* Rab25 Associates with α 5 β 1 Integrin to Promote Invasive

- Migration in 3D Microenvironments. *Dev. Cell* **13**, 496–510 (2007).
43. White, D. P., Caswell, P. T. & Norman, J. C. α 5 and β 1 integrin recycling pathways dictate downstream Rho kinase signaling to regulate persistent cell migration. *J. Cell Biol.* **177**, 515–525 (2007).
 44. Sinha, B. *et al.* Cells respond to mechanical stress by rapid disassembly of caveolae. *Cell* **144**, 402–13 (2011).
 45. Boulant, S., Kural, C., Zeeh, J.-C., Ubelmann, F. & Kirchhausen, T. Actin dynamics counteract membrane tension during clathrin-mediated endocytosis. *Nat. Cell Biol.* **13**, 1124–31 (2011).
 46. Bazinet, C., Katzen, A. L., Morgan, M., Mahowald, A. P. & Lemmon, S. K. The *Drosophila* clathrin heavy chain gene: clathrin function is essential in a multicellular organism. *Genetics* **134**, 1119–34 (1993).
 47. Grant, B. & Hirsh, D. Receptor-mediated Endocytosis in the *Caenorhabditis elegans* Oocyte. *Mol. Biol. Cell* **10**, 4311–4326 (1999).
 48. Mitsunari, T. *et al.* Clathrin Adaptor AP-2 Is Essential for Early Embryonal Development. *Mol. Cell. Biol.* **25**, 9318–9323 (2005).
 49. Monks, J. & Neville, M. C. Albumin transcytosis across the epithelium of the lactating mouse mammary gland. *J. Physiol.* **560**, 267–80 (2004).
 50. Pauloin, A., Tooze, S. A., Michelutti, I., Delpal, S. & Ollivier-Bousquet, M. The majority of clathrin coated vesicles from lactating rabbit mammary gland arises from the secretory pathway. *J. Cell Sci.* **112 (Pt 22)**, 4089–100 (1999).
 51. Ishikawa, Y. *et al.* Role of the clathrin adaptor PICALM in normal hematopoiesis and polycythemia vera pathophysiology. *Haematologica* **100**, 439–51 (2015).
 52. Jing, S. Q., Spencer, T., Miller, K., Hopkins, C. & Trowbridge, I. S. Role of the

- human transferrin receptor cytoplasmic domain in endocytosis: localization of a specific signal sequence for internalization. *J. Cell Biol.* **110**, 283–94 (1990).
53. Anderson, R. G., Brown, M. S. & Goldstein, J. L. Role of the coated endocytic vesicle in the uptake of receptor-bound low density lipoprotein in human fibroblasts. *Cell* **10**, 351–64 (1977).
54. Huang, F., Khvorova, A., Marshall, W. & Sorkin, A. Analysis of Clathrin-mediated Endocytosis of Epidermal Growth Factor Receptor by RNA Interference. *J. Biol. Chem.* **279**, 16657–16661 (2004).
55. Bryant, D. M. & Stow, J. L. The ins and outs of E-cadherin trafficking. *Trends in Cell Biology* **14**, 427–434 (2004).
56. Paterson, A. D., Parton, R. G., Ferguson, C., Stow, J. L. & Yap, A. S. Characterization of E-cadherin endocytosis in isolated MCF-7 and Chinese hamster ovary cells. The initial fate of unbound E-cadherin. *J. Biol. Chem.* **278**, 21050–21057 (2003).
57. Mellman, I. & Yarden, Y. Endocytosis and Cancer. *Endocytosis* 1–24 (2014). doi:10.1101/cshperspect.a016949
58. Tan, Y.-H. C. *et al.* CBL Is Frequently Altered in Lung Cancers: Its Relationship to Mutations in MET and EGFR Tyrosine Kinases. *PLoS One* **5**, e8972 (2010).
59. Grandal, M. V. *et al.* EGFRvIII escapes down-regulation due to impaired internalization and sorting to lysosomes. *Carcinogenesis* **28**, 1408–1417 (2007).
60. Tosoni, D. *et al.* The Numb/p53 circuitry couples replicative self-renewal and tumor suppression in mammary epithelial cells. *J. Cell Biol.* **211**, 845–862 (2015).

61. Tosoni, D. *et al.* Pre-clinical validation of a selective anti-cancer stem cell therapy for Numb-deficient human breast cancers. *EMBO Mol. Med.* **9**, 655–671 (2017).
62. Pece, S. *et al.* Loss of negative regulation by Numb over Notch is relevant to human breast carcinogenesis. *J. Cell Biol.* **167**, 215–221 (2004).
63. Colaluca, I. N. *et al.* NUMB controls p53 tumour suppressor activity. *Nature* **451**, 76–80 (2008).
64. Mosesson, Y., Mills, G. B. & Yarden, Y. Derailed endocytosis: an emerging feature of cancer. *Nat. Rev. Cancer* **8**, 835–50 (2008).
65. He, H. *et al.* Identification and characterization of nine novel human small GTPases showing variable expressions in liver cancer tissues. *Gene Expr.* **10**, 231–42 (2002).
66. Cheng, K. W., Lu, Y. & Mills, G. B. Assay of Rab25 Function in Ovarian and Breast Cancers. in *Methods in enzymology* **403**, 202–215 (2005).
67. Chen, H. *et al.* Epsin is an EH-domain-binding protein implicated in clathrin-mediated endocytosis. *Nature* **394**, 793–797 (1998).
68. 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. *J. Biol. Chem.* **274**, 33959–65 (1999).
69. Spradling, K. D., Burke, T. J., Lohi, J. & Pilcher, B. K. Cloning and Initial Characterization of Human Epsin 3, a Novel Matrix-Induced Keratinocyte Specific Transcript. *J. Invest. Dermatol.* **115**, 332 (2000).
70. De Camilli, P. *et al.* The ENTH domain. *FEBS Lett.* **513**, 11–8 (2002).
71. Horvath, C. A. J., Vanden Broeck, D., Boulet, G. A. V, Bogers, J. & De Wolf, M. J. S.

