The mesenchymal status of metastatic cancer cells promotes a stromal crosstalk leading to epithelial re-acquisition and metastatic colonisation

Yaiza del Pozo Martín

University College London and
Cancer Research UK - London Research Institute
PhD Supervisor: Ilaria Malanchi

A thesis submitted for the degree of
Doctor of Philosophy
University College London

September 2015

Declaration

I Yaiza del Pozo Martín confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Solid epithelial tumours are complex structures in which the associated stroma supports cancer cells (Quail and Joyce, 2013). During metastatic progression, cancer cells disseminate from their tissue of origin and recapitulate the tumour structure at distant organs, including the stromal compartment. Metastasis initiating cells (MICs) are functionally discriminated among the bulk of cancer cells for their high ability to establish metastasis (Malanchi et al., 2012, Baccelli et al., 2013). Additionally, efficient metastasis requires the expression of specific molecules within the local microenvironment (Oskarsson et al., 2014). Thus, a favourable microenvironment or niche is a crucial early step in metastatic progression. However what features of MICs mediate metastatic niche activation is poorly characterised.

One strategy adopted by metastatic cells to disseminate from primary tumours is the activation of the developmental programme epithelial-to-mesenchymal transition (EMT). However, EMT is a reversible programme that needs to be inhibited at the target site for tumour cells to re-acquire epithelial characteristics compatible with metastatic outgrowth (Nieto, 2013). To successfully metastasise cancer cells need to retain self-renewal and growth properties through epithelial plasticity. This implies that during metastasis 'stemness' should not be strictly coupled to EMT as previously suggested (Mani et al., 2008). To date, both the potential advantage of disseminated cancer cells mesenchymal status and the source of their epithelial plasticity at the metastatic site remain unknown.

In this thesis we use metastatic breast cancer models to elucidate the enhanced niche-induction ability of mesenchymal MICs, its relationship to EMT and the source of its epithelial modulation during metastatic colonisation. Importantly, we identify THBS2 as a novel effector linked to the EMT status of cancer cells that enhances stromal niche activation. Subsequently, the newly activated stroma triggers cancer cell BMP-dependent re-epithelialisation promoting metastatic outgrowth. Thereby, we describe a temporally controlled metastatic colonisation where the EMT status of cancer cells promotes its own inhibition via a cancer cell-stromal crosstalk that initially enhances metastatic niche formation, and ultimately favours a cancer cell proliferative state compatible with metastatic outgrowth.

Acknowledgements

First of all, I would like to thank Ilaria for being such a wonderful supervisor. Her vital guidance has allowed me to cultivate my scientific thinking, and her inspiring dedication and enthusiasm helped me to push myself and get the most out of my PhD. Above all, thanks for being an excellent person - always supportive and encouraging.

Next, I want to thank all the present and past members of the Malanchi Lab – Stefanie, Rob, Luigi, Misa, Laurie, Sathya, Mathilde, Anna, Ana and Freddie - for sharing extraordinary moments and making the lab a stimulating and yet fun place to work. A special thanks to Luigi for giving me always the good advice scientifically - I would still be cloning without your help - and personally for always listening to my grumbles. To Rob, thanks for that contagious positive attitude you always show. To Stefanie, my PhD companion from the first until these last thesis-writing days, only you can truly understand the experience of this PhD; specially thanks for your encouraging excitement about science and all those in lab late-night talking sessions.

A big thanks to the Sahai Lab for their constant help and advice, and their insightful scientific and non-scientific discussions. Particularly I am more than grateful to Erik for being a great mentor always willing to help and give critical advice. A special thanks to Danielle and Fernando for their crucial scientific contributions to this thesis. Thanks also to the Hill Lab, particularly Caroline and Ana for their valuable scientific discussions and insightful suggestions. I am also grateful to my thesis committee Axel Behrens and Julian Downward for their valuable input, support and guidance throughout my PhD.

Then, I want to thank all members of the FACS Lab, Biological Resources Unit and Histopathology Unit for sharing their expertise, teaching me the basis of their fundamental work and always being happy to help - even at indecent (non)-working hours! This thesis would not have been possible without their invaluable technical support.

To my friends, the ones left behind that are far but always there when I need them – specially Aroa, Azahara, Carmen, Pina, Marina and Manu - and the new ones that came my way during this ‘PhD adventure’ making it indeed worth living – particularly Martina, Nil and Rute for sharing unforgettable moments together inside and outside the LRI, Mauro for being always there willing to listen and give me a hug, and Carl for bringing all his music creating a perfect soundtrack for this final stage of the PhD.

To my family, all and each of you that made me the person I am today, thanks for your unconditional support and love since always; mainly to my mother Cristina and my brother Miguel, os quiero mucho.

To Alfonso for being the best partner I could have by my side, thanks for being always encouraging, optimistic and understanding, ‘oma et’.

Table of Contents

Table of figures

List of tables

Abbreviations

β1BA	Integrin β1 blocking antibody
ALK	Anaplastic lymphoma kinase
AML	Acute myeloid leukemia
ANGPTL4	Angiopoietin-like protein 4
BMP	Bone morphogenic proteins
BMDC	Bone marrow-derived cell
CAF	Cancer-associated fibroblast
CCL	C-C motif chemokine ligand
CCM	CAF conditioned media
CIC	Cancer-initiating cell
CSC	Cancer stem cell
CpG	C-phosphate-G
CTC	Circulating tumour cell
CTC-cluster	Circulating tumour cell cluster
CXCL10	C-X-C motif chemokine 10
CYR61	Cysteine-rich angiogenic inducer 61
DMSF	Distant metastasis free survival
DNMT3A	DNA methyl-transferase 3A
ECM	Extracellular matrix
EGF	Epidermal growth factor
EGFP	Enhanced green fluorescent protein
ELISA	Enzyme-linked immune-sorbent assay
EMT	Epithelial to mesenchymal transition
EMT-TF	Epithelial to mesenchymal transition transcription factor
EPCAM	Epithelial cell adhesion protein
ER	Estrogen receptor
ESC	Embryonic stem cell
FACS	Fluorescence activated cell sorting
FAP	Fibroblast activation protein
FGF	Fibroblast growth factor
GFP	Green fluorescent protein

GEMM	Genetically engineered mouse model
GO	Gene ontology
GSEA	Gene set enrichment analysis
GSK3β	Glycogen synthase kinase-3β
HGF	Hepatocyte growth factor
HNSCC	Head and neck squamous cell carcinoma
HSC	Hematopoietic stem cell
ID1	DNA binding protein inhibitor 1
ITGA	Integrin α
ITGB	Integrin β
IV	Intra-venous injection
K14	Keratin 14
LEF	Lymphoid enhancer-binding factor 1
LIN	Lineage negative
LOX	Lysyl oxidase
LTR	Long terminal repeat (lentiviral sequence)
MAPK	Mitogen activation protein kinase
MDA231	MDA-MB-231
MDSC	Myeloid-derived suppressor cell
MET	Mesenchymal to epithelial transition
MGMT	O-6-methylguanin-DNA methyl-transferase
MIC	Metastasis-initiating cell
MLH1	MutL homolog 1
MMP	Metalloprotease
MMTV-PyMT	Mouse mammary tumour virus – polyoma middle T antigen
MSC	Mesenchymal stem cell
NF	Normal lung primary fibroblasts
NFkB	Nuclear factor κappaB
NK	Natural killer cell
NLF3	Normal lung fibroblast 3
PDGF	Platelet-derived growth factor
PGK	Phosphoglycerate kinase
PI3K	Phosphoinositide 3-kinase
PR	Progesterone receptor

POSTN	Periostin
PTEN	Phosphatase and tensin homolog
PCR	Polymerase chain reaction
SMA	Smooth muscle actin
SSCC	Skin squamous cell carcinoma
TAM	Tumour-associated macrophage
TCF	T cell factor
TGFβ	Transforming growth factor β
TGFβ-RI	Transforming growth factor β receptor I
TIC	Tumour-initiating cell
THBS1	Thrombospondin-1
THBS2	Thrombospondin-2
TNC	Tenascin C
TP53	Tumour suppressor 53 gene
T-reg	T-regulatory cell
VEGF	Vascular endothelial growth factor
WGS	Whole genome sequencing
YAP	Yes-associated protein

Chapter 1. Introduction

1.1 Tumours as complex ecosystems

Originally, tumours were considered homogeneous masses of aberrant cell growth but in the last decades knowledge has built up on the current view of tumours as highly dynamic heterogeneous structures. Tumour cell heterogeneity is driven by both, genetic determinants that generate clonal evolution among the cancer cell mass through new mutations, and epigenetic modulations that generate further diversity and functionally distinct populations among those clones. Adding complexity to the intricate tumour cell mass, it is now acknowledged that cancer cell heterogeneity is highly modulated by the tumour associated microenvironment, a compendium of different tissue recruited non-cancer cells and extracellular components (Quail and Joyce, 2013). The dynamic relationship between tumour cells and the microenvironment throughout tumour progression creates an extraordinarily complex structure, where the crosstalk between the different components modulates the function of the tumour as a whole. Therefore, by functioning as a complex ecosystem, the intrinsic tumour cell properties are reinforced, ultimately leading to therapy failure as shown by recent reports (Hirata et al., 2015, Junttila and de Sauvage, 2013). In light of these findings, it has now become clear that in order to develop more effective therapies and avoid disease relapse, tumour cell heterogeneity must be studied in the context of its microenvironment.

1.1.1 Tumour cell heterogeneity

1.1.1.1 Genetic origin of cancer

Carcinogenesis is complex process that depends on multiple determinants, but it ultimately has a genetic origin. A founding principle in cancer research is that tumour initiation is driven from the sequential acquisition of genetic mutations, which leads to malignant transformation of normal human cells that acquire sustained chronic proliferation (Bruce Alberts, 2002). One of the first studies

published supporting this view analysed the newly generated mutations across different stages of colorectal carcinoma showing that genetic alterations can determine phenotypic changes (Vogelstein et al., 1988). Mutations in the cell genome occur spontaneously during DNA replication or can be triggered by carcinogens. Most mutations are repaired by the DNA repair machinery of the cell (Hanahan and Weinberg, 2000), but a very low frequency are missed by these repair mechanisms. The majority of mutations occurring in cells are innocuous and persist in the genome without major consequences, or are seriously deleterious directly inducing programmed cell death (apoptosis). Importantly, only multiple mutations in potentially oncogenic genes will generate a malignant cell of origin leading to cancer. Oncogenic genes are those that encode proteins with the potential of sustaining aberrant cell division directly by activating mitogenic signalling pathways, or indirectly inhibiting apoptosis (Hanahan and Weinberg, 2000).

Cell proliferation within a normal tissue is tightly regulated by the presence of growth factors and anti-growth signals maintaining tissue homeostasis. In this context, malignant tumour cells possess oncogenic mutations that make them unresponsive to anti-growth and pro-apoptotic signals obtaining a limitless replicative potential (Hanahan and Weinberg, 2011). Moreover, to circumvent the need of growth factor stimulation to proliferate, cancer cells acquire mutations that constitutively activate signalling pathways downstream growth factor receptors. For instance, 40% of human melanomas contain activating mutations affecting the B-Raf protein, which results in constitutive activation of the mitogen-activated protein (MAP)-kinase pathway and cell proliferation (Figure 1.4.B) (Davies and Samuels, 2010). Similarly, mutations in the catalytic subunit of phosphoinositide 3-kinase (PI3K), which hyperactivates the Akt signal transducer enhancing proliferation are usually detected in different tumour types (Figure 1.4.B) (Jiang and Liu, 2009, Yuan and Cantley, 2008). Tumour cells employ other indirect mechanisms that do not activate mitogenic pathways but disrupt negative feedback loops that normally attenuate proliferation. The prototypic example are the Ras oncoproteins, where mutations affecting the Ras GTPases transform their transitory activity in permanent leading to the sustained activation of multiple genes involved in proliferation and survival (Figure 1.4.B) (Buday and Downward, 2008). Another important example of proteins mediating negative-feedback loops that normally

regulate proliferation is the phosphatase PTEN. PTEN deactivates PI3K by degrading its ligand; therefore loss of function mutations in PTEN leads to PI3K ligand accumulation and hyperactivation of the pathway ultimately boosting proliferation (Jiang and Liu, 2009, Yuan and Cantley, 2008).

An additional strategy driving carcinogenesis is the inactivation of tumour suppressor proteins to bypass programmes that negatively regulate proliferation or activate cell death. The most common examples of tumour suppressor genes are the RB (retinoblastoma-associated) and TP53. The RB protein integrates several extracellular and intracellular signals deciding if the cell cycle should progress or stop (Burkhart and Sage, 2008). Therefore, inactivating mutations in the RB protein lead to the loss of function of this gatekeeper of the cell cycle, and permit unlimited proliferation. Likewise RB, under physiological conditions p53 integrates stress and abnormality signals from intracellular sensors and has the ability to halt the cell cycle until harmful conditions are normalised. When irreparable cell damage is detected, p53 triggers apoptosis to eliminate the aberrant cell. Therefore, inactivating mutations or deletions in TP53 allow tumour cells to proliferate under severe stress conditions and carrying ample DNA damage (Pflaum et al., 2014).
Any newly arisen hypermutant tumour cell is at risk of being lost until it acquires the first selectively advantageous mutation, also called driver mutation. These driver mutations, such as inactivating TP53 mutations, allow cell division without DNA repair and cell cycle checkpoints ultimately leading to genomic instability. Subsequently, genomic instability is carried along as a hitchhiking phenotype further propagating diversity through the accumulation of additional mutations and chromosomal rearrangements (Bakhoum and Swanton, 2014). One common genomic instability example in many tumour types is the loss of the protective chromosomal telomeric DNA, which leads to karyotypic diversification via amplification and deletion of chromosomal fragments in tumour cells (Artandi and DePinho, 2010). In fully developed tumours, parallel mutagenic processes generate countless mutations and chromosomal rearrangements in the genome. Whole genome sequencing (WGS) of tumours has enabled complex analysis of the selected mutations occurring during tumour progression in patients, but the high inter-tumour and intra-tumour heterogeneity makes it difficult to understand whether a particular polymorphism is a driver or a passenger mutation.

1.1.1.2 The Clonal Evolution model – genetic drivers of cancer

As briefly explained above, tumour initiation is driven by the sequential acquisition of genetic mutations that generate tumour cell heterogeneity through the origination of new clones. This leads to the idea of a Darwinian evolutionary process in tumours where the acquisition of new mutations will provide the new clones and its progeny with a beneficial advantage that the rest of the tumour cells did not share. Although most mutagenic processes are deleterious, over time advantageous mutations arise promoting selection and survival of a dominant clone and its progeny (Figure 1.1.A). These concepts set the basis of the Clonal Evolution model (Bakhoum and Swanton, 2014). Over the last years, genome sequencing of tumours disclosed many of the founding driver mutations in patients and revealed the coexistence of different subclones carrying distinct mutations within tumours. This suggests that a branching clonal diversification process rather than a linear clonal evolution operates throughout tumour progression (Stratton, 2011). The integration of these new findings refined the Clonal Evolution model that currently proposes that genetic heterogeneity within the same tumour provides the substrate for a Darwinian evolution that often occurs in a branched manner leading to genetic diversity. The multiple subclones that emerge trough branching evolution processes may differ not only genetically but also functionally, being only some of those subclones required to sustain tumour growth. Moreover, branched diversification might result from the adaptation of the clones to different microenvironments within the tumour suggesting that beyond genetic mutations there are other non-genetic determinants of tumorigenesis that can generate functional diversity. Indeed, epigenetic mechanisms controlling gene expression are more frequent than mutations in cancer genomes; on average only one extracellular matrix gene is mutated while four are hypermethylated in individual colon or breast tumours (Yi et al., 2011). Notably, although epigenetic modifications are not irreversible as mutations, some types of modifications are long lasting and heritable, being sufficient to generate distinct cellular states. Furthermore, the identification of subclones within single tumours provides strong evidence that intra-tumoural heterogeneity could be generated by the unique mutation spectrum present in an original clone, emphasising the key role of non-mutagenic processes in generating tumour cell heterogeneity (Kreso and Dick, 2014).

1.1.1.3 The Cancer Stem Cell model – epigenetic plasticity

The Cancer Stem Cell model proposes that beyond genetic diversity, cellular plasticity and functional diversity contribute to carcinogenesis, and hence cancer can be generated from the mutation spectrum present in an original clone displaying 'stem cell' characteristics. Stem cells in normal tissues are defined by their exclusive self-renewal abilities. Self-renewal is the biological process by which normal stem cells perpetuate themselves through symmetric division expanding the stem cell pool, or asymmetric cell division generating a stem cell and a more differentiated cell (Figure 1.1.B) (Reya et al., 2001). Molecularly, there is still considerable uncertainty about how self-renewal capacity and cell fate are determined, but studies in normal embryonic stem cells (ESCs) suggest a strong dependency on epigenetic modifiers such as the Polycomb complex (Sparmann and van Lohuizen, 2006, Scaffidi and Misteli, 2011). Likewise ESCs, cell fate decisions and self-renewal in cancer stem cells are likely to be controlled by numerous dynamic epigenetic changes rather than the permanent mutations that generate the cancerous phenotype (Wood et al., 2007). Indeed, next-generation sequencing techniques applied to mapping chromatin and DNA methylation have revealed that it is the balance between co-existing permissive and repressive chromatin modifications, a state termed poised chromatin, what maintains functionally distinct gene expression states in normal and cancer stem cells (Chaffer et al., 2013, Bernstein et al., 2006).

One of the most characterised mechanisms modelling the epigenome is the hypermethylation of CpG islands by the Polycomb repressive complex. CpG islands are DNA sequences localised in the promoter region of genes that become silenced upon Polycom recruitment and methylation (Deaton and Bird, 2011). Additionally, CpG islands contain multiple microRNAs that are efficiently repressed upon methylation. The silencing of these microRNAs can subsequently alter multiple target genes leading to upregulation of oncogenes (Saito et al., 2006), or constitutive activation of signalling pathways that promote EMT, invasion and metastatic activity (Figure 1.4.B – grey box) (Burk et al., 2008, Chi and Bernstein, 2009). These examples illustrate how epigenetic changes can generate cellular plasticity and functional diversity in cancer independently of mutagenic processes.

Accordingly, the Cancer Stem Cell (CSC) theory postulates that similar to the growth in normal proliferating tissues the growth of tumours is driven by a limited number of stem cells that are capable of self-renewal (Clarke et al., 2006). The hypothesis that developmental programmes usually underlying normal tissue stem cell function might operate in tumours, started back in the late 1970s. Several studies in different tumour types suggested that the bulk tumour cells were the differentiated progeny of tumour 'stem' cells (Bennett et al., 1978, Hager et al., 1981). Although the molecular mechanisms underlying self-renewal processes are still emerging, stem cells can be functionally identified when assayed in tissue regeneration assays *in vivo*, where they require self-renewal and differentiation capabilities to recapitulate the original tissue (Reya et al., 2001). Similarly, in cancer functional assays are used to identify cancer stem cells (CSCs), where a limited number of cells are tested for their ability to initiate and maintain long-term growth (Kreso and Dick, 2014).

Considering this functional definition, early evidence supporting the Cancer Stem Cell model came from leukemia studies. It was found that the majority of leukemic cells isolated from patients were post-mitotic, suggesting that there could be a small population of highly proliferative leukemic cells that would restock the bulk of leukemic cells in the blood (Clarkson, 1969). These results sparked the idea of tumours as hierarchically organised structures where a small subpool of cells with self-renewal capacity is responsible for the maintenance of the tumour structure. The development of flow cytometry and xenograft techniques to engraft patient cells in immune-compromised mice allowed the identification of $CD34^{+}CD38^{-}$ tumour-initiating cells (TICs) in leukemia (Lapidot et al., 1994). These were the only cells within the bulk of the tumour that could efficiently engraft in mice and recapitulate the disease heterogeneity observed in patients, being termed 'cancer stem cells' (Bonnet and Dick, 1997). These initial studies in leukemia were followed by many others identifying CSC populations in solid tumours. The first one was in human breast carcinoma, where a population expressing $CD44^{+}CD24^{-}$ was shown to sustain tumour growth in mouse xenografts. Remarkably the xenografts generated by this subset of cells could recapitulate to a certain extent the tumour heterogeneity observed in the patients (Al-Hajj et al., 2003).

To date, potential CSC populations have been described in most solid cancer types, skin (Malanchi et al., 2008), brain (Singh et al., 2004, Chen et al., 2012), colon (Dalerba et al., 2007, O'Brien et al., 2007) and prostate (Wang et al., 2009). Moreover, using mouse models and lineage tracing genetic tools to label tissue specific stem cells, it has been possible to determine that skin and intestinal carcinomas arise from within the stem cell compartment of these tissues (Malanchi et al., 2008, Barker et al., 2007). Supporting the functional evidence that cancer arises within the tissue stem cell compartment, epigenetic mechanisms that can induce genetic changes favouring cancer development are active in intestinal stem cells. For instance, *de novo* CpG methylations in the promoter of DNA-repair genes can create mutational hotspots. The methylation of the promoter and silencing of MLH1 leads to microsatellite instability (Shoemaker et al., 2000), and likewise silencing of o-6-methylguanin-DNA methyltransferase (MGMT) leads to increased G—A mutations (Esteller et al., 2000). Therefore, a working hypothesis in line with the Cancer Stem Cell model is that tumorigenesis starts with the abnormal expansion of the stem cell compartment of a given tissue (Figure 1.1.B), although the mechanisms underlying this process remain to be characterised (Ohm and Baylin, 2007, Cedar and Bergman, 2008).

In summary, the cancer stem cell model proposes that tumours like tissues are hierarchically organised structures, where only a small subpool of cells with limited proliferation and self-renewal capacity sustains tumour growth.

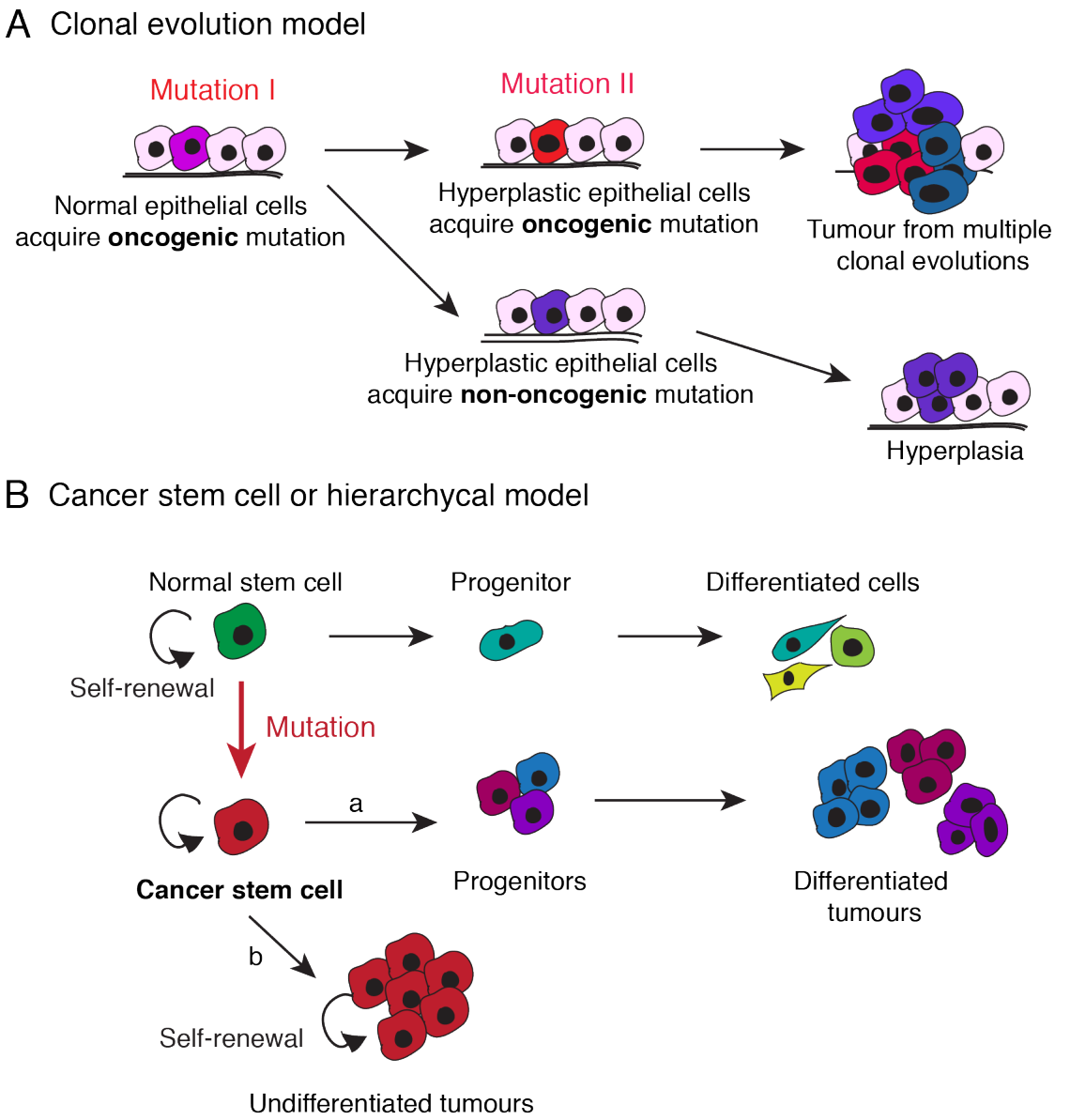


Figure 1.1 The Clonal Evolution and the Cancer Stem Cell Models of carcinogenesis
(A) The Clonal Evolution Model: healthy epithelial cells acquire an oncogenic mutation (mutation I) that forms a hyperplasia. Some of the cells will acquire additional oncogenic mutations becoming cells of origin, and after multiple clonal evolutions in parallel the tumour results from different clones (shown in different colours). Alternatively, those cells that acquire a few oncogenic mutations but not enough to initiate malignant growth will form hyperplasias. (B) The Cancer Stem Cell model: normal tissue stem cells have a limited proliferative capacity and give rise to progenitor cells that further differentiate in different cell types. When a normal stem cell gains oncogenic mutations generates a cancer cell with self-renewal capacity that can generate a) progenitors with higher proliferative rate, that will give rise to highly proliferative differentiated tumours; b) alternatively, when oncogenic mutations de-regulate self-renewal, the cancer stem cell pool will acquire a high proliferative rate giving rise to undifferentiated tumours.

1.1.1.4 Unified model for Cancer Evolution

The Clonal Evolution and the Cancer Stem Models have been traditionally viewed as antagonistic models to explain carcinogenesis. To date, the major caveat when trying to establish a general model for carcinogenesis has been the lack of integration between the genomic and functional properties of tumour-initiating cells. It remains to be determined which genetic clones can graft in mice and display stem-like properties. In an integrative effort, sequencing studies in breast cancer have shown how the common ancestor containing conserved driver mutations appears early during clonal evolution indicating that much of the time is spent driving subclonal diversification (branching evolution) (Nik-Zainal et al., 2012). Moreover, this study showed that driver mutations leading to clonal expansion occur rarely in long-lived populations, which rather passively accumulate mutations without expanding (Nik-Zainal et al., 2012). These findings suggest that rare clones with limited proliferative capacity within tumours maintain long-term growth. Additionally, studies in leukemia have defined the functional role in tumorigenesis of different subclones. Shlush et al. tracked somatic mutations in large numbers of single leukemia cells isolated from patients, and found a driver recurrent mutation (DNMT3A) in a fraction of the hematopoietic stem cell (HSC) pool. DNMT3A mutant HSCs showed a repopulation advantage over the non-mutated HSCs in xenografts, establishing their identity as pre-leukemic HSCs. Moreover, these DNMT3A mutated pre-leukemic HSCs were enriched in relapses compared to samples at diagnosis, indicating that they are chemoresistant. These findings demonstrate that an early appearing clone can drive the clonal expansion of HSCs from which Acute Myeloid Leukemia (AML) originates and potentially trigger relapse after persisting chemotherapy (Shlush et al., 2014).

In line with these preliminary studies integrating genomic and functional data, Kreso and Dick propose that the Clonal Evolution and CSC models can be harmonised: the acquisition of favourable mutations impacting on self-renewal properties (i.e. acquiring the DNMT3A mutation that drives AML) can result in clonal expansion of the original CSC population, and in parallel another cell can gain a different mutation that allows it to form a new stem-like subclone. Over time, genetic mutations accumulate and subclones evolve in parallel presenting CSCs as

dynamic entities that evolve over time (Figure 1.2 – X axis, genetic determinants). The CSC subclones generated will further divide giving rise to three types of scenarios: 1) hierarchically organised clonal populations where only a few cells retain self-renewal capacity; 2) an intermediate hierarchy where the number of CSCs is relatively high (symmetric divisions equal asymmetric divisions); 3) a homogeneous mass of CSCs, that due to the acquisition of mutations directly impacting on self-renewal, divides symmetrically giving rise to predominantly undifferentiated tumours (Figure 1.2 – pink CSC subclones) (Kreso and Dick, 2014). In parallel to this mutagenic evolution in cancer stem cells, context-dependent phenotypic plasticity can be gained through epigenetic modulations enabling cancer cells to gain/lose stem cell properties (Figure 1.2. – Y axis, epigenetic determinants) (Lim et al., 2009, Quail et al., 2012). This epigenetically driven plasticity is key to reconcile the Clonal Evolution and the Cancer Stem Cell models (Plaks et al., 2015): microenvironment-induced epigenetic regulations would determine whether CSCs will transiently amplify (self-renewal) (Figure 1.2 – grey curved arrows) or differentiate to nonCSCs (Figure 1.2 – black arrows). Moreover, recent reports claim that the gain of stem-like features is possible, and tumour cells have been proposed to interconvert between stem-like and differentiated states depending on the microenvironmental cues received (Chaffer et al., 2013, Gupta et al., 2009). Although still under debate, these studies propose nonCSCs could de-differentiate regaining stem-like properties (Figure 1.2 – blue arrows).

In conclusion, the unified model of carcinogenesis proposes CSCs as dynamic entities that evolve trough the acquisition of mutations generating new CSC subclones and in parallel undergo epigenetic modulations that ultimately control their fate. Given the complexity of the regulatory mechanisms driving cancer stem cell phenotypes during tumour progression, it is of central importance to develop assays to study their function *in situ*, which would allow examination of their fate decisions and clonal identity within their biological context. In the future, the integration of genomic and functional data acquired in the native microenvironment of cancer stem cells would allow the identification of definitive markers to identify and isolate these populations in different human cancer types.

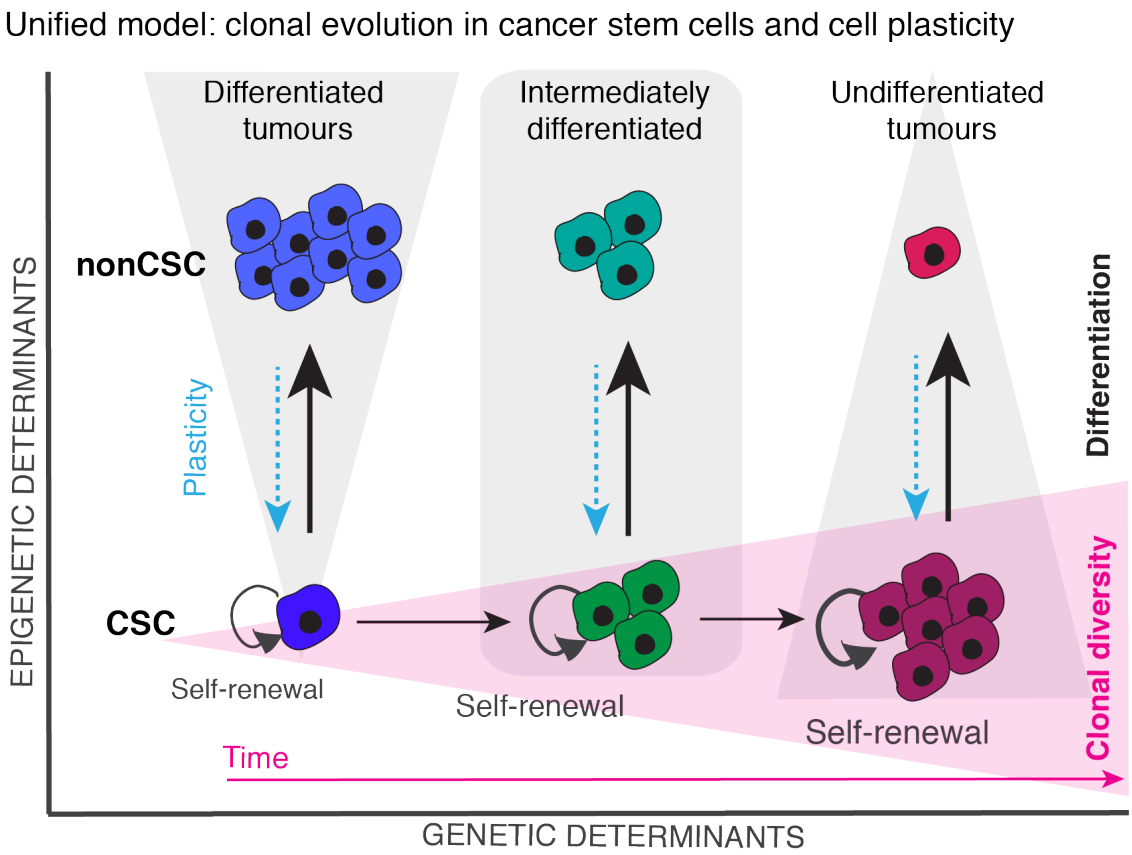


Figure 1.2 Unified Model of carcinogenesis: clonal evolution in cancer stem cells and cellular plasticity
Genetic determinants - The founder low cycling CSC population can gain new mutations over time that impact on their stem-cell-like features (pink triangle – clonal diversity), creating new CSCs subclones in the tumour with higher self-renewal capacity that evolve in parallel. Epigenetic determinants - CSCs are not static entities, and their proliferative status can be modulated extrinsically (differentiation/plasticity arrows) changing their relative frequency in the tumour mass. As a result of the combined genetic and epigenetic factors, some subclones will organise as differentiated tumours, with a hierarchical structure where CSCs are rare (blue CSC subclone); others will show a high frequency of CSCs but still display a hierarchy (green CSC subclone), while subclones with high self-renewal potential will form homogeneous undifferentiated tumours (pink CSC subclone). Reverse transitions can occur as well, differentiated nonCSCs can re-acquire a stem-like status (blue arrows – plasticity) through new mutations impacting on self-renewal or microenvironmental modulations regulating their epigenetic status.

1.1.1.5 Cancer stem cell: concept implications

To date, many studies have highlighted the unique properties of CSCs in different human xenografts and mouse models leading to the compilation of the 'CSC hallmarks' (Malanchi, 2013). The first hallmark of CSCs is their exclusive ability to self-renew and differentiate, being able to initiate tumour growth in clonal or limited dilution assays (Malanchi, 2013). The second hallmark is their ability to sustain tumour growth as demonstrated in skin carcinoma mouse models, where the ablation of the CSC population led to tumour regression in two independent studies (Malanchi et al., 2008, Beck et al., 2015); Lastly, CSCs display unique molecular mechanisms that make them chemoresistant compared to the bulk of the tumour cells in glioblastoma, colon cancer, breast cancer and other tumours (Bao et al., 2006, Ishikawa et al., 2007, Saito et al., 2010, Diehn et al., 2009), suggesting that they are the main drivers of relapse in patients. These hallmarks emphasise the functional relevance of CSCs in tumours.

Although the central role of CSCs in tumorigenesis is increasingly recognised, there is still a great controversy in the field due to the way they have been defined and the different tests that are used to assess their potential (Badve and Nakshatri, 2012, Barrett et al., 2012, Kreso and Dick, 2014, Nguyen et al., 2012).
First, regarding their definition, the term cancer 'stem' cell implies that as normal stem cells, CSCs possess self-renewal capacity. However, self-renewal is typically deregulated in CSCs. For instance, leukemic stem cells display abnormally elevated levels of the RNA-binding protein Musashi that ultimately triggers symmetric cell division expanding the leukemic stem cell pool (Ito et al., 2010). Also, we previously mentioned how mutations in the epigenetic modifier DNMT3A lead to leukemic stem cells clonal expansion in AML (Shlush et al., 2014). Second, a major problem with stem cells studies is that usually the assays used to assess self-renewal potential are performed in exogenous microenvironments after cell isolation, and therefore they do not consider the key microenvironmental factors that normally regulate stem cell function.

All of the above facts indicate that the term cancer 'stem cell' poses a problem and its use therefore should be limited to those cases where clonal self-renewing CSCs

can be prospectively purified from primary tumours for functional testing. Consequently, we define putative CSCs functionally according to the tests used to assess their intrinsic potential *in vivo* as tumour-initiating cells (TICs) or metastasis-initiating cells (MICs) (Oskarsson et al., 2014, Ombrato and Malanchi, 2014).

1.1.2 Microenvironmental heterogeneity

Traditionally the biology of tumours was studied as being ruled by the genetic and epigenetic modifications driving tumour cell heterogeneity, however over the last few decades the tumour microenvironment has emerged as an equally important component for disease onset and progression. It is now acknowledged that tumour heterogeneity is not only determined by the diversity in the cancer cell compartment but also by the miscellaneous composition of its microenvironment, and the proportion and activation status of the stromal cells within it (Hanahan and Weinberg, 2011). To date, it is acknowledged that the dynamic bidirectional communications between tumour cells and the microenvironment determine the overall tumour fitness, invasive potential and therapy outcome (Junttila and de Sauvage, 2013). Moreover, disease initiation and patient prognosis can be determined by the intricate mechanisms taking place in this complex tumour-stroma ecosystem (Farmer et al., 2009).

1.1.2.1 Stromal cells

The systemic interactions between tumour and stromal cells resemble other normal physiological processes, such as inflammation or wound healing, where the microenvironment transiently exerts a direct regulation on epithelial tissue homeostasis. However, in contrast to the transient activation of stromal cells during these physiological processes, tumour's recruitment and activation of stromal cells persists and constantly evolves throughout disease progression and never resolve (McAllister and Weinberg, 2014).

Immune cells are one of the best-characterised components of the microenvironment. Indeed, chronic inflammatory stress in tissues is known to have a direct effect in cancer incidence (Grivennikov and Karin, 2010). Generally, immune cells can contribute to tumorigenesis by directly regulating tumour cells function through paracrine signals or mediating more complex immune-suppressive responses. Some immune cells, such as macrophages and neutrophils exert direct pro-tumorigenic functions. Macrophages and neutrophils are critical effectors during immune defence, but in the tumour context there is evidence of their pro-tumorigenic role (Figure 1.3). It is not currently understood how these innate immune cells change from a tumour suppressor to a tumour promoting phenotype during tumour initiation. However, it is widely proven that during tumour progression tumour-associated macrophages (TAMs) and neutrophils can directly enable tumour cell invasion through paracrine signalling (Wyckoff et al., 2007, Glogauer et al., 2015) or secrete crucial extracellular components that will affect disease progression (Akkari et al., 2014). Consequently, depletion of macrophages and neutrophils during tumour progression can partially impair metastatic dissemination (Lohela et al., 2014, Gregory and Houghton, 2011). In addition to the pro-tumorigenic roles of some immune components, complex immunosuppressive mechanisms have been observed during tumour development. These mechanisms allow tumour cells to escape from the surveillance mechanisms implemented in normal physiological conditions and are usually mediated by myeloid-derived suppressor cells (MDSCs) recruited to the tumour site and regulatory T-cells (Figure 1.3). It is well characterised that infiltrating MDSCs in developing tumours impair different defence mechanisms, such as T-cell activation (Mazzoni et al., 2002), inhibition of NK cells cytotoxicity (Liu et al., 2007a) and antigen presentation by dendritic cells (Gabrilovich et al., 2012). Notably, depletion of MDSCs with neutralizing antibodies can reduce metastatic outcome in several animal models (Talmadge and Gabrilovich, 2013). In the same way, regulatory T-cells (T-reg) can also impair cytotoxic T-cells and natural killer (NK) cells function by disrupting immune-surveillance mechanisms (Gasteiger et al., 2013, Fu et al., 2007). Altogether, these findings exemplify the pro-tumorigenic role of the altered immune system in the context of the tumour microenvironment compare to normal physiological conditions.

Cellular and extracellular components of the tumour microenvironment

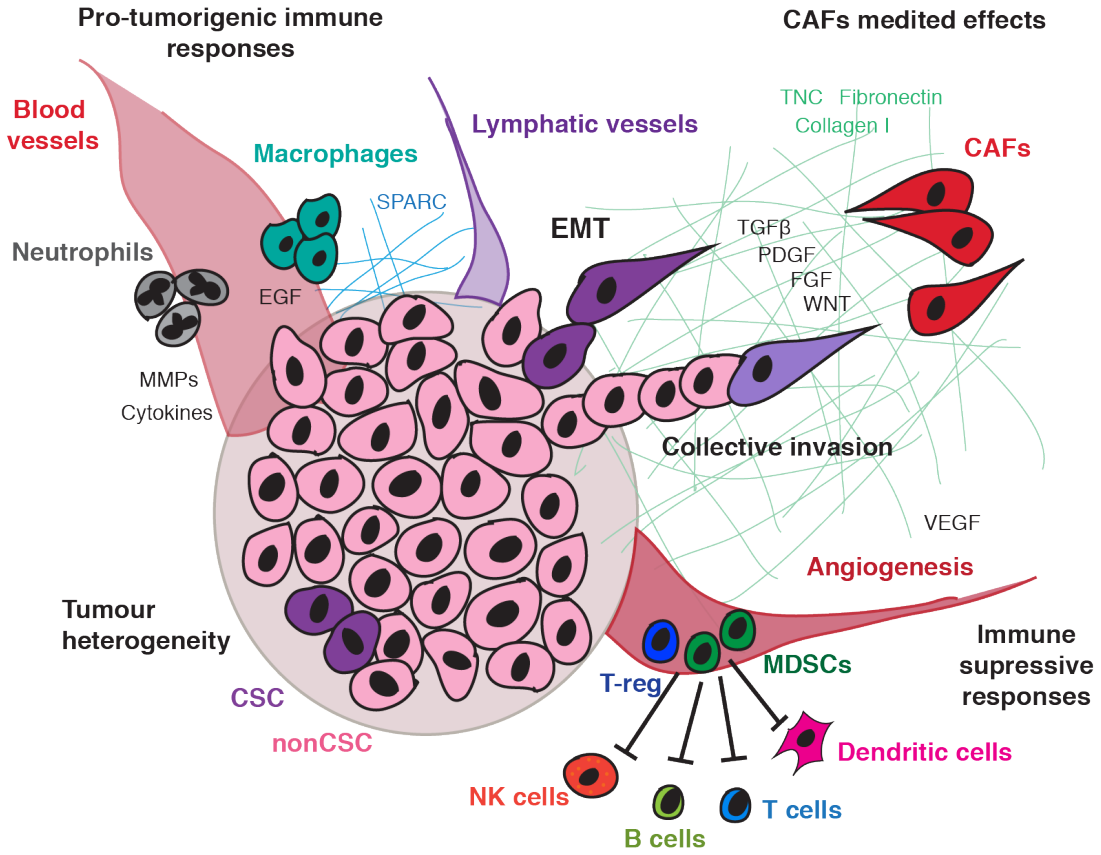


Figure 1.3 Main cellular and extracellular components of the tumour microenvironment
The cancer cell mass in tumours is heterogeneous, and hierarchically organised in CSC and nonCSC that recruit stromal components through the secretion of cytokines and ECM components. Recruited macrophages (i.e. by tumour cell derived SPARC) and neutrophils, exert pro-tumorigenic activities through the secretion of cytokines, growth factors such as EGF, and support tumour cell invasion by secreting MMPs. CAFs are activated by tumour-cell derived factors (i.e. TGFβ, FGF, PDGF) and have a crucial role in producing ECM matrix (Fibronectin, TCN and collagen-I) components. CAF-derived factors enhance cancer cell dissemination at the tumour-stromal interface (TGFβ) and stimulate tumour cell self-renewal and proliferation (WNT, FGF). Angiogenesis increases as macrophages and CAFs are recruited to the tumour site contributing to the secretion of angiogenic factors (i.e. VEGF). As tumours grow and vasculature develops, immune-suppressive cells such as T-reg cells and MDSCs are mobilised and disrupt the immune-surveillance mechanisms typically performed by NK, T, B and dendritic cells.