- Epsin: Inducing membrane curvature. *Int. J. Biochem. Cell Biol.* **39**, 1765–1770 (2007).
72. Sen, A., Madhivanan, K., Mukherjee, D. & Claudio Aguilar, R. The epsin protein family: Coordinators of endocytosis and signaling. *Biomol. Concepts* **3**, 117–126 (2012).
73. Hyman, J., Chen, H., Di Fiore, P. P., De Camilli, P. & Brunger, A. T. Epsin 1 undergoes nucleocytoplasmic 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). *J. Cell Biol.* **149**, 537–46 (2000).
74. Lohi, O. & Lehto, V. P. VHS domain marks a group of proteins involved in endocytosis and vesicular trafficking. *FEBS Lett.* **440**, 255–7 (1998).
75. Huber, A. H., Nelson, W. J. & Weis, W. I. Three-dimensional structure of the armadillo repeat region of beta-catenin. *Cell* **90**, 871–82 (1997).
76. Wendland, B., Steece, K. E. & Emr, S. D. Yeast epsins contain an essential N-terminal ENTH domain, bind clathrin and are required for endocytosis. *EMBO J.* **18**, 4383–4393 (1999).
77. Stukenberg, P. T. *et al.* Systematic identification of mitotic phosphoproteins. *Curr. Biol.* **7**, 338–48 (1997).
78. Cadavid, A. L., Ginzler, A. & Fischer, J. A. The function of the Drosophila fat facets deubiquitinating enzyme in limiting photoreceptor cell number is intimately associated with endocytosis. *Development* **127**, 1727–36 (2000).
79. 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. *Proc.*

- Natl. Acad. Sci. U. S. A.* **106**, 13838–43 (2009).
80. Chen, K. L. *et al.* Epsin Family of Endocytic Adaptor Proteins as Oncogenic Regulators of Cancer Progression. *J. Can. Res. Updates* **2**, 144–150 (2013).
81. Lai, E. C. Notch signaling: control of cell communication and cell fate. *Development* **131**, 965–973 (2004).
82. Yamamoto, S., Schulze, K. L. & Bellen, H. J. Introduction to Notch Signaling. in *Methods in molecular biology (Clifton, N.J.)* **1187**, 1–14 (2014).
83. Schadler, K. L., Zweidler-McKay, P. A., Guan, H. & Kleinerman, E. S. Delta-like ligand 4 plays a critical role in pericyte/vascular smooth muscle cell formation during vasculogenesis and tumor vessel expansion in Ewing's sarcoma. *Clin. Cancer Res.* **16**, 848–56 (2010).
84. Tian, X., Hansen, D., Schedl, T. & Skeath, J. B. Epsin potentiates Notch pathway activity in *Drosophila* and *C. elegans*. *Development* **131**, 5807–15 (2004).
85. Wang, W. & Struhl, G. *Drosophila* Epsin mediates a select endocytic pathway that DSL ligands must enter to activate Notch. *Development* **131**, 5367–80 (2004).
86. Langridge, P. D. & Struhl, G. Epsin-Dependent Ligand Endocytosis Activates Notch by Force. *Cell* **171**, 1383–1396.e12 (2017).
87. Mitsudomi, T. & Yatabe, Y. Epidermal growth factor receptor in relation to tumor development: EGFR gene and cancer. *FEBS J.* **277**, 301–308 (2010).
88. Seshacharyulu, P. *et al.* Targeting the EGFR signaling pathway in cancer therapy. *Expert Opin. Ther. Targets* **16**, 15–31 (2012).
89. Caldieri, G., Malabarba, M. G., Di Fiore, P. P. & Sigismund, S. EGFR Trafficking in Physiology and Cancer. in *Progress in molecular and subcellular biology* **57**,

- 235–272 (2018).
90. Stang, E. *et al.* Cbl-dependent Ubiquitination Is Required for Progression of EGF Receptors into Clathrin-coated Pits. *Mol. Biol. Cell* **15**, 3591–3604 (2004).
 91. Sigismund, S. *et al.* Clathrin-mediated internalization is essential for sustained EGFR signaling but dispensable for degradation. *Dev. Cell* **15**, 209–19 (2008).
 92. Sigismund, S. *et al.* Clathrin-independent endocytosis of ubiquitinated cargos. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 2760–5 (2005).
 93. Hawryluk, M. J. *et al.* Epsin 1 is a Polyubiquitin-Selective Clathrin-Associated Sorting Protein. *Traffic* **7**, 262–281 (2006).
 94. Sigismund, S. *et al.* Clathrin-independent endocytosis of ubiquitinated cargos. *Proc. Natl. Acad. Sci.* **102**, 2760–2765 (2005).
 95. Kazazic, M. *et al.* Epsin 1 is involved in recruitment of ubiquitinated EGF receptors into clathrin-coated pits. *Traffic* **10**, 235–45 (2009).
 96. Etienne-Manneville, S. Cdc42 - the centre of polarity. *J. Cell Sci.* **117**, 1291–1300 (2004).
 97. Aguilar, R. C. *et al.* Epsin N-terminal homology domains perform an essential function regulating Cdc42 through binding Cdc42 GTPase-activating proteins. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4116–21 (2006).
 98. Li, L. *et al.* Epsin2 promotes polarity establishment and meiotic division through activating Cdc42 in mouse oocyte. *Oncotarget* **7**, 50927–50936 (2016).
 99. Coon, B. G., Burgner, J., Camonis, J. H. & Aguilar, R. C. The epsin family of endocytic adaptors promotes fibrosarcoma migration and invasion. *J. Biol. Chem.* **285**, 33073–81 (2010).