Another key cellular component of the tumour microenvironment are fibroblasts. Traditionally, fibroblasts have been described as non-vascular, non-epithelial and non-inflammatory cells of the connective tissue (Tarin and Croft, 1969). These multifunctional cells have roles in ECM deposition, wound healing, epithelial differentiation and regulation of inflammation in physiological conditions (Kalluri and Zeisberg, 2006). Fibroblasts are key in maintaining tissue homeostasis through the deposition of specific ECM components that define the biomechanical properties of the tissue and form the basement membrane that confines tissue limits. Indeed, in tissues rich in ECM such as the skin, fibroblasts consist of a number of distinct mesenchymal cells types with different origins, locations and functions, giving an insight into the complexity of the fibroblastic stroma (Driskell and Watt, 2015). In physiological conditions, fibroblasts also secrete factors that directly regulate epithelial cell function in a paracrine manner (Luhr et al., 2012, Visco et al., 2009).

In the tumour microenvironment cancer-associated fibroblasts (CAFs) display distinct characteristics compared to those of normal tissue fibroblasts and are present in aberrantly high numbers. CAFs result from normal fibroblasts that become activated by pro-fibrotic tumour-derived factors such as TGFβ, platelet-derived growth factor (PDGF) and fibroblast growth factor 2 (FGF2) (Figure 1.3) (Kalluri and Zeisberg, 2006). In breast cancer, normal fibroblasts get activated as matrix stiffness in the tumour microenvironment increases as a result of TGFβ signalling and matrix deposition by tumour cells themselves and recruited normal fibroblasts (Calvo et al., 2013). The aberrantly stiff matrix promotes nuclear translocation of the transcriptional co-activator YAP (Yes associated protein-1) that governs further fibroblast activation. YAP signalling activation leads to CAFs morphological changes through the expression of cytoskeletal components such as smooth muscle action (SMA) and fibroblast activation protein (FAP). Once activated, CAFs secrete aberrantly high levels of ECM components, such as type I, III and IV collagen and fibronectin that increase further matrix stiffness sustaining this YAP-dependent mechanical signalling loop that maintains fibroblasts activated throughout disease progression (Calvo et al., 2013). During cancer progression, in mammary carcinomas CAFs have been shown to influence tumour initiation through the secretion of TGFβ and HGF. Moreover, when immortalised mammary epithelial cells were grafted into mice with either normal fibroblasts or CAFs

isolated from mammary gland or mammary carcinoma respectively, only in the presence of CAFs tumours were developed (Olumi et al., 1999).
Additionally, CAFs are known to confer mesenchymal-like features to malignant epithelial cells promoting their metastatic ability. This is mediated by the secretion of TGFβ, MMPs and ECM components that stimulate integrin signalling, and production of growth factors, such as EGF and FGF2 that can activate tyrosine-kinase receptor signalling on cancer cells (Dumont et al., 2013). Furthermore, in line with this mesenchymal induction ability, CAFs have been shown to promote both single cell invasion through TGFβ stimulation, and collective cell invasion by leading the way and in turn opening micro-tracks in the dense ECM (Gaggioli et al., 2007). Altogether, these findings illustrate the multiple pro-tumorigenic functions of CAFs throughout tumour progression.

Additional fundamental cellular components of the tumour microenvironment are the blood vessels that supply tumours with oxygen and nutrients from the bloodstream. The vasculature also determines the possibility of stromal cell recruitment to the tumour site, and the exit of tumour cells to reach circulation and metastasise (Figure 1.3). Tumour vasculature development requires the co-operation of different cell types: pericytes to provide vessel coverage, endothelial cells that conform the building blocks of vessels and bone-marrow progenitors. In addition to the cells forming the blood vessels, other stromal cells such as TAMs, CAFs and mesenchymal stem cells (MSCs) contribute to promote vascularisation by releasing pro-angiogenic molecules such as VEGF and FGF, or downregulating the expression of anti-angiogenic factors such as thrombospondin-1 or β-interferon (Hanahan and Weinberg, 2000). Angiogenesis is a hallmark of cancer, and therapeutic disruption of pro-angiogenic mechanisms can impair tumour progression in most cancer types (Hanahan and Weinberg, 2000).

There are other emerging stromal cell types that can influence tumorigenesis and are currently being studied. Mesenchymal stem cell (MSC) recruitment to the tumour site has been reported in several cancer types (Johann and Muller, 2015). MSCs are stromal progenitors that can differentiate into osteoclast, chondrocytes, myocytes and adipocytes (Pittenger et al., 1999). This differentiated progeny further influences the tumour microenvironment; for instance, adipocytes have

been shown to play a role in obesity-associated cancers (Morris et al., 2011, Nieman et al., 2011). Also, it is interesting the pro-tumorigenic impact of the resident microbiota in organs such as the skin (Hoste et al., 2015) or the gut (Zackular et al., 2013), where they directly regulate tumour cells though the secretion of factors in a paracrine manner, or exacerbate the pro-tumorigenic inflammatory responses in the tumour. Last, it is intriguing the emerging systemic role of tumour and stromal cell-derived exosomes as microenvironmental components regulating disease progression (Peinado et al., 2011). The increasing complexity of the tumour microenvironment requires further investigation to uncover potentially targetable mechanisms to efficiently complement cancer cell-directed therapies.

1.1.2.2 The extracellular matrix

Together with the different stromal cells discussed above, the extracellular matrix (ECM) is a crucial element of the tumour microenvironment. Although long considered as a steady structure supporting tissue morphology the ECM can actually exert both mechanical and biochemical instructions to all cellular components in tumours (Lu et al., 2012). Particularly in cancer, it is well characterised how an abnormal composition of the ECM can change the behaviour of both epithelial and stromal cells, impacting in the success of tumour initiation as well as triggering invasion during later disease development (Lu et al., 2012). Thus, abnormal ECM composition and dynamics are considered today a hallmark of cancer (Hanahan and Weinberg, 2011). ECM matrix components can be secreted by most cell types; as previously mentioned, in physiological conditions fibroblasts are the main producers and modellers of the tissue matrix (section 1.1.2.1), but in the tumour microenvironment other cell types such as tumour cells themselves and some immune components are known to secrete key ECM components supporting tumour progression (Quail and Joyce, 2013).

The ECM is composed of secreted proteins, glycoproteins, proteoglycans and polysaccharides with different physical and biochemical properties (Lu et al., 2012).

Its composition will determine the physical properties (rigidity, porosity, insolubility, spatial organisation) and biochemical properties that *per se* can enhance cell growth and the invasive behaviour of primary breast cancer cells *in vivo* (Lo et al., 2000). Physically, a stiffer microenvironment can be achieved by a single component variation in the ECM, having enormous effects in cell morphology and migration. Nguyen-Ngoc et al. have shown how a collagen-I rich stiff microenvironment is enough to induce a conserve dissemination response in normal and malignant mammary epithelium (Nguyen-Ngoc et al., 2012). On the other hand, the biochemical properties of the ECM define its signalling capabilities: ECM components will trigger signal transduction cascades initiated at the cell surface that ultimately will reach the nucleus resulting in dramatic changes in cell behaviour. Additionally, the ECM can bind different signalling ligands such as TGFβ, BMPs, WNTs, FGFs, limiting their diffusive range, accessibility and signal direction implementing an extra level of modulation to autocrine and paracrine signalling mechanisms (Hynes, 2009).

Integrins are the main mean by which ECM molecules initiate transduction cascades to change cell behaviour. Integrins are heterodimeric receptors formed by a α and β subunit that link the extracellular microenvironment to the intracellular actin cytoskeleton. The binding of integrin receptors to ECM ligands provides physical information of its location, adhesive state and surrounding matrix (Guo and Giancotti, 2004). Indeed, a stiff microenvironment can increase integrin β1 subunit expression and activity through mechanotransduction, enhancing cell migration through the formation of focal adhesions (Shibue et al., 2013). Integrin biochemical stimulation and downstream signalling can be achieved by different ECM glycoproteins present in the tumour microenvironment, such as fibronectin (FN), Vitronectin (VN), Tenascin-C (TNC) and latent forms of TGFβ. Additionally, integrins cooperate with other cell surface receptors to regulate biological processes such as cell proliferation and differentiation, cell shape and migration, and survival (Yoshida et al., 2015). Thus, preventing integrin biochemical stimulation by ECM components inhibits the activation of many downstream pathways required for breast cancer progression in both, tumour and stromal cells (Guo and Giancotti, 2004, Uberti et al., 2010, Seguin et al., 2014).

The biomechanical properties of the ECM also have profound effects in other crucial cellular components of the microenvironment. Particularly, angiogenesis is complex process where the ECM is involved in endothelial cell survival and proliferation (Sweet et al., 2012), lumen formation (Newman et al., 2011) and other aspects of tubulogenesis during blood vessel formation. Indeed, there is a fine regulation between pro and anti-angiogenic factors in tumours, and ECM glycoproteins are key in regulating the accessibility of these factors, or acting directly as endothelial cell regulators. For example, in breast cancer the ECM glycoprotein biglycan indirectly displays pro-angiogenic functions by sequestrating the angiogenic inducer VEGFA in the ECM, generating a reservoir of VEGF that will be released with ECM matrix degradation processes promoting angiogenesis (Berendsen and Olsen, 2014). While other glycoproteins, such as Trombospondin-1 (THBS1) and Trombospondin-2 (THBS2) exert anti-angiogenic effects directly on endothelial cell survival, inducing apoptosis through activation of CD36 receptor (Jimenez et al., 2000). Indeed, THBS2 overexpression results in decreased mammary tumour growth in human xenograft mouse models of breast cancer (Koch et al., 2011). Certainly, the abnormal ECM composition of the tumour microenvironment leads to the formation of functional but morphologically distinct tumour vasculature (Myers et al., 2011).

Additionally, the ECM matrix composition can regulate the recruitment of immune cells through the diffusivity of chemoattractants, or physically limiting the intersticial space. For instance, the ECM matrix glycoprotein SPARC (secreted protein acidic and rich in cysteine) is crucial for the production and organisation of the ECM and the recruitment of macrophages to the tumour site in Lewis Lung Carcinomas (LLC) (Figure 1.3) (Brekken et al., 2003).

Lastly, as emphasised by Lu et al. one of the most relevant functions of ECM-cells interactions is that they are reciprocal. Cells continuously produce and rearrange the ECM changing its properties, and in turn the new properties of the ECM will impact on cell behaviour. As discussed above, in cancer these permanent feedback loops will determine disease onset and progression (Lu et al. 2015). Indeed, in breast cancers tumours can be stratified according to their ECM composition, which is predictive of patient outcome (Bergamaschi et al., 2008).

Moreover, high expression of proteases and integrins, which denotes a high ability to remodel and interact with the ECM, are associated with poor prognosis and relapse (Bergamaschi et al., 2008). These findings complement the *in vivo* evidence that tumours with different metastatic potential differ in their ECM composition (Naba et al., 2012). Altogether, these results highlight the crucial role of the ECM during cancer progression and denote the prognostic value of analysing the ECM composition of tumours.

1.2 Overview of the metastatic process

Metastasis is the multistep process by which tumour cells confined in their tissue of origin spread via the circulatory systems of the body to other organs, successfully re-initiating tumour growth. In carcinomas, the metastatic process starts when epithelial cells acquire invasive characteristics that allow them to infiltrate the surrounding tissues and reach the vasculature (intravasation). Subsequently, tumour cells in the circulation will spread throughout the body to the microvasculature of distant tissues where they will arrest and exit circulation (extravasation). Once in the new tissue parenchyma, most cancer cells will die in the new environment. Survivors will either stay dormant or if the appropriate conditions are encountered they will colonise the naive tissue re-initiating tumour growth (Chaffer and Weinberg, 2011).

A general characteristic of the metastatic process is that it is extremely inefficient. Experimental metastasis studies in melanoma performing intravenous injection of tumour cells directly into the circulation have shown that around 20% of the cells that enter circulation undergo apoptosis. Upon extravasation into the secondary sites only around 2% of the cells originally injected will form micrometastasis. Moreover, not all these micrometastasis will progress towards metastatic growth, being estimated that only around 0.02% of the cells initially injected will successfully metastasise (Luzzi et al., 1998). Therefore, the step of metastatic colonisation at the secondary site is usually viewed as a series of difficulties that tumour cells need to overcome in order to grow, and represents the most promising step to be therapeutically targeted.

Metastases result from disseminated tumour cells and clinical evidence indicates that even when primary tumours are surgically removed with perfect marginal precision metastases can occur. Indeed, metastasis accounts for 90% for the mortality associated to cancer disease. Despite the unquestionable therapeutic relevance of targeting metastatic progression, many aspects of this complex multistep process remain insufficiently understood (Chaffer and Weinberg, 2011). In the next sections, I will summarise the main knowledge built over the past decades for each step of this intricate process and discuss the relevance of therapies targeting the different steps of the metastatic cascade to ultimately prevent metastatic outgrowth in patients.

1.2.1 Leaving the primary tumour: single cell *versus* collective invasion

As mentioned above, the metastatic cascade starts with local invasion of tumour cells that transform a locally growing tumour into a systemic live-threatening disease. Typically, solid epithelial tumours have been observed to invade in two major ways, as single cells when cell-cell junctions are absent, or collectively as cell clusters joined by the maintenance of cell-cell adhesions (Friedl et al., 2011). The commonality underlying both ways of invasion is the joint action of the cytoskeleton that serves as the cells' motor, with cell surface receptors that engage the surrounding ECM directing cell movement (Ridley et al., 2003). The activation of signalling pathways controlling cancer cell cytoskeletal dynamics and hence cell migration is sparked by microenvironmental changes occurring during late-stage tumour progression (Friedl and Alexander, 2011).

1.2.1.1 Single cell migration – a focus on mesenchymal motility

Although there are different mechanisms underlying single cell dissemination in carcinomas, the acquisition of a mesenchymal phenotype via Epithelial-to-mesenchymal-transition (EMT) is thought to be the main mechanism facilitating single cell invasion. This mesenchymal type of motility is based on integrin-mediated ECM adhesion and the use of proteases to generate traction force. Generally, mesenchymal migrating cells contain focalised proteases in the cell surface such as metalloproteases (MMPs) that generate microtracks in the ECM serving as roads for cells to migrate (Friedl and Alexander, 2011).

An alternative single cell invasion mechanism is the amoeboid migration characterised for the lack of protease activity and integrin-ECM contacts, constituting a mesenchymal independent mechanism by which cells 'squeeze' through tissues by propulsive cytoplasm forward flow cytoskeletal dynamics (Friedl et al., 2012). Indeed, integrin blockage in carcinoma cells induces mesenchymal-to-amoeboid transition (MAT) as an alternative single cell mediated invasion mechanism to EMT (Zaman et al., 2007). Also, in squamous cell carcinoma a podoplanin-mediated EMT independent single cell invasion mechanism has been described (Wicki et al., 2006).

Nonetheless, in most carcinomas single cell invasion results from the loss of epithelial characteristics usually driven by the induction of EMT, a developmental programme by which static polarised epithelial cells change into an invasive migratory phenotype. Upon EMT activation cells adopt an elongated morphology with focalised cell-matrix interactions and display enhanced ECM remodelling abilities (Figure 1.4.A) (Nieto, 2011, Nieto, 2013). Molecularly, the EMT programme is driven by the activation of the transcription factors Snail, Twist and Zeb that induce a cellular programme to disable cell-cell junctions through repression of proteins such as E-cadherin, ZO-1 and claudins. Alongside, EMT provides cells with invasive behaviour through changes in the cytoskeleton structure (expression of Vimentin) and ECM remodelling ability (expression of proteases) (Peinado et al., 2007, Tiwari et al., 2013). Nuclear localisation of the EMT transcription factors (EMT-TFs) together with E-cadherin loss, and gain of spindle-shaped morphology and Vimentin expression are common EMT hallmarks in embryogenesis and cancer (Cano et al., 2000, Vesuna et al., 2008, Nieto, 2013).

Inducers of EMT in the tumour microenvironment: focus on TGFβ signalling

Different stimuli present in the tumour microenvironment can trigger the activation of the EMT-TFs in cancer cells (Figure 1.4.B): 1) physical constrains imposed by the aberrant ECM are transduced via integrins inducing PI3K-AKT activation (Figure 1.4.B - grey arrows), which leads to NFkB and JUN mediated SNAIL upregulation (Desprat et al., 2008); 2) inflammatory signals such as CCL2 induces NFkB activation and Snail expression (Lopez-Novoa and Nieto, 2009); 3) metabolic stress, such as hypoxia induces HIF1 mediated Snail activation (Dong et al., 2013) and 4) abnormal activation of several signalling pathways such as WNT or NOTCH that can indirectly activate the EMT-TFs (Thiery et al., 2009). Moreover, different growth factors such as EGF and FGF represent a major way to activate EMT through RAS–RAF–ERK-MAPK signalling cascade in cancer cells (Castellano and Downward, 2011, Sakuma et al., 2012).

The main EMT inducer in the tumour microenvironment is TGFβ (Heldin et al., 2012). The TGFβ superfamily comprises in humans 30 different ligands being the best characterised the TGFβs, activins and NODAL that signal through the ALK4/7 type-I receptors (TGFβ branch of the pathway) and Bone-Morphogenic-Proteins

(BMPs) that signal through the ALK1/2/3/6 type-I receptors (BMP branch) (Wakefield and Hill, 2013). The canonical bases of TGFβ signalling are that BMPs activate SMAD1, SMAD5 and SMAD8 downstream whereas TGFβ, NODAL and activins signal through SMAD2 and SMAD3. Upon reception of TGFβ or BMPs signals, the corresponding downstream SMADs proteins are phosphorylated and form a complex binding to the common node SMAD4 shared by both branches of the pathway. These SMADs complexes translocate to the cell nuclei where they regulate gene expression (Wakefield and Hill, 2013). TGFβ stimulation directly induces EMT through pSMAD2-3 nuclear translocation, which activates the three families of EMT core transcription factors Snail, Twist and Zeb (Figure 1.4.B) (Xu et al., 2009). Moreover, TGFβ can also indirectly induce the expression of EMT-TFs in cancer cells through the activation of PI3K (Schlegel et al., 2015) and ROCK signalling pathways (Morin et al., 2011) (Figure 1.4.B - red arrows).

Functionally, TGFβ displays a dual role during tumour progression; in the initial stages of tumour development TGFβ arrests proliferation and can trigger tumour cell apoptosis (Levy and Hill, 2006) whereas in fully developed tumours stromal cell derived TGFβ enhances invasion and metastatic activity via activation of EMT (Thiery, 2009). Indeed, TGFβ in late-stage breast cancer promotes single cell dissemination through the bloodstream (Giampieri et al., 2009) suggesting that EMT is required for blood vessel intravasation regardless of the local cell invasion modality. Moreover, cells undergoing EMT display attenuated proliferation and enhanced resistance to cell death (Figure 1.4.B) (Roussos et al., 2010) properties that ensure efficient migration and adaptation to reach secondary sites.

In contrast to all these pro-invasive abilities triggered by TGFβ, BMP usually controls cell proliferation (Massague, 2012) and finely tunes the balance between self-renewal and differentiation in normal and cancer stem cells (He et al., 2004, Lonardo et al., 2011). Importantly, the two different branches of the pathway TGFβ and BMP are known to antagonise each other at the level of the SMADs complexes. This fits with the idea of a balance between these two pathways regulating EMT modulations and the functionally antagonistic migratory and proliferative states observed in tumour cells (Figure 1.4.A). For instance, antagonism BMP signalling by TGFβ have been reported to involve the formation of inhibitory complexes formed by pSMAD3 and pSMAD1 in response to combined TGFβ and BMP signalling (Gronroos et al., 2012). Also, the interplay between BMP and their

antagonists can determine the ability of disseminated tumour cells to exit dormancy and establish metastases at distant sites (Gao et al., 2012). All these emerging inter-regulatory mechanisms between BMP and TGFβ seem to be key during metastatic progression, where partial EMT states can be observed and are modulated in a context dependent manner. We will further discuss the role of TGFβ signalling during metastatic colonisation in section 1.4.

EMT in breast cancer: AXL-mesenchymal status

In breast cancer, phenotypic EMT involves E-cadherin loss (Vesuna et al., 2008) accompanied by *de novo* Vimentin and AXL expression. AXL, one of the main proteins analysed in this work, have been widely characterised as an EMT defining marker in breast cancer (Figure 1.4.A) (Vuoriluoto et al., 2011, Gjerdrum et al., 2010, Asiedu et al., 2013). AXL is a tyrosine kinase receptor that was first described to be overexpressed in leukemia, where it correlates with worse prognosis and overall survival in patients (Rochlitz et al., 1999). In breast cancer, AXL is typically overexpressed in luminal ER^+ and triple negative breast cancers (Meric et al., 2002) where as in leukemia represents a poor prognosis indicator (Berclaz et al., 2001). As other tyrosine kinase receptors, AXL can be activated by binding to its ligand Gas6, mediating signal transduction processes involved in epithelial differentiation, proliferation and survival (Korshunov, 2012). Nevertheless, in breast cancer AXL usually acts as a downstream effector of the EMT programme (Gjerdrum et al., 2010, Vuoriluoto et al., 2011) and its induction in this context can be independent from its tyrosine kinase receptor functions (Paccez et al., 2014). For instance, in a xenograft model using the basal breast cancer cell line MDA-MB-231, AXL mediated a mesenchymal-like phenotype providing increased single cell migration and metastatic capacity without affecting cell proliferation or survival *in vivo* (Gjerdrum et al., 2010), two typical downstream functions mediated by AXL as a tyrosine kinase upon Gas6 stimulation (Korshunov, 2012). Molecularly, the master EMT inducer TGFβ positively regulates AXL expression and in turn this TGFβ-AXL axis further maintains EMT through the activation of PI3K signalling (Figure 1.4.B - purple arrows) (Li et al., 2014b). Accordingly, AXL positively correlates with Snail, Slug and Twist1 expression in breast cancer cells (Gjerdrum et al., 2010). Moreover, in line with the inter-dependency of AXL and the EMT

phenotype, it has been shown that Vimentin expression functionally links both, EMT activation and AXL upregulation in breast cancer cells (Vuoriluoto et al., 2011). Furthermore, AXL expression is directly regulated by miR34a (Mudduluru et al., 2011), a microRNA that also targets Snail expression inhibiting EMT processes (Figure 1.4.B – grey box) (Siemens et al., 2011). Of therapeutic interest is the development of the small molecule inhibitor R428, also known as BGB324, which specifically blocks AXL and the downstream EMT-TFs in a dose-dependent manner in triple negative breast cancer. This AXL-EMT-TFs negative feedback loop ultimately downregulates the expression of the receptor at the cell surface (Holland et al., 2010). R428 is currently in phase-I clinical trials to prevent metastatic dissemination and circumvent AXL-mediated chemoresistant mechanisms in early stage carcinomas (Sheridan, 2013).

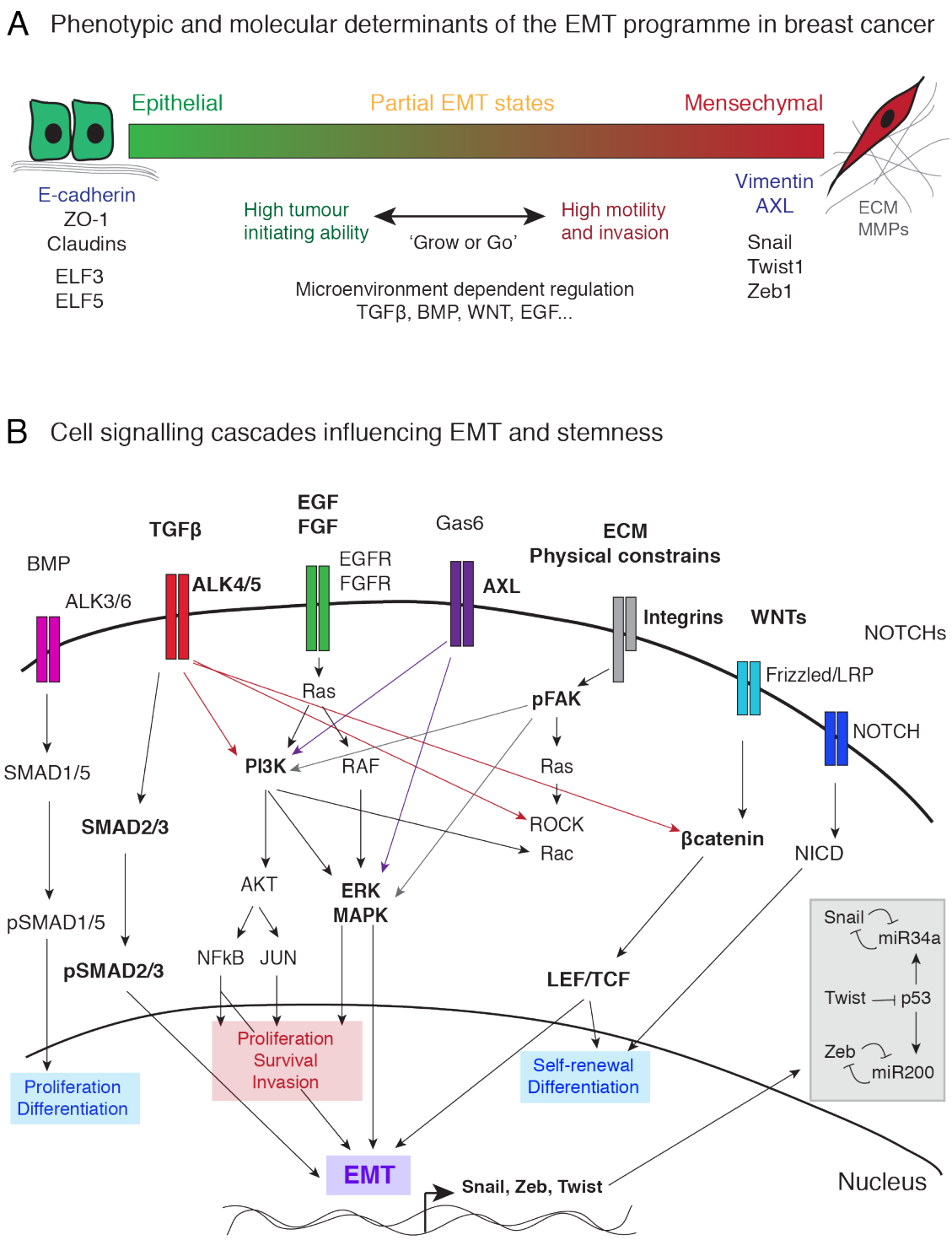


Figure 1.4 The EMT programme: phenotypic changes and molecular drivers
(A) Phenotypic characteristics and molecular changes of epithelial breast cancer cells undergoing epithelial-to-mesenchymal transitions. Epithelial cells in tissues maintain polarity thanks to the expression of cell-cell adhesion molecules such as E-cadherin, ZO-1 and claudins. As tumours develop the microenvironment changes and different signals (TGFβ, WNT, EGF) can trigger EMT through the activation of the core EMT-TFs (Snail, Twist1 and Zeb1). The expression of EMT-TFs drives the loss of epithelial characteristics, and the gain of mesenchymal markers (Vimentin,

AXL) and spindle-like morphology. Functionally, EMT induction implies the acquisition of ECM remodelling abilities (MMPs and ECM secretion), and changes in cytoskeletal regulation that enhances motility ultimately driving single cell dissemination. The multicolour scale represents an EMT gradient (green-epithelial to red-mesenchymal) where partial EMT states rather than an all-or-none regulation of the programme can occur in tumour cells. The mesenchymal/EMT features of tumour cells can be reverted also in a context dependent manner (BMP, lack of EMT inducers) driving cells re-epithelialisation, which functionally confers high proliferation. (B) Signalling pathways directly involved in the regulation of EMT (bold), and other related processes such as self-renewal, proliferation and survival that are molecularly closely related to EMT. TGFβ is the main upstream regulator of the EMT programme, directly regulating it through pSMAD2-3 that activates the expression of all core EMT-TFs, but also indirectly TGFβ activates PI3K–AKT, ERK MAPK and JUN pathways leading to EMT activation (red arrows). TGFβ in breast cancer also regulates the expression of the tyrosine kinase AXL, which is a downstream effector of the EMT programme that further maintains it through the activation of PI3K-ERK signalling (purple arrows). Different growth factors, such as EGF and FGF represent another major pathway to activate EMT through the RAS–RAF–MEK–ERK MAPK signalling cascade. Other signalling pathways such as WNT also participate in EMT; WNT signalling indirectly promotes EMT by inhibiting glycogen synthase kinase-3β (GSK3β) to stabilise β-catenin, which translocates to the nucleus to engage the transcription factors LEF and TCF promoting a gene expression programme that favours EMT. Last, biomechanical stimuli from the ECM are transduced via integrin-induced AKT activation regulating SNAIL through NFkB and JUN, or alternatively integrin-induced ERK activation (grey arrows). The grey box displays the connections between the inhibition of the tumour suppressor p53 and the EMT-TFs in cancer cells: Twist1 during tumour initiation represses p53, this leads to the repression of miR200 and miR34a, which trigger the expression of Zeb1 and Snail respectively exacerbating EMT in tumour cells.

EMT beyond single cell invasion: the link between EMT and stemness

The expression of EMT-TFs in carcinomas is evident and has been widely characterised at the invasive front of tumours, in both human samples and mouse models (De Craene and Berx, 2013). The link between EMT and the intrinsic stemness/tumour initiation abilities in cancer cells is not surprising since several common phenomena such as inflammation, physical constrains determined by the ECM, hypoxia and abnormal activation of TGFβ, WNT or NOTCH signalling are known to activate both (Figure 1.4.B). But the first study linking directly EMT and stem-like features related to malignant transformation was published in 2008; Morel et al. used human normal breast epithelial cells (HMLE) and exogenously expressed the oncogene Ras observing a gain in both, EMT mesenchymal features and tumorigenic activity (Morel et al., 2008). Conversely, Mani et al. induced the exogenous expression of the EMT transcription factor Twist1 in the same epithelial cell line (HMLE), resulting in the generation of cancer-stem like cells exhibiting self-renewal and tumour-initiating abilities (Mani et al., 2008). These findings imply that the EMT programme has additional roles beyond invasion, contributing to cancer cells tumorigenic abilities. Indeed, more recently Beck et al. used mouse genetics in a skin squamous cell carcinoma model where tumours arise from mutations in the Ras oncogene, previously shown to synergise with Twist1 (Morel et al., 2012), and revealed that Twist1 is required for skin tumour initiation and posterior tumour maintenance in a dosage-dependent manner (Beck et al., 2015). These findings corroborate the role of the EMT programme in tumorigenesis beyond tumour cell dissemination; and highlight the plasticity of cancer cells to transit between more epithelial or mesenchymal states during tumour initiation and progression.

Partial EMT states

Traditionally, the EMT process has been viewed as an all-or-none transition where once EMT transcription factors are induced epithelial cells will adopt a full mesenchymal state. Nevertheless, in many physiological processes such as wound healing where EMT is transiently induced in the damaged epithelium, partial EMT states can be observed. In these situations, cells can express EMT transcription factors but cell-cell contacts are maintained defining a partial EMT state (Arnoux et al., 2008). Additionally, in normal breast epithelial cells the induction of the EMT-TF

Twist1 leads to single cell dissemination but E-cadherin expression is maintained at the cell membrane (Shamir et al., 2014). This observation supports the existence of partial EMT states; in this case, Twist1 induces a more mesenchymal invasive phenotype in epithelial cells without the concomitant E-cadherin downregulation usually observed upon full EMT induction (Nieto, 2011). Indeed, during tumour initiation Twist1 is required in a dosage-dependent manner (Beck et al., 2015) suggesting that this EMT-TF can be finely regulated generating a gradient of Twist1-driven EMT states. Altogether, these observations indicate that a whole range of partial EMT states, rather than full epithelial/mesenchymal conversions occur in a context dependent manner.

In conclusion, in this section I describe how EMT processes are a favoured explanation to how single epithelial tumour cells initiate invade locally to ultimately disseminate throughout the body. Additionally, EMT is a complex programme that can induce global genomic changes affecting other cellular functions, such as cell survival or self-renewal that are key during tumour progression. I will further discuss the role of EMT in other steps of the metastatic process rather than dissemination in sections 1.2.2 and 1.4.

1.2.1.2 Collective cell migration

Collective invasion is the process by which cohesive groups of tumour cells that retain cell-cell junctions migrate into the surrounding stroma. The leading edge of the clusters generates the migratory force through ECM-cell contacts and cytoskeleton contractility pushing the attached neighbouring cells with them (Friedl et al., 2012). Cell-cell connections are mediated by cadherins, tight junction proteins and adhesion receptors; altogether they maintain front-rear directionality and synchronised cytoskeletal rearrangements (Hegerfeldt et al., 2002, Hidalgo-Carcedo et al., 2011). In breast cancer, collective cell invasion has been described to be led by fibroblasts, intrinsically mesenchymal cells, that generate microtracks in the ECM that then collective tumour cell strands follow to invade (Gaggioli et al., 2007). Alternatively, stromal independent collective invasion mechanisms occur and can be explained through cancer cell heterogeneity; a recent study by Cheung et al. describes $K14^+$ cells leading collective invasion across the different subtypes of breast cancers in human patient samples employing 3D invasion assays, and in mouse models *in vivo*. Interestingly, this $K14^+$ invasive leader cells do not undergo a molecular EMT programme. Changes in the ECM composition are suggested to contribute to the phenotypic invasive switch of these cells, inducing this K14 basal-epithelial programme that drives collective breast cancer invasion (Cheung et al., 2013). Over the last years, collective invasion is attracting increasing attention as a dissemination mechanism facilitating metastasis, supported by the clinical evidence of high numbers of cell clusters at the tumour-stromal interface in patient samples. However, although circulating tumour cell clusters can be isolated from patients blood, to date it remains unknown how clusters intravasate into the circulation. These findings will be further discussed in the next section (section 1.2.2)

In summary, changes in the cytoskeletal dynamics triggered by specific microenvironmental signals control cell behaviour in both migrating single cells or invasive leader cells driving collective invasion. In contrast, the two ways of invasion are distinctively characterised by the loss (single cell invasion) or maintenance (collective invasion) of adherens junctions between cells.

1.2.1.3 Single *versus* collective cell migration: one or multiple routes towards metastasis?

Xenograft models using breast cancer cell lines in mice were initially used to establish a causal relationship between EMT and metastatic dissemination. In this context, interfering with potent EMT inducers or the EMT-TFs has profound effects in dissemination (Heldin et al., 2012). Indeed, the first evidence of single cell invasion in tumours came from intravital multiphoton imaging in xenograft breast cancer models where single-cell dissemination at the stromal border was visualised (Wang et al., 2002). In the spontaneous transgenic MMTV-PyMT model, tumours display areas of high TGFβ-dependent pSMAD2-3 activity where local single cell migration and intravasation into lymphatic and blood vessels can be observed by sophisticated intravital imaging techniques (Forrester et al., 2005, Giampieri et al., 2009). The presence of this potent EMT inducer and single migrating cells predicts an active EMT programme, although a detailed examination of these pSMAD2-3 positive migrating cells regarding EMT markers has not been made. Importantly, overexpression of potent EMT inducers, such as TGFβ in late-stage MMTV-PyMT tumours dramatically increases metastatic burden, reinforcing the idea of a causal relationship between EMT at the primary tumour and metastatic dissemination (Muraoka-Cook et al., 2006). Overall, there is convincing evidence that mesenchymal tumour cells can be found in metastatic breast carcinomas, however its presence does not allow the conclusion that EMT is a pre-requisite for metastasis.

Despite this substantial body of evidence reporting EMT in mouse models as the main mean of single cell invasion at the primary site, its relevance is still debated due to the lack of available tools to observe this process in human samples (Garber, 2008, Ledford, 2011). Conventional histological analysis of human carcinoma samples does not allow the discrimination between mesenchymal stromal cells and epithelial cells that underwent EMT, as epithelial adhesion molecules typically used to recognise carcinoma cells (EPCAM, E-cadherin, claudins) are downregulated or lost during epithelial-to-mesenchymal transitions (Figure 1.4.A). The absence of incontestable EMT evidence in patients, together with the fact that histological analyses of tumours reveal the presence of all, single invading cells, compact

clusters of disseminated cells and elongated invasive strands of tumour cells still connected to the main tumour body (Figure 1.5) (Friedl et al., 2012), question whether EMT is the preferred mode of dissemination in human primary tumours.

Relevant to collective invasion is that in human carcinoma histological samples the most frequent units found at the invasive front of tumours are cancer epithelial cells organised as multicellular units (Friedl et al., 2012). Therefore, it has been proposed that collective invasion could be an important alternative mechanism to single cell EMT intravasation to metastasise, although to date little is known about this process. In a cooperative model, Giampieri et al. using intravital imaging in breast cancer mouse models found that cell clusters invade through the lymphatic system while single cells can infiltrate both lymphatic and vascular vessels (Giampieri et al., 2009). If EMT would be required for vascular intravasation, collectively invading cells could intravasate following TGFβ transient pulses mediated by the surrounding stromal cells, mainly TAMs (Tumour Associated Macrophages) and CAFs (Cancer Associated Fibroblasts) that secrete TGFβ and could mediate the switch from cohesive to single cell invasion at the blood vessel periphery. Indeed, as we will cover in the next section, there is now ample evidence of an active EMT programme in circulating tumour cells isolated from patients' blood (Armstrong et al., 2011). This provides convincing proof of EMT in patients after intravasation (section 1.2.2).

To date, the current evidence suggests that EMT certainly constitutes one of the preferred invasion modes leading to metastasis, but it is still under discussion whether other routes towards metastasis, such as collective invasion are required as well. The development of new experimental systems such as primary tumour organoids derived from patients' samples will allow elucidating this question.

Regarding therapeutic approaches targeting cancer cell dissemination, they can be useful in patients where surgical removal of primary tumours is not possible in order to prevent further tumour cell spreading and delay disease progression. However, the real problem is the presence of disseminated tumour cells at secondary sites, and therefore current therapeutic efforts are directed to develop therapies to prevent metastatic outgrowth from disseminated cells.

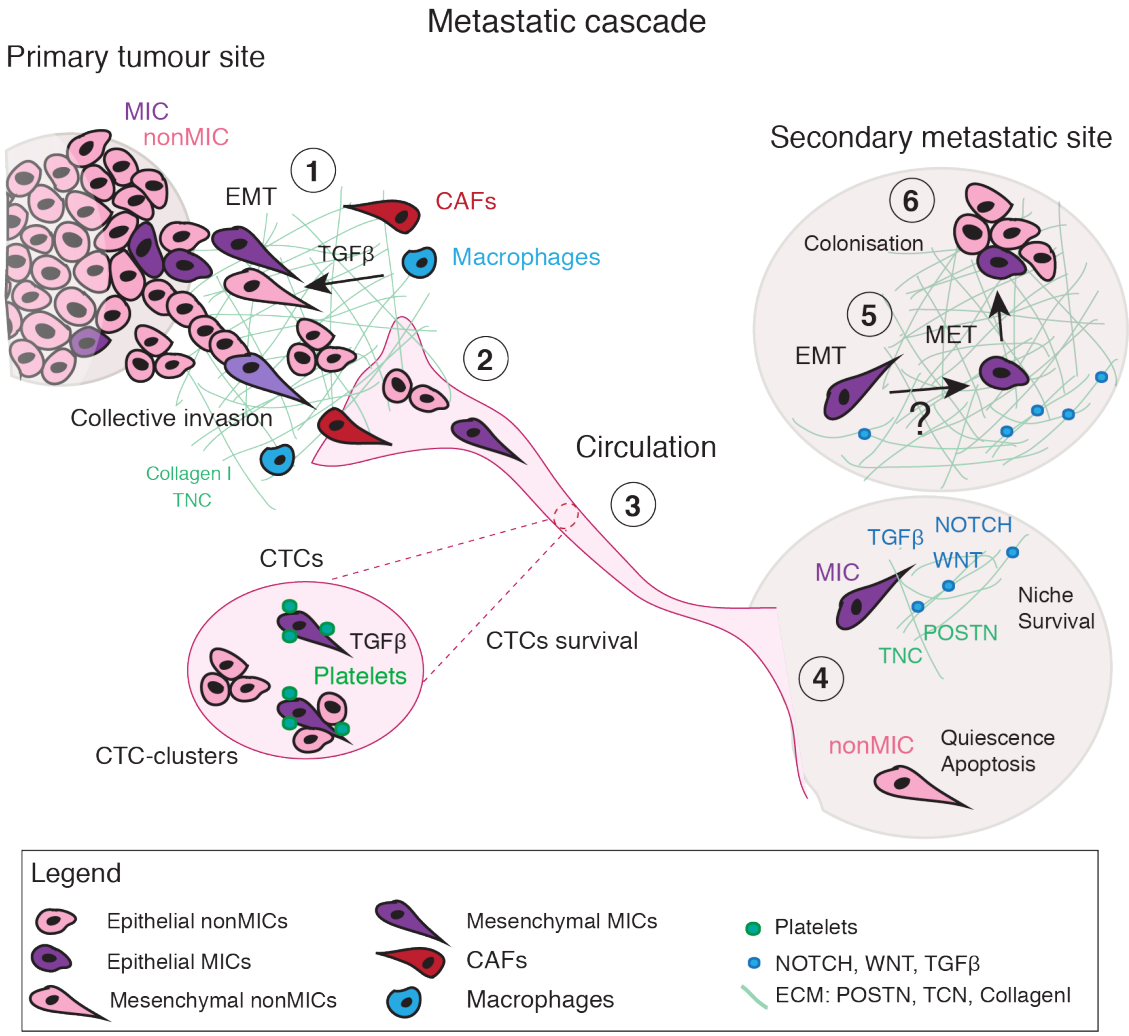


Figure 1.5 The multistep metastatic cascade
Metastasis is a complex multistep process by which tumour cells leave the primary site and reach distant organs of the body through the circulation. The process is usually divided in the following steps: (1) Local invasion: at the tumour-stromal border the presence of CAFs and macrophages and their secreted factors assist local collective and single cell invasion. (2) Intravasation. (3) Travelling in the blood stream CTCs are protected by platelets that secrete high levels of TGFβ enhancing survival. Both, single mesenchymal CTCs or clusters of CTCs containing epithelial and mesenchymal cells can be observed. (4) Extravasation at distant sites: in the foreign tissue only those disseminated cells with an intrinsic potential to metastasise will successfully overcome quiescence and induce a niche that will support their survival. (5) Early colonisation: a switch from a migratory to a proliferative phenotype needs to occur, although the triggers and molecular basis of this process are not understood. (6) Late colonisation and metastatic outgrowth: tumour cells in a proliferative state actively grow at the newly colonised tissue.

1.2.2 Travelling in the bloodstream: EMT in circulating tumour cells

Once the local invasion steps described above are completed, the first indication of dissemination is the presence of tumour cells in the draining lymph nodes that are in close proximity to the primary tumour. But the spread through the lymphatic system usually remains local, and dissemination to distant sites of the body occurs almost entirely through the blood (Hanahan and Weinberg, 2000). Intravasation is the process by which cancer cells migrate through the endothelial cells that constitute blood vessels reaching the blood stream (Wyckoff et al., 2007). Migrating cancer cells transitioning between the primary tumour and the metastatic sites isolated from the blood stream are termed Circulating Tumour Cells (CTCs). CTCs are widely studied as they have the potential to inform about intra-tumour heterogeneity and clonal evolution dynamics throughout the different stages of the disease and course of treatments in patients (Krebs et al., 2014).

In line with the different types of invasion observed in primary tumours, both single cells (CTCs) and cohesive cell clusters of at least 3 cells (CTC-clusters) have been isolated from the circulation (Figure 1.5) (Hou et al., 2012). However, one limitation when examining CTC-clusters is that the possibility of cell clumping during blood sample processing cannot be entirely excluded (Hou et al., 2012). Importantly Aceto et al. transplanting different fluorescently label tumours in a spontaneous metastatic mouse model excluded that clusters are formed by cell clumping in the circulation *in vivo*, but rather leave the primary tumour as cohesive multicellular units that can metastasise (Aceto et al., 2014). Nevertheless, although CTC-clusters are rare compare to single cells in circulation, they have been shown to possess enhanced survival and thereby metastatic potential contributing to poor prognosis in lung and breast cancer patients (Hou et al., 2012, Aceto et al., 2014). But as previously indicated, it remains unknown how epithelial cell clusters intravasate and extravasate during metastasis.

Regarding the EMT status of cells in the blood stream, traditionally CTCs were isolated using biased antigen specific techniques, i.e. EPCAM-based isolations that would only isolate tumour cells in an epithelial state. However, in recent years technological improvements have allowed the development of non-antigen

dependent techniques (Krebs et al., 2014) that have demonstrated the EMT heterogeneity displayed by CTCs in lung, breast, prostate and other cancer types (Kalluri and Weinberg, 2009, Kallergi et al., 2011, Armstrong et al., 2011). Indeed, heterogeneity in the expression of EMT markers has been observed in single CTCs, that can be found in epithelial, mesenchymal or partial EMT states (Armstrong et al., 2011). CTC-clusters reveal an even more interesting scenario: despite the fact that cells adhere together suggesting a complete epithelial phenotype, clusters contain cells that display partial EMT phenotypes. Remarkably, Vimentin-expressing cells form part of these cohesive cell groups (Hou et al., 2011) arguing against the mandatory cell-cell contact loss described when undergoing mesenchymal transitions (Nieto, 2011). Additionally, in breast cancer patients Yu et al. reported CTCs to be in a mixed epithelial-mesenchymal phenotype. Using *in situ* hybridisation probes for a panel of EMT markers they uncover a gradient of EMT states along with TGFβ activation in CTCs (Yu et al., 2013). Moreover, they described an association between the expression of mesenchymal markers and clusters of circulating tumour cells, which further supports the idea of partial EMT states being compatible with cell-cell adhesion maintenance (Yu et al., 2013). Also, most CTCs analysed in the bone marrow are in a non-proliferative quiescence state as they are travelling (Muller et al., 2005, Pantel et al., 1993), which correlates with their TGF β -driven increased expression in EMT markers (Yu et al., 2013) and the activation of migratory pathways (Patsialou et al., 2015), illustrating the previously mentioned 'grow or go' alternative states adopted by cancer cells in different moments of the metastatic cascade. Notably, in this study they also observed that after chemotherapy more mesenchymal cells were enriched in patients (Yu et al., 2013). This highlights the reported correlation between EMT and stem-like characteristics such as chemotherapy resistance (Hennessy et al., 2009, Asiedu et al., 2013).

In summary, CTCs studies have been crucial to provide evidence of epithelial-to-mesenchymal transitions, collective cell migration and cells with stem-like properties within the circulation (Figure 1.5). As CTCs can be easily acquired through non-invasive techniques from patients compared to tumour biopsies, they can be used to design personalised therapies throughout the course of disease.

Analysing the dynamic biological properties of CTCs and the evolution of their mutational landscape post-treatment can help the prevention of relapses.

1.2.3 Reaching the target site: the bottleneck of metastasis

As we analyse every step of the metastatic process it becomes clear that metastatic success is a function of many variables. As previously discussed (section 1.2.1), metastasis is an extremely inefficient process where there is a very small probability for cancer cells to succeed. To proceed towards colonisation, disseminated cancer cells within the foreign microenvironment of the metastatic site will need to evade the immune-surveillance mechanisms in place and avoid dormancy in the absence of a supportive microenvironment to grow (Valiente et al., 2014, Luzzi et al., 1998, Nguyen et al., 2009). Moreover, if these initial difficulties are overcome, metastatic cells will still need to switch the migratory phenotype they display upon arrival to the secondary site towards a more epithelial phenotype compatible with proliferation. All these steps make metastatic colonisation the bottleneck of the metastatic process; nevertheless it is the most life-threatening as well. Therefore, it is of great interest to understand the mechanisms underlying colonisation for each cancer type, trying to find the commonalities among these programs to develop effective broad-spectrum anti-metastatic therapies.

CTCs will exit circulation at distant organs, as they are physically restrained in the microvasculature. Therefore, in each cancer type the favoured metastatic sites are influenced by proximity and circulation patterns (Nguyen et al., 2009). However, beyond the passive role of circulation, it is known that different tumour types have preferred metastatic sites. As proposed by the ‘seed and soil’ theory (Paget, 1989), metastasis depends on the combination of the intrinsic tumour cell features (the seed), such as the ability to form invadopodia (Leong et al., 2014), secrete specific proteases (Sevenich et al., 2014) or recruit stromal components that will facilitate extravasation (Labelle et al., 2011, Reymond et al., 2013), as well as on extrinsic aspects such as the vasculature composition and the microenvironmental factors

present in a given site (the soil) (Nguyen et al., 2009). In agreement with this theory, breast cancer typically metastasises to bones, lungs and brain, and several studies have identified in xenograft models the organ-specific traits expressed by breast tumour cells to successfully colonise those tissues. Massague's laboratory created a model of tissue-specific experimental metastasis by intracardially injecting MDA-MB-231 cells into immunocompromised mice, and isolating them afterwards from the different metastatic sites. After several rounds of experimental metastasis injection and isolation, they generated functionally distinct MDA-MB-231 clones that preferably metastasise to the lung, brain or bones. Using the, breast to bone MDA-MB-231 pre-selected cells, Zhang et al. showed that the expression of a Src-associated gene set is required for breast to bone metastasis (Zhang et al., 2013). In the specific breast to brain metastatic variant, the proteolytic enzyme Cathepsin-S has been described to be crucial mediating the blood-brain barrier transmigration of MDA-MB-231 cells allowing extravasation to the brain (Sevenich et al., 2014). These findings highlight the relevance of tumour cell heterogeneity to generate cells with different organ-specific metastatic traits (seeds) within the same primary tumour, which allows adaptation to the requirements needed to metastasise to different organs.

Typically, extravasated tumour cells at metastatic sites exist in two modalities, solitary dormant cells or actively growing colonies (Figure 1.5). Cellular dormancy describes a state of mitotic arrest in which cells exit cell cycle and enter G_o phase (Naumov et al., 2002). It is known that disseminated cells can remain latent or dormant for years after primary tumour excision in patients, and despite the clinical relevance of this fact little is known about how cells enter in this dormant state and what triggers the eventual exit (Sosa et al., 2014). What determines the entrance of disseminated tumour cells into dormancy upon extravasation is not fully understood, but growing evidence shows that the particular microenvironment composition can determine the initial tumour cell survival and later outgrowth (Oskarsson et al., 2014). In a recent study in head and neck squamous cell carcinoma (HNSCC), Bragado et al. showed how the local microenvironment (soil) to which tumour cells arrive is also a key determinant for metastatic success. They found TGFβ2 as a key microenvironmental determinant of cell dormancy in the restrictive environment of the bone marrow compared to the permissive metastatic lung soil. Systemic

inhibition of TGFβ-RI awakens dormant disseminated cells and fuels multi-organ metastasis (Bragado et al., 2013). In a similar study in breast cancer, Gao et al. reported that BMP signals in the lung parenchyma maintain breast tumour cell in a dormant state upon extravasation. Only certain tumour cells that express the BMP antagonist Coco overcome dormancy and successfully grow metastases (Gao et al., 2012). Altogether, the above studies support the long suggested seed and soil hypothesis (Paget, 1989) as a prerequisite to metastasise.

In conclusion, successful metastatic initiation depends on both, the intrinsic cancer cell abilities to extravasate and adapt to the foreign microenvironment at distant sites, and the organ-specific microenvironmental factors that will help or hinder each of these steps (Figure 1.5). In line with the critical role of the microenvironment, it is known that in order to successfully metastasise, disseminated tumour cells need to 'educate' the microenvironment they arrive to creating a favourable surrounding to grow (Malanchi et al., 2012, Irmisch and Huelsken, 2013). These concept of an active 'education' of the microenvironment by metastatic cells leads to the idea of Metastasis-initiating Cells (MICs), a distinct population of tumour cells that possess unique characteristics to initiate metastasis actively inducing a supportive microenvironment or 'metastatic niche' upon extravasation. Due to the relevance of these ideas for the work presented in this thesis, I will further develop them in the next section (section 1.3).

1.3 Metastatic colonisation: MICs and metastatic niches

As described in the previous section, the metastatic process consists of a series of sequential complex challenges that tumour cells need to overcome to successfully grow at distant sites. In this section, we will summarise the recent discoveries indicating that only certain populations within tumours possess the intrinsic abilities required to achieve metastatic outgrowth at distant sites, the so-called metastasis-initiating cells (MICs). As illustrated throughout this introductory chapter, tumour progression certainly depends on a favourable microenvironment (section 1.1.2). This dependency becomes even more critical during metastasis when disseminated cells arrive into a naive tissue and need to re-initiate a supportive stroma, the metastatic niche, which will trigger and sustain tumour cell growth.

Following on the founding ideas of the 'seed and soil' theory that illustrate the dependency of the metastatic cells (seeds) on the microenvironment (soil) (section 1.2.3), the MIC concept implies an active role of this distinct population of cells to promote a favourable microenvironment (or niche) to grow (Malanchi, 2013). Although several metastatic populations might coexist during dissemination exhibiting site-specific extravasation and seeding capabilities, ultimately their metastatic outgrowth will depend on their capacity to directly or indirectly generate a metastatic niche. In light of these notions, the next sections aim to revise the essential role of the tumour-driven stromal crosstalks required upon arrival to distant tissues to trigger a metastatic niche that will further support tumour growth.

1.3.1 Metastasis-initiating cells (MICs)

As previously described (section 1.1.1.3), tumours are organised hierarchically as their tissue of origin, where a small subpool of cells with self-renewal capacity, termed cancer stem cells (CSCs), is responsible for the initiation and maintenance of the tumour mass (Malanchi et al., 2008, Singh et al., 2004, O'Brien et al., 2007, Wang et al., 2009). However, it remains to be determined if metastases display a similar hierarchy to the one observed in primary tumours.

Several studies have shown that there are specific populations within tumours that display high metastasis-initiating ability compare to the bulk of tumour cells, and

they are accordingly termed metastasis-initiating cells (MICs) (Hermann et al., 2007, Pang et al., 2010, Malanchi et al., 2012). This functional definition refers to those disseminated tumour cells that are capable of re-initiating macroscopic tumour growth at secondary sites (Oskarsson et al., 2014). Importantly, evidence from the clinic indicates that high expression of adult stem cell markers in primary tumours correlate with poor survival and relapse in patients (Dalerba et al., 2011, Pece et al., 2010).

Studies in animal models have allowed the identification of cell populations that drive metastatic colonisation. In the breast cancer metastatic mouse model MMTV-PyMT, Malanchi et al. described a population of MICs within the primary tumours defined by the co-expression of CD24 and CD90. CD24 was previously characterised as an epithelial marker in the mammary gland tissue, where lineage negative, $CD24^-$ cells where functionally characterised as non-epithelial (pan-cytokeratin negative cells), $CD24^{low}$ contains the myoepithelial mammary gland cells (co-expressing cytokeratin-5) and $CD24^{high}$ contains the luminal epithelial cells (co-expressing cytokeatin-8) (Sleeman et al., 2006). The marker CD90, also known as Thy-1, is typically expressed in the immune cell compartment of the tumour ($CD45^+$), mainly in T cells (Haeryfar and Hoskin, 2004). Also some adipocytic and fibroblastic-like populations of mesenchymal origin can express CD90 in the mammary gland, but they do not display the epithelial CD24 marker (Paunescu et al., 2011). In the mammary epithelium, CD90 is expressed in bipotent normal mammary epithelial progenitor cells in human (Raouf et al., 2008), and by a small subpool, 1-4% of the $CD24^+$ mammary epithelial cells in the MMTV-Wnt1 mouse model of breast cancer (Cho et al., 2008). Interestingly, this rare $CD24^+CD90^+$ population displayed high tumour-initiating abilities when isolated from MMTV-Wnt1 tumours and tested in serial dilution transplantation assays (Cho et al., 2008). Moreover, in primary human breast carcinomas CD90 has been recently shown to be expressed in a small subpool of epithelial cells with increased tumour-initiating abilities, where CD90 functionally mediates an interaction with TAMs (tumour-associated macrophages) that facilitate the dissemination of this $CD90^+$ tumour cells (Lu et al., 2014). Malanchi et al. functionally defined this lineage negative ($CD45^-$, $Ter119^-$ and $CD31^-$), $CD24^+CD90^+$ subpool in the MMTV-PyMT tumours by their exclusive ability to initiate metastatic growth in experimental metastasis assays: $CD24^+CD90^+$ and $CD24^+CD90^-$ tumour cells were freshly isolated from

primary late-stage PyMT carcinomas by cell sorting, and injected intravenously directly into the circulation of immune-deficient mice. In this setting, only the $CD24^{+}CD90^{+}$ population was capable of efficiently metastasising, defining a MIC population responsible for metastatic colonisation in this model (Malanchi et al., 2012). Further characterisation of this rare $CD24^{+}CD90^{+}$ population with exclusive metastatic ability in the MMTV-PyMT model, in order to find novel mechanisms that allow breast to lung metastatic colonisation, is the starting point of thesis.
Additionally, Baccelli et al. detected MICs expressing stem cell markers in breast cancer patient blood that when isolated and inoculated into immune-deficient recipient mice generated bone, liver and lung metastasis. The same study showed that high numbers of this MIC population ($EPCAM^{+}CD44^{+}CD47^{+}MET^{+}$) in the blood of patients correlates with lower survival and increased number of metastatic sites (Baccelli et al., 2013), underlining the importance of characterising the mechanisms by which these cells mediate metastatic colonisation.

One central issue that remains elusive is whether MICs originate from tumour-initiating cells (TICs) that acquire new priming mutations to metastasise, maintaining the intrinsic stemness capacity. In TICs it has been shown that mutations impacting on proliferation, self-renewal and survival pathways drive tumour initiation (Vogelstein et al., 2013), so the expectation would be that further clonal evolution would generate new subclones that acquire pro-metastatic mutations. However, genome sequencing analysis of metastases and matched tumours in pancreatic cancer patients have revealed little evidence for metastasis-specific driver mutations other than the mutations found in classic oncogenes that are already present at the primary site (Yachida et al., 2010). Therefore metastases are unlikely to be the direct consequence of the acquisition of pro-metastatic mutations, but rather the consequence of epigenetic programmes that modulate the activation of invasive and survival programs, such as EMT (section 1.1.1.3) (Scheel and Weinberg, 2011, Quail et al., 2012). In line with this idea, MICs are likely to arise from TICs clones that acquire metastatic capacities through epigenetic modulations that transitorily activate invasion and survival programmes that allow them to disseminate, survive in the circulation and extravasate at distant metastatic sites.

1.3.2 Emerging metastatic niches

1.3.2.1 The concept of niche

In developmental biology, niches are functionally distinct microenvironments formed by specialised cells and ECM components that regulate normal stem cell function, controlling the required equilibrium between quiescence, and symmetric and asymmetric division to maintain tissue homeostasis (Sneddon and Werb, 2007). In some organs, such as brain or skin, the availability of bona-fide stem cell markers combined with lineage tracing techniques have allowed the visualisation of the normal stem cell niche, and its evolution during tumour initiation processes (Beck et al., 2011, Calabrese et al., 2007). However, the concept of niche in cancer to define a specific location where CSCs reside as in normal stem cell biology cannot be applied due to the disorganised nature of tumours. Therefore, in cancer the term niche is generally used to describe the compendium of microenvironmental cues that are essentially required by CSCs to initiate and maintain tumour growth.

In line with this definition of niche in the tumour context, increasing similitudes between normal and tumour stem cells microenvironments continue to be described. For instance, many factors produced by fibroblasts that comprise the normal stem cell niche, as such BMP antagonists that limit differentiation are also found in the tumour cell niche (Sneddon et al., 2006). In normal skin, hair follicle stem cell (HFSC) niche soluble factors need to be critically regulated to maintain the balance between quiescence and proliferative states such as BMPs and TGFβ2 respectively; the same factors are found to be deregulated in squamous skin carcinoma (SSC) mediating the aberrant proliferation of CSCs (Oshimori et al., 2015). Likewise, there are many ECM factors that are known components of the adult stem cell niche and can be found in the tumour microenvironment as well (Wong and Rustgi, 2013). The glycoprotein Tenascin-C (TNC) modulates EGF signalling in adult brain stem cells through limiting the availability of soluble FGF2 and BMP4 signals in the stem cell niche. This EGF regulation is crucial to ensure adult stem cell function and tissue homeostasis in the brain (Garcion et al., 2004).

Equally, TNC have been found to be a critical autocrine regulator of CSCs in glioma (Nie et al., 2015). Therefore, as their normal counterparts, cancer stem cells are thought to reside in niches within the tumour microenvironment that stimulate their self-renewal ability, induce angiogenesis and contribute to the recruitment of other stromal cells to the initiation site (Plaks et al., 2015).

1.3.2.2 The metastatic niche

We have analysed how after extravasation disseminated tumour cells first need to overcome dormancy and survive in the foreign tissue (section 1.2.3). Subsequentially those tumour cells with the intrinsic abilities required to colonise the new site (MICs), will trigger a crosstalk of signals with the naïve tissue microenvironment inducing a metastatic niche that will ultimately enable successful metastatic growth. In contrast to the situation in primary tumours during metastatic colonisation single disseminated MICs can be visualised at the metastatic tissue, and therefore in this setting the niche induction and its evolution can be monitored.

The first support from the microenvironment after extravasation at the target tissue comes from different stromal cells that promote tumour cell survival. In breast cancer metastasis to the lungs, survival at the target site has been shown to depend on a paracrine interaction with resident lung macrophages that are stimulated by tumour-derived colony-stimulating factor-1 (CSF1). In turn, macrophages produce epidermal growth factor (EGF) that promotes tumour cell survival through AKT signalling activation (Patsialou et al., 2009). Similarly, endothelial cells offer a perivascular niche to breast cancer cells upon extravasation to the brain promoting survival. Valiente et al. have visualised how single tumour cells expressing serpins tightly adhere to the blood vessel wall upon extravasation to the brain. In this perivascular niche, endothelial cells provide the oxygen and nutrient support that determines tumour cell survival in this new environment (Valiente et al., 2014).

Next, the surrounding microenvironment will determine whether tumour cells undergo dormancy or start proliferating. In breast cancer metastasis to the lungs, the perivascular niche has been shown to regulate tumour cell dormancy after extravasation in a dynamic fashion. Ghajar et al. have shown how immediately after extravasation, the endothelial cells from the naïve lung vasculature secrete Thrombospondin-1 (THBS1) promoting breast cancer cells quiescence. However, in those vessels where tumour cells trigger angiogenic pathways, vessel sprouting is accompanied by TGFβ and periostin secretion, leading to micrometastatic growth (Ghajar et al., 2013).

Those cells that survive and are equipped with the intrinsic machinery required to overcome pro-dormancy signals and induce a favourable microenvironment, will succeed and metastasise. Several studies have described how disseminated metastatic cells or their recruited stroma induce ECM changes that are crucial to sustain cancer cells self-renewal capacity and re-initiate tumour growth. In mouse models of breast cancer metastasis to the lungs, Malanchi et al. described the crucial role of Periostin (POSTN) in regulating the self-renewal and differentiation of MICs driving metastasis. POSTN is an ECM proteoglycan secreted by fibroblastic stroma that serves as a trap for Wnt ligands, amplifying the limited stromal derived Wnt and creating a Wnt rich niche around disseminated MICs that supports self-renewal maintenance. Importantly, a POSTN free microenvironment (POSTN knockout mice) impaired wild-type MICs ability to colonise the lungs, highlighting the crucial role of this niche ECM protein for metastatic outgrowth (Malanchi et al., 2012). A similar ECM mediated mechanism was described by Oskarsson et al. where Tenascin-C (TNC) secreted by both, cancer cells and fibroblast, creates a niche were Notch signalling is amplified activating tumour cell self-renewal in breast to lung metastatic colonisation (Oskarsson et al., 2011). These studies highlight the decisive role of the ECM in regulating the availability of key soluble signals that maintain self-renewal capacity at the metastatic site, allowing colonisation.

In conclusion, the microenvironmental factors present in the naïve tissue microenvironment upon extravasation and the intrinsic abilities of cancer cells to create a supportive niche determine successful metastatic colonisation. The role of the metastatic niche and its relationship with the fact that only some cancer cells

(MICs) can successfully metastasise, is currently under systematic investigation to elucidate key unanswered questions:

- Do MICs have an enhanced ability to recruit and induce a favourable metastatic niche that will support their growth compared to the non-metastatic cancer cells of the tumour?

- In that case, what are the molecular programmes underlying the crosstalk initiated by MICs to educate the stroma?

- Does the ability to induce a niche upon arrival to the target site determine the entrance in dormancy or the switch to proliferation?

1.4 MICs plasticity: focus on EMT modulation during the metastatic cascade

Since colonisation abilities are required as soon as cancer cells reach the metastatic site (section 1.2.3) it is likely that pro-metastatic programmes are already induced at the primary site. Indeed, we have discussed earlier how the unique stromal composition at the invasive tumour front, rich in ECM and signals such as TGFβ, Wnt, Notch and pro-inflammatory cytokines, triggers cell dissemination though the activation of different downstream pathways (section 1.2.1.1) (Figure 1.5). Interestingly, EMT is one of the induced programmes at the primary tumour border that mediates invasion but also confers disseminating tumour cells with additional roles. For instance, TGFβ-mediated EMT induction is active in CTCs as they navigate in the vascular system and enhances their survival (Figure 1.6.B) (section 1.2.2). Moreover, induction of the EMT transcription factors induces epigenetic plasticity in tumour cells boosting stem-like features and tumour-initiating properties (section 1.2.1.1) (Chaffer et al., 2013). In general, activation of the EMT programme can induce phenotypic, metabolic and functional cell plasticity in MICs through epigenetic changes (Tam and Weinberg, 2013). Therefore, it is reasonable to hypothesise this EMT-driven plasticity increases cell adaptability to the microenvironmental changes encountered throughout the complex multistep metastatic process.

As stated above, many cells leave the primary tumours displaying EMT features (section 1.2.1.1) that are maintained within the circulation (section 1.2.2). Therefore, the establishment of a direct link between EMT and tumour-initiating abilities would imply that disseminated EMT cells with higher stemness abilities had an increased metastatic potential. However, carcinoma metastases typically display epithelial differentiated features questioning the role of the EMT status of cancer cells in metastatic progression (Brabletz, 2012b). This dichotomy can be explained by the re-acquisition of epithelial traits during the metastatic colonisation step: EMT is induced at the primary site for tumour cells to disseminate, and once at the target site the reverse process MET (mesenchymal-to-epithelial transition) will induce a

cell status more compatible with proliferation (Figure 1.6.B). MET is a process that typically occurs during embryogenesis to reverse EMT allowing re-epithelialization after cell migration, and differentiation into multiple cell types in the embryo (Nieto, 2013). Although many of the molecular bases of EMT have been elucidated (Thiery et al., 2009), the signals involved in the induction of MET have not been well characterised. The most studied MET process in embryogenesis is the epithelialisation of the paraxial and intermediate mesoderm to form somites and precursors of the renal system, respectively. During this process, an increasing gradient of BMPs, the most prominent epithelialising agents, counteracts TGFβ signalling (Nieto, 2013). Several processes have been described where BMP7 antagonises TGFβ inducing re-epithelialisation, such as kidney development (Dudley et al., 1995), reprogramming of fibroblasts (Samavarchi-Tehrani et al., 2010) and in prostate cancer cells (Buijs et al., 2007). Although the molecular drivers of MET remain to be determined, what it is generally required for re-epithelialisation is the downregulation of the corresponding EMT-TFs (Nieto, 2011). Therefore, to date MET might be more precisely referred as an EMT inhibition or reversion.

In metastasis two key studies support the idea that EMT is reverted at the target site. Tsai et al. using an inducible Twist1 system in a spontaneous model of squamous cell carcinoma showed that exogenous Twist1 expression boosted cell dissemination as expected, but did not allow efficient metastatic colonisation; when after tumour cell dissemination Twist1 expression was turned off metastatic colonisation occurred, demonstrating *in vivo* the requirement of an EMT reversion at the target site (Tsai et al., 2012). In the second study, Ocaña et al. used the highly mesenchymal breast cancer cell line BT549 that expresses the potent EMT inducer Prrx1. When injected in a mouse model of experimental metastasis via tail vein injection, BT549 cells reached the metastatic site but did not grow. Inhibiting Prrx1 in BT549 cells was sufficient for metastatic colonisation to happen (Ocana et al., 2012). Both studies showed the requirement of an EMT reversion process to achieve metastatic colonisation.

Interestingly, the later study by Ocana et al. revealed that the potent EMT induction mediated by Prrx1 inversely correlated with stemness, questioning the previous

described direct link between EMT and stemness (Mani et al., 2008). These conflicting results can be harmonised by the idea of a gradient EMT model where existing partial EMT states are compatible with stemness (Ombrato and Malanchi, 2014). The EMT gradient model proposes that EMT modulations in tumour-initiating cells would be controlled in a dosage dependent manner to be compatible with stem-like features throughout tumorigenic evolution (Figure 1.6.A). As previously described (section 1.2.1.1), at the molecular level the EMT programme is induced by the expression of some of the core EMT-TFs (Twist1/2, Snail1, Slug and Zeb1/2). Functionally, the EMT-TFs trigger increased motility and invasion along with stemness abilities in epithelial cells that undergo EMT for the first time (Figure 1.6.A). According to this EMT gradient model, a further activation of the EMT programme in partially mesenchymal cells that already express Snail, Zeb or Twist would exacerbate their mesenchymal features and push these tumour cells outside the 'stemness window' (Figure 1.6.A - late EMT programme). This hypothesis can explain the controversial results described by Ocana et al. in relationship to an inverse correlation between EMT and stemness in breast cancer cells: the expression of the potent EMT inducer Prrx1 exacerbates the mesenchymal features of cancer cells driving them outside the 'stemness window' limits, therefore Prrx1 does not provide a gain, but a loss in stemness and cell plasticity (Figure 1.6.A) (Ocana et al., 2012).
Therefore, according to this EMT gradient model (Ombrato and Malanchi, 2014) the same tumour-initiating cells (TICs) defined depending on their stemness potential would transit from intermediate epithelial to mesenchymal states preserving stemness within some limits in response to the different microenvironmental cues (Figure 1.6.A). Consequently, TICs preserving stem-like features could be found in a motile mesenchymal state or in a stationary proliferative phenotype at the primary site depending on their location.

About the cellular effectors driving EMT reversion, as discussed above it has been shown that Twist1 needs to be down-regulated (Tsai et al., 2012) to allow metastatic colonisation. Moreover, supporting the role of inhibitor of differentiation (ID) proteins as master regulators of stemness in normal tissue stem cells and brain CSCs (Niola et al., 2012, Niola et al., 2013), it has been described that ID1 expression is required for efficient lung metastatic colonisation of breast cancer

cells (Gupta et al., 2007). Indeed, a recent study suggested that ID1 expression opposes to Twist1 in mesenchymal breast cancer cell lines *in vitro*, but the molecular mechanism underlying this potential ID1-mediated Twist1 downregulation remains unclear (Stankic et al., 2013). Other potential EMT inhibitors that could antagonise the EMT-TFs are the microRNAs miR200 (Dykxhoorn et al., 2009) and miR34a (Siemens et al., 2011) that antagonise Zeb2 and Snail respectively (Figure 1.6.A and 1.4.B – grey box). In line with the EMT gradient model, the activation of negative EMT regulators, such as ID1, miR200 and miR34a could potentially drive a mesenchymal reversion or MET that would switch MICs back to a more epithelial state compatible with proliferation, but different to the original epithelial state before the cells underwent EMT, now inside the 'stemness window' (Figure 1.6.A).

During the metastatic cascade, EMT is triggered at the invasive front of tumours where the stroma provides high levels of the EMT inducer TGFβ (Figure 1.6.B, Primary site) (section 1.2.1.1). In line with the hypothetical EMT gradient model, MICs undergo EMT acquiring partially mesenchymal features and an enhanced stemness state (i.e. poised chromatin status as defined in section 1.1.1.3). This is in agreement with the observation that mesenchymal CSCs are subsequently found in the blood stream (Yu et al., 2013), where EMT is further sustained by platelet-derived TGFβ (Figure 1.6.B, Circulation) (Labelle et al., 2011). In line with the reported EMT downregulation required for metastatic colonisation, extravasated CTCs need to leave the transient migratory/quiescence state they exhibit in the circulation and re-acquire a proliferative status compatible with metastatic outgrowth (Gupta et al., 2007). This illustrates the previously proposed 'go or grow' dichotomy (section 1.2.2) observed at the primary site, where an autocrine CSF1R signalling downstream of TGFβ is essential for inhibiting cell proliferation, while it activates single cell invasion in a human xenograft breast cancer model (Patsialou et al., 2015). This TGFβ-dependent mechanism exemplifies the migratory and proliferative cellular states as antagonistic, and linked to a more mesenchymal and a more epithelial phenotype respectively (Figure 1.4.A) (Patsialou et al., 2015).

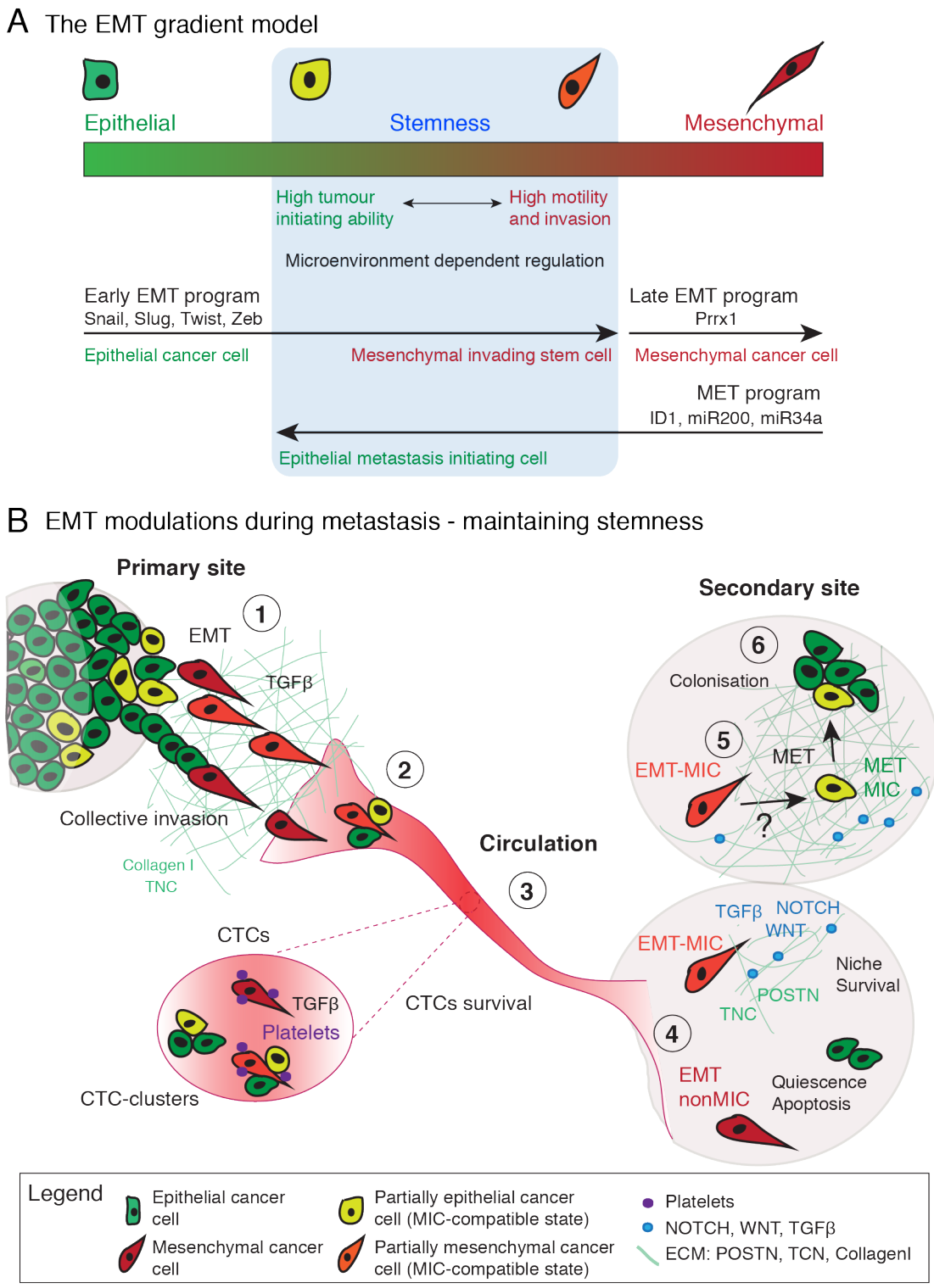


Figure 1.6 The EMT gradient model and its implications in the metastatic cascade

(A) EMT gradient model (adapted from Ombrato and Malanchi, 2014). Metastasis-initiating cells (MICs) transit between partial more epithelial or more mesenchymal states within a 'stemness window' where cells maintain epithelial plasticity preserving stem-like features. The different phenotypes adopted by the stem cells

within this ‘stemness window’ are regulated by the microenvironment. EMT gradient model explanation: (1) epithelial cells undergo EMT in a biphasic fashion. A first induction of the EMT programme (driven by Snail, Twist and Zeb) confers partial mesenchymal features compatible with stem-like characteristics. If a further exacerbation of this EMT phenotype (driven by potent EMT activators, such as Prrx1) occurs, cells would acquire a fully mesenchymal state incompatible with epithelial plasticity and would imply the loss of stem-like characteristics. (2) The reverse process, an EMT inhibition or MET, allows cancer stem cells in a mesenchymal motile state to re-gain a more epithelial phenotype compatible with proliferation (driven by miR200, miR34a and ID1), and the stem cell properties of self-renewal and differentiation. (B) EMT modulations during metastasis. (1) Local invasion: MICs and nonMICs disseminate from primary tumours as single partially mesenchymal cells (nonMICs-red, MICs-orange), or collectively as epithelial clusters (nonMICs-green, MICS-yellow). (2) Intravasation. (3) Travelling in the blood stream CTCs are protected by platelets that secrete high levels of TGFβ enhancing their survival and EMT features. Both, single epithelial and mesenchymal CTCs or CTC-clusters containing epithelial and mesenchymal cells can be observed in blood patient samples. (4) Extravasation at distant sites: in the foreign tissue only those cells with an intrinsic potential to metastasise will successfully overcome quiescence and induce a niche that will support their survival. (5) Early colonisation: the switch from a migratory (EMT) to a more epithelial proliferative phenotype (through MET or a mesenchymal inhibition) needs to occur for the cells to actively divide, although the triggers and molecular basis of this process are not understood. (6) Late colonisation: tumour cells in a more epithelial proliferative state actively grow at the newly colonised tissue, forming epithelial macro-metastases.

In light of these results, the current working model during metastasis is that once at the secondary site metastatic cells displaying mesenchymal features need to switch back to a more epithelial phenotype compatible with proliferation (Figure 1.6.B – 6 Secondary site). In agreement with the EMT gradient model (Figure 1.6.A), the same MICs defined depending on their stemness potential would transit from intermediate mesenchymal to a more epithelial state inside the 'stemness window' in response to the different microenvironmental cues received throughout the metastatic process (Figure 1.6.B). Accordingly, MICs could be found in a motile mesenchymal state or in a stationary proliferative phenotype (Figure 1.6.A) (Ombrato and Malanchi, 2014).

Molecularly, little is known about the microenvironmental factors triggering EMT reversion, as well as the downstream effectors that mediate the change towards a more epithelial phenotype. As analysed throughout this introduction, the dynamic bidirectional communications between tumour cells and the microenvironment determine overall disease progression (section 1.1.2). Particularly, metastatic colonisation depends on the microenvironmental cues as much as on the cell intrinsic properties (section 1.3). Accordingly, it is sensible to hypothesise that the lack of microenvironmental pro-EMT signals, such as TGFβ, or the presence of epithelializing factors, such as BMPs, at the metastatic niche could modulate the EMT phenotype of disseminated metastatic cells, but to date the triggers of this process remain unknown.

Despite the therapeutic relevance of targeting MICs and preventing their outgrowth at secondary sites, there are still many open questions regarding the molecular determinants of epithelial plasticity during the metastatic process:

- How is mesenchymal inhibition molecularly modulated in MICs? Is it an attenuation of the mesenchymal features via the downregulation of the EMT-TFs, or a complete re-epithelialisation (MET) as observed during physiological processes?

- Are EMT modulations intrinsically controlled by MICs or do they depend on microenvironmental cues? And in the assumption of the later, what microenvironmental signals trigger epithelial re-acquisition at the metastatic site allowing colonisation?

- Can the metastatic niche extrinsically modulate the epithelial plasticity (EMT-MET transitions) of MICs, guaranteeing the maintenance of stemness?

- Does the mesenchymal reversion need to occur in a temporally controlled manner during metastatic colonisation as suggested by the need of Twist1 loss at the target site (Tsai et al., 2012)?

Shedding some light into these complex questions has been one of the main focuses of this thesis.

1.5 Mouse models to study the metastatic process

As a complex multistep systemic process, understanding the means by which tumour cells metastasise requires the development of mouse models that closely resemble the evolution of the human disease.

Xenograft models, where human cell lines are grafted into immune-compromised mice have been widely used but in the absence of a fully competent immune system they tend to grow rapidly not mimicking the slow doubling time of most human cancers. Additionally, tumours develop lacking essential microenvironment components (immune cells) and in the absence of histocompatible human epitopes in the mouse stroma, which might limit tumour-stromal interactions. Altogether, these facts emphasise the limited value of xenograft models when trying to understand disease evolution in the context of the essential microenvironmental crosstalks that define it (Francia et al., 2011).

The alternative in order to overcome these limitations is the development of genetically engineered mouse models (GEMMs) that allow the generation of orthotopic tumours in immune competent hosts, containing the stroma and vasculature of the same specie. These models are specifically created by overexpression of oncogenes or deletion of tumour-suppressor genes relevant in a particular cancer type (Ottewell et al., 2006).

In this thesis, we focus on the study of the previously described metastasis-initiating cell (MIC) population in the metastatic breast cancer mouse model MMTV-PyMT (section 1.3.1) (Malanchi et al., 2012). Therefore, in this section we will describe the advantages and limitations of this model regarding its resemblance to the human disease, and its use to study the mechanisms underlying breast cancer metastatic colonisation.

1.5.1 Mouse models of breast cancer metastasis

Genetic mouse model are specifically created by activation of oncogenes or deletion of tumour-suppressor genes relevant to the human tumorigenic processes in a particular cancer type (Ottewell et al., 2006). In breast cancers, it is now appreciated that a small number of genes (i.e. PTEN, TP53 and PI3K) are

recurrently mutated across the different subtypes of breast cancer being the main genetic drivers of the disease (Cancer Genome Atlas, 2012).

Importantly, breast cancers are typically classified in three different molecular subtypes according to the cellular expression of the hormone receptors binding estrogen (ER) and progesterone (PR), and the oncogene HER2 in: luminal A (ER^+, PR^+, low HER2 expression), luminal B ($ER^{+/-}$, $PR^{+/-}$, high HER2 expression) and basal (ER^-, PR^-, low HER2 expression) (Schnitt, 2010). A recent comprehensive molecular portrait of human primary breast cancers elaborated employing Whole Genome Sequencing (WGS) revealed numerous subtype-associated driver mutations that are summarised below (Table 1.1) (Cancer Genome Atlas, 2012). Notably, the tumour suppressor TP53 is mutated at double frequencies in the luminal B subtype compared to A, and basal-like tumours show a high frequency (80%) of TP53 mutations. This TP53 mutational pattern in breast cancer correlates with tumour grade and patient prognosis (Cancer Genome Atlas, 2012).

Table 1.1 Main genomic features of breast cancer

Subtype	Luminal A	Luminal B	Basal-like
ER+/HER2- (%)	87	82	10
HER2+ (%)	7	15	2
TNBCs (%)	2	1	80
TP53 pathway	TP53 mut (12%)	TP53 mut (32%)	TP53 mut (84%)
PI3K/PTEN pathway	PI3K mut (49%); PTEN mut (13%)	PI3K mut (32%); PTEN mut (24%)	PI3K mut (7%); PTEN mut (35%)
Additional DNA mutations	GATA3 (14%); MAP3K1 (14%)	MAP3K1 (5%)	
DNA methylation		Hypermethylated	Hypomethylated

Many mouse breast cancer models have been created based on genetic alterations that typically occur in human breast carcinomas to mimic the morphological and mechanistic evolution of the disease. However, most failed to metastasise, as exemplified by the mouse model carrying inactivated BRCA1 and BRCA2 genes in the mammary epithelium, the most common mutations found in women with inherited breast cancers (Lin et al., 2003). Therefore murine models that efficiently

develop metastases have been created by the overexpression of the oncogenes Her2, Neu and the polyoma middle T antigen (PyMT) under the control of the mouse mammary tumour virus LTR (MMTV LTR) – MMTV-Her2, MMTV-Neu and MMTV-PyMT – achieving a potent expression of these oncogenes specifically in the mammary epithelium that leads to high metastatic burden to the lungs and other organs (Herzig and Christofori, 2002).