100. Brady, R. J., Damer, C. K., Heuser, J. E. & O'Halloran, T. J. Regulation of Hip1r by epsin controls the temporal and spatial coupling of actin filaments to clathrin-coated pits. *J. Cell Sci.* **123**, 3652–3661 (2010).
101. Kerbel, R. S. Tumor Angiogenesis. *N. Engl. J. Med.* **358**, 2039–2049 (2008).
102. Lobov, I. B. *et al.* Delta-like ligand 4 (Dll4) is induced by VEGF as a negative regulator of angiogenic sprouting. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3219–24 (2007).
103. Pasula, S. *et al.* Endothelial epsin deficiency decreases tumor growth by enhancing VEGF signaling. *J. Clin. Invest.* **122**, 4424–4438 (2012).
104. Rahman, H. N. A. *et al.* Selective targeting of a novel Epsin-VEGFR2 interaction promotes VEGF-mediated angiogenesis. *Circ. Res.* **118**, 957–969 (2016).
105. Tessneer, K. L. *et al.* Genetic Reduction of Vascular Endothelial Growth Factor Receptor 2 Rescues Aberrant Angiogenesis Caused by Epsin Deficiency. *Arterioscler. Thromb. Vasc. Biol.* **34**, 331–337 (2014).
106. Messa, M. *et al.* Epsin deficiency impairs endocytosis by stalling the actin-dependent invagination of endocytic clathrin-coated pits. *Elife* **3**, e03311 (2014).
107. Wu, H. *et al.* Epsin deficiency promotes lymphangiogenesis through regulation of VEGFR3 degradation in diabetes. *J. Clin. Invest.* **128**, 4025–4043 (2018).
108. Clevers, H. Wnt/ β -Catenin Signaling in Development and Disease. *Cell* **127**, 469–480 (2006).
109. Clevers, H. & Nusse, R. Wnt/ β -Catenin Signaling and Disease. *Cell* **149**, 1192–1205 (2012).
110. Ueno, K. *et al.* Down-regulation of frizzled-7 expression decreases survival,

- invasion and metastatic capabilities of colon cancer cells. *Br. J. Cancer* **101**, 1374–81 (2009).
111. Chang, B. *et al.* Epsin is required for Dishevelled stability and Wnt signalling activation in colon cancer development. *Nat. Commun.* **6**, 6380 (2015).
112. 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.* **107**, 21511–21516 (2010).
113. Spradling, K. D., McDaniel, A. E., Lohi, J. & Pilcher, B. K. Epsin 3 Is a Novel Extracellular Matrix-induced Transcript Specific to Wounded Epithelia. *J. Biol. Chem.* **276**, 29257–29267 (2001).
114. Overstreet, E., Fitch, E. & Fischer, J. A. Fat facets and Liquid facets promote Delta endocytosis and Delta signaling in the signaling cells. *Development* **131**, 5355–66 (2004).
115. Wang, Y., Dai, Z., Sadee, W. & Hancock, W. S. A pharmacoproteomics study of the cancer cell line EKVX using capillary-LC/MS/MS. *Mol. Pharm.* **3**, 566–578 (2006).
116. Hellwig, B. *et al.* Epsin Family Member 3 and Ribosome-Related Genes Are Associated with Late Metastasis in Estrogen Receptor-Positive Breast Cancer and Long-Term Survival in Non-Small Cell Lung Cancer Using a Genome-Wide Identification and Validation Strategy. *PLoS One* **11**, e0167585 (2016).
117. Pawlowski, K. M. *et al.* Comparison of cellular and tissue transcriptional profiles in canine mammary tumor. *J. Physiol. Pharmacol.* **60 Suppl 1**, 85–94 (2009).
118. Tessneer, K. L., Cai, X., Pasula, S., Dong, Y. & Liu, X. NIH Public Access. **2**, 144–150 (2014).

119. Coon, B. G., Drenzo, D. M., Konieczny, S. F. & Aguilar, R. C. Epsins' novel role in cancer cell invasion. *Commun. Integr. Biol.* **4**, 95–7 (2011).
120. Wang, Y. *et al.* Overexpression of Epsin3 enhances migration and invasion of glioma cells by inducing epithelial-mesenchymal transition. *Oncol. Rep.* **40**, 3049–3059 (2018).
121. Mori, J. *et al.* EPSIN 3, A Novel p53 Target, Regulates the Apoptotic Pathway and Gastric Carcinogenesis. *Neoplasia* **19**, 185–195 (2017).
122. Hecht, A., Litterst, C. M., Huber, O. & Kemler, R. Functional characterization of multiple transactivating elements in beta-catenin, some of which interact with the TATA-binding protein in vitro. *J. Biol. Chem.* **274**, 18017–25 (1999).
123. le Duc, Q. *et al.* Vinculin potentiates E-cadherin mechanosensing and is recruited to actin-anchored sites within adherens junctions in a myosin II-dependent manner. *J. Cell Biol.* **189**, 1107–15 (2010).
124. Yonemura, S., Wada, Y., Watanabe, T., Nagafuchi, A. & Shibata, M. α -Catenin as a tension transducer that induces adherens junction development. *Nat. Cell Biol.* **12**, 533–542 (2010).
125. Etienne-Manneville, S. Adherens Junctions During Cell Migration. in *Sub-cellular biochemistry* **60**, 225–249 (2012).
126. Graham, T. A., Weaver, C., Mao, F., Kimelman, D. & Xu, W. Crystal structure of a beta-catenin/Tcf complex. *Cell* **103**, 885–96 (2000).
127. Cold Spring Harbor Laboratory. Press. *Cold Spring Harbor perspectives in biology*. (Cold Spring Harbor Laboratory Press, 2009).
128. Furuse, M. Molecular basis of the core structure of tight junctions. *Cold Spring Harb. Perspect. Biol.* **2**, a002907 (2010).

129. Hartsock, A. & Nelson, W. J. Adherens and tight junctions : Structure , function and connections to the actin cytoskeleton. **1778**, 660–669 (2008).
130. Zhang, J. *et al.* TGF- β -induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci. Signal.* **7**, ra91-ra91 (2014).
131. Slorach, E. M., Chou, J. & Werb, Z. Zeppo1 is a novel metastasis promoter that represses E-cadherin expression and regulates p120-catenin isoform expression and localization. *Genes Dev.* **25**, 471–484 (2011).
132. Buckley, C. D. *et al.* Cell adhesion. The minimal cadherin-catenin complex binds to actin filaments under force. *Science* **346**, 1254211 (2014).
133. Yap, A. S., Niessen, C. M. & Gumbiner, B. M. The juxtamembrane region of the cadherin cytoplasmic tail supports lateral clustering, adhesive strengthening, and interaction with p120ctn. *J. Cell Biol.* **141**, 779–89 (1998).
134. Aberle, H. *et al.* Assembly of the cadherin-catenin complex in vitro with recombinant proteins. *J. Cell Sci.* **107 (Pt 12)**, 3655–63 (1994).
135. Franke, W. W. Discovering the molecular components of intercellular junctions- a historical view. *Cold Spring Harb. Perspect. Biol.* **1**, a003061 (2009).
136. Stahley, S. N. & Kowalczyk, A. P. Desmosomes in acquired disease. *Cell Tissue Res.* **360**, 439–456 (2015).
137. Nielsen, M. S. *et al.* Gap Junctions. in *Comprehensive Physiology* **2**, 1981–2035 (John Wiley & Sons, Inc., 2012).
138. Haase, K., Al-Rekabi, Z. & Pelling, A. E. Mechanical Cues Direct Focal Adhesion Dynamics. in *Progress in molecular biology and translational science* **126**, 103–134 (2014).

139. Eke, I. & Cordes, N. Focal adhesion signaling and therapy resistance in cancer. *Semin. Cancer Biol.* **31**, 65–75 (2015).
140. Walko, G., Castañón, M. J. & Wiche, G. Molecular architecture and function of the hemidesmosome. *Cell Tissue Res.* **360**, 363–378 (2015).
141. Shilova, O. N., Shilov, E. S., Lieber, A. & Deyev, S. M. Disassembling a cancer puzzle: Cell junctions and plasma membrane as targets for anticancer therapy. *J. Control. Release* **286**, 125–136 (2018).
142. Takeichi, M., Hatta, K., Nose, A. & Nagafuchi, A. Identification of a gene family of cadherin cell adhesion molecules. *Cell Differ. Dev.* **25 Suppl**, 91–4 (1988).
143. Pokutta, S., Herrenknecht, K., Kemler, R. & Engel, J. Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. *Eur. J. Biochem.* **223**, 1019–26 (1994).
144. Patel, S. D. *et al.* Type II Cadherin Ectodomain Structures: Implications for Classical Cadherin Specificity. *Cell* **124**, 1255–1268 (2006).
145. Ozawa, M., Hoschützky, H., Herrenknecht, K. & Kemler, R. A possible new adhesive site in the cell-adhesion molecule uvomorulin. *Mech. Dev.* **33**, 49–56 (1990).
146. Vestweber, D. & Kemler, R. Identification of a putative cell adhesion domain of uvomorulin. *EMBO J.* **4**, 3393–8 (1985).
147. Nelson, W. J. & Nusse, R. Convergence of Wnt, beta-catenin, and cadherin pathways. *Science* **303**, 1483–7 (2004).
148. WU, C. *et al.* Interaction between Wnt/ β -catenin pathway and microRNAs regulates epithelial-mesenchymal transition in gastric cancer (Review). *Int. J. Oncol.* **48**, 2236–2246 (2016).

149. Ishiyama, N. & Ikura, M. The Three-Dimensional Structure of the Cadherin–Catenin Complex. in *Sub-cellular biochemistry* **60**, 39–62 (2012).
150. Buckley, C. D. *et al.* The minimal cadherin-catenin complex binds to actin filaments under force. *Science (80-.)*. **346**, 1254211–1254211 (2014).
151. Kourtidis, A., Lu, R., Pence, L. J. & Anastasiadis, P. Z. A central role for cadherin signaling in cancer. *Exp. Cell Res.* **358**, 78–85 (2017).
152. Peifer, M. Cell adhesion and signal transduction: the Armadillo connection. *Trends Cell Biol.* **5**, 224–9 (1995).
153. Cadigan, K. M. & Peifer, M. Wnt signaling from development to disease: insights from model systems. *Cold Spring Harb. Perspect. Biol.* **1**, a002881 (2009).
154. McEwen, A. E., Escobar, D. E. & Gottardi, C. J. Signaling from the Adherens Junction. in *Sub-cellular biochemistry* **60**, 171–196 (2012).
155. Reynolds, A. B., Roesel, D. J., Kanner, S. B. & Parsons, J. T. Transformation-specific tyrosine phosphorylation of a novel cellular protein in chicken cells expressing oncogenic variants of the avian cellular src gene. *Mol. Cell. Biol.* **9**, 629–38 (1989).
156. Reynolds, A. B. *et al.* Identification of a new catenin: the tyrosine kinase substrate p120cas associates with E-cadherin complexes. *Mol. Cell. Biol.* **14**, 8333–42 (1994).
157. Aghib, D. F. & McCrea, P. D. The E-cadherin complex contains the src substrate p120. *Exp. Cell Res.* **218**, 359–69 (1995).
158. Hong, J. Y., Oh, I.-H. & McCrea, P. D. Phosphorylation and isoform use in p120-catenin during development and tumorigenesis. *Biochim. Biophys. Acta - Mol. Cell Res.* **1863**, 102–114 (2016).

159. Keirsebilck, A. *et al.* Molecular Cloning of the Human p120^{ctn}Catenin Gene (CTNND1): Expression of Multiple Alternatively Spliced Isoforms. *Genomics* **50**, 129–146 (1998).
160. Sarrio, D. & Gamallo, C. Cytoplasmic localization of p120^{ctn} and E-cadherin loss characterize lobular breast carcinoma from preinvasive to metastatic lesions. 3272–3283 (2004). doi:10.1038/sj.onc.1207439
161. Castaño, J. *et al.* Specific phosphorylation of p120-catenin regulatory domain differently modulates its binding to RhoA. *Mol. Cell. Biol.* **27**, 1745–57 (2007).
162. Yanagisawa, M. *et al.* A novel interaction between kinesin and p120 modulates p120 localization and function. *J. Biol. Chem.* **279**, 9512–21 (2004).
163. Valls, G. *et al.* Upon Wnt stimulation, Rac1 activation requires Rac1 and Vav2 binding to p120-catenin. *J. Cell Sci.* **129**, 2120–2123 (2016).
164. Hong, J. Y. *et al.* Shared molecular mechanisms regulate multiple catenin proteins: canonical Wnt signals and components modulate p120-catenin isoform-1 and additional p120 subfamily members. *J. Cell Sci.* **123**, 4351–65 (2010).
165. Daniel, J. M. Dancing in and out of the nucleus: p120^(ctn) and the transcription factor Kaiso. *Biochim. Biophys. Acta* **1773**, 59–68 (2007).
166. Chen, X., Kojima, S., Borisy, G. G. & Green, K. J. p120 catenin associates with kinesin and facilitates the transport of cadherin–catenin complexes to intercellular junctions. *J. Cell Biol.* **163**, 547–557 (2003).
167. Roczniak-Ferguson, A. & Reynolds, A. B. Regulation of p120-catenin nucleocytoplasmic shuttling activity. *J. Cell Sci.* **116**, 4201–12 (2003).
168. Abe, K. & Takeichi, M. EPLIN mediates linkage of the cadherin catenin complex