The PyMT is a membrane-attached protein that, although is not expressed by human cells, acts as a potent oncogene because its product binds several transduction pathways, including Src kinases and the RAS and PI3K pathways (section 1.1.1.1) (Figure 1.4.A), which are all altered in the most common subtype of human breast cancers, the luminal B (Dankort, 2000) (Cancer Genome Atlas, 2012). Lin et al. published a meticulous study analysing disease progression in this model, and comparing it histologically and molecularly to the evolution of human breast cancers. They found that in the MMTV-PyMT model disease progression occurs over a multistage process where early mammary gland hyperplasias appear first, and then evolve to adenoma, and early and late carcinoma possibly through the microenvironmental changes and new mutations acquired during the different stages. As in human breast cancers, each of these four stages is accompanied by distinct cell morphological and molecular changes as summarised in Figure 1.7. The loss of estrogen receptor (ER) and progesterone receptor (PR) in late stage carcinomas resembles the human disease progression. Moreover, the infiltration of leukocytes during the adenoma stage correlates with the shift in proliferation and the loss of basement membrane integrity, which stresses the role of the microenvironment in the disease progression towards the metastatic phase in this model, as observed in humans (Lin et al., 2003).
Sequencing analysis has allowed further in-depth comparisons between this model and the different subtypes of human breast cancers. Consistent with the data published by Lin et al., a comparative sequencing analysis between human tumours and mouse models clustered the MMTV-PyMT model with a group of 33/44 human $Her2^{+}/ER^{-}$ tumours, classifying it as a luminal B subtype (Herschkowitz, 2007).

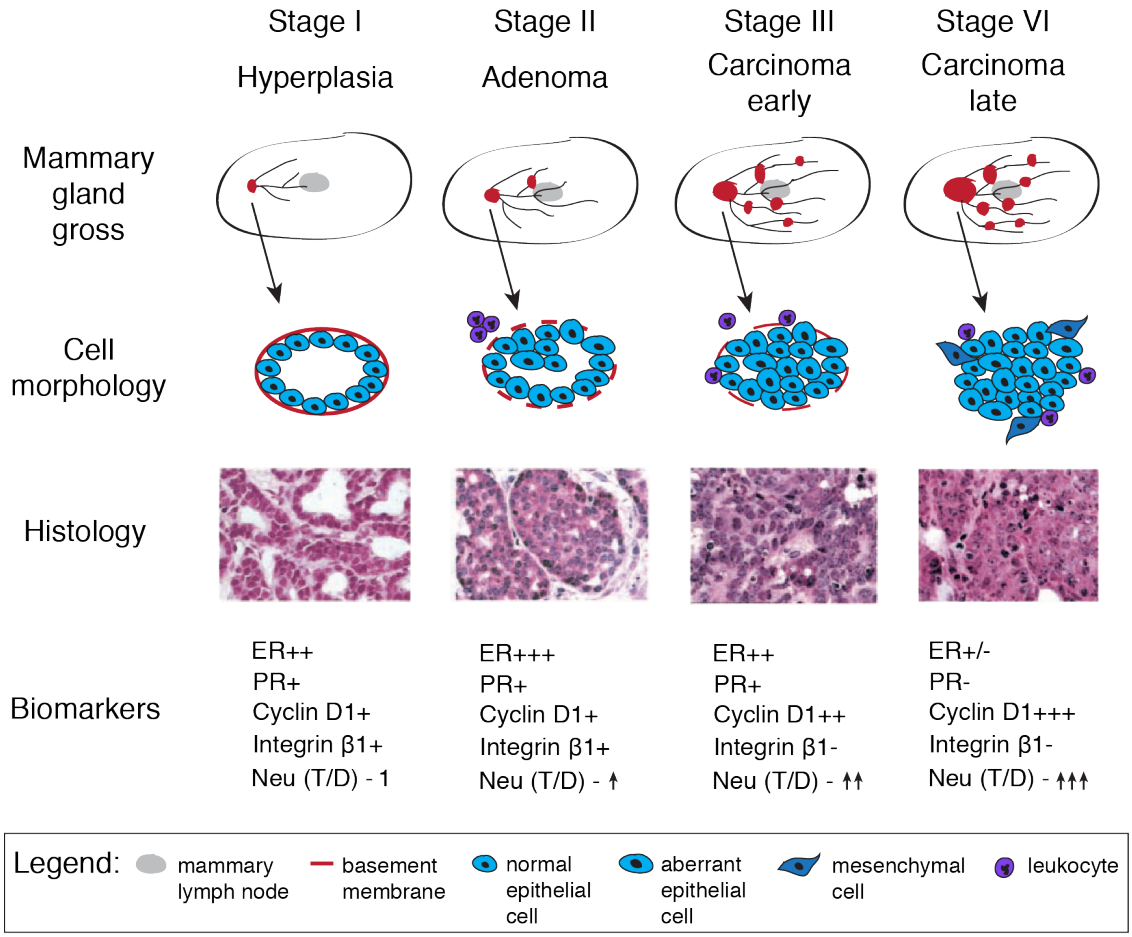


Figure 1.7 Disease progression in the MMTV-PyMT model
Adapted from Lin et al. 2003. In the top panel, the scheme shows the gross development of lesions in the mammary gland (lesions indicated by red dots), and the epithelial cell morphology, basement membrane integrity and leukocyte infiltration on each stage of tumour development. The histological H&E stains below illustrate the corresponding morphological changes in the developing tumours described in the schematics above. Last, the changes in biomarkers presence on each stage of tumour development are summarised in the lower panel, where (T/D) is the ratio of Neu/Her2 expression between tumorigenic lesions (T) and normal ducts (D) in age-matched mammary glands.

1.5.2 Experimental *versus* spontaneous models to study metastasis: advantages and limitations

As a systemic process, metastasis should be ideally studied in spontaneous models that allow the asynchronic development of the disease and a similar therapeutic setting to test preclinical therapies. Nevertheless, there are many challenges associated with their uses as tumours have to grow to a considerably big size (2-4 cm^3) before being surgically resected, and at that stage several complications might arise; for example the fact that very big blood vessels are connected to the tumours which makes the surgery laborious and reduces the number of mice available for the experiment after such a complicated procedure. Also the fact that spontaneous tumours develop asynchronically requires a careful randomisation of the mice to ensure equal metastatic burden before proceeding with the experiments (Francia et al., 2011).

Therefore, alternatively experimental models of metastasis have been widely used to study the colonisation step of the metastasis and undercover the molecular mechanisms underlying this complex step that constitutes the bottleneck of the metastatic cascade, as previously analysed (section 1.2.3). Experimental metastatic models consist of the direct injection of tumour cells in the circulatory system, via intracardiac or tail vein injections. Therefore, one of the main disadvantages of experimental metastasis models is that tumour cells do not undergo many challenges occurring during spontaneous tumour dissemination (i.e. local invasion, intravasation) as they are seeded directly into the circulation. Additionally, experimental metastasis assays are usually carried in the absence of an instigating primary tumour, missing the systemic effects of the disease that can positively impact on the ability of the cells to metastasise. Systemic tumour responses include immune cell recruitment to distant sites or release of certain factors to the circulation that will promote metastatic seeding in specific organs through the creation of ‘pre-metastatic niches’ before tumour cells reach the target tissues (McAllister and Weinberg, 2014). Using spontaneous mouse models of breast cancer, it has been shown that tumour-derived IL1β-IL17 results in systemic GCSF release that mobilises neutrophils from the bone marrow promoting breast to lung metastasis (Coffelt et al., 2015). In prostate cancer, CAFs implanted into mice

have been shown to secrete GDF15 systemically into the circulation, promoting the metastatic ability of distally implanted benign prostate cancer cells (Bruzzese et al., 2014). ECM components secreted by tumour cells, such as the lysyl oxidase (LOX) enzyme have been shown to act systemically and create a pre-metastatic niche favourable for breast cancer lung metastasis. Also, hypoxia-induced LOX secretion at the primary site will induce collagen fibers crosslinking increasing tissue stiffness and BMDCs recruitment to the lungs, creating a supportive niche before disseminated cells arrive (Erler and Giaccia, 2006). These findings are contributing to the view of cancer as a systemic disease, questioning the study of the metastatic processes in the absence of an instigating primary tumour.

In contrast, experimental metastasis models offer the advantages of synchronising the metastatic process circumventing differences in primary tumour development and allowing the analysis of the temporal evolution of the colonisation step. Indeed, in breast cancer, experimental metastasis assays have proven to be highly successful for investigating the molecular basis of organ-specific metastasis and their potential therapeutic targeting (Valiente et al., 2014, Sevenich et al., 2014, Zhang et al., 2013), as well as for defining specific populations within tumours with high metastasis initiating ability (Malanchi et al., 2012, Oskarsson et al., 2011). Certainly, as previously described (section 1.3.1), Malanchi et al. have identified a population of MICs ($CD24^{+}CD90^{+}$) in the MMTV-PyMT model that when tested in experimental metastasis assays compared to the bulk of tumour cells, retain exclusive metastatic colonisation ability (Malanchi et al., 2012).
Furthermore, experimental metastasis assays are key in the study of metastatic colonisation, allowing the synchronisation of the process to study the molecular dynamic changes between early and late colonisation. For instance, as we analysed in the previous section, EMT is known to be active in circulating tumour cells but it is inhibited after extravasation leading to metastatic outgrowth. Understanding the molecular basis underlying the dynamic EMT modulations during colonisation requires the synchronisation and temporal control of the metastatic process, therefore in this case experimental metastasis assays become the preferred model. The other advantage of experimental metastasis is that in the absence of the systemic pro-tumourigenic responses induced by a primary tumour, the pre-metastatic niche is not formed, and the minimum cancer cell intrinsic

features required for colonisation can be exposed in light of the naïve tissue microenvironment.

In conclusion, the MMTV-PyMT model of breast cancer is a suitable model to study the metastatic process: first, disease progression occurs in a slow progressive stage-by-stage manner, mimicking the cellular morphology and biomarkers changes observed in humans at every stage (section 1.3.1); second, the close similarity of this model to human breast cancer is also exemplified by the gradual loss of steroid hormone receptors (estrogen and progesterone) and β1-integrin, which is associated with overexpression of Neu and cyclin D1 in late-stage metastatic cancers (Figure 1.7) (Lin et al., 2003). Importantly, at the molecular level the PyMT oncogene activates multiple signalling cascades required for the development of human mammary tumours, therefore oncogenic evolution occurs through the same processes (section 1.3.1). The high metastatic incidence displayed by these mice makes them the most widely used mouse model to study metastatic progression in breast cancer (Cheung and Ewald, 2014). Finally, a distinct specialised population of metastasis-initiating cells (MICs) have been already identified in these tumours (Malanchi et al., 2012), providing the possibility of further exploiting this model to study the biology and potential therapeutics features of stem-like populations during metastasis.

1.6 Aim of the thesis

The major goal of this work was to assess what differential features of MICs make them the main population driving breast cancer metastatic colonisation to the lungs. As previously discussed (section 1.3.2), the metastatic niche plays a key role supporting tumour cell growth at the newly colonised tissue (Oskarsson et al., 2011, Malanchi et al., 2012), but to date knowledge about how and when tumour cells induce their metastatic niche is limited. In this project, we wanted to shed light into the strategies employed by disseminating tumour cells to induce their metastatic niche, and analyse the consequences of the presence of this newly activated stroma on metastatic progression. Specifically, we generated the gene expression profile of MICs and focused on those features that could contribute to their metastatic ability upon dissemination to a distant naïve tissue. We set out to answer the question: what intrinsic characteristics allow MICs to generate a favourable microenvironment determining metastatic success?

Chapter 2. Materials & Methods

2.1 Statistical Analysis

Unless otherwise stated, statistical analyses were performed using Prism software (GraphPad Software). Mean values and standard errors (sem) are shown in most graphs generated from several biological repetitions of experiments. P values were obtained by Student t-tests with paired or unpaired samples when analysing the differences in one variable between two experimental groups. Two-way ANOVA was used where indicated to perform multiple comparisons or multiple variable analyses between experimental groups. Significance was set at $P < 0.05$. Graphs show either the actual P value or symbols describing it (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p< 0.0001$).

2.2 Microarray samples and data analysis

2.2.1 Accession Numbers

Microarray expression data were deposited in the Gene Expression Omnibus (GEO) database, and are publicly available at http://www.ncbi.nlm.nih.gov/geo/ with accession number GEO:GSE63558. In this thesis the top hits expressed in MICs compared to nonMICs are included in the section 7.1. Appendix.

2.2.2 Microarray samples preparation

MICs and nonMICs were freshly isolated by FACS from late-stage PyMT tumours according to CD24 and CD90 expression. RNA isolation was performed using MagMax-96 Total RNA Isolation Kit (AM1830, Ambion) that allows high quality RNA extraction from samples with low cell numbers, < 10,000 cells. RNA quality for each sample was assessed using Agilent RNA 6000 Pico Kit (5067-1511, Aligent Technologies), and only those samples with a RNA Integrity Number (RIN) > 9 were considered for processing. RNA was amplified and analysed on Illumina gene expression arrays (mouse) at the Barts and London Genome Centre.

2.2.3 Gene expression data analysis

Gene expression data were analysed by the Bioinformatics Unit at the LRI using Bioconductor version 2.2, available at http://bioconductor.org, running on R version 2.7.1 (R Core Team, 2013), available at http://www.R-project.org/. Normalised probe set expression values were calculated using log2 transformation and quantile normalization using the 'lumi' package (Du et al., 2008). To determine significant differences of expression levels between the different groups, a moderated Student's t test was computed on a probe-by-probe basis using the empirical Bayes statistics in the 'Limma' package (Smyth, 2005). The resultant p-values were adjusted for multiple testing using the False Discovery Rate (FDR) Benjamini and Hochberg method (Benjamini et al., 1995), where any probe sets that exhibited an adjusted p-value FDR $q<0.05$ were called differentially expressed.

Differentially expressed genes identified in MICs *versus* nonMICs (MIC signature) were used to generate a hierarchical Heatmap (Figure 3.1.A). Genes were clustered using a Euclidean distance matrix and average linkage clustering. Red indicates higher expression and blue indicates lower expression relative to the mean expression of probes across the four samples.

Differentially expressed genes were used to determine enrichment of pathways and processes using MetaCore^{TM} integrated software, version 6.19 developed by Thomson Reuters (2014), available at http://thomsonreuters.com/metacore/. This analysis employs a hypergeometric distribution to determine whether a list of differentially expressed genes is enriched in a pathway or process gene set relative to all genes on the array. Pathways or processes that showed a FDR of less than 0.05 were called enriched. Gene ontology analysis of cellular compartments and biological processes are shown in Figure 3.1.B-C. The circos plot was generated plotting the top overexpressed pathways in MICs and a selection of genes known to contribute to breast cancer cells metastatic potential (Figure 3.2.B).

2.2.4 GSEA

The MIC signature was analysed using the GSEA software developed by the Broad Institute of MIT and Harvard (USA) and available at www.broadinstitute.org following the programme guidelines to determine the correlation of previous published signatures with the MIC signature. Microarray data was adjusted to contain only the gene symbols, a description column and numerical expression values and saved in .txt format. A phenotype file was created to describe the columns in the expression data file and the layout of the experiment i.e number of repeats. Genesets were obtained from the MSigDB (The Molecular Signatures Database), curated by The Broad Institute and also generated from publicly available microarray data (as indicated in Figure 3.2). The specific settings applied in all analyses are: number of permutations - 1,000, permutation type - gene set, enrichment statistic -weighted, metric for ranking genes - t-test. The rest of the fields were left as defaulted.

The free software ranks the genes in the microarray data file according to expression changes between conditions. It subsequently tests how a list of genes (geneset) is distributed within the ranked data and calculates an enrichment score that represents the degree to which the geneset is represented at the top or bottom of the ranked list. The score is calculated by walking down the ranked list and measuring a running-sum statistic. The score is increased on encountering a gene within the geneset and decreased on encountering genes not in the geneset. The enrichment score is then the maximum deviation score achieved. An estimate of statistical significance is obtained.

2.3 *In silico* data analysis

2.3.1 cBioPortal for Cancer Genomics

The correlation analysis between the expression of THBS2 and different EMT markers was performed using cBioPortal for Cancer Genomics, develop by Memorial Sloan Kettering Cancer Center (Gao et al., 2013, Cerami et al., 2012), and available at http://www.cbioportal.org/public-portal/. The tumour data set used

was The Cancer Genome Atlas (TCGA provisional), Breast Invasive Carcinoma project (Cancer Genome Atlas, 2012).

2.3.2 GOBO

THBS2 Distant Metastasis Free Survival (DMFS) prognosis prediction according to tumour grade from breast cancer patients was done using GOBO (Gene expression-based Outcome for Breast cancer Online) 1.0.1 database, develop by the Department of Oncology of Lund University and Skåne University Hospital (Lund, Sweden), and available at http://co.bmc.lu.se/gobo/. The tumour data set consists of 1881 samples (Ringner et al., 2011). All tumour types were selected for the analysis, splitting data in 3 quantiles and selecting 5 year for censoring data (Figure 3.22).
GOBO database allowed the Gene Co-expression analysis of AXL in breast cancer patient data (Table 4.1), and was also used to assess AXL expression levels in a panel of 30 breast cancer cell lines, classified as Luminal, Basal A or Basal B subtype according to the sequencing data published in Neve et al., 2006 (Figure 4.3.A).

2.3.3 String database

The search for possible interacting partners for THBS2 was done using STRING 9.1 (Search Tool for the Retrieval of Interacting Genes/Proteins) database, developed by CPR, EMBL, SIB, KU, TUD and UZH and available at http://string-db.org/ following the database guidelines. Confidence interval was set at the highest score (0.90) and no more than the highest 10 ‘interactors’ were shown. All prediction methods were left selected as defaulted considering co-expression, gene-fusion, databases, co-ocurrence, neighboorhood, textmining and experimental data. Confidence view of the resulting ‘THBS2 interactome’ where stronger associations are represented by thicker lines, is displayed in the diagram shown in Figure 3.15.A.

2.4 Buffers, solutions and chemicals

Some common buffers and solutions are listed below.

Phosphate buffered saline (PBS)

1.5 mM KH_2PO_4, 8 mM Na_2HPO_4, 2.7 mM KCl, 137 mM NaCl, pH 7.4

TE buffer

10 mM Tris HCl pH 7.5, 1 mM EDTA

50 x TAE

2 M Tris-acetate (242 g Tris base + 57.1 ml glacial acetic acid per litre), 0.05 M EDTA pH 8.0

6 x DNA loading buffer

0.25 % bromophenol blue, 0.25 % xylene cyanol FF, 30 % glycerol

Chemicals: Potassium chloride (KCl), Tris base, Tris HCl, potassium phosphate (KH_2PO_4), sodium phosphate (Na_2HPO_4), ethylene diamine tetraacetic acid (EDTA), glycine, β-mercaptoethanol, sodium hydroxide (NaOH), hydrochloric acid (HCl), tetramethylethylenediamine (TEMED), ethanol (EtOH), methanol (MeOH), dimethylsulfoxide (DMSO), Tween-20, Triton-X 100, bromophenol blue, bovine serum albumin (BSA), deoxynucleotides (dNTPs) and paraformaldehyde (PFA) were all purchased from SIGMA. Acetic acid, glycerol, 40 % acrylamide solution, sodiumdodecyl sulphate (SDS), sodium acetate, sodium chloride (NaCl), and magnesium chloride ($MgCl_2$) were purchased from Fisher Scientific. UltraPure agarose was purchased from Invitrogen.

2.5 DNA purification and agarose gels

Agarose gels

Agarose was melted in 1 x TAE buffer in the microwave at concentrations between 0.5 and 1 % depending on the DNA fragments to be separated. Gel red nucleic acid (Biotium) was added at a concentration of 10,000X form the stock provided before casting the gel. DNA samples were mixed with 6 x DNA loading buffer and loaded. A 1 kb ladder was also loaded to help identify DNA fragment sizes. Gel electrophoresis was performed in TAE buffer at 100V. DNA was visualised using a UV transilluminator and if required the band of interest was cut out from the gel.

DNA purification

DNA was purified from agarose gels using QIAquick Gel extraction kit (Qiagen). The agarose band containing the DNA was dissolved at 65 °C in 3 volumes of Buffer QG. Then, one volume of isoproponol was added and the solution applied to a Qiaquick column. The column was subsequently washed, dried and the DNA eluted in dH2O.

2.6 Quantitative real time PCR

For quantitative RT-PCR, RNA preparation was performed using either MagMax-96 Total RNA Isolation Kit (AM1830, Ambion) when RNA was isolated from less than 10^5 cells per sample, or RNeasy Mini Kit (QIAGEN) for higher cell numbers per sample. cDNA synthesis was performed using and SuperScript III First-Strand Synthesis System (Invitrogen), according to the manufacturer's protocol.

Quantitative real-time PCR samples were prepared with 25 ng total cDNA for each PCR reaction. Reactions were set-up on ice in a 96 well optical reaction plate (Applied Biosystems) as outlined below:

10 µl Platinum SYBR Green qPCR SuperMix-UDG with Rox
0.5 µl Forward primer (0.5 pmol)
0.5 µl Reverse primer (0.5 pmol)

5 µl cDNA (5 ng/µl)
4 µl dH2O

The data generation and analysis were performed on a 7500 FAST Real-Time PCR System (Life Technologies) using the following PCR cycling conditions:

50 °C 20 s
95 °C 10 mins
95 °C 15 s - Denaturing
60 °C 60 s - Annealing/Extension return to denaturing x 40 cycles
4 °C hold

All human primers pairs used for each of the mRNA targets are shown in Table 2.1 and mouse primer in Table 2.2.

Table 2.1 Human RT-PCR primers

GENE SYMBOL	Forward (5' – 3')	Reverse (3' – 5')
GAPDH	AGGGGTCTACATGGC AACTG	CGACCACTTTGTCAAGCTCA
AXL	AACCT TCAACTCCTG CCTTC TCG	CAGCTTCTCCTTCAGCTCTTCAC
THBS2	GCTACATCAGAGTCTTAGTGC	AGAGAAGACAAATAGACCCAG

Table 2.2 Mouse RT-PCR primers

GENE SYMBOL	Forward (5' – 3')	Reverse (3' – 5')
GAPDH	CGTGTTCCTACCCCCAATGT	TGTCATCATACTTGGCAGGTTTCT
SNAIL	CTCTGAAGATGCACATCCGAA	GGCTTCTCACCAGTGTGGGT
SLUG	TGTGTCTGCAAGATCTGTGGC	TCCCCAGTGTGAGTTCTAATGTG
TWIST1	GCCGGAGACCTAGATGTCATTG	CACGCCCTGATTCTTGTGAA
ZEB1	GCCAGCAGTCATGATGAAAA	TATCACAATACGGGCAGGTG
E-CAD	CGACCCTGCCTCTGAATCC	TACACGCTGGGAAACATGAGC
VIM	CCAACCTTTTCTTCCCTGAA	TTGAGTGGGTGTCAACCAGA
MYL9	AGGCCTCAGGCTTCATCCACGA	ATGGGGTCTAGGCACTGGGGC
AMOTL2	AACCGCCACCTGGCAAGCAA	GGTCCTCGATGGCACCACGC
ANKRD1	AAACGGACGGCACTCCACCG	CGCTGTGCTGAGAAGCTTGTCTCT
CCL2	CAGGTCCCTGTCATGCTTCT	GTCAGCACAGACCTCTCTCT
CTGF	GCTTCGCAGGGCCTGACCAT	GCCGTGGGCTGCATTCCTCT
ICAM1	GAAGCTTCTTTTGCTCTGCC	AGCAGTACTGGCACCAGAAT
PLAC8	TCAGTGACTGCGGAGTCTGCC	CAGAGCTCTTGCCATCCAGCTCC
SPDR	GCCCAGCAGGTGCGCTATGA	CGGGGTGGCTTCCACGAGGT
TAGNL	AGGCGGCCTTTAAACCCCTCACC	CGGCCTACATCAGGGCCACAC
THBS1	GCGTTGCCAGGCTCCGAGTT	GGTGCGCAGGCCCTTCAGTT
ANGTL4	AAGATGCACAGCATCACAGG	ATGGATGGGAAATTGGAGC
CCL7	AGCTACAGAAGGATCACCAG	CACATTCCTACAGACAGCTC
CCL11	CAGATGCACCCTGAAAGCCATA	TGCTTTGTGGCATCCTGGAC
CXCL10	GACGGTCCGCTGCAACTG	GCTTCCCTATGGCCCTCATT
CYR61	GCCGTGGGCTGCATTCCTCT	GCGGTTCGGTGCCAAAGACAGG
WNT5A	GGTGCCATGTCTTCCAAGTT	ATCACCATGCCAAAGACAGA
COL1A1*	TCGTGGCTTCTCTGGTCTC	CCGTTGAGTCCGTCTTTGC
COL4A1*	ATAGCCAAAGCCAAACCCATT	CGCAGAGCAGAAGCAAGAA
LAMA2*	GTATCACCAAATTATCCAAGACCAT	TCGCTGGGCATATTTCTCATC
FN1*	AAGAGGACGTTGCAGAGCTA	AGACACTGGAGACACTGACTAA
MMP2*	GGACAAGAACCAGATCACATACA	GCCCGAGCAAAAGCATCAT
MMP9*	TGATGCTATTGCTGAGATCCAG	CCTGTAATGGGCTTCCTCTATG
MMP13*	CCAACCCTAAGCATCCCAAA	TCCTCGGAGACTGGTAATGG
MMP14*	CCATCAATACTGCCTACGAGAG	CTCCTTAATGTGCTTGGGGTAC
FSCN1*	AAAGACTCCACGGGCAAGTA	GTCACAGAACTCAAGGAAGAAATC
LIMKI*	TCTACAAGGACAAGCGGCTAA	CTCTGACTCCACGGGTACTG
MYHL10*	TCAGATGACGACACAGAAAGTAA	AGACTGGTGCTGGGATGAG
PFN1*	ACCGTTCCCTTTGGCTTTTG	ATGTGGTTTTGGCAGCAATAAG
MYLK*	CCACCAGCACCATCACTCT	CCCAGGAGAGAAGGACCCTA
THBS2*	GCCACCAGAACAACCAAGAC	CACCATCATTGTCATCATCAGAGT
ACTA2*	CTCTTCCAGCCATCTTTCATTG	GTTGTTAGCATAGAGATCCTTCCT

* Design and validated by PrimerDesign Ltd, UK.

2.7 Animal procedures

2.7.1 Mouse strains

Transgenic FVB/n mice expressing the polyoma middle T antigen oncogene under the mouse mammary tumour virus promoter (MMTV-PyMT) (Guy et al., 1992) were the sources of primary cancer cells. Cells were always isolated from late stage carcinomas, when tumours are at least 2cm. Age matched 6 to 10 weeks $Rag1^{-/-}$ FVB/n female mice (Mombaerts et al., 1992) were used as recipients for all orthotopic tumour transplantations and tail vein injections experiments, unless otherwise specified. Nude CD1-Foxn1 mice (Flanagan, 1966) were bought from Harlan UK and used in the experimental metastasis assays to obtain the samples for the microarray analysis. All mice were kept in accordance with UK regulations under project license PPL 80/2531.

2.7.2 Tumour initiation and metastatic colonisation assays

In tumour initiation assays (collective challenge), unless otherwise specified, either $3x10^5$ PyMT primary cells or 10^6 MDA-MB-231 cells were mixed in 100 µl of Growth Factor Reduced Matrigel (356230, Corning), and injected into the mammary fat pad of anesthetised mice. Tumour growth was evaluated 2 weeks after injection for early tumour initiation, and 6 weeks post-transplantation to evaluate differences in primary tumour growth and microenvironment.

For experimental metastasis assays (single cell challenge), unless otherwise stated, between $5x10^5$ – 10^6 PyMT primary cells or MDA-MB-231 cells, were resuspended in 100 µl PBS and injected into the tail vein of mice. All metastatic assays were performed using endogenously expressing, or lentivirus induced, GFP^+ or $VENUS^+$ cells. Macrometastasis were evaluated quantifying the number of superficial metastatic nodules in a fluorescent stereomicroscope (Lumar V.12, Zeiss), and micrometastasis by FACS analysis of lung cell preparations, or histological analysis after H&E staining and sectioning of the lungs.

2.7.3 Drug treatments

R428, also known as BGB324 (1946, Axon MedChem), was dissolved in DMSO at a final concentration of 21.15 mM or 12.5 mg/ml. *In vivo* treatments were performed as previously described (Holland et al., 2010). Briefly, for the 24 hours treatment time point, cells were treated *in vitro* with either vehicle (DMSO) or 1.5 uM R428 for 3 hours, 24 hours before injection. For treatment during the first week of metastasis, cells were pre-treated in culture as described above, and after tail vein injection mice received a daily dose of 12.5 mg/kg R428 by oral administration. Treatments between 10-15, and 25-35 days post-injection were performed as described above, but cells were not pre-treated in culture before injecting them.

LDN193189 was dissolved in citric acid – sodium citrate buffer solution pH=3.1 at a final concentration of 6.25 mg/ml. Mice were injected with 10^6 PyMT cells intravenously, and treated daily with an intraperitoneal dose of 35 mg/kg LDN193189 from day 5 to 12 post metastatic seeding.

2.8 Histology

2.8.1 Lung immunofluorescence

Immunofluorescence to assess *in vivo* niche activation was performed in 4% PFA fixed, paraffin embedded lungs. 4 µm thick lung sections were cut, de-paraffinased and re-hydratated using standard methods. After antigen retrival, sections were stained with DAPI, an anti-GFP antibody to label tumour cells and either SMA to detect activated fibroblasts, or endomucin to detect endothelial cells (antibodies description and working concentrations in Table 2.5). Sections were treated with Sudan Black B (SBS) to quench background auto-fluorescence.
Images were acquired in an upright Zeiss LSM780 confocal microscope. Quantitative analysis of images was performed manually counting the number of SMA^+ or $Endomucin^+$ cells in direct contact with the GFP-labelled tumour cells in the tissue.

2.8.2 Immunohistochemistry (IHC-P)

Tumours were fixed in 4% PFA for 24h, and then embedded in paraffin blocks. 4 µm thick tumour sections were cutted, deparraffinased and rehydratated using standard methods. After antigen retrival, sections were stained with haematoxylin and eosin (H&E) solution, or alternatively with DAPI and antibodies against endomucin (endothelial cells) and/or Ki67 (proliferating cells).
Images were acquired in an upright Nikon Eclipse Ni-U light microscope.

2.9 Lentivirus

2.9.1 Lentiviral constructs generation

All lentiviral vectors used in this study were generated using Gateway® technology (Life Technologies).
The lentivirus constructs expressing short hairpin RNAs (shRNAs) were obtained in a two-step approach: first, the shRNA paired primers were designed with XhoI, BglII restriction sites on the edges to be cloned into a Gateway pENTR plasmid under H1 promoter (attL1_H1promoter_cloning-site_attL2). Second, we performed a Gateway LR recombination between the pENTR plasmid containing the shRNA and the lentiviral vector p4300 (cPPT_CMVenh_attR_hPGK_EGFP_WPRE). Construct scheme in Figure 3.17 and Figure 4.5.

Table 2.3 shRNA sequences

shRNA	TARGET SEQUENCE (5' – 3')
shAXL F	GACATCTTCTTTCTCATGTGAAGCCCATCTGGTCATGGGCTTCACATGAGAAAGAAGATGTC TTTTT
shAXL R	AAAAAGACATCTTCTTTCTCATGTGAAGCCCATGACCAGATGGGCTTCACATGAGAAAGAAG ATGTC
shTHBS2 F	GCTGTAGGTTTCGACGAGTTTCTCGAGAAACTCGTCGAAACCTACAGCTTTTT
shTHBS2 R	AAAAAGCTGTAGGTTTCGACGAGTTTCTCGAGAAACTCGTCGAAACCTACAGC
shTHBS2 II F	CCACGTCAAGGACACTTCATTCTCGAGAATGAAGTGTCCTTGACGTGGTTTTT
shTHBS2 II R	AAAAACCACGTCAAGGACACTTCATTCTCGAGAATGAAGTGTCCTTGACGTGG

The lentiviral vectors expressing protein-coding DNA sequences were obtained by Gateway LR recombination between a pDONR plasmid containing the DNA coding sequence and a pDEST lentiviral backbone. For THBS2 overexpression, we engineered a lentiviral reporter construct recombining the pDONR plasmid containing the mouse DNA coding sequence of THBS2 (Addgene plasmid 12411; pcDNA3_mTSP2) with a pDEST lentiviral backbone that drives its expression under a mouse PGK promoter. The lentiviral backbone contains as well a truncated isoform of the human CD2 receptor (hCD2) and a VENUS reporter gene (Figure 3.19.A). For a mild AXL expression, we recombined the pDONR plasmid containing the DNA coding sequence (Addgene plasmid 23945; pDONR223_AXLORF) with a pDEST lentiviral backbone that only has the lentivirus LTR driving regions and a CMV enhancer. The lentiviral backbone contains as well an EGFP reporter gene (Figure 5.6.A). Below, in Table 2.4 are the detailed descriptions of the expression constructs used in this study.

To generate the lentiviral vectors expressing shRNAs, we used hairpin oligomers with the target sequences shown in Table 2.3 (shRNA sequences). The shTHBS2 validated target sequences were obtained from Sigma Mission's website in collaboration with the RNAi consortium (TRC) scientist (http://www.sigmaaldrich.com/life-science/functional-genomics-and-rnai/shrna.html). The target sequence for shAXL was described previously (Gjerdrum et al., 2010). All these sequences do not match any cDNA sequence of the corresponding specie (mouse) registered in BLAST other than that of their own target gene.

Note: all restriction enzymes used to clone were purchased from New England BioLabs (NEB) and were supplied and used with the appropriate buffers. DNA polymerase 1 (Klenow fragment) was purchased from NEB and was used to blunt DNA ends. Ligation reactions were carried out using T4 DNA Ligase, also from NEB.

Table 2.4 Lentiviral plasmids description

Name	Promoter /driving sequence	Marker	Notes
GFP	LTR	hPGK-EGFP	Mild expression
LTRAXL	LTR-hAXLORF	hPGK-EGFP	Mild expression
shCONTROL	H1	hPGK-EGFP	High expression
shAXL	H1-mshAXL	hPGK-EGFP	High expression
shTHBS2 I	H1-mshTHBS2 I	hPGK-EGFP	High expression
shTHBS2 II	H1-mshTHBS2 II	hPGK-EGFP	High expression
VENUS	mPGK	hPGK-VENUS	High expression
THBS2ORF	mPGK-mTHBS2ORF	hPGK-VENUS	High expression

2.9.2 Lentiviral constructs sequencing

DNA constructs were checked by sequencing using the Big Dye terminator (BDT) kit (Applied Biosystems). For each sequencing reaction the following components were used:

8 µl BDT reaction mix
1 µl DNA (100ng)
1 µl Primer (4pmol)
10 µl dH_2O

The PCR settings used to amplify the single stranded DNA for sequencing were the following:

96 °C 1 min
96 °C 10 sec - Denaturing
50 °C 5 sec - Annealing
60 °C 4 min - Extension and return to denaturing x 25 cycles
4 °C hold

DNA for sequencing was purified using Performa DTR Gel filtration cartridges (EdgeBio). The purified DNA was then dried using vacuum centrifugation and analysed by the in house Equipment Park facility.

2.9.3 Lentivirus production

Third-generation VSV-G pseudo-typed high titers lentiviruses were generated by transient co-transfection of 293T cells with a four-plasmid combination as follows: One 10 cm dish containing $1x10^6$ 293ft cells was transfected using 2M CaCl2, and 2x HBS (51558, Sigma) with 15 µg lentiviral vector, 6 µg pMD2-VSVG, 6 µg pRSV-rev and 6 µg pMDL-g/p-RRE. Supernatants were collected 48h and 72h hours after transfection.

2.9.4 Primary cell lentiviral infections

For lentiviral transduction, 10^5 primary PyMT cells/well were seeded in 6-well tissue culture dish and infected with EGFP or VENUS reporter lentiviruses added in suspension using 10 µg/ml Polybrene (107689, Sigma). After 48 hours, successfully transduced cells were trypsinised and FACS sorted according to their EGFP/VENUS expression.

2.10 Cell culture

2.10.1 Culture conditions and reagents

The human mammary carcinoma cell line MDA-MB-231 and the mouse carcinoma 4T1 cell line were obtained from London Research Institute - Cell services, and maintained in 2D cultures in DMEM with 5% FCS at 37C and 5% CO_2.

Mouse mammary carcinoma PyMT cells were seeded 2D in plates coated with collagen-solution (100 mg/ml BSA, 1M HEPES, 1:100 PureCol (5005-B, Advanced BioMatrix) in HBSS), and maintained in MEM media (DMEM/F12 supplemented

with 2% FCS, 10µg/ml Insulin (I9278, Sigma), 20ng/ml EGF (PHG0313, Invitrogen) and 1:50 L-Glutamax (35050-061, Life Technologies), and grown at 37C and 5% CO_2.

Mouse primary lung fibroblasts were isolated from a single-cell suspension of lung tissue that was depleted from $Ter119^+$ red blood cells, immune $CD45^+$ cells and endothelial $CD31^+$ cells by MACS sorting. The remaining fraction was cultured overnight on plastic in DMEM with 10% FCS. Unless otherwise specified, all fibroblast experiments were performed on cells seeded on the top of a thin layer of a 2:1 mixture of collagen I (354249; BD Biosciences) and Matrigel (354234; BD Biosciences) as described previously (Calvo et al., 2013).

Fibroblast cell lines: for generating the NLF3 cell line, normal lung fibroblasts were isolated from healthy lungs of aGFP/FVB/n mice. Lung CAFs were isolated from metastatic lungs of MMTV-PyMT FVB/n mice, and the mcherry CAFs (mcCAFs) were isolated from primary MMTV-PyMT FVB/n tumours in carcinoma stage. All fibroblast cell lines were generated according to the published Bio-protocol e1097 (Calvo, F., Hooper, S. and Sahai, E. (2014). Isolation and Immortalization of Fibroblasts from Different Tumoral Stages. Bio-protocol 4(7): e1097), available at http://www.bio-protocol.org/e1097. Briefly, tissue pieces are placed into dishes where they are compressed under a 20 mm coverslip in DMEM media under sterile conditions. After 7 days, fibroblasts have come out of the tissue into the coverslip and the dish and can be trypsinased with Trypsin/EDTA and transferred into a new dish. Once they reach 80% confluency, primary fibroblasts are immortalised by infection with pBABE-HPV-E6-puromycin retroviruses. Positively infected fibroblasts are selected by addition of puromycin (2 µg/ml final concentration) in the complete culture media for 7-10 days, when resistant clones expressing HPV-E6 (i.e. immortalised cell lines) are visible.

2.10.2 *In vitro* drug treatments

R428 (1946, Axon MedChem) *in vitro* treatment was performed as previously described (Holland et al., 2010). Briefly, PyMT primary cells plated 2D were treated

with either vehicle (DMSO) or 1.5 μM R428 for 3 hours before washing and adding fresh media. 24 hours after *in vitro* assays were performed. MDA231 cells were treated with 3 μM R428 for 24 hours. Next day, fresh media was added and *in vitro* assays were performed.

LDN193189 and SB431542 were dissolved in DMSO to 10 mM and use at a final concentration of 10 μM and 1 μM respectively in the cell culture media. A single treatment was used as experiments were done over 24 - 48h.

2.11 Tumour sphere assay

Cells prepared as a single-cell suspension were plated in Ultra Low Attachment 96-well plates (Constar) with Sphere Media (DMEM/F12 supplemented with 1:50 B27 (Invitrogen), 20ng/ml EGF (PHG0313, Invitrogen), 20ng/ml FGF (PHG0026, Invitrogen), 4mg/ml heparin (H3393, Sigma)), and kept at 37C and 5% CO_2.
PyMT primary cells spheres were established from total tumour cell preparations. After tumour dissociation, cells were plated on collagen overnight; trypsinased and 10^4 cells/ 200ul per well were seeded. Only the characteristic hollow spheres formed by these cells were quantified as tumour-spheres.
For passaging, tumour-spheres were collected by centrifugation and dissociated both enzymatically by trypsin treatment (for 10 min at 37 °C) and mechanically by using a 16G needle, and the resulting single-cell suspension was counted and re-plated at the original density.
Quantitative analysis of images was performed using ImageJ software, and the Sphere Formation Index (SFI) was calculated considering both, number and size of the spheres, and was obtained by summing the area of all spheres formed divided by the number of single cells initially plated.

2.12 Western blotting

$CD24^{+}AXL^{+}$ cells were FACS sorted and cultured in collagen-I coated dishes overnight, and incubated for 24h with the pertinent media and treatments. Protein lysates were obtained after cell lysis with RIPA buffer (50mM Tris-HCl, 1% NP-40, 0.5% Na-deoxycholate, 0.1% SDS, 150 mM NaCl, 2 mM EDTA, 50 mM NaF) supplemented with Roche cOmplete protease inhibitors cocktail tablets following standard procedures. Protein lysates were run in 12% polyacrylamide gels, and blotted into Amersham™ Polyvinyllidine difluoride membranes (10-6000-29). Membranes blocking and antibody incubations were performed in 5% BSA/PBS. Antibody description and working concentration used can be found in Table 2.5.

2.13 BrdU proliferation assay

Cancer cells proliferation rate was assessed measuring their nuclear BrdU incorporation after culturing them for 3 hours in 30 μM BrdU supplemented media. Then, cells were washed with PBS, trypsinased, fixed with 70% ethanol for 20 mins, permebealised with 2N HCl, 0.5% TritonX-100 for 30 mins, and treated with 0.1M tetraboric acid for 2 mins. After washing twice with FACS buffer and cells were stained with anti-BrdU antibody (Table 2.5).

2.14 ELISA

To measured THBS2 secretion, tumour cells were plated either in sphere or adherent conditions at a cell density of 50,000 cells per 100 μl. The resulting conditioned-culture media was collected 24h after plating and frozen at -80C.
To quantify the levels of secreted protein in the conditioned-media, we used two different sandwich-ELISA kits, THBS2 ELISA Kit (SED822Mu, Cloud-Clone Corp) and Mouse THBS2 ELISA Kit (MBS942662, MyBioSource), and proceed according to the manufacturer's protocols. The resulting colorimetric reaction was measured

at 450nm using an automated ELISA plate reader (SpectreMax 190 Microplate Reader, Molecular Devices).

2.15 Flow cytometry

2.15.1 FACS

Flow cytometry analyses were carried out on a BD LSR-Fortessa system (BD Biosciences), and PI (Propidium Iodide) was used as a viability dye unless otherwise indicated.

To examine the relationship between MICs, AXL and Sca1 in primary PyMT cells, PyMT tumours were dissociated and plated overnight on collagen-I coated dishes. Next day, cells were trypsinased and stained for CD24 as an epithelial marker in combination with Sca1, AXL and CD90 in MACS buffer (0.5% BSA and 1mM EDTA in PBS). Antibody description and working dilutions can be found in Table 2.5.

.