- to F-actin and stabilizes the circumferential actin belt. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13–9 (2008).
169. Le Bras, G. F., Taubenslag, K. J. & Andl, C. D. The regulation of cell-cell adhesion during epithelial-mesenchymal transition, motility and tumor progression. *Cell Adh. Migr.* **6**, 365–373 (2012).
170. Peinado, H., Olmeda, D. & Cano, A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat. Rev. Cancer* **7**, 415–28 (2007).
171. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **139**, 871–890 (2009).
172. Kowalczyk, A. P. & Nanes, B. A. Adherens Junction Turnover: Regulating Adhesion Through Cadherin Endocytosis, Degradation, and Recycling. in *Subcellular biochemistry* **60**, 197–222 (2012).
173. Cadwell, C. M., Su, W. & Kowalczyk, A. P. Cadherin tales: Regulation of cadherin function by endocytic membrane trafficking. *Traffic* **17**, 1262–1271 (2016).
174. Ishiyama, N. *et al.* Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion. *Cell* **141**, 117–28 (2010).
175. Perez-Moreno, M. & Fuchs, E. Guilt by association: What p120-catenin has to hide: Figure 1. *J. Cell Biol.* **199**, 211–214 (2012).
176. Miyashita, Y. & Ozawa, M. Increased internalization of p120-uncoupled E-cadherin and a requirement for a dileucine motif in the cytoplasmic domain for endocytosis of the protein. *J. Biol. Chem.* **282**, 11540–8 (2007).
177. Nanes, B. A. *et al.* p120-catenin binding masks an endocytic signal conserved in classical cadherins. *J. Cell Biol.* **199**, 365–380 (2012).

178. Fujita, Y. *et al.* Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. *Nat. Cell Biol.* **4**, 222–231 (2002).
179. Mansouri, M., Rose, P. P., Moses, A. V & Früh, K. Remodeling of endothelial adherens junctions by Kaposi's sarcoma-associated herpesvirus. *J. Virol.* **82**, 9615–28 (2008).
180. Bryant, D. M. *et al.* EGF induces macropinocytosis and SNX1-modulated recycling of E-cadherin. *J. Cell Sci.* **120**, 1818–1828 (2007).
181. Kamei, T. *et al.* Coendocytosis of cadherin and c-Met coupled to disruption of cell-cell adhesion in MDCK cells--regulation by Rho, Rac and Rab small G proteins. *Oncogene* **18**, 6776–84 (1999).
182. Bryant, D. M., Wylie, F. G. & Stow, J. L. Regulation of Endocytosis, Nuclear Translocation, and Signaling of Fibroblast Growth Factor Receptor 1 by E-Cadherin. *Mol. Biol. Cell* **16**, 14–23 (2005).
183. Janda, E. *et al.* Raf plus TGF β -dependent EMT is initiated by endocytosis and lysosomal degradation of E-cadherin. *Oncogene* **25**, 7117–7130 (2006).
184. Nieto, M. A., Huang, R. Y. Y. J., Jackson, R. A. A. & Thiery, J. P. P. Emt: 2016. *Cell* **166**, 21–45 (2016).
185. Roxanis, I. Occurrence and significance of epithelial-mesenchymal transition in breast cancer. *J. Clin. Pathol.* **66**, 517–21 (2013).
186. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **139**, 871–890 (2009).
187. Ye, X. *et al.* Upholding a role for EMT in breast cancer metastasis. *Nature* **547**, E1–E3 (2017).
188. Jolly, M. K. Implications of the Hybrid Epithelial/Mesenchymal Phenotype in

- Metastasis. *Front. Oncol.* **5**, 1–19 (2015).
189. Grigore, A., Jolly, M., Jia, D., Farach-Carson, M. & Levine, H. Tumor Budding: The Name is EMT. Partial EMT. *J. Clin. Med.* **5**, 51 (2016).
190. Singh, A. & Settleman, J. EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* **29**, 4741–51 (2010).
191. Micalizzi, D. S., Farabaugh, S. M. & Ford, H. L. Epithelial-mesenchymal transition in cancer: Parallels between normal development and tumor progression. *J. Mammary Gland Biol. Neoplasia* **15**, 117–134 (2010).
192. Peinado, H., Portillo, F. & Cano, A. Transcriptional regulation of cadherins during development and carcinogenesis. *Int. J. Dev. Biol.* **48**, 365–375 (2004).
193. Ikenouchi, J., Matsuda, M., Furuse, M. & Tsukita, S. Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. *J. Cell Sci.* **116**, 1959–67 (2003).
194. Grande, M. T. *et al.* Snail1-induced partial epithelial-to-mesenchymal transition drives renal fibrosis in mice and can be targeted to reverse established disease. *Nat. Med.* **21**, 989–97 (2015).
195. Serrano-Gomez, S. J., Maziveyi, M. & Alahari, S. K. Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. *Mol. Cancer* **15**, 18 (2016).
196. Gregory, P. A., Bracken, C. P., Bert, A. G. & Goodall, G. J. MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle* **7**, 3112–3117 (2008).
197. Aiello, N. M. *et al.* EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. *Dev. Cell* **45**, 681–695.e4 (2018).
198. Kim, K., Lu, Z. & Hay, E. D. Direct evidence for a role of beta-catenin/LEF-1

- signaling pathway in induction of EMT. *Cell Biol. Int.* **26**, 463–76 (2002).
199. Leroy, P. & Mostov, K. E. Slug is required for cell survival during partial epithelial-mesenchymal transition of HGF-induced tubulogenesis. *Mol. Biol. Cell* **18**, 1943–52 (2007).
200. Lim, J. & Thiery, J. P. Epithelial-mesenchymal transitions: insights from development. *Development* **139**, 3471–86 (2012).
201. Heldin, C.-H., Vanlandewijck, M. & Moustakas, A. Regulation of EMT by TGF β in cancer. *FEBS Lett.* **586**, 1959–1970 (2012).
202. Schmierer, B. & Hill, C. S. TGF β -SMAD signal transduction: molecular specificity and functional flexibility. *Nat. Rev. Mol. Cell Biol.* **8**, 970–82 (2007).
203. Shi, Y. & Massagué, J. Mechanisms of TGF- β signaling from cell membrane to the nucleus. *Cell* **113**, 685–700 (2003).
204. Vander Ark, A., Cao, J. & Li, X. TGF- β receptors: In and beyond TGF- β signaling. *Cell. Signal.* **52**, 112–120 (2018).
205. Thuault, S. *et al.* Transforming growth factor- β employs HMGA2 to elicit epithelial-mesenchymal transition. *J. Cell Biol.* **174**, 175–83 (2006).
206. Tian, X. J., Zhang, H. & Xing, J. Coupled reversible and irreversible bistable switches underlying TGF β -induced epithelial to mesenchymal transition. *Biophys. J.* **105**, 1079–1089 (2013).
207. Thakur, N. *et al.* TGF β -induced invasion of prostate cancer cells is promoted by c-Jun-dependent transcriptional activation of Snail1. *Cell Cycle* **13**, 2400–2414 (2014).
208. Yamashita, M. *et al.* TRAF6 Mediates Smad-Independent Activation of JNK and p38 by TGF- β . *Mol. Cell* **31**, 918–924 (2008).

209. Ozdamar, B. *et al.* Regulation of the Polarity Protein Par6 by TGF Receptors Controls Epithelial Cell Plasticity. *Science (80-.)*. **307**, 1603–1609 (2005).
210. Cicchini, C. *et al.* TGFbeta-induced EMT requires focal adhesion kinase (FAK) signaling. *Exp. Cell Res.* **314**, 143–52 (2008).
211. Coluccia, A. M. L. *et al.* SKI-606 Decreases Growth and Motility of Colorectal Cancer Cells by Preventing pp60(c-Src)–Dependent Tyrosine Phosphorylation of β -Catenin and Its Nuclear Signaling. *Cancer Res.* **66**, 2279–2286 (2006).
212. Siegel, P. M., Shu, W., Cardiff, R. D., Muller, W. J. & Massagué, J. Transforming growth factor β signaling impairs Neu-induced mammary tumorigenesis while promoting pulmonary metastasis. *Proc. Natl. Acad. Sci.* **100**, 8430–8435 (2003).
213. Ciruna, B. & Rossant, J. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Dev. Cell* **1**, 37–49 (2001).
214. Oda, H., Tsukita, S. & Takeichi, M. Dynamic behavior of the cadherin-based cell-cell adhesion system during *Drosophila* gastrulation. *Dev. Biol.* **203**, 435–50 (1998).
215. Cheung, M. *et al.* The transcriptional control of trunk neural crest induction, survival, and delamination. *Dev. Cell* **8**, 179–92 (2005).
216. Teddy, J. M. & Kulesa, P. M. In vivo evidence for short- and long-range cell communication in cranial neural crest cells. *Development* **131**, 6141–51 (2004).
217. Azhar, M. *et al.* Transforming growth factor beta in cardiovascular development and function. *Cytokine Growth Factor Rev.* **14**, 391–407 (2003).
218. Macias, H. & Hinck, L. Mammary gland development. *Wiley Interdiscip. Rev. Dev. Biol.* **1**, 533–57 (2012).

219. McBryan, J. & Howlin, J. Pubertal Mammary Gland Development: Elucidation of In Vivo Morphogenesis Using Murine Models. *Methods Mol. Biol.* **1501**, 77–114 (2017).
220. Lu, P., Sternlicht, M. D. & Werb, Z. Comparative Mechanisms of Branching Morphogenesis in Diverse Systems. *J. Mammary Gland Biol. Neoplasia* **11**, 213–228 (2006).
221. Propper, A. Y. Wandering epithelial cells in the rabbit embryo milk line. A preliminary scanning electron microscope study. *Dev. Biol.* **67**, 225–31 (1978).
222. Williams, J. M. & Daniel, C. W. Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. *Dev. Biol.* **97**, 274–90 (1983).
223. Ruan, W. & Kleinberg, D. L. Insulin-Like Growth Factor I Is Essential for Terminal End Bud Formation and Ductal Morphogenesis during Mammary Development¹. *Endocrinology* **140**, 5075–5081 (1999).
224. Joshi, P. A. *et al.* Progesterone induces adult mammary stem cell expansion. *Nature* **465**, 803–807 (2010).
225. Horseman, N. D. *et al.* Defective mammopoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. *EMBO J.* **16**, 6926–6935 (1997).
226. Ormandy, C. J. *et al.* Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. *Genes Dev.* **11**, 167–78 (1997).
227. Ewald, A. J., Brenot, A., Duong, M., Chan, B. S. & Werb, Z. Collective Epithelial Migration and Cell Rearrangements Drive Mammary Branching Morphogenesis. *Dev. Cell* **14**, 570–581 (2008).