To assess AXL expression in metastatic lungs containing GFP labeled cancer cells, the tissue was dissociated and prepared as a single cell suspension as previously described (Malanchi et al., 2012). Subsequently, it was stained for lineage negative markers (CD31, CD45 and Ter119), CD24, CD90 (when expression was targeted to MICs) and mAXL-biotin or hAXL in MACS buffer. Antibody description and working dilutions can be found in Table 2.5.

The expression of different protein assessed by FACS in NLF3, MDA231 and 4T1 cell lines was done after trypsinisation of cultured cells and staining with the pertinent antibodies in MACS buffer.

Differences in extravasation to the lungs by R428-treated or untreated PyMT cells was assessed by perfusing the lungs with PBS to flush out the blood, and then the clean tissue was dissociated in single-cell suspension and the number of GFP^+ MICs present in the lungs were quantified.

2.15.2 Cell sorting

All cell-sorting experiments were carried out on either a BD Influx or FACSAria cell sorter systems (BD bioscience).
To sort the MICs from PyMT primary tumours or lungs containing GFP labeled cancer cells, we used the staining indicated above (2.11.1). $CD24^{+}AXL^{+}$ and $CD24^{+}AXL^{-}$ cells were sorted from primary PyMT tumour single-cell preparations. The staining protocol previously described (Malanchi et al., 2012), was adapted and a mAXL-biotinylated antibody was incorporated (antibody description and working dilutions can be found in Table 2.5).
$EGFP^{+}$ or $VENUS^{+}$ cells expressing the lentiviral constructs were trypsinased, washed with FACS buffer and sort according to the expression of the reporter gene.

2.15.3 ImageStream

ImageStream analyses were carried out in an ImageStream®^{X} Mark II Imaging Flow Cytometer (Amnis-Millipore).

For the whole lung analysis, cells were dissociated into a single cells suspension and depleted by MACS sorting from lineage negative cells (Ter119, CD45 and CD31). Next, the remaining fraction was stained with a live/dead fixable Near-Infrared dye (Invitrogen, L10119), and after, fixed with 2% PFA for 30 min and permeabilised with 0.1% Trinton-X for 10 min. Next, cells were resuspended in FACS buffer (2% FCS in PBS) and stain with primary antibodies against Twist1, pSMAD1-5, ID1 or pSMAD2-3 at 4C overnight (antibody description and working dilutions can be found in Table 2.5). Subsequently, cells were washed twice in MACs buffer (0.5% BSA and 1mM EDTA in PBS) and incubated for 45 min with secondary antibodies at room temperature. DAPI was added at 1 µg/ml to cells before analysis for nuclear staining. Tumour cells were identified by GFP expression (Figures 5.1.B-D and Figure 5.4.A-F).

In vitro, ImageStream analysis of non-labeled MICs co-cultured with NLF3 for 4-5 days (as specified in the scheme in Figure 5.1.F) was performed from a single cell

suspension after gel digestion. NLF3 were excluded by GFP expression, and MICS were analysed for Twist1 expression changes (Figure5.1.F-I).

The acquired data was analysed using Amnis IDEA management software to compensate for multi-colour spectral overlap; plot and graphical gating of the cells as well as obtaining representative images of cells.

2.16 Cell immunofluorescence

All cancer cell immunofluorescence microscopy was performed on cells seeded on glass-bottom 35 mm MatTek dish (P35-1.5-14-C, MatTek) or glass coverslips coated (Thermo Scientific) with collagen solution (Cell culture procedure). Samples were fixed with 4% PFA for 10 min, permeabilised with 0.5% TritonX-100 for 30 min and 0.1% TritonX-100 for 20 min. Samples were blocked for 45 min at room temperature in blocking solution: 4% BSA in PBS with 0.05% Tween20 (Sigma). Then, cells were incubated with primary antibody in blocking solution overnight at 4 C. After 3 washes of 5 min in PBS, secondary antibody in blocking solution was added. After 3 washes of 5 min in PBS, samples were mounted using Dako Fluorescent Mounting Media (S3023, Dako), and analysed using either upright or inverted Zeiss LSM710.
Quantitative analysis of images was performed using Imaris 7.6.4. Tumour cells were identified using an automated threshold based on GFP expression or a homogeneous cell body staining. The mean intensity of E-cadherin, AXL and Vimentin was then measured.

All fibroblast immunofluorescence, unless stated otherwise, were performed on cells seeded on the top of a thin layer of a mixture of 2:1 collagen-I:Matrigel as described previously (Calvo et al., 2013). Briefly, after 2 days on the gel and the pertinent treatments, cells were fixed in 4% paraformaldehyde for 1 h. Cells were permeabilised by incubation in PBS 0.5% NP-40 (Sigma) at 4C for 20 min, in PBS with 0.3% Triton X-100 (Sigma) at room temperature for 20 min and in PBS with 0.1% Triton X-100 at room temperature for 15 min. Samples were blocked for 60

min at room temperature in blocking solution: 4% BSA in PBS with 0.05% Tween20 (Sigma). Then, cells were incubated with primary antibody in blocking solution in a wet chamber overnight at 4C. After 3 washes of 15 min in PBS, secondary antibody in blocking solution was added. After 3 washes of 15 min in PBS, samples were mounted and analysed using an inverted Zeiss LSM710.
Quantitative analysis of images of CAF markers was performed using Imaris 7.6.4. To quantify YAP nuclear localization, we calculated for each cell the nuclear YAP intensity mean (determined by the intensity within the region delimited by DAPI staining) and the cytosolic YAP intensity (determined by the intensity mean of the staining within the region delimited by GFP) for at least 50 cells in triplicate. Plots represent the ratio of nuclear *versus* cytosolic YAP staining. To quantify FAP, fibroblasts were identified using an automated threshold based on their GFP expression. The mean intensity of FAP was then measured.

Antibody description and working dilutions can be found in Table 2.5.

2.17 DQ collagen assay

To assess the collagenase activity of tumour cells degradation of DQ-collagen-I by living cells was measured. A single cells suspension of tumour cells was embedded in 100 µl of 2:1 collagen-I:Matrigel mixture, being 1/3 of the total collagen DQ (D12060, Invitrogen), and seeded on a 24-well glass-bottom MatTek dish (P24G-1.5-10F, MatTek). Once the gel was set, cells were kept at 37C and 5% CO2 in MEM media for 48 hours. Next, the gels were fixed and permeabilised as previously described (Calvo et al., 2013), and cells were stain with DAPI and Phalloidin-AF555 (A34055, Life Technologies).
Quantitative analysis of images was preformed using ImageJ software. The intensity mean signal of the fluorescent products of DQ collagen I degradation was measured drawing a proportional square area to the cell cluster size.

2.18 Live cell imaging

2.18.1 Cell tracking (2D motility)

For two dimensional image acquisition, cells were plated sub-confluent into 24-well glass bottom MatTek plates (P24G-1.5-10F, MatTek) coated with collagen solution. Once attached, cells were kept in the microscope chamber at 37C and 5% CO_2 and images were acquired every 5 mins for 12 hours. The acquired digital images were merged using MetaMorph Automation and Imaging Analysis Software (Molecular Devices Inc.). Cell tracking speed analysis was performed using Mathematica 8.0 (Wolfram).

2.18.2 Cell growth curves

Primary PyMT $CD24^{+}AXL^{+}$ cells were FACS sorted and plated sub-confluent into 96-well high content imaging black plates (CLS4580, Corning) coated with collagen solution. Cell growth was monitored over 3-4 days using IncuCyteZOOM® Live Cell Imaging. All conditions were performed in triplicate and 3 images per well were acquired every 3 hours. The total area covered by cells was automatically measured, and the percentage of confluency over time was calculated as the average area covered by cells relative to the total well area of the triplicates.

2.19 CAF induction assays

To assess fibroblast activation via their ability to remodel an extracellular matrix, normal primary lung fibroblasts or the NLF3 cell line were embedded in 100 µl of 2:1 collagen-I:Matrigel and seeded on a 35-mm glass-bottom MatTek dish (P35-1.5-14-C, MatTek). Once the gel was set, we introduced a low height co-culture insert (PICM-ORG-50, Millipore) were tumour cells were seeded in a 1:3 fibroblast:tumour cells ratio. The co-culture was maintained in MEM media (DMEM/F12 supplemented with 2% FCS, 10µg/ml Insulin (I9278, Sigma), 20ng/ml EGF (PHG0313, Invitrogen) and 1:50 L-Glutamax (35050-061, Life Technologies). Gel contraction was monitored daily by taking photographs of the gels. Unless

stated otherwise, the gel contraction value refers to the contraction observed after 3 days. To obtain the gel contraction value the relative diameters of the well and the gel were measured using ImageJ software, and the contraction was calculated as the gel area relative to the well area in arbitrary units (A.U.).

2.20 3D spheroid assays

2.20.1 Mixed spheroid invasion assay

GFP^+ MICs were freshly isolated from an actin-GFP MMTV-PyMT mouse and GFP^- nonMICs were isolated from unlabelled MMTV-PyMT mouse by cell sorting according to their CD24/CD90 expression (see 2.16).
Next, GFP^+ MICs and GFP^- nonMICs were re-aggregated in ultra-low attachment conditions overnight in 1:9 ratio, in line with their relative distribution in primary tumours *in vivo*. The micro-spheroids formed were embedded in: pure collagen-I BD Biosciences cat. no. 354249), a 2:1 mixture of collagen-I and Matrigel (BD Biosciences cat. no. 354234), yielding a final collagen concentration of ~4.6 mg/ml and a final Matrigel concentration of ~2.2 mg/ml, or pure Matrigel gels. Microspheroids were allowed to invade for 48 hours before fixation with 4% PFA and staining with Phalloidin-TRITC (Sigma P-1951) and DAPI. Spheroids were imaged using a Zeiss LSM780 confocal microscope and invasive cells quantified using Volocity 3D Image Analysis Software by PerkinElmer.

2.20.2 Spheroids-CAFs invasion assay

FACS sorted GFP^+ MICs or nonMICs were re-suspended in DMEM with 10% FBS and 0.2% methyl cellulose at 2.5 x 10^5 cells/ml with or without 2.5 x 10^4 mcherry labelled PyMT-CAF. Spheroids were formed by hanging drop overnight and then embedded in 2:1 collagen-I:Matrigel mix. Spheroids were allowed to invade for 48 hours before fixation with 4% PFA, followed by DAPI staining. Spheroids were imaged using a Zeiss LSM780 confocal microscope and invasive cells quantified using Volocity 3D Image Analysis Software by PerkinElmer.

2.20.3 Organoid assay in Matrigel

A single cell suspension of FACS sorted MICs was directly seeded in pure Matrigel and maintain in Sphere Media (DMEM/F12 supplemented with 1:50 B27 (Invitrogen), 20ng/ml EGF (PHG0313, Invitrogen), 20ng/ml FGF (PHG0026, Invitrogen), 4mg/ml heparin (H3393, Sigma)) at 37C and 5% CO_2 for 10 days. After 3 days cells are grouped into clusters. Quantitative analysis of images was performed 10 days after seeding using ImageJ software. The Organoid Formation Index (OFI) was determined considering both, number and size of the organoids, and was obtained by summing the area of all organoids formed divided by the number of single cells initially plated.

Table 2.5 Antibodies

ANTIBODY	COMPANY	REF CODE	CLONAL (CLONE)	DILUTION (TECHNIQUE)
Endomucin	Santa Cruz	sc-65495	Rat monoclonal (V.7C7)	1:200 (IHC-P)
Ki67 (human)	Dako	A047	Rabbit polyclonal	1:500 (IHC-P)
Ki67	Novocastra	NCL-Ki67p	Rabbit polyclonal	1:200 (IHC-P)
FAP-α	Santa Cruz	sc-54539	Goat polyclonal	1:100 (IF)
YAP	Santa Cruz	sc-101199	Mouse monoclonal (63.7)	1:50 (IF)
Vimentin (human)	Abcam	ab8978	Mouse monoclonal	1:100 (IF)
Vimentin	Sigma	V2258	Mouse monoclonal (LN-6)	1:200 (IF)
E-cadherin	Abcam	ab11512	Rat monoclonal (DECMA-1)	1:200 (IF)
Twist	Santa Cruz	sc-81417	Mouse monoclonal	1:50 (FC)
Twist1	Abcam	ab49254	Rabbit polyclonal	1:200 (IF)
Zeb1	Santa Cruz	Sc-25388	Rabbit polyclonal	1:100 (IF)
Snail	Abcam	ab82846	Rabbit polyclonal	1:100 (IF)
AXL-biotinylated	R&D systems	BAF854	Goat polyclonal	1:100 (FC)
AXL	R&D systems	MAB8541	Rat polyclonal	1:100 (IF)
AXL (human)	R&D systems	MAB154	Mouse polyclonal	1:100 (IF/FC)
Integrin-α4 - PE	eBioscience	12-0492-81	Rat monoclonal (R1-2)	1:300 (FC)
Integrin-α6 - PE	eBioscience	12-0495-82	Rat monoclonal (eBioGoH3)	1:300 (FC)
Integrin-α7 - PE	LS Bio	LS-C106132	Hamster monoclonal (HMb1-1)	1:100 (FC)
Integrin-β1 active	BD bioscience	550531	Rat monoclonal (9EG7)	1:50 (IF)
Integrin-β1	Abcam	ab24693	Mouse monoclonal (P5D2)	1:100 (B/IF)
Integrin-β1 - APC	eBioscience	17-0291-82	Hamster monoclonal (HMb1-1)	1:300 (FC)
Integrin-β2	BD bioscience	557440	Rat monoclonal (GAME-46)	1:300 (FC)
Integrin-β4	eBioscience	50-1049-82	Rat monoclonal (439-9B)	1:300 (FC)
CD24-AF450	eBioscience	48-0242-82	Rat monoclonal (M1/69)	1:100 (FC)
CD90-APC	eBioscience	17-0900-82	Mouse monoclonal (HIS51)	1:100 (FC)
CD31-PE	eBioscience	12-0311-82	Rat monoclonal (390)	1:300 (FC)
CD45-PE	eBioscience	12-0451-	Rat monoclonal (30-F11)	1:300 (FC)

		82		
Ter119-PE	eBioscience	12-5921-82	Rat monoclonal (Ter-120)	1:300 (FC)
FAK	Santa Cruz	sc-558	Rabbit polyclonal	1:300 (IF)
P-FAK	Santa Cruz	sc-16563-R	Rabbit polyclonal	1:50 (IF)
THBS2	Santa Cruz	sc-12313	Goat polyclonal	1:100 (IF)
ID1	Biocheck	BCH-1/195-14	Rabbit monoclonal (195-14)	1:100 (IF) / 1:5000 (WB)
ID1	GeneTex	GTX53624	Rabbit polyclonal	1:20 (FC)
pSMAD1-5	Cell signal	13820	Rabbit monoclonal (D5B10)	1:100 (IF)
pSMAD1-5	Cell signal	9511	Rabbit monoclonal (D5B10)	1:100 (IF) / 1:20 (FC)
pSMAD2-3	Cell signal	8828	Rabbit monoclonal (D27F4)	1:100 (IF) / 1:20 (FC)
BrdU	Abcam	ab6326	Rat monoclonal (BU1/75 ICR1)	1:100 (FC)

- All antibodies are mouse specific, unless otherwise specified.
- IF: Immunofluorescence
- FC: Flow cytometry
- WB: Western Blot
- IHC-P: Immunohistochemistry – paraffin embedded tissue
- B: Blocking antibody

Chapter 3. MICs display a highly secretory mesenchymal status that enhances fibroblasts activation and metastasis initiation

3.1 MICs isolated from primary breast tumours display a highly secretory mesenchymal phenotype

As discussed in section 1.3.1, MICs were previously described in the MMTV-PyMT mouse model of breast cancer that spontaneously metastasises to the lungs. MICs are characterised by the co-expression of the epithelial marker CD24, previously shown to stain all epithelial cells in the mammary gland tissue (Sleeman et al., 2006) and the CD90 marker, which was previously shown to be highly expressed by mammary bipotent epithelial progenitor cells in human (Raouf et al., 2008), and a small subpool of $CD24^+$ epithelial cells with tumour-initiating abilities in both, mouse and human (Cho et al., 2008, Lu et al., 2014); see introduction – section 1.3.1 for more details about these markers. This lineage negative ($CD45^-$, $Ter119^-$ and $CD31^-$), $CD24^+CD90^+$ subpool in the MMTV-PyMT tumours was functionally defined as metastasis-initiating cells (MICs) by their exclusive ability to initiate metastatic growth in experimental metastasis assays (Malanchi et al., 2012). In this thesis, we used this functionally validated markers, CD24 and CD90, to further characterise this population of MICs.

3.1.1 MIC gene expression profile

In order to investigate which mechanisms allow MICs to successfully seed metastases we analysed their gene expression profile using microarrays. MICs account for about 2-3% of the total tumour cell mass at the primary site and when directly isolated and tested *in vivo* show a higher metastatic behaviour. To gain some insight on the intrinsic characteristics discriminating MICs from the bulk of the tumour, we freshly isolated MICs ($CD24^+CD90^+$) and nonMICs ($CD24^+CD90^-$) from late-stage MMTV-PyMT tumours by cell sorting (Figure 3.1.A). In order to identify

MICs in total PyMT tumour cell preparations, we excluded small particles and cell doublets (scatter), death cells (viability dye) and lineage negative cells (immune cells (CD45) endothelial cells (CD31) and erythrocytes (Ter119)) by the gating strategy shown in Figure 3.1.A. Next, CD24 allows the discrimination of the total mammary epithelial cells, $CD24^+$, compared to other non-epithelial lineages, $CD24^-$ (Figure 3.1.A), such as fibroblasts that cannot be excluded otherwise. The epithelial populations of MICs ($CD24^+CD90^+$) and nonMICs ($CD24^+CD90^-$) were freshly isolated from different PyMT tumours and their gene expression profiles were generated and compared to create a MIC signature (Figure 3.1.B and Appendix 7.1).

In collaboration with Probir Chakravarty from the Bioinformatics Unit at the London Research Institute, we performed a global gene ontology (GO) analysis to determine the cellular compartments and biological processes most represented in the MIC signature. Regarding the cellular compartment analysis we found that most gene products upregulated in MICs are localised within the cell membrane or in the extracellular region (Figure 3.2.A). This highly secretory status suggests that these cells are more prone to interact with their surrounding microenvironment. The GO analysis on biological processes, consistent with previous reports confirmed high Wnt signalling (Malanchi et al., 2012, Malanchi et al., 2008, Reya and Clevers, 2005) and TGFβ signalling (Oshimori et al., 2015) to be hallmarks of tumour-initiating cells. Moreover, we observed highly upregulated cell adhesion and extracellular matrix (ECM) remodelling gene sets and a protease signature (Figure 3.2.B), reaffirming the previous cellular compartment analysis that indicated an extracellular localisation for most MIC upregulated proteins. Interestingly, we also found the MIC signature to highly correlate with an epithelial-to-mesenchymal transition (EMT) programme (Figure 3.2.B) that could partially explain the high levels of secreted components expressed by these cells (Chong et al., 2012).

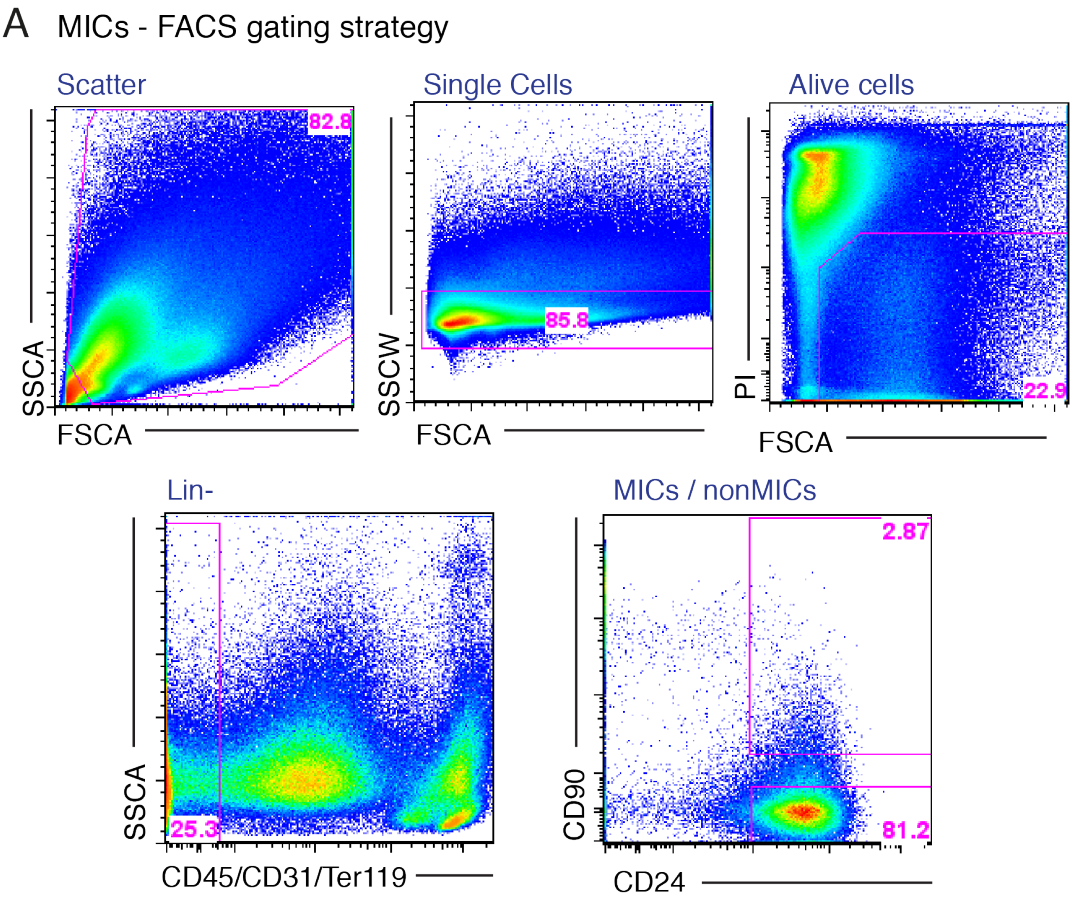
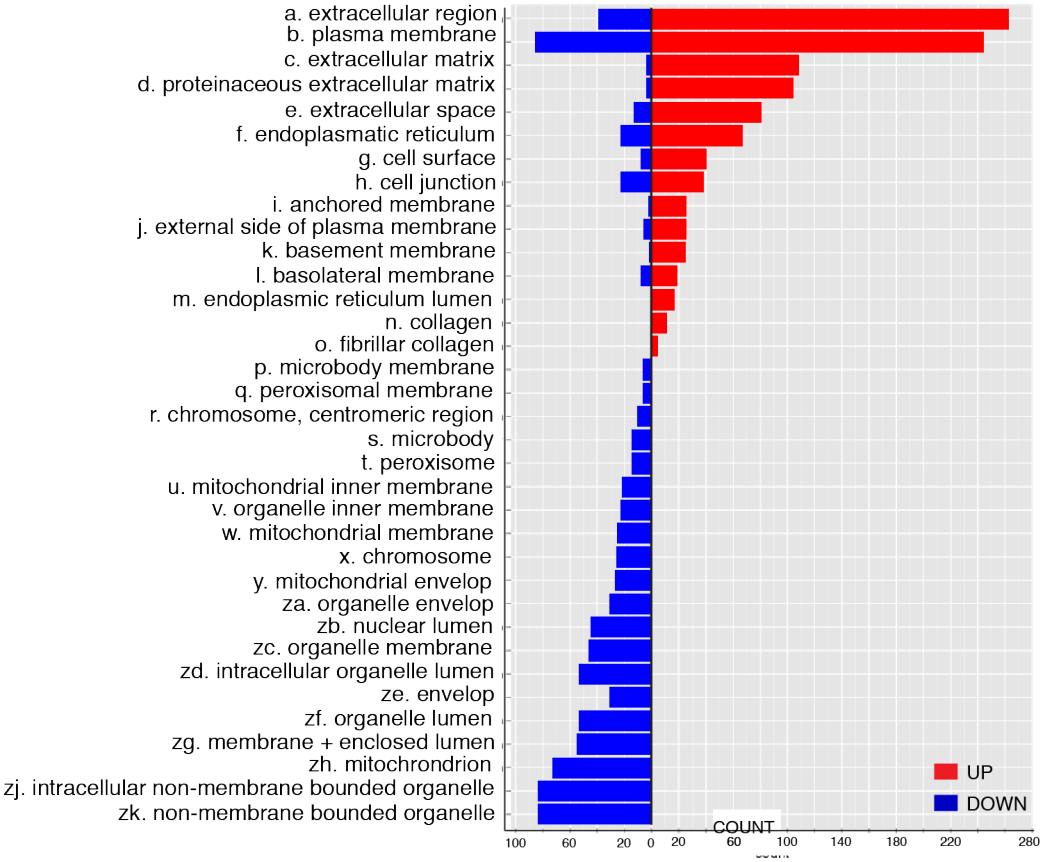


Figure 3.1 MIC isolation strategy and generation of the MIC signature
(A) FACS gating strategy to isolate MICs and nonMICs from PyMT tumours. First, forward/side scatter; second, single cells; third, exclude death cells (PI-positive); forth, gate on lineage negative (Lin-: CD45-, CD31- and Ter119-; excludes immune cells, endothelial cells and erythrocytes respectively); fifth, plot CD24 against CD90, where double positive cells are MICs and CD24 single positives are nonMICs. (B) Heatmap compares the gene expression profiles of MICs and nonMICs isolated from late-stage MMTV-PyMT primary tumours.

A GO cellular compartment - MIC signature



B GO biological processes - MIC signature

Process name	0 2 4 6 8 10 12 -log (pValue)	p value
Cell adhesion_ECM remodelling		1.842e-16
Epithelial to mesenchymal transition		2.419e-08
Proteases in hematopoietic stem cell mobilisation		2.000e-07
TGF-β, WNT and cytoskeletal remodelling		4.059e-07
Cytoskeleton remodelling		5.724e-07

Figure 3.2 MIC signature: a global analysis

(A) GO analysis of cellular compartments indicates the cellular localisation of genes upregulated in MICs. (C) GO analysis of the top biological processes significantly upregulated in MICs.

Next, we performed a more detailed pathway analysis of the MIC gene expression profile validating by Gene Set Enrichment Analysis (GSEA) the observed general features described in the initial GO analysis (Figure 3.2). We found the transcriptome of MICs highly correlating with previously published signatures for Wnt signalling (available at Wnt homepage Stanford University), TGFβ signalling and EMT (Figure 3.3.A) (Farmer et al., 2009) as suggested by the GO analysis. Moreover, the MIC signature highly correlated also with YAP/TAZ signalling (Figure 3.3.A) (Dupont et al., 2011), a growth control regulatory pathway with emerging potent roles in cancer stem cells and mechanotransduction (Piccolo et al., 2013).

Subsequently, we generated a circos plot in collaboration with Probir Chakravarty summarising the main pathways found in the MIC signature (Figure 3.3.B). We established 3 colour-coded categories for these pathways, the first including Wnt signalling that is related to self-renewal and tumour-initiating potential (grey), the second including YAP/TAZ, TGFβ signalling and EMT is related to cancer cells' invasion (red), and the third category groups different pathways related to microenvironmental regulation, secretion and cell-ECM interactions (blue). Additionally, we highlighted in the plot several genes related to these pathways that are highly expressed in MICs (Appendix 7.1) and were previously characterised to contribute to lung metastatic colonisation in breast cancer, such as SPARC, TNC, LOX, MMPs and CCL2 (section 1.1.2.2, 1.3.2 and 1.5.2).

To find the commonalities between these three pathway categories, we decided to further characterise the EMT signature found in MICs since EMT have been previously described to contribute to all these general features in cancer cells: microenvironmental regulation (Taube et al., 2010, Lopez-Novoa and Nieto, 2009), invasion (Rhim et al., 2012, Tsai et al., 2012) and stemness (Mani et al., 2008, Beck et al., 2015).

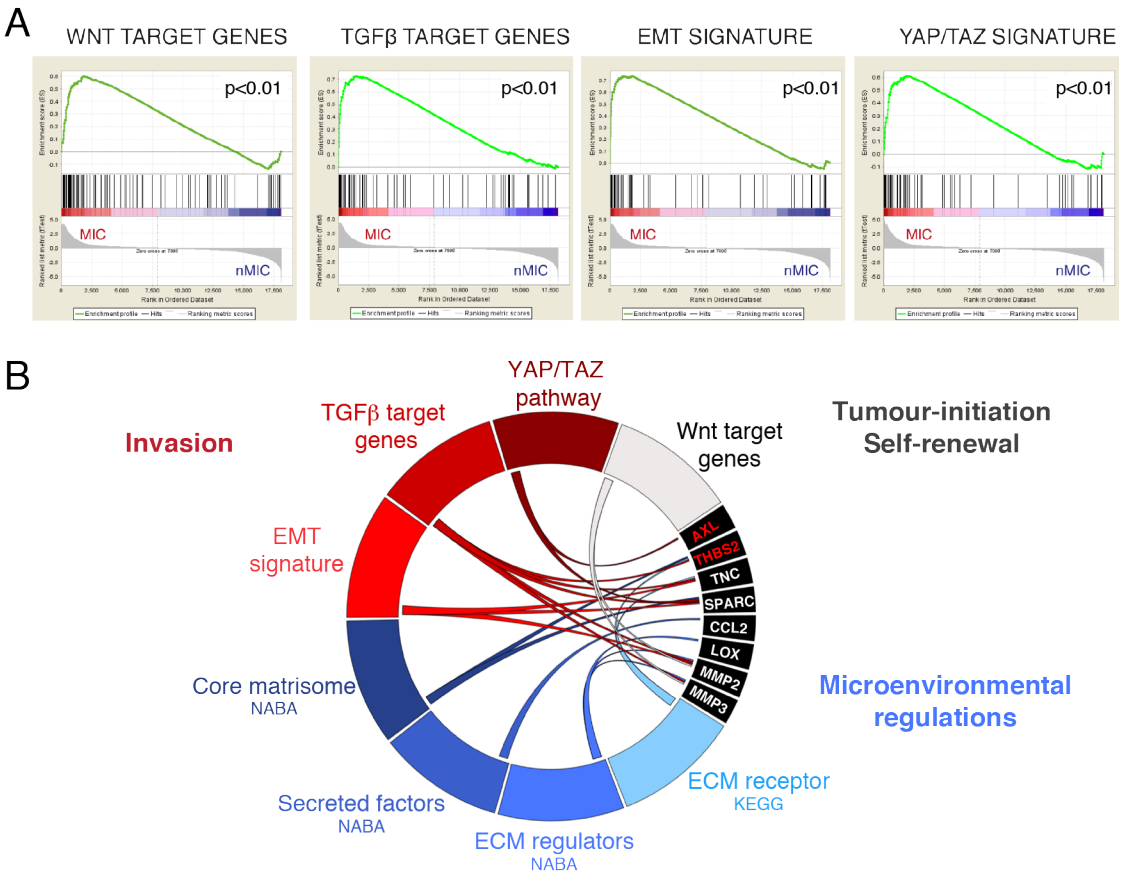


Figure 3.3 MIC signature: detailed analysis of relevant processes and genes involved in metastasis
(A) GSEA plots show the correlation of the MIC signature with Wnt target genes, TGFβ target genes, an EMT and a YAP/TAZ signature. (B) Circos plot displays the top up-regulated pathways in MICs: Wnt, related to self-renewal and tumour-initiation (grey), TGFβ, EMT and YAP/TAZ related to invasion mechanisms in cancer cells (red), and pathways related to microenvironmental regulation and cell-ECM interactions (blue). Relevant genes related to these pathways and previously described to contribute to the metastatic abilities of cancer cells such as SPARC, TNC or LOX are indicated in white, whereas other genes related to these pathways and analysed in this thesis are highlighted in red.

In summary, MICs express many previously characterised factors that could contribute to their metastatic potential, supporting the described high metastatic ability of these cells (Malanchi et al., 2012). Importantly, we found different pathway categories, all related to a highly secretory EMT status defining the gene expression profile of MICs.

3.1.2 MICs isolated from primary breast tumours display a partial mesenchymal state

In order to functionally validate the highly secretory EMT status of MICs, we freshly isolated MICs and nonMICs from late-stage MMTV-PyMT tumours and analysed the presence of different well-characterised EMT markers. Despite the high level of heterogeneity, we found the MIC subpool to be in a more mesenchymal state compared to nonMICs, defined by the expression of low E-cadherin and high Vimentin (Figure 3.4.A-C). However, we could observe that around 25% of the cells within the MIC subpool showed mild E-cadherin expression at the membrane (Figure 3.4.B). The fact that MICs maintain the epithelial marker E-cadherin suggests that these cells are in an intermediate EMT state rather than undergoing a full mesenchymal conversion.

The more mesenchymal status of MICs compared to the bulk of the tumour was further confirmed by the higher expression of the EMT core transcription factors at the mRNA level, as well as the clear nuclear localisation of Zeb1 in MICs compared to nonMICs (Figure 3.5).

Consistently with the higher expression of EMT-TFs MICs express many genes related to a single cell invasion programme, the mesenchymal type of motility based on integrin-mediated ECM adhesion and the use of proteases to generate traction force (section 1.2.1.1). The first set of genes is related to the ECM via the secretion of ECM components or metalloproteases conferring ECM remodelling abilities (Figure 3.6.A). The second gene set is directly linked to motility through the regulation of actin cytoskeleton components (Figure 3.6.B). The expression of these genes strongly suggests differential ECM remodelling and invasive abilities in MICs.

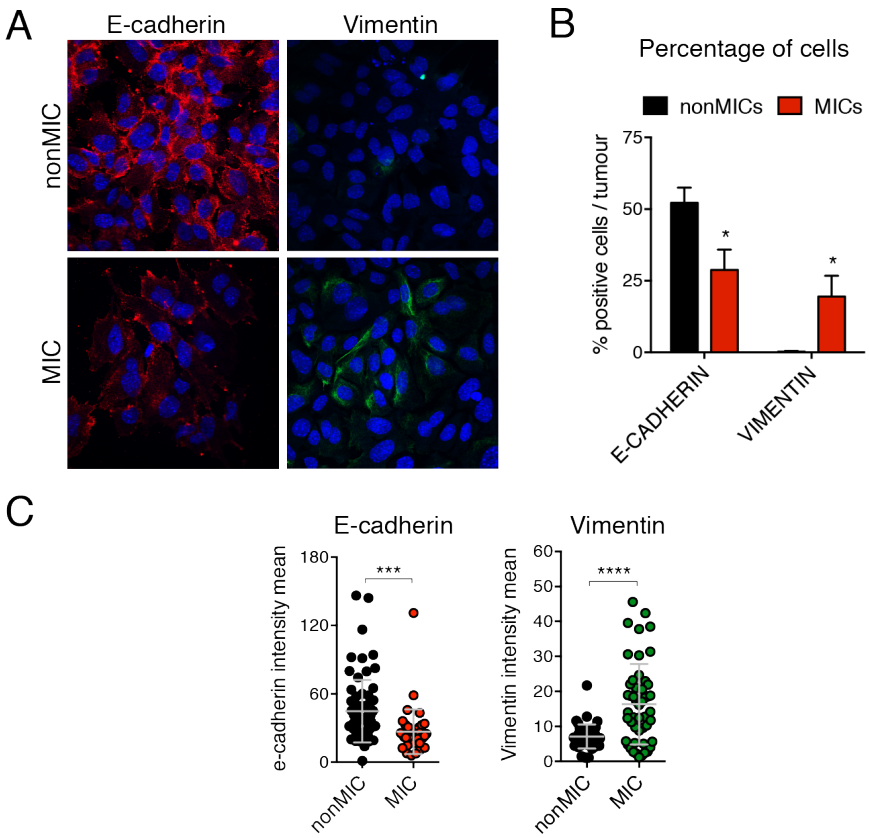


Figure 3.4. Mesenchymal phenotypic marker expression in MICs
(A) Representative images show the expression of E-cadherin and Vimentin in freshly isolated nonMICs ($CD24^{+}CD90^{-}$) and MICs ($CD24^{+}CD90^{+}$). Cells were stained for analysis 12 hours post-plating in collagen-coated dishes. Scale bar, 40 µm. (B) Graph indicates the percentage of E-cadherin and Vimentin positive cells in the MIC and nonMIC populations (n= 4 tumours). (C) Charts indicate the expression levels of E-cadherin and Vimentin in freshly isolated nonMICs and MICs. Data from one representative experiment from a total of 5 are shown. Line and error bars indicate mean ± SD of the population.

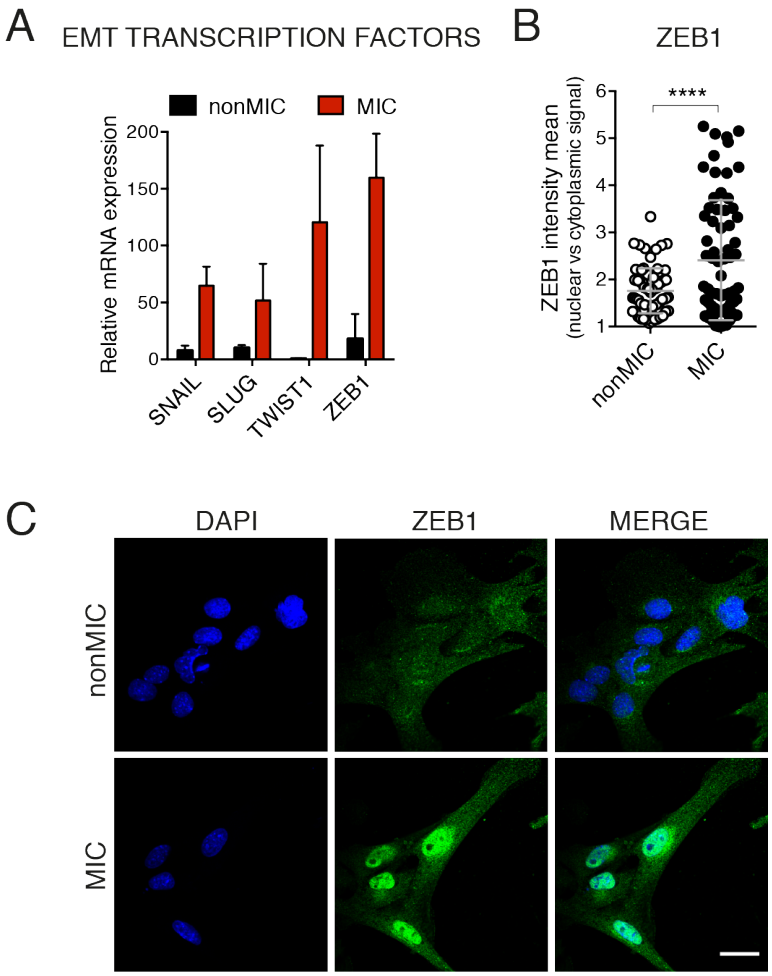


Figure 3.5 EMT core transcription factor expression in MICs
(A) Quantitative real time PCR analysis compares the gene expression levels of the core EMT transcription factors. Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bar represents mean ± sem. $p<0.0001$ by ANOVA comparing the gene sets of MIC *versus* nonMIC. (B) Charts indicate the level of nuclear *versus* cytoplasmic intensity mean of ZEB1. Data from one representative experiment from a total of 2 are shown. Line and error bar indicate mean ± SD of the population. (C) Images show ZEB1 cellular localisation (green) and DAPI (blue). Scale bars, 30 and 20 µm respectively.

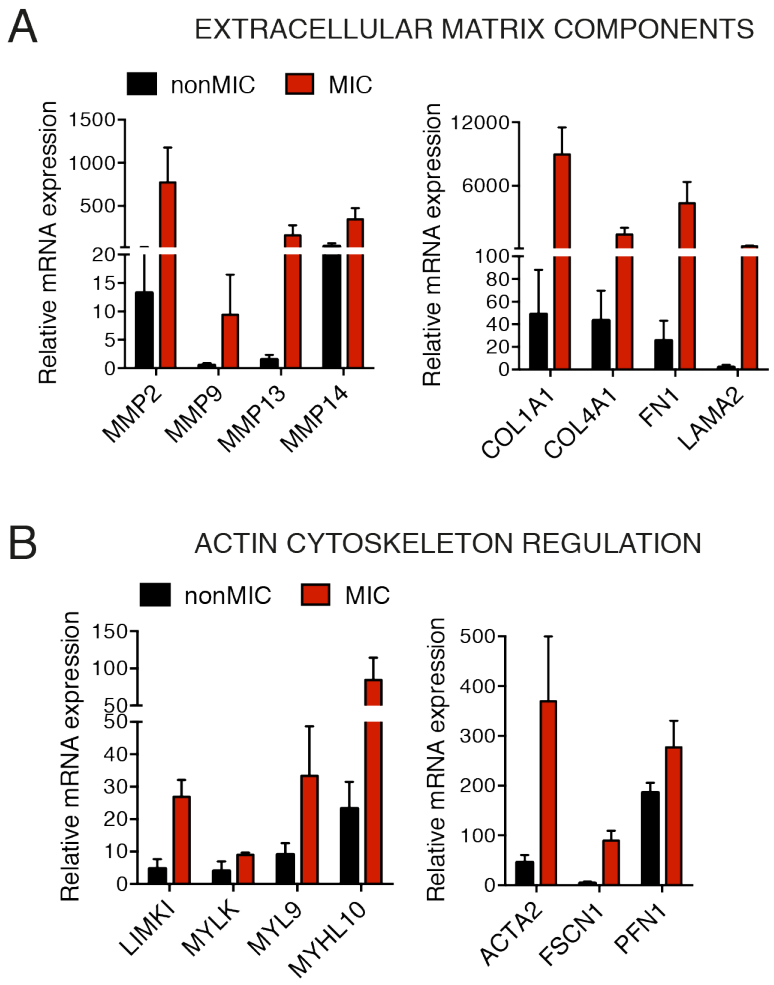


Figure 3.6 Expression of invasion-related genes in MICs
Quantitative real time PCR analysis compares the gene expression levels of extracellular matrix (A) and actin cytoskeleton (B) components in MICs and nonMICs. Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bars represent mean ± sem. $p<0.0002$ for each plot by ANOVA comparing the gene sets in MIC *versus* nonMIC.

Consequently, we next validated these observations by performing functional assays *ex vivo* after MICs isolation from primary tumours. First, in order to corroborate the high secretion of ECM remodelling enzymes expected from MIC gene expression analysis we performed a DQ-collagen assay. MICs and nonMICs were embedded after FACS isolation into DQ-collagen-Matrigel matrices where the collagenase-driven hydrolysis of the DQ substrate results in separation of the quenched dye molecules from one another increasing the fluorescence signal (Figure 3.7.A). Although all tumour cells can degrade collagen fibers, MICs showed higher collagenase activity as reported by the DQ signal (Figure 3.7.B).

Second, to validate the invasive phenotype suggested by the gene expression analysis and the partial mesenchymal phenotype, we performed three-dimensional invasion assays. GFP-MICs and non-labelled nonMICs isolated from actin-GFP or unlabelled MMTV-PyMT tumours were mixed in a 1:9 ratio forming spheroids where the *in vivo* MICs:nonMICs frequency is maintained (Figure 3.7.C). We collaborated with Danielle Park, from the Sahai Lab at the London Research Institute to perform the 3D invasion assays. Danielle embedded the mixed spheroids in matrices with different permissiveness towards invasion: pure Matrigel rich in laminins, collagen IV and other basement membrane components that do not facilitate invasion; pure collagen-I matrices, a stiffer environment similar to the stromal border of tumours that promotes cancer cell invasion; and 2:1 collagen-I:Matrigel mixes, an intermediate condition that allows mammary cells *in vitro* to closely reproduce their *in vivo* behaviour (Nguyen-Ngoc et al., 2012). In these three-dimensional organotypic assays, PyMT cells in the spheroids can invade into the matrices as collective protrusions or as single cells (mesenchymal-like invasion). Therefore, to evaluate the mesenchymal features of labelled (MICs) or unlabelled (nonMICs) cells in each spheroid, we analysed the invasion of single cell elongated protrusions. As expected from the matrices properties all cancer cells could invade as elongated protrusions in collagen-I pure matrices whereas only some MICs formed elongated protrusions in pure Matrigel (Figure 3.7.D). Interestingly, in the 2:1 collagen-I:Matrigel mixture that more closely reproduces the *in vivo* microenvironment we could observe a differential phenotype: nonMICs invaded collectively forming bulbous protrusions while MICs clearly invaded as single cell elongated protrusions (Figure 3.7.D-E). This result reaffirms that MICs display a more mesenchymal invasive phenotype.

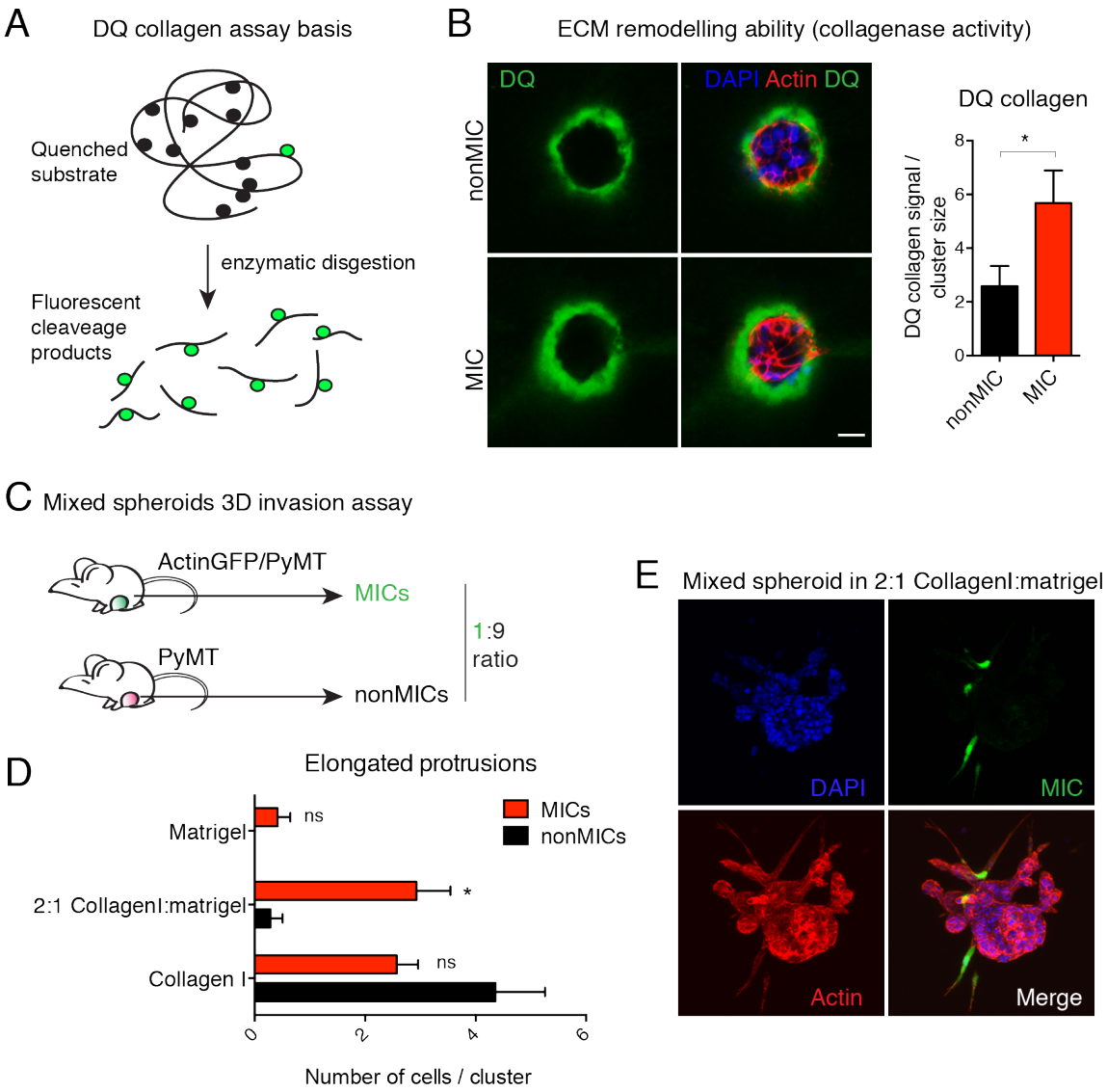


Figure 3.7 Secretory activity and invasive properties of MICs
(A) Schematic shows the working principle of the DQ collagen assay. The collagen fibers contain a fluorescent substrate that is quenched by the close proximity of the molecules within the three-dimensional structure of the collagen (black dots). When the collagen fibers are cleaved, substrate fluorescent increases (green dots). (B) DQ collagen assay. MICs and nonMICs were freshly isolated from MMTV-PyMT tumours and seeded into 2:1 DQ-collagen:Matrigel gels. After 48h, the formed cell clusters were fixed and stained, and their enzymatic activity analysed. (Left) Images show the DQ collagen signal (green) surrounding nonMIC or MIC clusters delimited by F-actin (red) and DAPI (blue) staining. Scale bar, 15 µm. (Right) Histogram indicates mean ± sem of the DQ collagen signal relative to the cell cluster size. (C) Mixed spheroids invasion assay. Experimental set up: MICs from PyMT/actin-GFP tumours, and nonMICs from non-labelled PyMT tumours were freshly isolated by FACS, and re-aggregated overnight in suspension in a 1:9

MIC:nonMIC ratio. The spheroids formed were embedded in different matrices: only collagen-I, 2:1 collagen-I:Matrigel and only Matrigel. (D) Histogram shows mean ± sem of cells invading as elongated protrusions within the clusters in the different matrices. (E) Representative fluorescent images of one micro-spheroid seeded in 2:1 collagen-I:Matrigel mixture stained with DAPI (blue) and F-actin (red). Note the invasive MIC cells identified by GFP (green).

Taken together, these functional data indicate that MICs are in a highly secretory, more mesenchymal state in comparison to the bulk cancer cells in the tumour mass.

3.1.3 MIC mesenchymal features actively contribute to their metastatic colonisation potential at the target site

Next, we aimed to investigate the contribution of the observed mesenchymal phenotype to the metastatic colonisation potential of MICs. Therefore we analysed the generated microarray data searching for differentially expressed genes in MICs, focusing on those directly related to EMT – in Appendix 7.1 of this thesis we display the list of genes that showed a fold increment expression (MICs/nonMICs) ≥ 3, and below a summary list containing the top 100 upregulated genes in the MIC signature can be found (Table 3.1). We found several previously reported EMT-related genes highly expressed in MICs: Fibronectin1 (FN1) and different metalloproteases (MMP13, MMP2, MMP3, MMP14), validated at the mRNA level in Figure 3.6; and the mesenchymal intermediate filament protein Vimentin, validated at the protein level in Figure 3.4. Among these highly expressed genes in MICs related to EMT (Table 3.1 – blue), we found the tyrosine kinase receptor AXL that have been previously characterised as a downstream effector of the EMT programme (Asiedu et al., 2013). Moreover, AXL expression is also dependent on TGFβ and YAP/TAZ pathways in solid cancers, enhancing tumour cell dissemination from the primary site (Li et al., 2014b, Li et al., 2014a). Accordingly, AXL have been shown to correlate with a higher metastatic risk and poor prognosis in breast cancer patients (Gjerdrum et al., 2010) – a more detailed description of AXL functions in breast cancer is included in section 1.2.1.1.

We decided to further validate AXL as an EMT marker in MICs because of 1) its correlation with poor prognosis in breast cancer patients, 2) the fact that 3 different pathways (EMT, TGFβ and YAP/TAZ signalling) found to be overexpressed in MICs can control AXL expression, and 3) the fact that it is a transmembrane receptor that facilitates monitoring the EMT status of cancer cells *in vivo*. First, using cBioPortal for Cancer Genomics database, we confirmed in human breast cancer patient data the previously published observations in breast cancer animal models showing that AXL presence correlates with the expression of different EMT markers (Figure 3.8.A). Second, we validated the array data performing quantitative real time PCR and immunofluorescence in freshly isolated MICs and nonMICs, confirming the higher expression of AXL in the MIC compartment of the PyMT tumours (Figure 3.8.B-C).

No	GENE SYMBOL	FOLD INCREMENT	DEFINITION
1	Fn1	73.60964635	Mus musculus fibronectin 1 (Fn1), mRNA.
2	Serping1	60.46920116	Mus musculus serine (or cysteine) peptidase inhibitor, clade G, member 1 (Serping1), mRNA.
3	Lum	60.06918146	Mus musculus lumican (Lum), mRNA.
4	Thy1	59.32025165	Mus musculus thymus cell antigen 1, theta (Thy1) or CD90, mRNA.
5	Mmp3	56.8142886	Mus musculus matrix metallopeptidase 3 (Mmp3), mRNA.
6	Serping1	54.48133702	
7	Mmp13	50.80097994	Mus musculus matrix metallopeptidase 13 (Mmp13), mRNA.
8	Loxl1	49.87156687	Mus musculus lysyl oxidase-like 1 (Loxl1), mRNA.
9	Col1a1	47.68421086	
10	Ctsk	44.40448725	Mus musculus cathepsin K (Ctsk), mRNA.
11	Dcn	43.61584256	Mus musculus decorin (Dcn), mRNA.
12	Col6a1	43.02501894	Mus musculus procollagen, type VI, alpha 1 (Col6a1), mRNA.
13	Mmp2	42.58322206	Mus musculus matrix metallopeptidase 2 (Mmp2), mRNA.
14	Adamts2	41.75490813	Mus musculus a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 2 (Adamts2), mRNA.
15	Lrrc15	41.68029907	
16	Mfap5	41.64301515	Mus musculus microfibrillar associated protein 5 (Mfap5), mRNA.
17	Col5a1	41.04235801	Mus musculus procollagen, type V, alpha 1 (Col5a1), mRNA.
18	Rnase4	40.18861336	Mus musculus ribonuclease, RNase A family 4 (Rnase4), transcript variant 1, mRNA.
19	Pi16	39.30356586	Mus musculus peptidase inhibitor 16 (Pi16), mRNA.
20	Rarres2	38.82640932	Mus musculus retinoic acid receptor responder (tazarotene induced) 2 (Rarres2), mRNA.
21	Adamts2	37.03161082	Mus musculus a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 2 (Adamts2), mRNA.
22	Pla1a	35.33017006	Mus musculus phospholipase A1 member A (Pla1a), mRNA.
23	Cpxm1	34.56071134	Mus musculus carboxypeptidase X 1 (M14 family) (Cpxm1), mRNA.
24	Cygb	34.15923397	Mus musculus cytoglobin (Cygb), mRNA.
25	Ccdc80	34.05706562	Mus musculus coiled-coil domain containing 80 (Ccdc80), mRNA.
26	Ccl11	33.36436321	Mus musculus small chemokine (C-C motif) ligand 11 (Ccl11), mRNA.
27	Fstl1	32.97793885	Mus musculus follistatin-like 1 (Fstl1), mRNA.
28	Lox	32.66522761	Mus musculus lysyl oxidase (Lox), mRNA.
29	Col6a1	32.32838235	Mus musculus procollagen, type VI, alpha 1 (Col6a1), mRNA.
30	Col6a2	32.24366031	Mus musculus procollagen, type VI, alpha 2 (Col6a2), mRNA.
31	Serpinf1	32.22446558	Mus musculus serine (or cysteine) peptidase inhibitor, clade F, member 1 (Serpinf1), mRNA.
32	Pcolce	32.16013333	Mus musculus procollagen C-endopeptidase enhancer protein (Pcolce), mRNA.
33	Scara3	31.79005897	Mus musculus scavenger receptor class A, member 3 (Scara3), mRNA.
34	Aebp1	31.55275609	Mus musculus AE binding protein 1 (Aebp1), mRNA.
35	Fbln1	31.30597931	Mus musculus fibulin 1 (Fbln1), mRNA.
36	Mmp3	31.29927231	Mus musculus matrix metallopeptidase 3 (Mmp3), mRNA.
37	Fbn1	30.84942594	
38	Ogn	30.62286362	Mus musculus osteoglycin (Ogn), mRNA.
39	Mmp3	30.22038636	Mus musculus matrix metallopeptidase 3 (Mmp3), mRNA.
40	Tnxb	29.18681299	Mus musculus tenascin XB (Tnxb), mRNA.
41	Dpt	28.81349432	Mus musculus dermatopontin (Dpt), mRNA.

42	**Thbs2**	**27.92699494**	**Mus musculus thrombospondin 2 (Thbs2), mRNA.**
43	LOC100047583	27.8890865	PREDICTED: Mus musculus similar to apolipoprotein D (LOC100047583), mRNA.
44	Serpina3n	27.67160631	Mus musculus serine (or cysteine) peptidase inhibitor, clade A, member 3N (Serpina3n), mRNA.
45	Col12a1	27.39104496	Mus musculus collagen, type XII, alpha 1 (Col12a1), mRNA.
46	Knsl5	26.95470532	
47	Svep1	26.90604697	Mus musculus sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1 (Svep1), mRNA.
48	Spon2	26.83853041	Mus musculus spondin 2, extracellular matrix protein (Spon2), mRNA.
49	Plat	26.34455797	
50	Cd248	26.15887428	Mus musculus CD248 antigen, endosialin (Cd248), mRNA.
51	Col14a1	25.92587138	
52	LOC638301	25.55375508	PREDICTED: Mus musculus similar to interferon activated gene 204 (LOC638301), mRNA.
53	Bicc1	25.1432103	Mus musculus bicaudal C homolog 1 (Drosophila) (Bicc1), mRNA.
54	Srpx2	24.90728486	Mus musculus sushi-repeat-containing protein, X-linked 2 (Srpx2), transcript variant 2, mRNA.
55	LOC100044430	24.55997339	PREDICTED: Mus musculus similar to Interferon activated gene 205 (LOC100044430), mRNA.
56	Tnxb	24.27757116	Mus musculus tenascin XB (Tnxb), mRNA.
57	Sparc	24.11332411	Mus musculus secreted acidic cysteine rich glycoprotein (Sparc), mRNA.
58	Igfbp4	23.58079355	
59	scl0001849.1_2273	23.17580957	
60	6330406I15Rik	22.6360887	Mus musculus RIKEN cDNA 6330406I15 gene (6330406I15Rik), mRNA.
61	Mfap2	22.34155105	Mus musculus microfibrillar-associated protein 2 (Mfap2), mRNA.
62	Pdgfra	22.29305127	Mus musculus platelet derived growth factor receptor, alpha polypeptide (Pdgfra), transcript variant 1, mRNA.
63	Itga11	22.19115202	Mus musculus integrin alpha 11 (Itga11), mRNA.
64	Fbln1	22.09933361	
65	Htra3	21.72598387	
66	Cxcl14	21.70535391	Mus musculus chemokine (C-X-C motif) ligand 14 (Cxcl14), mRNA.
67	Nid1	21.42034706	Mus musculus nidogen 1 (Nid1), mRNA.
68	Prelp	21.29436558	Mus musculus proline arginine-rich end leucine-rich repeat (Prelp), mRNA.
69	6330406I15Rik	20.90904317	Mus musculus RIKEN cDNA 6330406I15 gene (6330406I15Rik), mRNA.
70	Gpx8	20.8173571	Mus musculus glutathione peroxidase 8 (putative) (Gpx8), mRNA.
71	Scarf2	20.81035285	Mus musculus scavenger receptor class F, member 2 (Scarf2), mRNA.
72	Sparc	20.49187503	Mus musculus secreted acidic cysteine rich glycoprotein (Sparc), mRNA.
73	Ccl7	20.40200837	
74	Lum	20.27088078	
75	Bicc1	20.12116588	Mus musculus bicaudal C homolog 1 (Drosophila) (Bicc1), mRNA.
76	Osr2	19.93992721	Mus musculus odd-skipped related 2 (Drosophila) (Osr2), mRNA.
77	Olfml2b	19.85429131	Mus musculus olfactomedin-like 2B (Olfml2b), mRNA.
78	Serpine2	19.81592693	
79	Itgbl1	19.72979134	Mus musculus integrin, beta-like 1 (Itgbl1), mRNA.
80	Col8a1	19.66821469	
81	ENSMUSG00000043795	19.60575329	PREDICTED: Mus musculus predicted gene, ENSMUSG00000043795 (ENSMUSG00000043795), mRNA.

82	Timp1	19.04713677	Mus musculus tissue inhibitor of metalloproteinase 1 (Timp1), transcript variant 2, mRNA.
83	Dcn	18.84139087	Mus musculus decorin (Dcn), mRNA.
84	Timp1	18.41986061	
85	C1qtnf3	18.41977383	Mus musculus C1q and tumor necrosis factor related protein 3 (C1qtnf3), mRNA.
86	Igfbp4	18.36908763	Mus musculus insulin-like growth factor binding protein 4 (Igfbp4), mRNA.
87	Ly6c1	18.30177212	Mus musculus lymphocyte antigen 6 complex, locus C1 (Ly6c1), mRNA.
88	Serpina3h	18.14573816	Mus musculus serine (or cysteine) peptidase inhibitor, clade A, member 3H (Serpina3h), mRNA.
89	Gp38	17.90928912	
90	scl0002507.1_236	17.80258841	
91	Ppic	17.68622567	Mus musculus peptidylprolyl isomerase C (Ppic), mRNA.
92	Prg4	17.59252918	
93	Col3a1	17.5644643	Mus musculus collagen, type III, alpha 1 (Col3a1), mRNA.
94	C2	17.55475936	Mus musculus complement component 2 (within H-2S) (C2), mRNA.
95	2310016C16Rik	17.26052088	Mus musculus RIKEN cDNA 2310016C16 gene (2310016C16Rik), mRNA.
96	Has1	17.09402607	Mus musculus hyaluronan synthase1 (Has1), mRNA.
97	Gpx3	17.07563602	Mus musculus glutathione peroxidase 3 (Gpx3), transcript variant 2, mRNA.
98	Olfml3	16.73507945	Mus musculus olfactomedin-like 3 (Olfml3), mRNA.
99	Rarres2	16.7212608	Mus musculus retinoic acid receptor responder (tazarotene induced) 2 (Rarres2), mRNA.
100	**Axl**	**16.71865497**	**Mus musculus AXL receptor tyrosine kinase (Axl), mRNA.**

Table 3.1 Top 100 upregulated genes in the MIC signature

List contains a summary of the MIC gene expression profile data, top upregulated genes in MICs compared to nonMICs. Red – genes characterised in this study. Blue – genes related to Epithelial-to-Mesenchymal Transition (EMT).

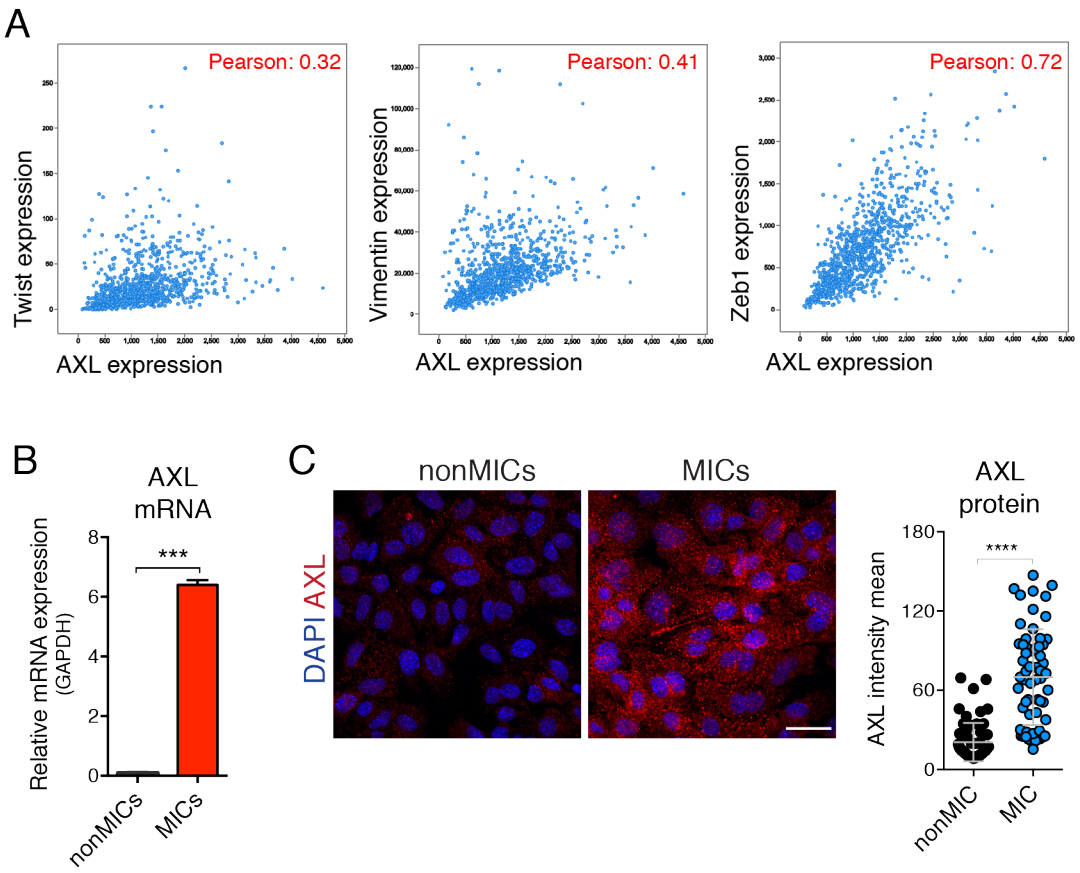


Figure 3.8 AXL correlates with the EMT status of cancer cells in human breast tumour data and is highly express by MICs
(A) Gene expression correlation analysis between AXL and Twist, Vimentin and Zeb1 mesenchymal markers in human breast carcinoma patient samples - cBioportal database. (B) Quantitative real time PCR analysis compares the gene expression levels of AXL. Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bar represents mean ± sem. (C) Representative images show AXL staining (red) and nuclear DAPI (blue) in MICs and nonMICs. Scale bar, 20 µm. Chart indicates the levels of AXL. Data from one representative experiment from a total of 3 are shown. Line and error bar indicate mean ± SD of the population.

Confirmed the high expression level of AXL in MICs, we decided to focus on the AXL^+-MICs in order to specifically address the potential contribution of the mesenchymal features to metastatic colonisation, independently of the rest of stem-like features associated to MICs (i.e. high Wnt signalling, Figure 3.3). We started analysing the relationship between the MIC marker ($CD90^+$) and AXL among the epithelial tumour cells ($CD24^+$) by flow cytometry in freshly isolated MMTV-PyMT tumour cells. The MIC subpool ($CD24^+CD90^+$) accounts for 2-3% of the alive, lineage negative (LIN^-: $CD45^-$, $Ter119^-$ and $CD31^-$) cells, of which around 60% are AXL^+ (Figure 3.9.A). The MIC compartment of the tumours displayed a 2-fold increment in AXL-expressing cells compare to nonMICs (Figure 3.9.A). Conversely, the mesenchymal $CD24^+AXL^+$ population constitutes around 28% of the total alive, lineage negative cells in PyMT tumours, of which around 6.93% of the are MICs ($CD24^+CD90^+$); the $CD24^+AXL^+$ population contains 3 times more MICs than the more epithelial $CD24^+AXL^-$ population (Figure 3.9.B). The relative frequency and percentage of overlap between these populations is represented in the Venn diagram, where the mesenchymal MICs (triple $CD24^+CD90^+AXL^+$ - dark blue) is indicated, and constitutes around 1% of the alive, lineage negative tumour cells (Figure 3.9.C). These results show that MICs expressed higher levels of AXL than nonMICs, and the $CD24^+AXL^+$ mesenchymal population is enriched in MICs.

As shown in the FACS plots, the MICs and $CD24^+AXL^+$ mesenchymal populations do not show a linear relationship and only partially overlap (Figure 3.9.C). In order to specifically address the contribution of the AXL mesenchymal features to the metastatic potential of MICs, we focused on the $CD24^+AXL^+$ mesenchymal population - evaluating the triple $CD24^+AXL^+CD90^+$ was not technically possible as this small population constitutes around 1% of the total tumour cells (Figure 3.9.C), and *in vivo* experimental metastatic assays were not technically possible to perform with such low cells numbers. Therefore we proceeded to functionally validate the mesenchymal $CD24^+AXL^+$ population, which is highly enriched in MICs compared to the $CD24^+AXL^-$ subpool.

Immunofluorescence analysis revealed, as expected by the presence of the mesenchymal marker AXL, that both MICs and the $CD24^+AXL^+$ population display

a partial mesenchymal phenotype, with significantly higher number of Vimentin-positive cells and lower E-cadherin compared to the more epithelial nonMICs and $CD24^+AXL^-$ populations, respectively (Figure 3.10.A).

Next, in order to specifically address the contribution of the mesenchymal features to the metastatic potential of MICs, we assayed the metastasis-initiating potential of the mesenchymal $CD24^+AXL^+$ population in parallel to the more epithelial $CD24^+AXL^-$ tumour fraction. To exclude any advantage to grow a primary tumour mass *in vivo*, we performed in parallel primary tumour formation assays with both subpools. We used tail vein injection, or orthotopic transplantation respectively (Figure 3.10.B-C). Both assays require the intrinsic ability to self-renew and grow either in a collective challenge when locally injected in the mammary fat pad, or in a single cell challenge when directly seeded into the circulation as a model of experimental metastatic colonisation. Interestingly, we observed a mild decrease (not significant) in the primary tumour growth ability of the mesenchymal $CD24^+AXL^+$ cells compared to the $CD24^+AXL^-$ more epithelial population (Figure 3.10.D). Next, we assessed the metastatic colonisation potential of the mesenchymal $CD24^+AXL^+$ population compared to the rest of the tumour cells by intravenous injection into mice. Strikingly, in contrast to the results observed in the primary tumour initiation assays only the mesenchymal $CD24^+AXL^+$ cancer cells, which are highly enriched in MICs, were capable of lung colonisation (Figure 3.10.E).

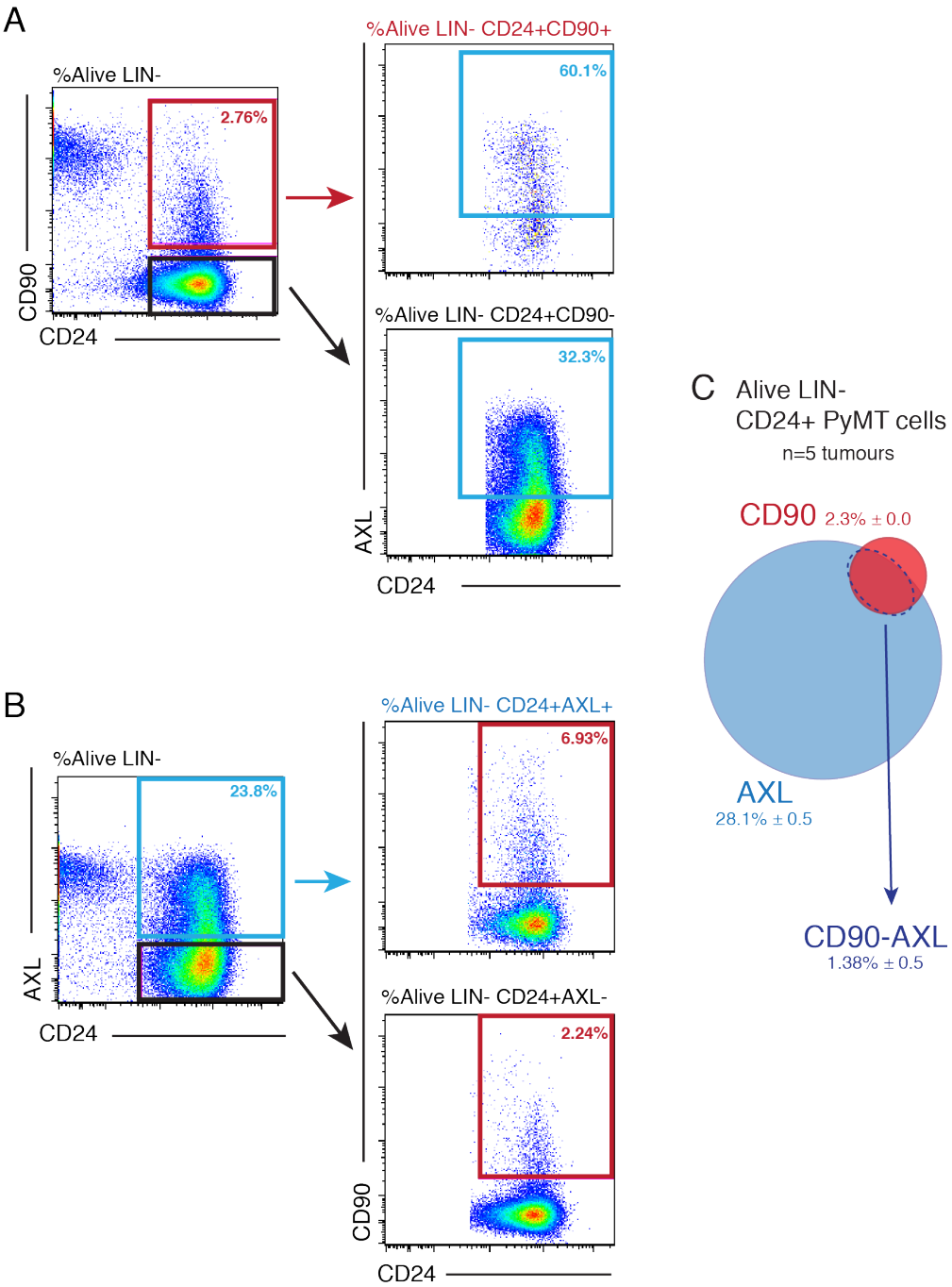


Figure 3.9 Relationship between MICs and the mesenchymal $CD24^{+}AXL^{+}$ population in PyMT tumour cells

(A) Representative FACS analysis plots of the percentage of AXL-expressing cells (blue gates) in the MIC (red gate) and nonMIC (black gate) compartment of the PyMT tumours. (B) Representative FACS analysis plots of the percentage of MICs (red gates) in the $CD24^{+}AXL^{+}$ (blue gate) and $CD24^{+}AXL^{-}$ (black gate) compartment of the PyMT tumours. (C) Venn diagram illustrates the relationship between MICs and $CD24^{+}AXL^{+}$ cells, and their relative frequency in PyMT tumours. Mean ± sem for each population is indicated (n=5 tumours).

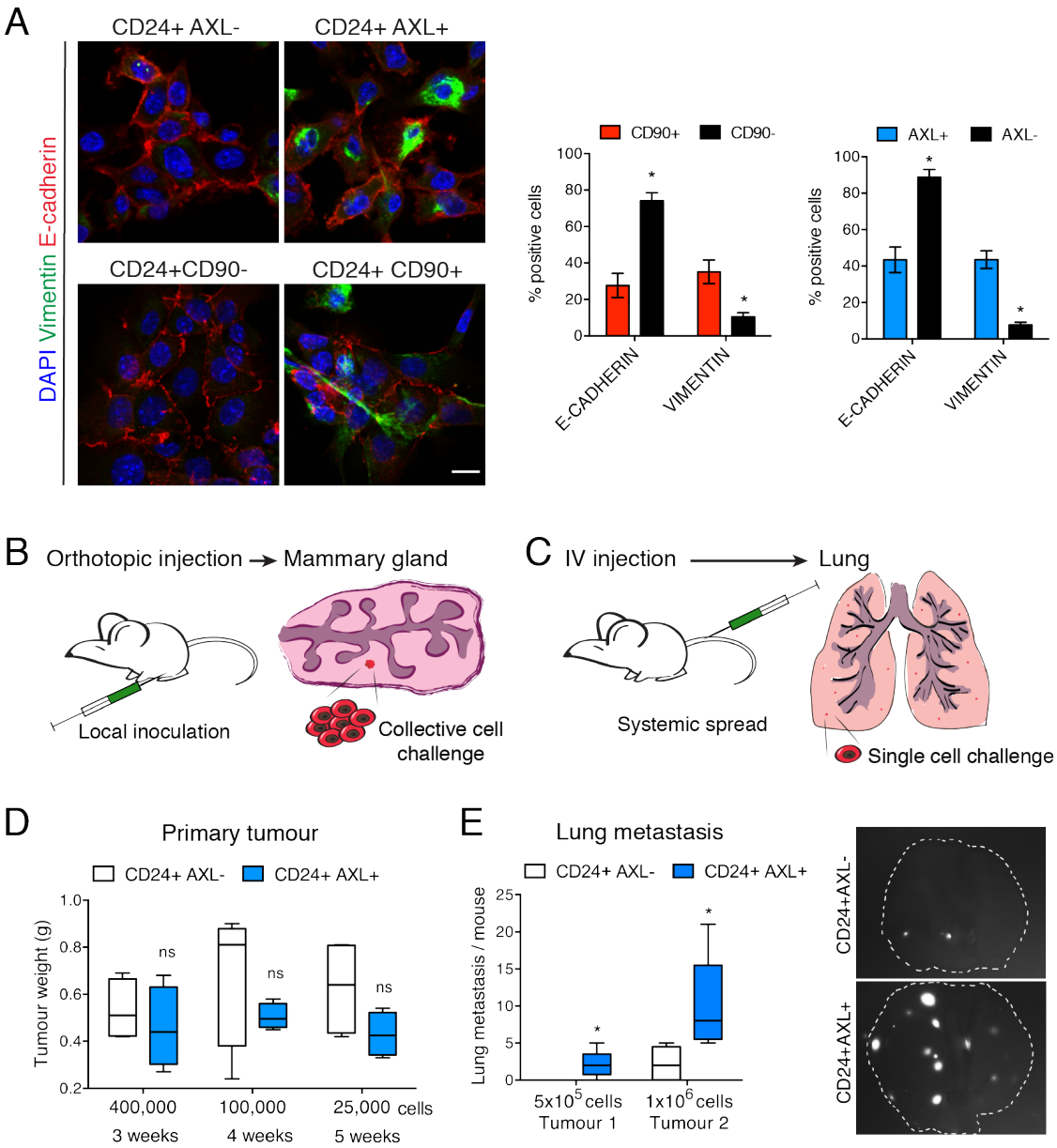


Figure 3.10 Tumour-initiating and metastatic potential of the mesenchymal CD24$^+$AXL$^+$ population in PyMT tumour cells

(A) (Left) Representative pictures show E-cadherin and Vimentin expression in the indicated MMTV-PyMT subpools. Cells were isolated by FACS sorting from the same primary tumour and plated overnight on collagen-coated dishes prior to analysis. (Right) Bar graphs indicate the percentage of E-cadherin and Vimentin positive cells in each population. Data from one representative experiment from a total of two are shown. Line and error bar indicate mean ± sem of the population. (B) Schematic shows a collective cell challenge for primary tumour growth. Tumour cells are injected into the mammary fat pad into a Matrigel plug. (C) Schematic represents a single cell challenge for metastatic growth. A single tumour cell suspension dissolved in PBS is directly injected into the tail vein of mice. (D) Box plot shows the primary tumour burden of primary PyMT CD24$^+$AXL$^-$ *versus* CD24$^+$AXL$^+$ cells in a collective cell challenge. PyMT cells were freshly isolated from primary tumour and plated for 16h. The subpools were isolated by FACS

sorting and the indicated cell numbers were injected into the mammary fat pad in Matrigel plugs (n=4 mice per group). (E) (Left) Box plot shows metastatic potential of primary PyMT/aGFP $CD24^+AXL^-$ *versus* $CD24^+AXL^+$ cells in a single cell challenge. $5x10^5$ cells were prepared as a single cell suspension and injected in the tail vein directly after cell sorting (2 independent experiments shown with n=5-6 mice per group). (Right) Representative images of superficial GFP^+ lung metastases formed by $CD24^+AXL^-$ or $CD24^+AXL^+$.

Subsequently, we wanted to assess whether MICs mesenchymal features, as suggested by the enhanced lung colonisation ability of the $CD24^+AXL^+$ cells, are essential for metastatic colonisation specifically but not for primary tumour initiation. Interestingly, in contrast to the previously reported linear correlation between mesenchymal features and stemness in breast cancer (Mani et al., 2008), we observed a decrease in tumour growth when mesenchymal AXL^+ primary PyMT cells were compared to the more epithelial AXL^- subpool (Figure 3.10.D). These results suggest that the mesenchymal features in the MMTV-PyMT model do not correlate with higher tumour-initiating ability or stemness, but confer an advantage exclusively for metastatic growth at distant sites (Figure 3.10.E).

To test this hypothesis, we analysed in parallel the expression of stem-cell antigen-1 (Sca1), previously described as a functional mammary epithelial stem/progenitor cell marker to study tumour initiation in genetically engineered mouse models (Welm et al., 2002, Grange et al., 2008). Specifically, the epithelial $CD24^+Sca1^+$ subpool has been previously reported to define a tumour-initiating population in the MMTV-Neu mouse model (Liu et al., 2007b). Therefore, we next checked Sca1 expression in the MMTV-PyMT tumours as a potential tumour-initiating cell marker, and its relationship to the CD90 and AXL markers, and the mesenchymal features that define metastasis-initiating properties. The $CD24^+Sca1^+$ population accounts for 60% of the total tumour cells in the PyMT model (Figure 3.11.A). The majority of the Sca1 population (around 75%) is negative for the mesenchymal marker AXL, therefore the $Sca1^+$ subpool is not enriched in mesenchymal characteristics (Figure 3.11.A-B). In line with this observation, both the $CD24^+Sca1^+$ and its depleted fraction show a similar expression of E-cadherin and Vimentin (Figure 3.11.D-E). Furthermore, as expected by its low overlap with the mesenchymal AXL population (Figure 3.11.B), the $CD24^+Sca1^+$ population is in a more epithelial state compared

to MICs and the $CD24^+AXL^+$ population (Figure 3.11.F). Moreover, in contrast to the observed enrichment in MICs within the AXL^+ fraction of tumour cells (Figure 3.9.B), the Sca1 compartment of the tumour is not enriched in MICs (Figure 3.11.A).

Next, we tested in parallel all these more epithelial or more mesenchymal PyMT populations for their ability to initiate a primary tumour mass. We isolated the three populations MICs, $CD24^+AXL^+$ and $CD24^+Sca1^+$ cells from the same primary tumour and injected either 10^3 or 10^4 cells sub-cutaneously. Interestingly, all subpools could give rise to primary tumour masses. Notably, MICs generated significantly bigger tumours than the likewise mesenchymal $CD24^+AXL^+$ subpool (Figure 3.12.A-C). Additionally, MICs tumours grew as efficiently as the more epithelial $CD24^+Sca1^+$ tumours. These data indicate that the mesenchymal features *per se* are not sufficient to boost primary tumour growth in this model, and suggest that in this assay other stem-like features of MICs not shared by the mesenchymal $CD24^+AXL^+$ population, such as high Wnt signalling (Figure 3.3) define their tumour-initiating properties.

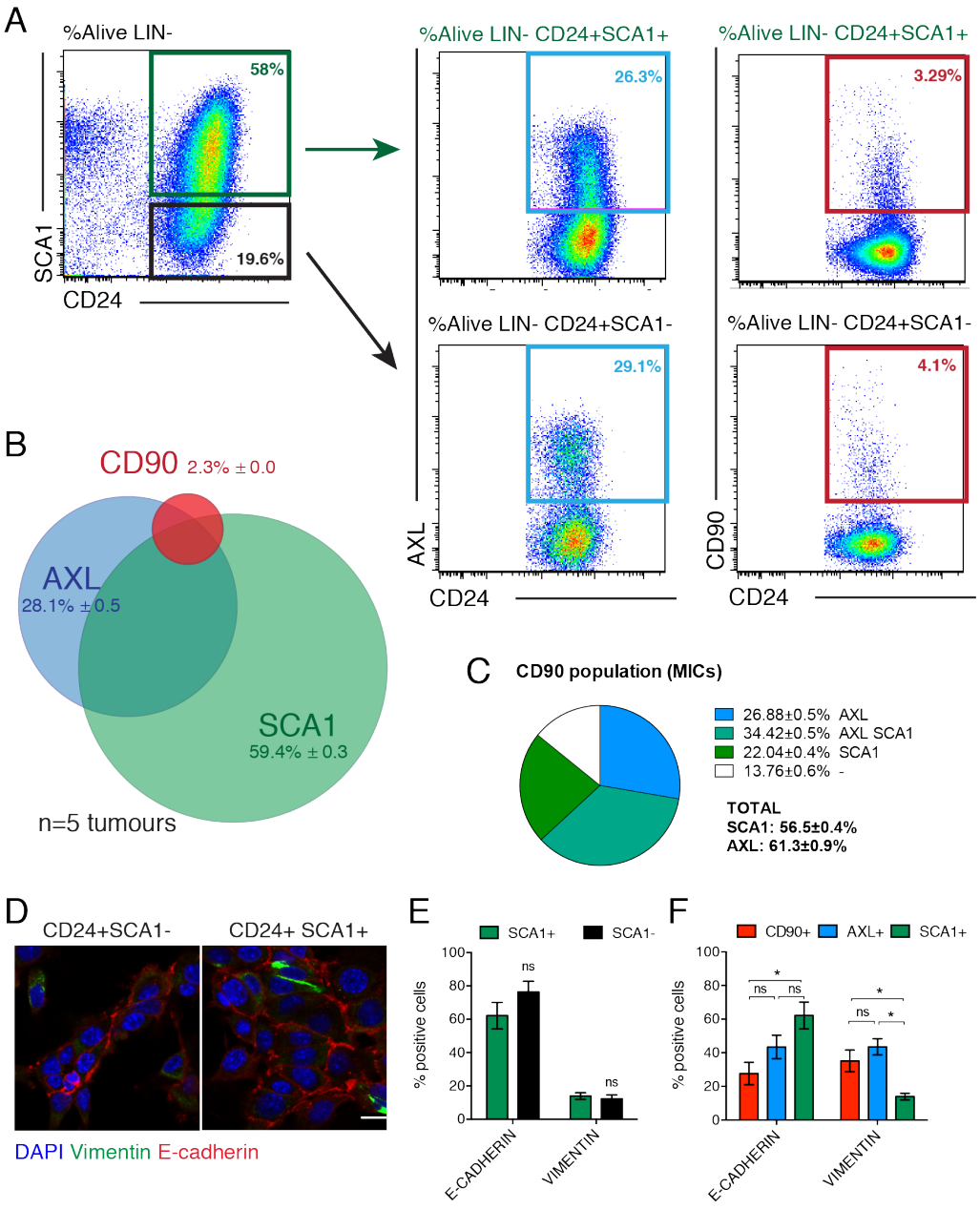


Figure 3.11 Relationship between MICs, AXL$^+$ and Sca1$^+$ epithelial populations in PyMT tumours

(A) Representative FACS analysis plots show the percentage of MICs (red gates) or CD24$^+$AXL$^+$ (blue gates) within the CD24$^+$Sca1$^+$ and CD24$^+$Sca1$^-$ compartments of the PyMT tumours. (B) Venn diagram illustrates the relationship between the different PyMT populations, and their relative frequency in PyMT tumours. Mean ± sem for each population is indicated (n=5 tumours). (C) Analysis of the percentage of MICs (CD90$^+$) expressing AXL and Sca1 markers. Pie graphs display the mean ± sem of each population (n=5 tumour cell preparations). (D) Representative pictures show E-cadherin and Vimentin expression in the indicated MMTV-PyMT subpools. (E-F) Graphs indicate the percentage of E-cadherin and Vimentin positive cells in each population, all isolated in parallel from the same primary tumour. Data from one representative experiment from a total of two are shown. Line and error bar indicate mean ± sem of the population.