228. Sternlicht, M. D. *et al.* The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. *Cell* **98**, 137–46 (1999).
229. Lochter, A. *et al.* Matrix metalloproteinase stromelysin-1 triggers a cascade of molecular alterations that leads to stable epithelial-to-mesenchymal conversion and a premalignant phenotype in mammary epithelial cells. *J. Cell Biol.* **139**, 1861–72 (1997).
230. Wechselberger, C. *et al.* Cripto-1 Enhances Migration and Branching Morphogenesis of Mouse Mammary Epithelial Cells. *Exp. Cell Res.* **266**, 95–105 (2001).
231. Pierce, D., Johnson, M., ... Y. M.-G. & & 1993, undefined. Inhibition of mammary duct development but not alveolar outgrowth during pregnancy in transgenic mice expressing active TGF-beta 1. *genesdev.cshlp.org*
232. Ewan, K. B. *et al.* Latent transforming growth factor-beta activation in mammary gland: regulation by ovarian hormones affects ductal and alveolar proliferation. *Am. J. Pathol.* **160**, 2081–93 (2002).
233. Kouros-Mehr, H. & Werb, Z. Candidate regulators of mammary branching morphogenesis identified by genome-wide transcript analysis. *Dev. Dyn.* **235**, 3404–3412 (2006).
234. Ingman, W. V. & Robertson, S. A. Mammary Gland Development in Transforming Growth Factor Beta1 Null Mutant Mice: Systemic and Epithelial Effects¹. *Biol. Reprod.* **79**, 711–717 (2008).
235. Cheng, N. *et al.* Loss of TGF- β type II receptor in fibroblasts promotes mammary carcinoma growth and invasion through upregulation of TGF- α -, MSP- and HGF-mediated signaling networks. *Oncogene* **24**, 5053–5068 (2005).

236. Joseph, H., Gorska, A. E., Sohn, P., Moses, H. L. & Serra, R. Overexpression of a Kinase-deficient Transforming Growth Factor- β Type II Receptor in Mouse Mammary Stroma Results in Increased Epithelial Branching. *Mol. Biol. Cell* **10**, 1221–1234 (1999).
237. Rastaldi, M. P. *et al.* Epithelial-mesenchymal transition of tubular epithelial cells in human renal biopsies. *Kidney Int.* **62**, 137–46 (2002).
238. Zeisberg, M. *et al.* Fibroblasts Derive from Hepatocytes in Liver Fibrosis via Epithelial to Mesenchymal Transition. *J. Biol. Chem.* **282**, 23337–23347 (2007).
239. Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* **119**, 1420–1428 (2009).
240. Arnoux, V., Nassour, M., L’Helgoualc’h, A., Hipskind, R. A. & Savagner, P. Erk5 controls Slug expression and keratinocyte activation during wound healing. *Mol. Biol. Cell* **19**, 4738–49 (2008).
241. Hudson, L. G. *et al.* Cutaneous wound reepithelialization is compromised in mice lacking functional Slug (Snai2). *J. Dermatol. Sci.* **56**, 19–26 (2009).
242. Yang, F. *et al.* Essential role for Smad3 in angiotensin II-induced tubular epithelial-mesenchymal transition. *J. Pathol.* n/a-n/a (2010).
doi:10.1002/path.2721
243. Boutet, A. *et al.* Snail activation disrupts tissue homeostasis and induces fibrosis in the adult kidney. *EMBO J.* **25**, 5603–13 (2006).
244. Sarrió, D. *et al.* Cytoplasmic localization of p120ctn and E-cadherin loss characterize lobular breast carcinoma from preinvasive to metastatic lesions. *Oncogene* **23**, 3272–83 (2004).
245. Birchmeier, W. & Behrens, J. Cadherin expression in carcinomas: role in the

- formation of cell junctions and the prevention of invasiveness. *Biochim. Biophys. Acta* **1198**, 11–26 (1994).
246. Pastushenko, I. *et al.* Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463–468 (2018).
247. Reichert, M. *et al.* Regulation of Epithelial Plasticity Determines Metastatic Organotropism in Pancreatic Cancer. *Dev. Cell* **45**, 696–711.e8 (2018).
248. Dai, X. *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* **5**, 2929–43 (2015).
249. Hon, J. D. C. *et al.* Breast cancer molecular subtypes: from TNBC to QNBC. *Am. J. Cancer Res.* **6**, 1864–1872 (2016).
250. Cowell, C. F. *et al.* Progression from ductal carcinoma *in situ* to invasive breast cancer: Revisited. *Mol. Oncol.* **7**, 859–869 (2013).
251. Knudsen, E. S. *et al.* Progression of ductal carcinoma in situ to invasive breast cancer is associated with gene expression programs of EMT and myoepithelia. *Breast Cancer Res. Treat.* **133**, 1009–24 (2012).
252. Burstein, H. J., Polyak, K., Wong, J. S., Lester, S. C. & Kaelin, C. M. Ductal carcinoma in situ of the breast. *N. Engl. J. Med.* **350**, 1430–41 (2004).
253. Lopez-Garcia, M. A., Geyer, F. C., Lacroix-Triki, M., Marchiό, C. & Reis-Filho, J. S. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology* **57**, 171–92 (2010).
254. O'Connell, P. *et al.* Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J. Natl. Cancer Inst.* **90**, 697–703 (1998).
255. Giardina, C. *et al.* Pure ductal carcinoma in situ and in situ component of ductal invasive carcinoma of the breast. A preliminary morphometric study. *J. Exp.*

- Clin. Cancer Res.* **22**, 279–88 (2003).
256. Ottesen, G. L. Carcinoma in situ of the female breast. A clinico-pathological, immunohistological, and DNA ploidy study. *APMIS. Suppl.* 1–67 (2003).
257. Zingarello, A., Mazouni, C., Rivera, S., Mokdad-Adi, M. & Pistilli, B. Prognostic assessment and systemic treatments of invasive local relapses of hormone receptor-positive breast cancer. *The Breast* **35**, 162–168 (2017).
258. Turashvili, G., Bouchalova, K., Bouchal, J. & Kolar, Z. Expression of E-cadherin and c-erbB-2/HER-2/neu oncoprotein in high-grade breast cancer. *Cesk. Patol.* **43**, 87–92 (2007).
259. Radisky, E. S. & Radisky, D. C. Matrix Metalloproteinase-Induced Epithelial-Mesenchymal Transition in Breast Cancer. *J. Mammary Gland Biol. Neoplasia* **15**, 201–212 (2010).
260. Huber, M. A., Kraut, N. & Beug, H. Molecular requirements for epithelial-mesenchymal transition during tumor progression. *Curr. Opin. Cell Biol.* **17**, 548–558 (2005).
261. Savagner, P., Yamada, K. M. & Thiery, J. P. The zinc-finger protein slug causes desmosome dissociation, an initial and necessary step for growth factor-induced epithelial-mesenchymal transition. *J. Cell Biol.* **137**, 1403–19 (1997).
262. Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).
263. Klockenbusch, C. & Kast, J. Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin beta1. *J. Biomed. Biotechnol.* **2010**, 927585 (2010).
264. Ramirez, A. *et al.* A keratin K5Cre transgenic line appropriate for tissue-specific