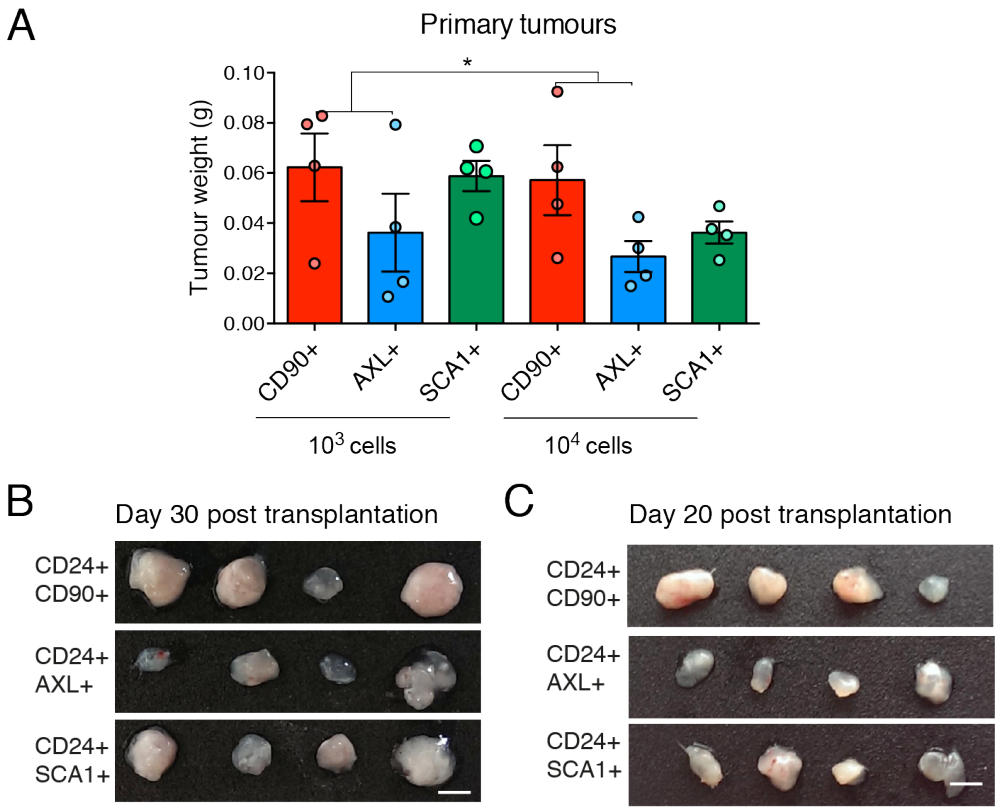


Figure 3.12 Tumour-initiation ability of different primary PyMT populations
(A) Sub-cutaneous tumour initiation assay. The indicated PyMT populations were FACS sorted from the same primary tumour cell preparation after overnight plating on collagen-I coated dishes. Either 10^3 cells or 10^4 cells from each population were transplanted sub-cutaneously into RAG1 mice. (B-C) Images of the tumours harvested either 30 or 20 days after transplanting 10^3 or 10^4 cells, respectively. Scale bars, ~ 5 mm.

In conclusion, in the MMTV-PyMT mouse model there is no evidence for a more mesenchymal tumour cell pool sustaining primary tumour growth. Nevertheless, the mesenchymal characteristics observed in MICs are sufficient to define metastatic competency as indicated by the high colonisation ability of the AXL^+ subpool (Figure 3.10.E), but yet independent of the intrinsic stemness potential required for tumour initiation (Figure 3.12). In other words, these results suggest that the mesenchymal characteristics observed in MICs are specifically defining metastatic competency rather than tumour-initiation abilities.

3.2 MICs have an increased ability to activate lung fibroblasts

In the previous section, we have shown that MICs are in a highly secretory mesenchymal state as defined by high AXL expression (Figure 3.3 and 3.8), which confers a metastatic advantage independently of stemness (Figure 3.10). Therefore, we hypothesised that the secretory mesenchymal features might confer MICs an increased ability to generate a favourable environment or niche, as they extravasate into the lungs. Notably, this characteristic would be crucial during metastatic colonisation when disseminated single cells in a foreign tissue need to create a favourable niche (section 1.3.2).

To test this hypothesis, we used the mesenchymal $CD24^{+}AXL^{+}$ population that likewise MICs displays enhanced metastatic ability (Figure 3.10.E), and we analysed their microenvironment for the presence of activated stroma just after extravasation into the metastatic tissue. Actin-GFP/$CD24^{+}$ cells positive or negative for AXL were isolated by FACS from primary tumours and intravenously injected into mice. Lungs were analysed 3 days post-injection, when single disseminated tumour cells have extravasated into the lung parenchyma (Figure 3.13.A). We observed that $CD24^{+}AXL^{+}$ cells were surrounded by a higher number of smooth muscle actin (SMA) activated fibroblasts compared to $CD24^{+}AXL^{-}$ cells suggesting their exclusive ability to induce a fibroblastic niche (Figure 3.13.B-C). Remarkably, no difference in their tissue localisation was detected as measured by their proximity to lung endothelial cells stained with endomucin, excluding that $CD24^{+}AXL^{+}$ and $CD24^{+}AXL^{-}$ cells occupy different pre-existing niches upon extravasation to the lung (Figure 3.13.B-C). These results indicate that tumour cells displaying an AXL-mesenchymal status, in contrast to more epithelial tumour cells, show an increased ability to activate fibroblastic stroma upon arrival into the lungs.

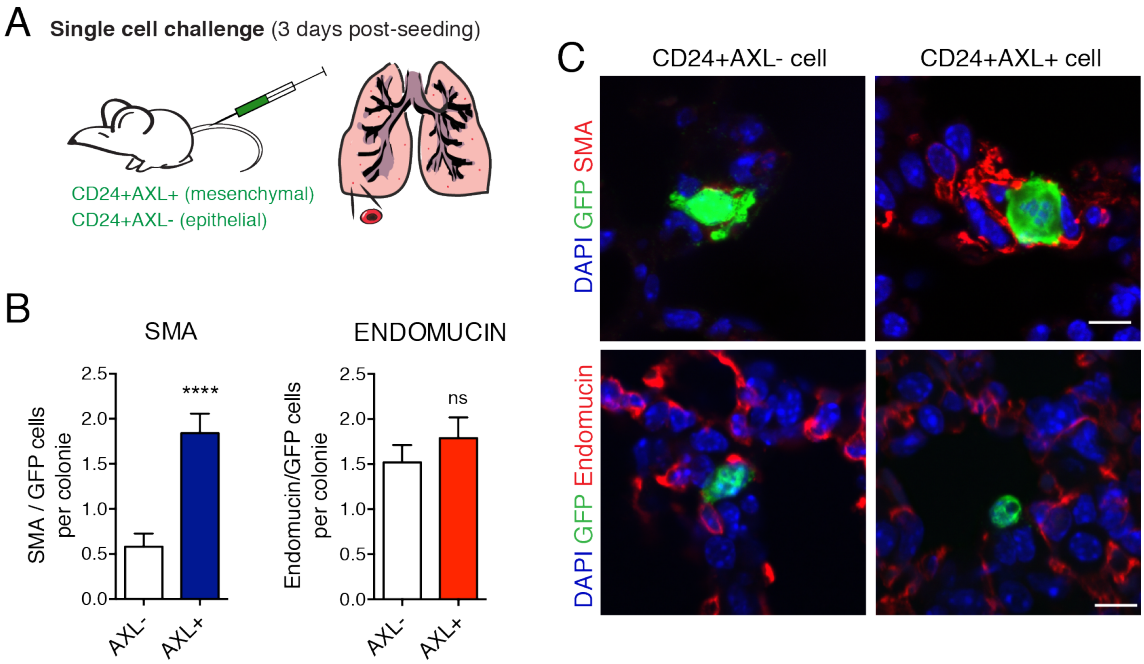


Figure 3.13 AXL-mesenchymal cancer cells activate a fibroblastic metastatic niche upon extravasation to the lungs
(A) Experimental set up: Actin-GFP/CD24^{+}AXL^{+} and CD24^{+}AXL^{-} cells were FACS sorted, and 10^{6} cells were injected intravenously into immunodeficient RAG1 mice. Lungs were analysed 72h post-seeding. (B) Bar graphs display the number of SMA (left) and Endomucin (right) positive cells in direct contact with the GFP^{+} tumour cells (n≥50 cells). (C) Representative pictures show single Actin-GFP/CD24^{+}AXL^{+} and CD24^{+}AXL^{-} cells (green) in the lung, DAPI staining (blue) and SMA and Endomucin (red). Scale bars, 8 µm and 10 µm respectively.

Subsequently, we validated whether the observed *in vivo* fibroblast activation by the AXL mesenchymal pool is a characteristic of MICs during early metastatic colonisation. Due to MICs low frequency within the cancer cells, further *in vivo* testing with this subpool was not technically possible. Therefore we set out to test *ex vivo* the ability of MICs to activate lung fibroblasts. As described beforehand, fibroblasts are key components of the metastatic niche and secrete different molecules, such as POSTN, that support metastatic colonisation (section 1.3.2) (Malanchi et al., 2012). This supportive role of fibroblasts is observed once they are activated and display characteristic features of pro-tumorigenic cancer-associated fibroblasts (CAFs) (Dumont et al., 2013), such as nuclear YAP translocation and increased FAP and SMA expression (Calvo et al., 2013). However, it remains unknown how the resident lung fibroblasts are activated during early metastatic colonisation. To shed light into this question, we tested whether the exclusive metastatic ability of MICs could depend on their higher ability to activate fibroblasts, as suggested by the disseminated mesenchymal AXL^+ subpool *in vivo* (Figure 3.13). We directly cultured FACS sorted cancer cells; either MICS or nonMICs, with a GFP labelled normal lung fibroblast cell line (NLF3) and monitored the induction of activation markers in the fibroblasts *ex vivo* (Figure 3.14.A). All tumour cells can induce fibroblast activation, therefore we designed a experimental set up where low numbers of tumour cells in a short period of time are tested for their differential ability to induce fibroblast activation markers. Nuclear YAP translocation, required for CAF induction (Calvo et al., 2013), as well as fibroblast activation protein (FAP) expression were highly induced in fibroblasts co-cultured with MICs after only 24h, while nonMICs were only able to trigger mild FAP expression (Figure 3.14.B-D). This result suggests that limited numbers of MICs can trigger fibroblast activation more efficiently that the rest of tumour cells.

To confirm that the presence of these markers represent features of CAFs, we performed a functional gel contraction assay. This assay gives an indirect measurement of the enhanced ability of CAFs to physically remodel an ECM compared to non-activated normal fibroblasts. Low numbers of MICs or nonMICs

isolated by FACS from primary mammary tumours were seeded in the upper chamber of a trans-well coculture dish. In the lower chamber of the coculture primary lung fibroblasts freshly isolated from wild type mouse lungs (NF) were embedded in 2:1 collagen-I:Matrigel matrices (Figure 3.15.A). Remarkably, MICs were able to induce stromal cell activation more efficiently compared than the rest of primary cancer cells as monitored by the increased fibroblast gel contraction, as well as the gain of a CAF gene expression signature (Figure 3.15.B-C).

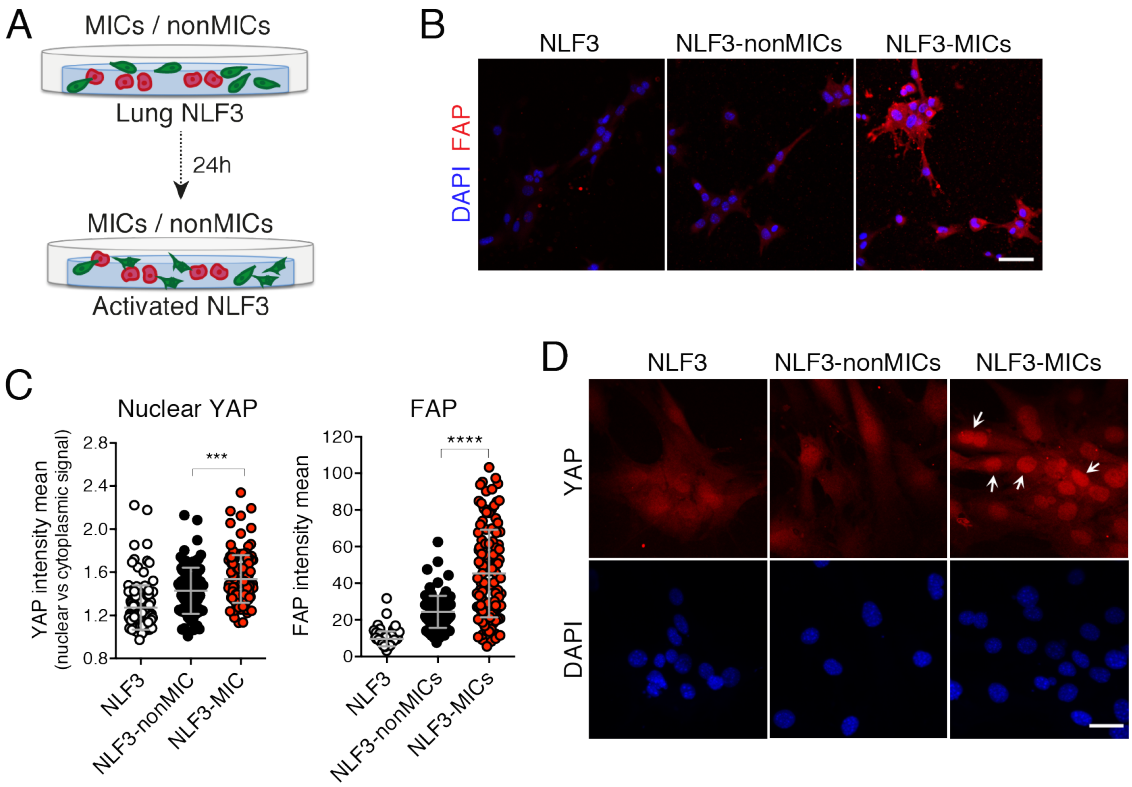


Figure 3.14 MICs display enhanced ability to induce lung fibroblast activation phenotypically

(A) Schematic shows the coculture setting where nonMICs or MICs are seeded on top of a thin layer of Matrigel:Collagen-I with GFP^+ labelled normal lung fibroblast cell line (NLF3). After 24 hours fibroblasts were monitored by immunofluorescence for changes in the expression of CAF-defining markers. (B) Representative images of FAP levels in NLF3 cells cultured in the presence of nonMICs or MICs for 24h. Scale bar, 70 µm. (C) Chart indicates the quantification of the nuclear levels of YAP and FAP. Data from one representative experiment from a total of 3 are shown. Line and error bar indicate mean ± SD of the populations. (D) Representative images of YAP nuclear translocation in NLF3 cells cultured in the presence of nonMIC or MIC for 24h. Scale bar, 40 µm.

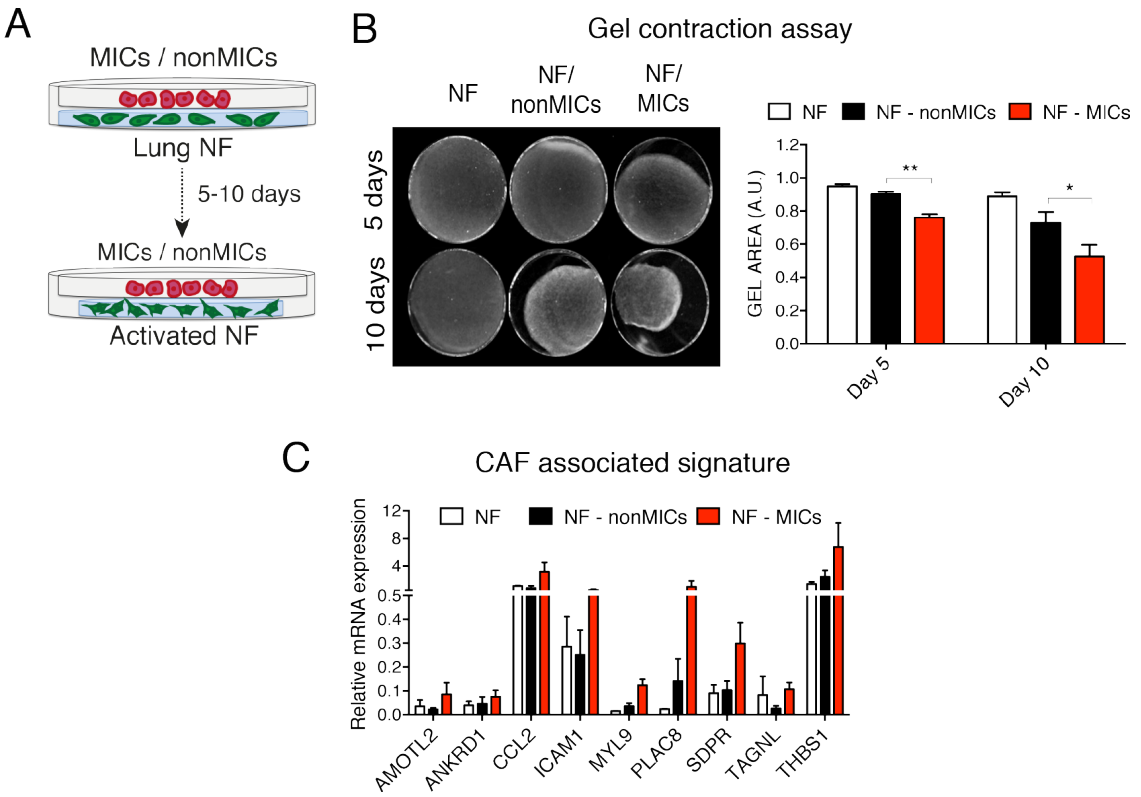


Figure 3.15 MICs display enhanced ability to induce lung fibroblast activation functionally

(A) Schematic shows the coculture setting where primary normal lung fibroblasts (NF) are seeded in the lower compartment embedded in Matrigel:Collagen-I. MICs and nonMICs are added to a separate compartment on top. Fibroblast activation-dependent gel contraction is monitored over 5 to 10 days. (B) (Left) Representative images of the gel contraction assay 5 and 10 days after coculture. (Right) Histogram displays mean ± sem (n=8, number of gels) assessed over 3 independent experiments. (C) Quantitative real time PCR analysis compares the gene expression levels of different CAF-defining genes induced by MICs and nonMICs. Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bars represent mean ± sem. p=0.0235 by ANOVA comparing the signatures of NLF3-nonMIC *versus* NLF3-MIC.

Collectively these data provide direct experimental evidence that cancer cells in an AXL-mesenchymal state possess an enhanced ability to trigger fibroblast activation, generating their metastatic niche.

3.3 The increased ability of metastatic cells to activate lung fibroblasts depends on the secretion of Thrombospondin-2, a key step for efficient metastasis

3.3.1 MIC-derived THBS2 activates lung fibroblasts

We have shown in the previous section how AXL defines a mesenchymal state in cancer cells conferring high metastatic potential (Figure 3.10.E), and these cells also display enhanced ability to induce activated fibroblasts early after extravasation into the naïve lung tissue (Figure 3.13.B-C). We sought next to elucidate the mechanisms by which MICs activate lung fibroblasts. For this purpose, we analysed the secretome observed in the MIC signature (Figure 3.3.B) and validated the top differentially expressed secreted factors in MICs compared to nonMICs by quantitative PCR (Figure 3.16.A-B). Next, we functionally tested these MIC-secreted proteins for their effect to activate normal lung fibroblasts in gel contraction assays. We used the recombinant proteins of the different factors identified at concentrations previously reported to exert their activity *in vitro,* THBS2 (5 µg/ml) (Hirose et al., 2008), ANGPTL4 (5 µg/ml) (Cazes et al., 2006), CCL7 (0.3 µg/ml) (Gouwy et al., 2008), CCL11 (0.1 µg/ml) (Saunders et al., 2008), CYR61 (1.5 µg/ml) (Francischetti et al., 2010) and WNT5a (0.5 µg/ml) (Li et al., 2013), and added them to normal lung fibroblasts in gel contraction assays. We found that two MIC-derived secreted factors, Thromspondin-2 (THBS2) and the Wnt ligand Wnt5a were able to boost the gel contraction ability of lung fibroblasts embedded in Matrigel:collagen-I matrices (Figure 3.16.C-D). We decided to focus on THBS2 rather than Wnt5a due to its higher differential expression in MICs compared to nonMICs (Figure 3.16.A). THBS2 is normally secreted on a per cell basis at a mean value of ~5 µg/l per 10^5 cells *in vitro* (Hankenson and Bornstein, 2002). It is therefore to acknowledge that the used of higher concentrations (5 µg/ml) of recombinant THBS2 could explain the potent effect and sufficiency of this factor to trigger fibroblast activation *in vitro*, while at physiological levels *in vivo* the synergistic effects of several factors, such as TGFβ, or the other validated MIC-

derived factors described, might be required. We found purified THBS2 to be sufficient to trigger YAP nuclear translocation and FAP expression in normal lung fibroblasts 24h after treatment *in vitro* (Figure 3.16.E), in line with the previous results showing YAP/FAP induction in normal lung fibroblasts after 24h coculture with MICs.

A

MIC secreted factors	Fold change (MIC vs nonMIC)
THBS2	27.93
CCL7	10.41
CCL11	6.42
ANGTL4	5.59
CXCL10	1.83
CYR61	1.62
WNT5A	1.45

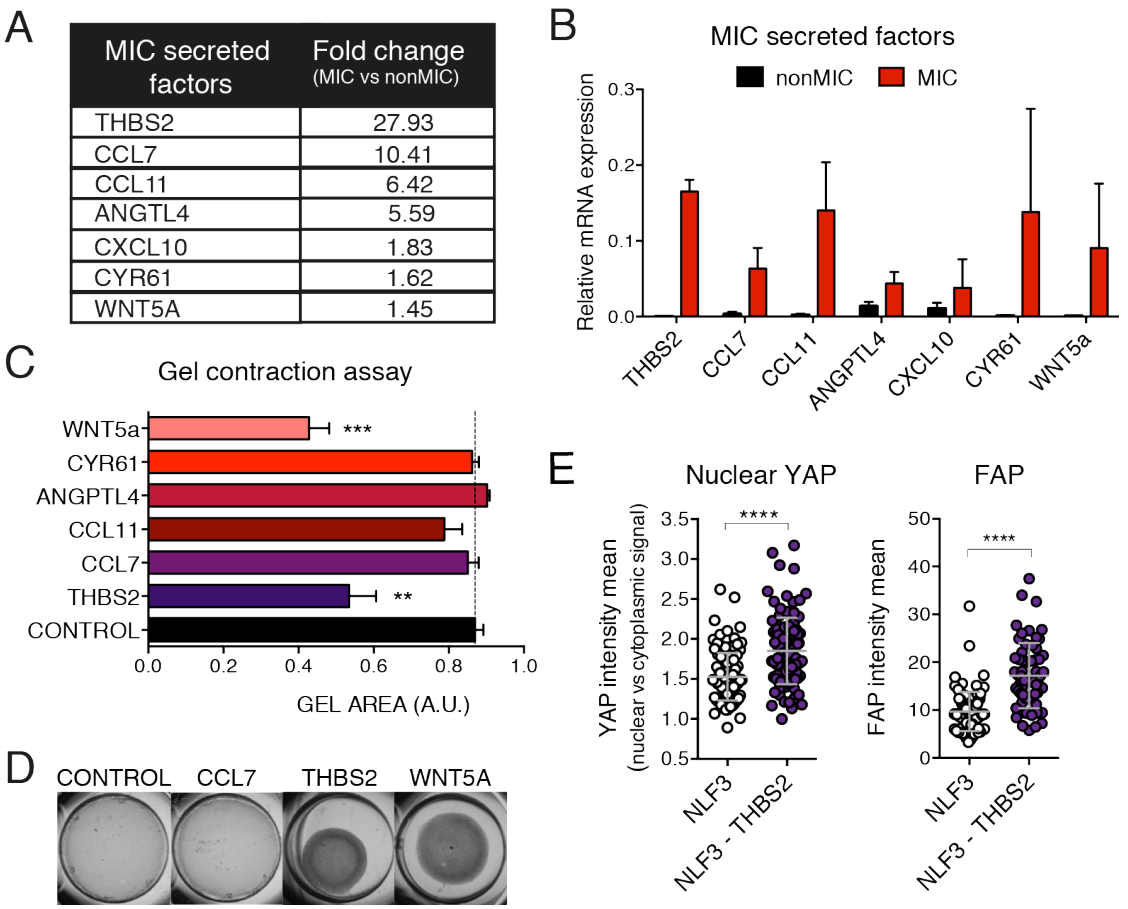


Figure 3.16 THBS2 is differentially expressed in MICs and triggers fibroblast activation

(A) List of some of the top secreted factors found in the MIC signature. Fold change expression of each factor obtained by comparing the expression levels in MIC *versus* nonMICs. (B) Quantitative real time PCR analysis shows the expression of the MIC secreted factors. Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bar represent mean ± sem. p=0.0057 by ANOVA comparing the gene set in MICs *versus* nonMICs. (C) Gel contraction assay by NLF3 3 days after addition of the THBS2 (5 µg/ml), ANGPTL4 (5 µg/ml), CCL7 (0.3 µg/ml), CCL11 (0.1 µg/ml), CYR61 (1.5 µg/ml) and WNT5a (0.5 µg/ml). (D) Representative images show the gel contraction quantified in (C) for CCL7, THBS2 and Wnt5a. (E) Charts indicate the nuclear levels of YAP and total FAP expression after addition of THBS2 (5 µg/ml) for 24h. Data from one representative experiment from a total of 3 are shown. Line and error bar indicate mean ± SD of the populations.

These data propose THBS2 as a specific MIC-secreted factor mediating their enhanced fibroblast activation ability.

3.3.2 THBS2 triggers integrin β1 signalling leading to fibroblast activation

After finding that THBS2 can enhance fibroblast activation, we investigated its mechanism of action. We searched *in silico* potential THBS2 interacting partners using the String database to predict THBS2 mediated protein-protein interactions. From the predicted interacting partners found many were other secreted proteins such as metalloproteases, other thrombospondins and the PDGFBB ligand (Figure 3.17.A). We focus on the transmembrane receptors identified by this analysis that could transduce THBS2 signal directly and found different integrins (Figure 3.17.A). As discussed earlier, integrins usually represent the main bridge for cell-ECM interactions (section 1.1.2.2). Among the potential THBS2 interacting partners (integrins α4, α6, α7 and β1) we found that only integrins α4, α6 and β1 are expressed by normal lung fibroblasts (Figure 3.17.B). As integrin complexes signal through the β subunit (section 1.1.2.2), we functionally tested the direct interaction of THBS2 with integrin β1. By using immunofluorescence and confocal microscopy we could detect the colocalisation of the exogenously added recombinant THBS2 with the endogenous lung fibroblasts integrin β1 (Figure 3.17.C). The images revealed that THBS2 addition to the fibroblasts causes integrin clustering and potential internalisation. Integrin β1 internalisation is a mechanism required at the leading edge of migrating cells to transiently disassemble focal adhesions, further facilitating their directional migration (Nishimura and Kaibuchi, 2007). Hence, the observed integrin internalisation suggests an enhanced fibroblast migration that would go in line with fibroblast activation (Calvo et al., 2013).

We further tested the functional relevance of THBS2-integrin β1 interaction by examining normal lung fibroblast integrin β1 activation by using a specific antibody against the active conformation of this integrin. Simultaneously, we monitored the phosphorylation of its downstream effector focal adhesion kinase (pFAK) that binds to and associates with the integrin-adapter protein–cytoskeleton complex, forming the basis of focal adhesions (section 1.1.2.2 and Figure 1.4.B). Both, the active isoform of integrin β1 and the phosphorylation levels of pFAK increased 24h after THBS2 stimulation. Notably, preventing integrin activation by using a specific integrin β1-blocking antibody (β1BA) before THBS2 stimulation halted THBS2-dependent integrin activation (Figure 3.17.D-E).

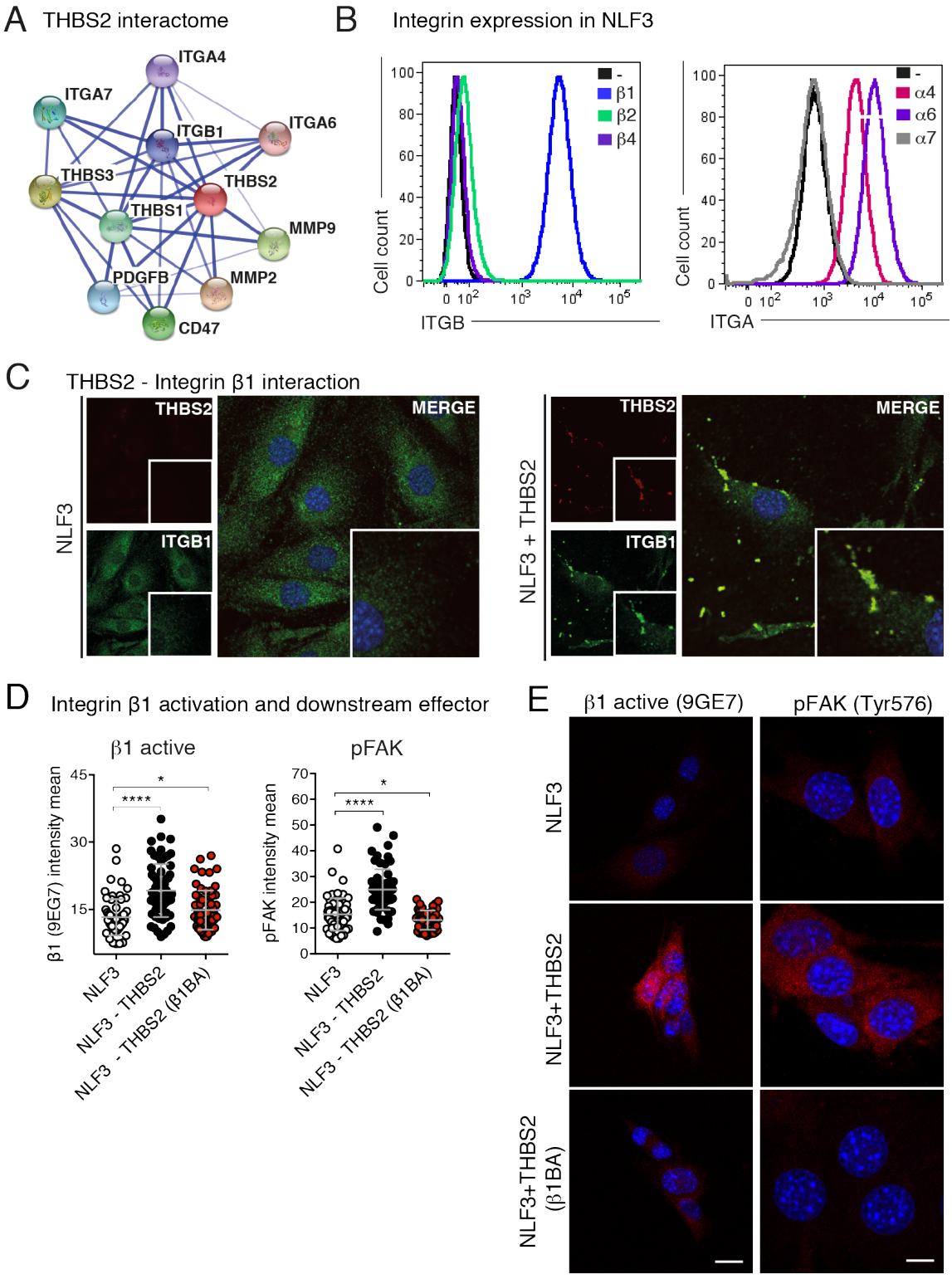


Figure 3.17 THBS2 binds integrin β1 in normal lung fibroblasts activating its downstream signalling

(A) THBS2 interactome: confidence view represents stronger associations by thicker lines. (B) Normal lung fibroblast cell line (NLF3) integrin expression assessed by FACS analysis. (C) Confocal images show recombinant THBS2 (red),

integrin β1 (green) and DAPI (blue). Note the colocalisation of the THBS2 and the integrin β1 clusters signals. (D) Charts indicate the levels of the active isoform of integrin β1 and its phosphorylated downstream effector FAK in NLF3, NLF3 + THBS2 (5 μg/ml) and NLF3 + THBS2 (5 μg/ml) in the presence of β1 integrin blocking antibody (β1BA) (20 μg/ml). Data from one representative experiment from a total of 2 are shown. Line and error bar indicate mean ± SD of the populations. (E) Representative images of quantification in (D). Scale bars, 20 μm and 10 μm, respectively.

We next checked the functional relevance of this THBS2-mediated signalling cascade for fibroblast activation. Blocking THBS2-mediated integrin β1 activation halted the induction of the early effectors of fibroblast activation as measured by YAP nuclear translocation and FAP expression over 24h (Figure 3.18.A-B). Moreover, blocking integrin β1 prevented the THBS2 mediated enhanced gel contraction ability of fibroblasts (Figure 3.18.C).

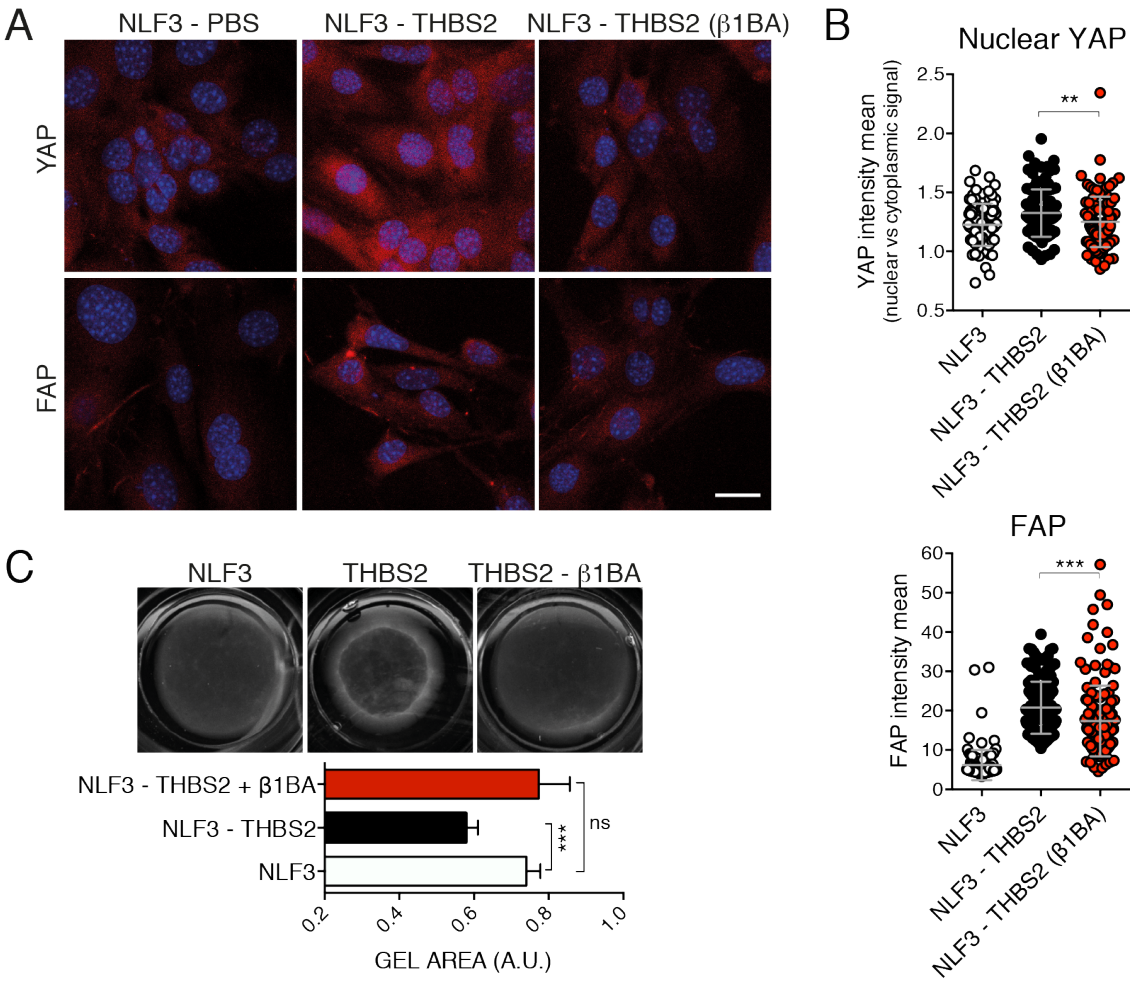


Figure 3.18 Preventing integrin β1 activation halts THBS2-mediated fibroblast activation

(A) Representative images of YAP nuclear translocation and FAP expression induction in NLF3 cultured in the presence of recombinant THBS2 (5 μg/ml) for 24h ± integrin β1 blocking antibody (20 μg/ml), added 2 hours before THBS2 treatment. Scale bar, 20 μm. (B) Chart indicates the quantification of nuclear YAP levels and total FAP levels in (A). Data from one representative of four experiments are shown. Line and error bar indicate mean ± SD of the populations. (C) Upper panel, images show gel contraction by NLF3 cultured in the presence of recombinant THBS2 (5 μg/ml) for 24h ± integrin β1 blocking antibody (20 μg/ml), added 2 hours before THBS2 treatment. Lower panel, histogram shows the mean ± sem (n= number of gels: NLF3, 16; NLF3 + THBS2, 16; NLF3 + THBS2 (β1BA), 4), assessed over multiple experiments.

Altogether, these data reveal a novel mechanism by which THBS2 enhances normal fibroblasts activation through integrin β1 signalling.

3.3.3 THBS2-mediated fibroblast activation is crucial for lung metastatic colonisation

In light of the previous results, we sought to investigate the relevance of THBS2 expression for metastasis. We used a lentiviral EGFP-reporter construct to express specific short-hairpin RNAs targeting THBS2 mRNA upon lentiviral infection (Figure 3.19.A); infected primary PyMT cancer cells were selected by cell sorting based on their EGFP expression. We used two different short-hairpin sequences and evaluated their corresponding knockdown efficiency at the protein level by ELISA (Figure 3.19.B). The first shRNA sequence (shTHBS2 I) reduced THBS2 levels by a third of its original expression, while the second (shTHBS2 II) was more efficient and no THBS2 was detected in the ELISA assay (Figure 3.19.B). Importantly, cancer cells with depleted THBS2 reduced their ability to activate lung fibroblasts in gel contraction assays (Figure 3.19.C). Together, these data propose THBS2 as a cancer cell secreted co-factor facilitating fibroblast activation.

To assess the role of MIC-derived THBS2 secretion during metastatic colonisation we tested shTHSB2 cancer cells to grow *in vitro* and *in vivo*. *In vitro*, we excluded that THBS2 could have a direct effect on the intrinsic potential of the cancer cells by measuring the ability of shTHSB2 cells to form spheres as a measurement of their self-renewal capacity *in vitro*. No significant effects were observed on the ability of shTHBS2 cells to form spheres in non-adherent conditions (Figure 3.20.A). Then, we performed the gold standard *in vivo* test for self-renewal assessing shTHBS2 cells to initiate tumour growth in orthotopic transplantation assays. Early primary tumour initiation ability was not affected in the absence of THBS2 (Figure 3.20.B). Collectively, these results indicate that THBS2 does not affect the intrinsic ability of PyMT tumour cells to self-renew and initiate a tumour mass.

Given that THBS2 depletion does not impact in the intrinsic potential of tumour cells to grow, we tested the metastatic colonisation potential of shTHBS2 PyMT cells. We used experimental metastasis assays where $EGFP^+$ sorted cells expressing shTHBS2 are directly injected into the tail vein of mice (as in Figure 3.10.C). Strikingly, THBS2 depletion impairs the metastatic colonisation ability of

primary PyMT cells (Figure 3.20.C). These results highlight the crucial role of MIC-derived THBS2 for metastatic colonisation.

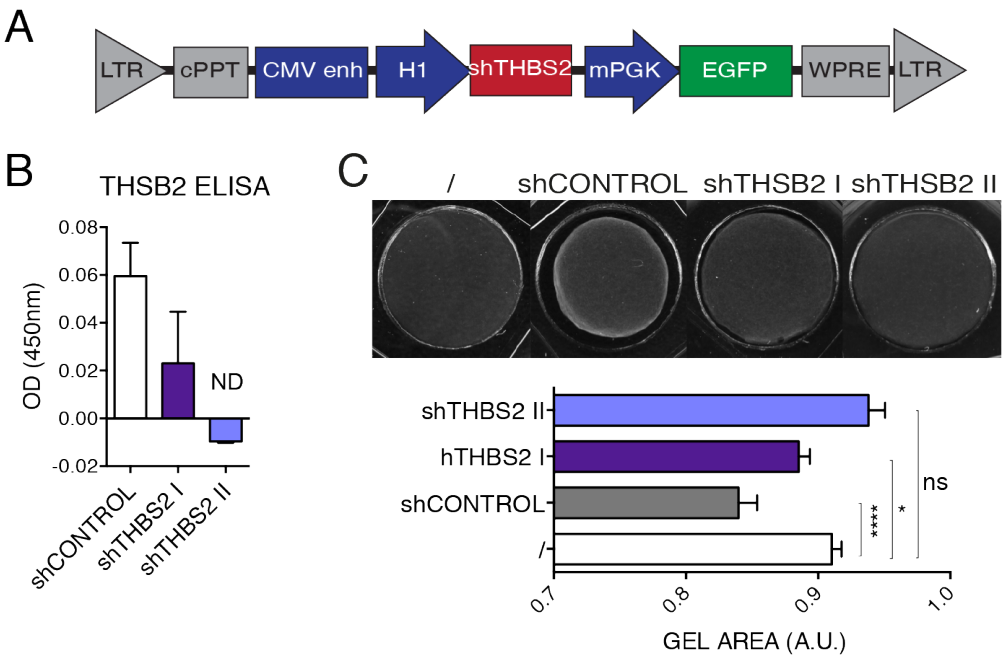


Figure 3.19 THBS2 knockdown in PyMT cells reduces their fibroblast activation ability
(A) Schematic shows the lentiviral reporter used to express the shRNA sequences. The shRNA sequences were cloned under the expression of a CMV enhancer followed by the H1 promoter. An independent mouse PGK promoter controls EGFP expression. (B) Histogram shows the relative levels of secreted THBS2 measured by ELISA present in the media conditioned by PyMT cells expressing shCONTROL or shTHBS2 lentiviral constructs. Data from 3 different experiments performed in duplicate. Bars represent mean ± sem. (B) (Upper panel) Images show gel contraction by normal lung fibroblasts (NLF3) cocultured with primary PyMT cells expressing shCONTROL or shTHBS2 lentiviral constructs. (Lower panel) Histogram shows the mean ± sem (n=4, number of gels).