- or generalized Cre-mediated recombination. *Genesis* **39**, 52–7 (2004).
265. Cicalese, A. *et al.* The Tumor Suppressor p53 Regulates Polarity of Self-Renewing Divisions in Mammary Stem Cells. *Cell* **138**, 1083–1095 (2009).
266. Aiello, N. M. *et al.* Upholding a role for EMT in pancreatic cancer metastasis. *Nature* **547**, E7–E8 (2017).
267. Zheng, H.-C. The molecular mechanisms of chemoresistance in cancers. *Oncotarget* (2017). doi:10.18632/oncotarget.19048
268. Voon, D. C., Huang, R. Y., Jackson, R. A. & Thiery, J. P. The EMT spectrum and therapeutic opportunities. *Mol. Oncol.* **11**, 878–891 (2017).
269. Shibue, T. & Weinberg, R. A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* **14**, 611–629 (2017).
270. Kourtidis, A., Lu, R., Pence, L. J. & Anastasiadis, P. Z. A central role for cadherin signaling in cancer. *Exp. Cell Res.* **358**, 78–85 (2017).
271. Clinical Development Success Rates. (2006).
272. Matsuzaki, T. *et al.* Regulation of endocytosis of activin type II receptors by a novel PDZ protein through Ral/Ral-binding protein 1-dependent pathway. *J. Biol. Chem.* **277**, 19008–18 (2002).
273. Guo, Z. *et al.* E-cadherin interactome complexity and robustness resolved by quantitative proteomics. *Sci. Signal.* **7**, rs7-rs7 (2014).
274. Guy, C. T. *et al.* Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10578–82 (1992).
275. Ursini-Siegel, J., Schade, B., Cardiff, R. D. & Muller, W. J. Insights from transgenic

- mouse models of ERBB2-induced breast cancer. *Nat. Rev. Cancer* **7**, 389–397 (2007).
276. Petrova, Y. I., Spano, M. M. & Gumbiner, B. M. Conformational epitopes at cadherin calcium-binding sites and p120-catenin phosphorylation regulate cell adhesion. *Mol. Biol. Cell* **23**, 2092–2108 (2012).
277. Cozzolino, M. *et al.* p120 Catenin is required for growth factor-dependent cell motility and scattering in epithelial cells. *Mol. Biol. Cell* **14**, 1964–77 (2003).
278. Aono, S., Nakagawa, S., Reynolds, A. B. & Takeichi, M. p120(ctn) Acts as an inhibitory regulator of cadherin function in colon carcinoma cells. *J. Cell Biol.* **145**, 551–562 (1999).
279. Dunphy, K. A., Tao, L. & Jerry, D. J. Mammary Epithelial Transplant Procedure. *J. Vis. Exp.* (2010). doi:10.3791/1849
280. Rolland, Y. *et al.* The CDC42-Interacting Protein 4 Controls Epithelial Cell Cohesion and Tumor Dissemination. *Dev. Cell* **30**, 553–568 (2014).
281. Shevchenko, A., Tomas, H., Havli\[\sbreve], J., Olsen, J. V & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860 (2007).
282. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
283. Stanko, J. P., Easterling, M. R. & Fenton, S. E. Application of Sholl analysis to quantify changes in growth and development in rat mammary gland whole mounts. *Reprod. Toxicol.* **54**, 129–35 (2015).
284. Kroner, A. *et al.* TNF and increased intracellular iron alter macrophage

polarization to a detrimental M1 phenotype in the injured spinal cord. *Neuron* **83**, 1098–116 (2014).

ACKNOWLEDGEMENTS

The conclusion of this PhD has been possible thanks to the help of many people.

First of all, I want to thank my PhD supervisor, Prof. Di Fiore Pier Paolo, because four years ago he choose me and gave me the possibility to work in this exciting environment. But, most importantly, he stimulated me to be creative and persistent.

I want also to thank my supervisor, Dr. Sigismund Sara, for her daily help and for all the scientific discussions. She helped me to go beyond bad results and never give up.

An important thanks goes also to Dr. Vecchi Manuela and Dr. Iavarone Claudia, which started the project that I had the opportunity to work on. I want also to thank the molecular pathology unit, Dr. Bertalot Giovanni, Dr. Luise Chiara and Dr. Lo Iodice Giovanna, because they helped me a lot in the analyses of murine mammary glands. I thank Prof. Pece Salvatore, because he noticed me for first. I acknowledge Dr. Confalonieri Stefano for his help the data analyses.

I am very grateful for the suggestions and feedback of my external advisor Dr. Lamaze Christophe, from the Curie Institute, and my internal advisor Prof. Scita Giorgio, from IFOM. I also acknowledge the help of Rosalind Gunby for critically reading my thesis.

I want to thank Irene Schiano Lomoriello, because we shared and collaborated in this project along these years trying to give our best. I want also to thank all my colleagues: Alexia for the biosketch; Chiara for the M&M's; Elisa for the discovery of Prince; Giusi for her grandmaternal advices; Stefano for the fantacalcio victories. But I need to acknowledge also all the people that helped me during these years, both scientifically and not: Simone for all the hours

ACKNOWLEDGMENTS

spent to discuss about everything; Matteo for the nicknames; Elisa and Ambra because we shared together a lot of incredible experiences. I want to acknowledge also people that I found here and contributed to turn out exciting this years: Valentina, Corey, Valentina, Giuseppe, Elena, Carlos, Marco.

Ringrazio la mia famiglia. Loro sanno la difficoltà di aver dovuto lasciare Salerno per la seconda volta. A mia madre e mio padre, si sono sempre fidati di me. A mia nonna, sempre presente nella mia vita. A mio fratello, per me rappresenta un esempio ed un modello.

Saluto mio nonno, che durante questi quattro anni è mancato.

Infine, all'amore della mia vita. A Carmen.

Ricett' 'o pappece vicin 'a noce: "Ramm 'o tiemp ca te spertose."