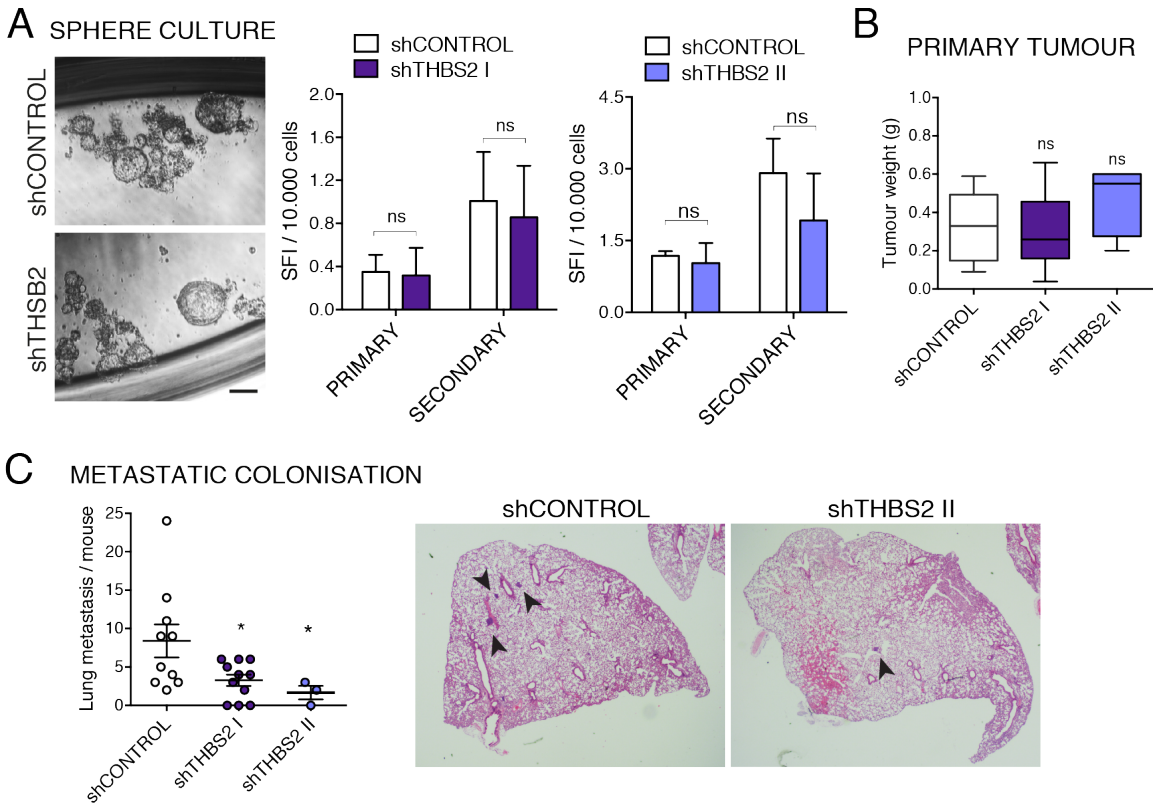


Figure 3.20 THBS2 depletion does not affect self-renewal and tumour initiation in primary PyMT cells but impairs metastatic colonisation ability
(A) Images show spheres formed by PyMT cells in suspension expressing shCONTROL and shTHBS2 lentiviral constructs. Histogram shows the mean ± sem of 2 independent experiments. Spheres were quantified 10 days after plating a single PyMT suspension (primary) and 10 days after the first passage (secondary). (B) Early primary tumour growth potential. Cells were FACS sorted to select for lentiviral expression and 50,000 shCONTROL or shTHBS2 PyMT cells were injected into the mammary fad pat. Tumour weight was evaluated 20 days post-injection (n=4-10). (C) Metastatic colonisation ability of shCONTROL and shTHBS2 PyMT cells. Positively infected PyMT cells were sorted as GFP^+, and $5x10^6$ cells were directly injected into the tail vein of mice. (Left) Micrometastasis evaluation 30 days post-injection. (Right) Representative images of lung sections with micrometastasis.

Next, to further assess the role of THBS2 during metastasis we engineered a lentiviral reporter construct containing the open reading frame (ORF) of the mouse THBS2 gene under the expression of a mouse PGK promoter. The lentiviral construct contained also a truncated isoform of the human CD2 receptor (hCD2) and a VENUS reporter gene (3.21.A). We confirmed THBS2 overexpression by real time PCR in the mouse 4T1 cell line after infection and EGFP selection, detecting a 15-fold increase in THBS2 secretion (Figure 3.21.B). Normally, in PyMT primary cells THBS2 expression is limited to the small MIC population within the tumour; by infecting total PyMT cells with this construct we broadened the number of cells that secrete THBS2 in the tumour cell mass. Subsequently, we evaluated the metastatic colonisation ability of THBS2ORF PyMT cells after intravenous injection. FACS analysis of the tumour cell content in the lungs 20 days post-intravenous injection revealed a higher cancer cell number in the lungs of mice injected with THBS2ORF PyMT cells, indicating an increased metastatic colonisation ability by PyMT cells when THBS2 is expressed by a larger cell number (Figure 3.21.C).

Using GOBO (Gene expression-based Outcome for Breast cancer Online) we performed a Distant Metastasis-Free Survival (DMFS) analysis over 5 years in breast carcinoma patient data, correlating THBS2 expression with tumour grade. We found that in tumours classified as grade 3, THBS2 high expression dramatically reduces the patients' DMFS over 5 years, indicating a correlation between high THBS2 levels and increased metastatic risk in advanced breast carcinoma (Figure 3.22). Notably, these clinical data reaffirm our findings indicating that THBS2 expression correlates with high metastatic risk in advanced stages of human breast carcinoma.

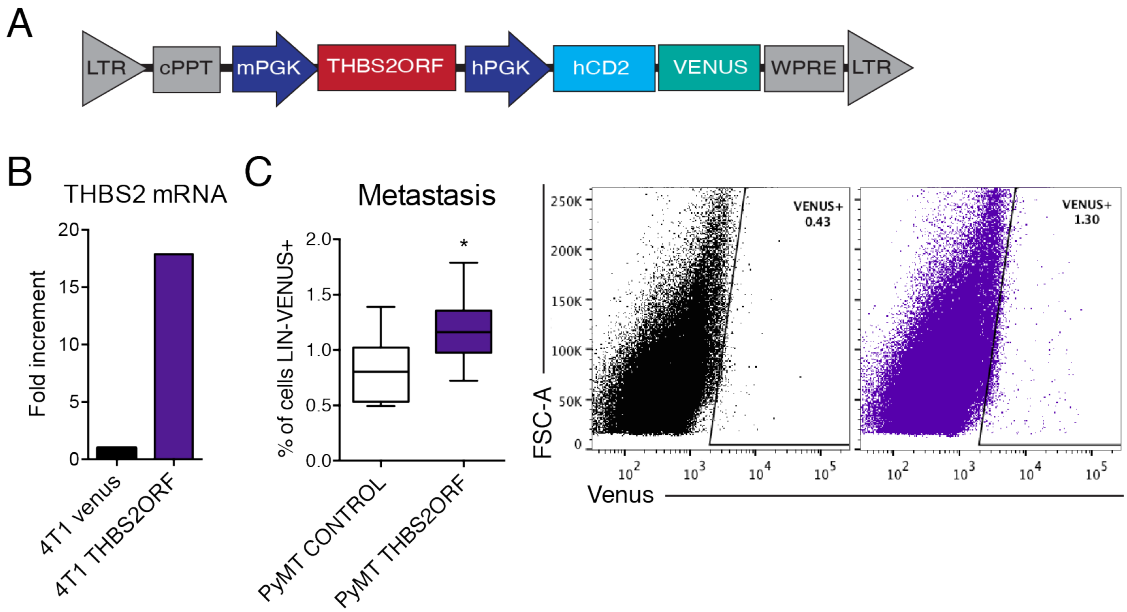


Figure 3.21 THBS2 overexpression enhances early metastatic colonisation
(A) Lentiviral construct expresses the mouse THBS2 coding sequence (THBS2ORF) under the mouse PGK promoter. An independent human PGK promoter controls VENUS and hCD2 reporter genes expression. (B) Quantitative real time PCR showing a 15-fold THBS2 mRNA expression in 4T1 cell line after infection with the THBS2ORF construct and hCD2^{+} selection. (C) FACS quantification of tumour cell content in the lungs 20 days after intravenous injection of VENUS^{+} selected CONTROL or THBS2ORF PyMT cells (n=8-9).

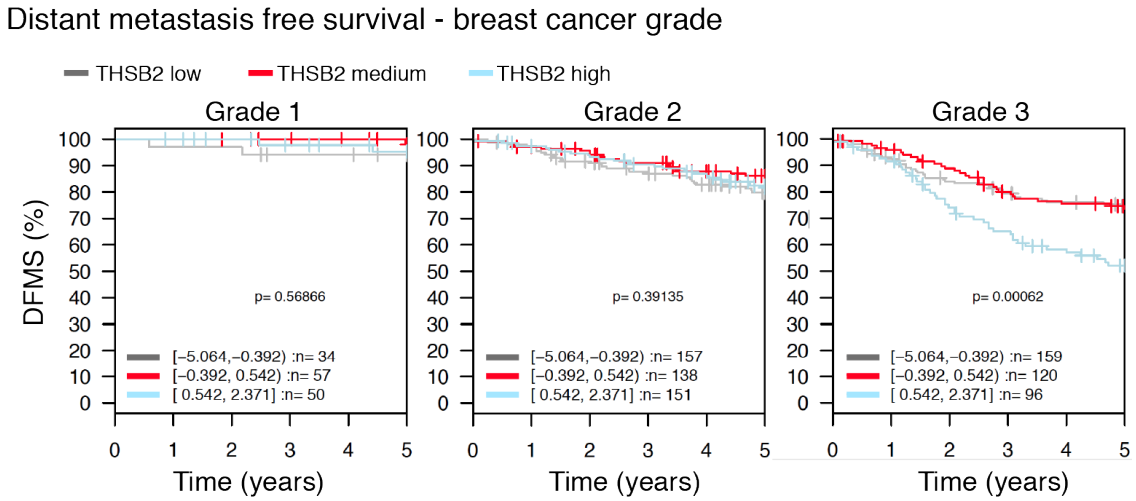


Figure 3.22 High THBS2 expression correlates with higher metastatic risk in breast carcinoma
Data generated using GOBO (Gene expression-based Outcome for Breast cancer Online). Plots show the percentage of Distant Metastasis-Free Survival (DMFS) over 5 years in breast cancer patients divided according to tumour grade. Within each grade, patients are further subdivided in low (grey), medium (red) or high (blue) THBS2 expression. Note the significant positive correlation between high THBS2 expression and decreased DMFS in grade 3 invasive breast carcinomas.

Overall, the data presented in this chapter corroborates our initial hypothesis that MICs induce a favourable microenvironment at the newly colonised tissue. Specifically, we have identified a novel role of THBS2 to mediate metastatic niche induction contributing to the metastatic potential of MICs.

3.3.4 THBS2 is a potent microenvironmental regulator during mammary carcinoma development

Next, we wanted to further analyse the role of THBS2 in tumorigenesis. We have shown that depleting THBS2 did not affect early tumour development (Figure 3.18.B, note tumour weight around 0.2 g). Strikingly, at later stages of tumour growth the absence of THBS2 led to increased tumour burden (Figure 3.22.A). We previously reviewed how changes in the ECM composition have profound effects in the tumour microenvironment (section 1.1.2.2). In particular, complex processes such as angiogenesis are tightly regulated by the surrounding ECM. Indeed, THBS2 have been reported as an antiangiogenic factor in a xenograft model of breast cancer using the mesenchymal cell line MDA-MB-435 (Koch et al., 2011). Therefore, we performed some histological analysis of shTHBS2 tumours and analyse specifically potential changes in the vasculature staining blood vessels with Endomucin, a membrane glycoprotein specifically expressed in the vascular endothelium (Kuhn et al., 2002). We found that THBS2 depletion lead to the development of longer and more interconnected vessels (Figure 3.22.A), which might increase the supply of oxygen and nutrients to tumour cells explaining the increased tumour growth.

Additionally, we generated primary PyMT tumour cells that exogenously expressed THBS2ORF by infecting them with the lentiviral construct in Figure 3.20.A, and tested their ability to generate primary tumours. In line with the previous observed increased tumour growth when depleting THBS2, increasing its expression dramatically reduced PyMT tumour burden (Figure 3.22.B). Histological analysis of these tumours revealed profound microenvironmental changes that affect both the vasculature and tumour cells spatial organisation (Figure 3.22.B). THBS2ORF tumours show decreased blood vessels density and reduced endothelial wall thickness, but also a different tumour cell organisation that suggests additional roles of THBS2 during tumour development. Interestingly, despite the reduced vascularisation we did not observe an evident increase in tumour necrotic areas.

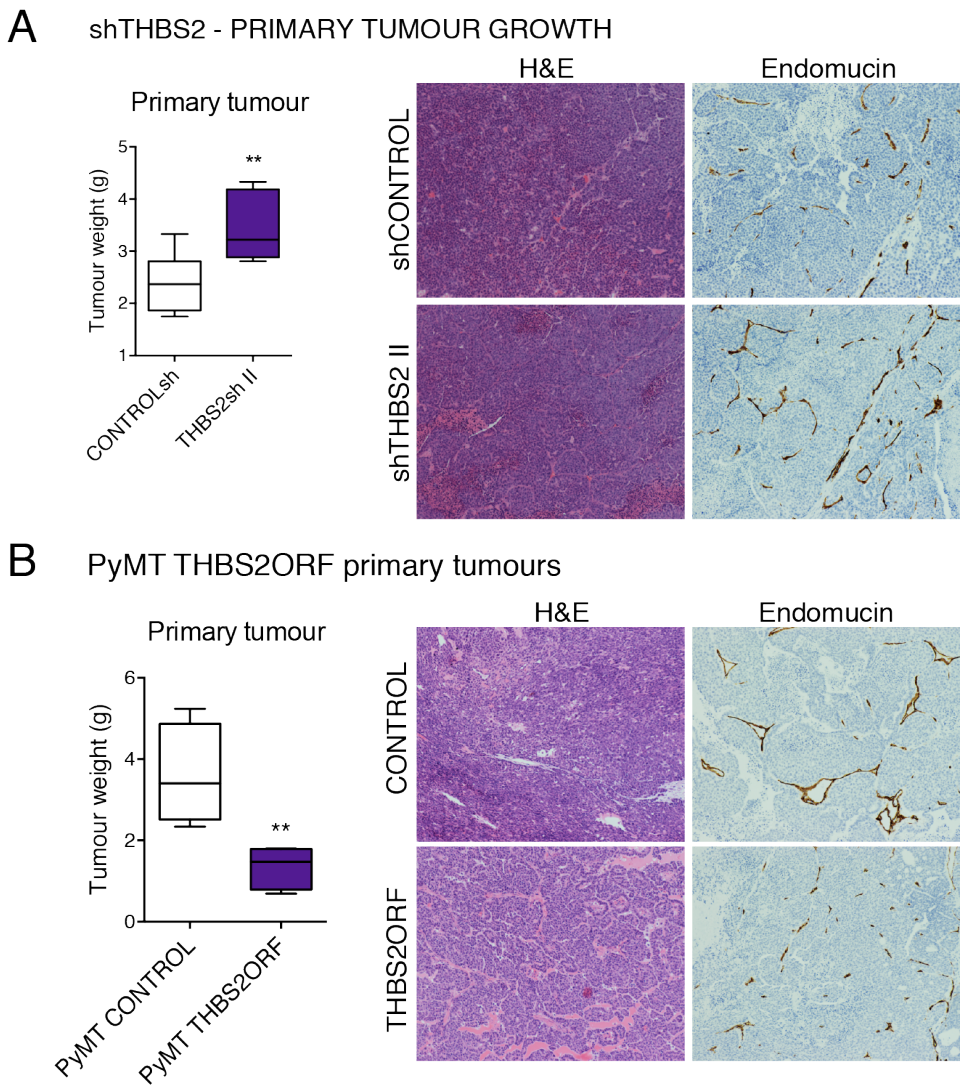


Figure 3.23 THBS2 effects in the primary tumour microenvironment

(A) Box plot shows late primary tumour growth: $3x10^5$ shCONTROL or shTHBS2 PyMT cells were injected into the fad mammary pat and tumour weight was evaluated 42 days post-injection (n=8). Histological analysis of shCONTROL or shTHBS2 PyMT tumours; (Left) H&E stain and (Right) Endomucin stain, endothelial cells (brown) and DAPI (blue). (B) Box plot shows late primary tumour growth evaluation, $3x10^5$ CONTROL or THBS2ORF PyMT cells were injected into the fad mammary pat and tumour weight was evaluated 42 days post-injection (n=6). Histological analysis of CONTROL or THBS2ORF PyMT tumours: (Right) H&E stain and (Left) Endomucin stain, endothelial cells (brown) and DAPI (blue).

Altogether, these findings suggest that different THBS2-mediated mechanisms operate during tumour initiation, progression and metastatic initiation, and propose THBS2 as a key microenvironmental regulator throughout disease progression.

Chapter 4. Early metastatic colonisation relies on the AXL-mesenchymal status and niche induction ability of cancer cells

4.1 Cancer cell niche induction ability is linked to their AXL-mesenchymal phenotype

In the previous chapter we described the metastatic niche activation ability as a crucial characteristic of mesenchymal cancer cells (Figure 3.13 and 3.15). In addition, we found the mesenchymal status of cancer cells, defined by AXL expression, to be sufficient to define their metastatic competence (Figure 3.10.E). Therefore, we next addressed whether the enhanced THBS2-dependent niche activation ability of metastatic could be linked to their AXL-mesenchymal phenotype. We checked in breast carcinoma patient samples whether THBS2 expression directly correlates with the mesenchymal features of breast cancer cells. First, using GOBO database (http://co.bmc.lu.se/gobo/coexpressed_genes.pl) we performed a gene co-expression analysis of the mesenchymal gene AXL in human breast carcinoma samples. We found that AXL is highly co-expressed with a signature containing THBS2 (Table 4.1). Importantly, this AXL co-expression gene list also contained the MIC marker (CD90 or Thy1), confirming in human breast tumours the link CD90-AXL that we observed in MICs in mice (Table 4.1). Note the perfect positive linear correlation, Pearson correlation coefficient r=1, of these genes with AXL (Table 4.1). Second, we analysed the specific correlation of THBS2 expression with genes used in this study to define a more mesenchymal status in breast cancer. Using the cBioPortal database for Cancer Genomics (http://www.cbioportal.org/) we found a direct correlation between THBS2 expression and the EMT-associated markers AXL, Vimentin, Twist1 and Zeb1 (Figure 4.1).

Gene symbol	Locus link	Correlation Pearson	STDdev	Gene name	Gene correlation network (all tumours)
AXL	558	1	0.77	AXL Tyrosine kinase receptor	ABL1;ADAM12;AEBP1;AKAP12;ANGPTL2;BGN;BMP1;BNC2;C10orf72;C1R;C1S; C5orf13;CALD1;CD200; CD248;CD81;CD93;CD99; CDH11;CILP;CLEC11A; CLIP3;CNN1;COL15A1; COL1A1;COL4A2;COL5A1;COL5A2;COL5A3;COL6A1;COL6A2;COL6A3;COL8A2;COMP;CPZ;CRISPLD2; CSF1R;CTGF;CTSK;CTSO;CUGBP2;CXCL12;CYR61; DAB2;DCN;DDR2;DIO2; DPYSL3;DVL3;ECM2; EDNRA;EFEMP2;EGR1; EHD2;ELN;EMILIN1;EMP3;ENG;ENTPD1;F13A1;FAP; FBLN1;FBLN2;FBLN5;FBN1;FERMT2;FHL1;FILIP1L; FLRT2;FNTB;FOXF2; FOXN3;FST;FSTL1;FYN; GAS1;GAS6;GAS7;GFPT2;GLG1;GPR124;GPX3;GSN;HEG1;HSPG2;HTRA1;IGFBP3;IGFBP6;ILK;ISLR;ITGA5;ITGBL1;JAM3;JUNB;KCTD12;KIAA1462;LAMB1;LOX; LOXL1;LRP1;LRRC15; LRRC17;LRRC32;LTBP2;LY86;MFAP2;MFAP4;MITF;MMP11; MMP14;MMP19; MMP2;MN1;MSN;MXRA5;MXRA8;MYH9;MYL9;NAP1L3;NBL1;NDN;NID1;NID2;ODZ4;OLFML1;OLFML2B;OLFML3;OMD;PALLD;PCOLCE;PDGFRB;PDGFRL;PECAM1;PKD2;PLAU;PLVAP;PLXNC1;PPAP2B;PRKCDBP;PRRX1;PTRF;RCAN2;RHOG;RUNX1T1;SEPT11;SERPINF1;SERPING1;SFRP4;SH2B3;SH3PXD2A;SHOX2;SLIT2;SNAI2;SPARC;SPOCK1; SPON1;SRPX;SSH1;SULF1;SYNPO;TAGLN;TCF4;TGFB1I1; TGFB3;**THBS2;THY1**;TIMP3;TRAM2;TSPAN4;UNC5B;VCAN; WIPF1;WISP1;WISP2;WNT2;ZCCHC24;ZEB1;ZFHX4

Table 4.1 AXL co-expressed genes analysis
AXL co-expressed genes in human breast carcinoma patient samples was analysed using GOBO database. The Co-expressed Genes algorithm in GOBO for a given gene, in this case AXL, uses Pearson as a correlation method, and returns gene lists that highly correlate with AXL expression in human breast tumours (correlation cut-off, 0.4). The list displayed in this table was the top co-expressed gene list with a Pearson correlation coefficient of 1 (absolute positive correlation).

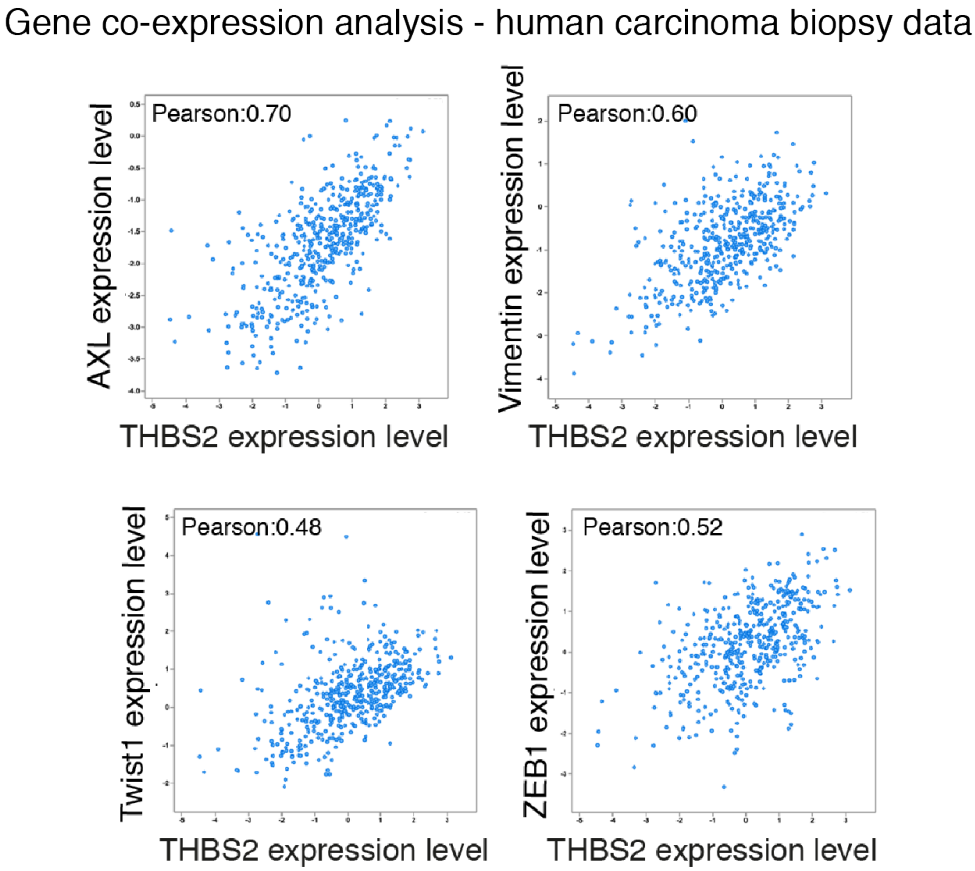


Figure 4.1 THBS2 expression correlates with different mesenchymal markers in human breast carcinoma patient data
Correlation plots show the gene expression correlation analysis between THBS2 and AXL, Vimentin, Twist1 and Zeb1 in human breast carcinoma patient samples using cBioportal data. The gene expression correlation algorithm in cBioPortal for Cancer Genomics uses Pearson as a correlation method comparing the expression of two given genes in a selected breast cancer patient data cohort (correlation cut-off, 0.4). The patient data cohort analysed was the Cancer Genome Atlas (TCGA provisional) containing 1098 patient samples, published in the Breast Invasive Carcinoma project (Cancer Genome Atlas, 2012).

These data strongly suggest that in human breast cancer there is a direct link between AXL expression and THBS2 secretion in breast cancer cells.

Next, we wanted to functionally validate the observed link between THBS2 and AXL expression. In order to do this, we inhibited the mesenchymal phenotype of tumour cells by targeting the EMT-associated receptor AXL in cancer cells and checked the impact of inhibiting the mesenchymal status on THBS2 secretion. AXL has been previously reported to directly interfere with the mesenchymal status of breast cancer cells (Asiedu et al., 2013, Vuoriluoto et al., 2011). We used two different approaches to block AXL; the specific small molecule inhibitor R428 that blocks the receptor activation and leads to its downregulation (Holland et al., 2010). This inhibitor allowed us to perform temporally controlled *in vivo* experiments. The second approach consisted of knocking down AXL expression by using previously described shRNA sequences against AXL delivered via lentiviral infection (Gjerdrum et al., 2010).

First, we validated that the specific inhibitor R428 interferes with AXL levels and leads to the inhibition of the EMT phenotype in MICs. We treated MICs with 1.5 µm R428 for 3 hours, and phenotypically evaluated the expression of the EMT-associated markers E-cadherin, AXL and Vimentin by immunofluorescence (Figure 4.2.A), as well as the levels of the EMT core transcription factors by real quantitative PCR (Figure 4.2.B). We could confirm that inhibiting AXL attenuates the mesenchymal status of MICs donwregulating AXL, Vimentin, the EMT transcription factors and upregulating E-cadherin. Second, we tested that this phenotypic changes translated into functional differences in cell behaviour by assessing MIC motility in 2D after R428 treatment. Remarkably, AXL downregulation was enough to decrease the overall motility of MICs (Figure 4.2.C).

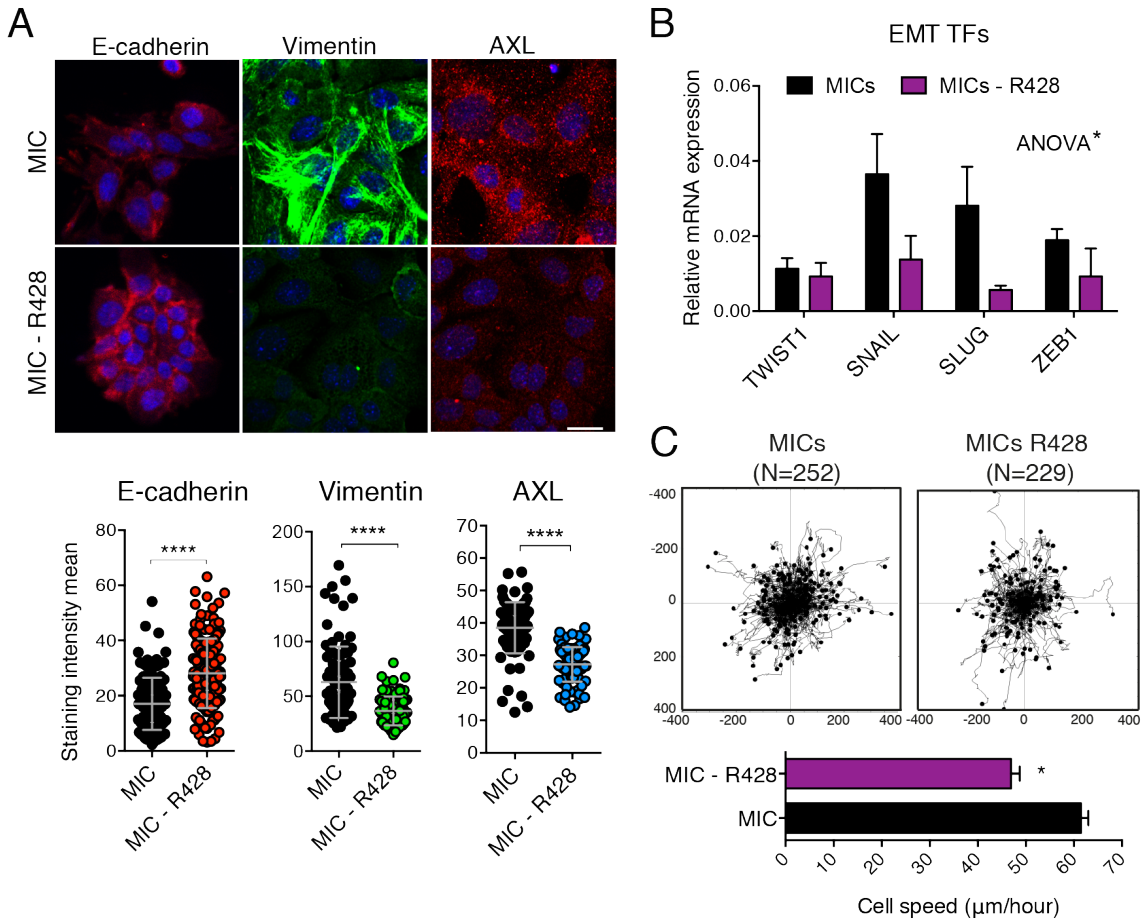


Figure 4.2 AXL inhibition attenuates the mesenchymal status of MICs
(A) Representative images show the expression of E-cadherin, Vimentin and AXL in MICs and MICs cultured overnight after FACS isolation and treated with DMSO or 1.5 µM R428. Data from one representative experiment from a total of 5 are shown. Chart line and error bars indicate mean ± SD of the population. Scale bar, 20 µm. (B) Quantitative real time PCR analysis compares the gene expression levels of different EMT transcription factors in MICs and MICs treated with R428. Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bar represent mean ± sem. p=0.0308 by ANOVA comparing the gene set in MICs *versus* MICs-R428. (C) Rose plots show the tracking path of MICs and MICs treated with 1.5 µM R428. (Lower panel) Histogram shows the mean ± sem of the cell speed quantified in 2 independent experiments.

In order to corroborate these results in a human model, we decided to search for human cell lines that highly express AXL. We examined AXL expression in a panel of breast carcinoma human cell lines using the GOBO database that compiles the data from a study where a panel of 30 human breast carcinoma cell lines were sequenced (Neve et al., 2006). In line with our data showing phenotypic expression of mesenchymal markers in MICs, high AXL expression was clearly restricted to basal B cell lines that cluster together based on their mesenchymal-like status, displaying Vimentin expression and loss of cytokeratins (Figure 4.3.A). We chose the broadly used MDA-MB-231 cell line, as it is one of the two with highest AXL expression. Then, we inhibited AXL in the MDA-MB-231 (MDA231) in a time course manner to assess how long is AXL downregulation sustained. The MDA231 cell line is homogeneously positive for AXL (Figure 4.3.B). After 24h treatment with R428 the overall expression of the receptor in the population is reduced by 50%, and 20% of cells completely lost AXL expression (Figure 4.3.B). AXL levels were nearly restored 48h after a single 3h treatment with the inhibitor (Figure 4.3.C). Alongside AXL we observed Vimentin downregulation (Figure 4.3.D), and functionally the overall cell motility of the MDA231 population decreased 24h after treatment with R428 (Figure 4.3.E).

Once we validated that AXL inhibition attenuates the mesenchymal status of both MICs and the MDA231 cells, we tested whether this would have a direct impact on the secretion of THBS2, establishing a direct link between the mesenchymal status of the tumour cells and their niche induction ability mediated by THBS2. First, we observed that the overall secretory status of MICs was compromised after AXL inhibition. We detected a general decrease in the expression of MIC secreted factors 24h after attenuating their mesenchymal features by blocking AXL (Figure 4.4.A). Particularly, we observed THBS2 secretion to decrease in both MICs and MDA231 cells after AXL mesenchymal inhibition (Figure 4.4.B). The level of secreted THBS2 into the cell culture media was measured by ELISA.

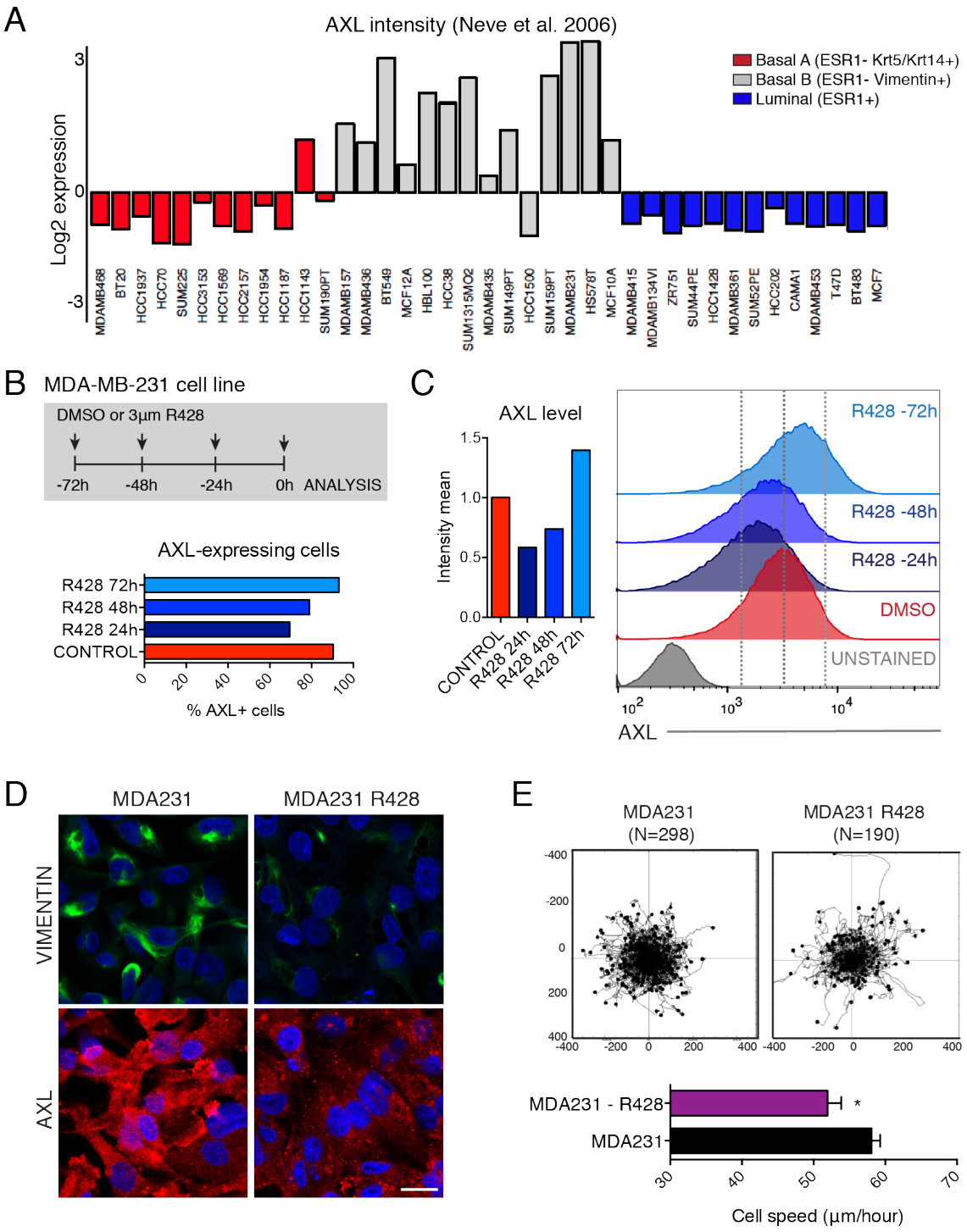


Figure 4.3 AXL inhibition attenuates the mesenchymal features of the MDA-MB-231 human cell line

(A) Analysis of AXL expression levels in a panel of 39 human cell lines, classified as luminal (ESR1$^+$) or basal (ESR1$^-$), the basal subtype was further divided into basal A (Krt5$^+$, Krt14$^+$) and basal B (Vimentin$^+$), using GOBO database. (B) Schematic shows the time course of AXL inhibition (R428) in the MDA231 cell line. Cells were treated with R428 3 µM 72, 48 or 24h prior to FACS analysis when AXL expression was determined. Bar graph shows the percentage of AXL$^+$ cells after

R428 treatment. (C) FACS analysis of AXL expression levels in MDA231 treated as in (B). Bar graph shows AXL intensity mean signal. FACS histogram plot displays the overall intensity of AXL in the MDA231 cell line treated as in (B). (D) Images show the expression levels of Vimentin and AXL in MDA231 and MDA231 treated with R428 3 µM for 24h. (E) Rose plots show the tracking path of MDA231 and MDA231 treated with 3 µM R428. (Lower panel) Histogram shows the mean ± sem of the cell speed quantified in 2 independent experiments.

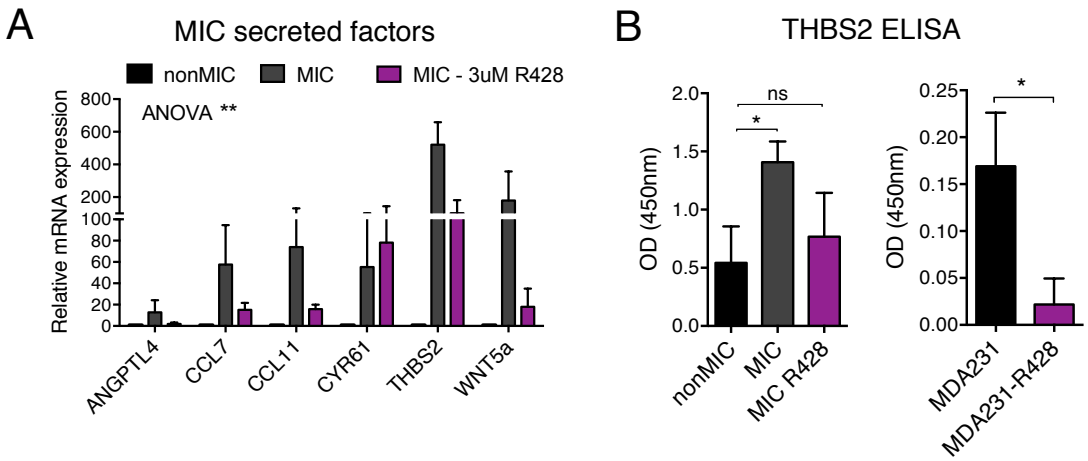


Figure 4.4 THBS2 secretion depends on the AXL-mesenchymal features of cancer cells

(A) Quantitative real time PCR analysis compares the gene expression levels of MIC secreted factors in nonMICs, MICs and MICs treated with 1.5 µM R428 for 24h. Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bar represent mean ± sem. p=0.0041 by ANOVA comparing the gene set in MICs *versus* MICs-R428. (B) ELISA assay shows the levels of secreted THBS2 in (Left) nonMICs, MICs and MICs treated with 1.5 µM R428, and (Right) MDA231 and MDA231 treated with R428 3 µM for 24h.

Considering the crucial role of THBS2 secretion in mediating MICs stromal activation advantage (Figure 3.19), we next assessed the consequences of inhibiting the mesenchymal features and therefore THBS2 secretion, on fibroblasts activation. We pre-treated freshly isolated MICs with R428 for 3h, and 24h after we assessed their lung fibroblast activation ability in coculture assays. Given that AXL inhibition effects last for up to 48h (Figure 4.3.B-C) we monitored the induction of early fibroblast activation markers in normal lung fibroblasts (NLF3). We observed that pre-treating MICs with R428, in line with the reduced secretion of THBS2 (Figure 4.4.B) attenuates their ability to induce nuclear YAP translocation and FAP expression in normal fibroblasts (Figure 4.5.A-B). We also confirmed these results using the MDA231 cell line. In this case, without the limitation of isolating high numbers of MICs from primary tumours we increased the number of cells per assay to decrease the time of the coculture assay. This allowed us to pre-treat the MDA231 cell line with R428 for 3h, and 24h after we washed the inhibitor and set the coculture assay as in Figure 3.13 to assess the gel contraction ability of the fibroblasts. As previously shown, in R428 pre-treated MDA231 cells AXL inhibition last up to 48-72h (Figure 4.3.A-B), which is the total duration of the gel contraction assay performed. R428 pre-treated MDA231 cells showed a delayed induction of fibroblast-driven gel contraction (Figure 4.5.C) in line with their loss of mesenchymal features (Figure 4.3.B) and THBS2 expression (Figure 4.4.B).

To confirm that the above effects are the result of a targeted inhibition and downregulation of AXL and not unspecific effects from the small molecule inhibitor R428, we engineered a lentiviral reporter construct containing a validated shRNA sequence against AXL. The lentiviral vector contained an EGFP reporter gene that allows us to select the infected cells by FACS (Figure 4.6.A). First, we validated the knockdown efficiency in primary PyMT tumour cells after lentiviral infection (Figure 4.6.B). Second, we verified that upon AXL knockdown the mesenchymal status of the shAXL cells was attenuated as observed by the reduced Vimentin expression and the upregulation of E-cadherin in primary infected PyMT cells (Figure 4.6.C). In line with the results obtained with the AXL inhibitor, we could detect a decrease in THBS2 secretion levels in shAXL PyMT cells (Figure 4.6.D). Last, we functionally tested the ability of PyMT shAXL cells to induce fibroblast activation in a gel

contraction assay. As expected from the reduced secretion of THBS2, knocking down AXL in PyMT cells reduced their ability to activate lung fibroblasts (Figure 4.6.E).

To determine the relevance of AXL expression during metastatic colonisation we tested shAXL cancer cells to grow *in vitro* and *in vivo*. We excluded that depleting AXL could have a direct effect on the intrinsic stemness potential of cancer cells by measuring the ability of shAXL cells to form spheres *in vitro*. No significant effects were observed in the sphere formation ability of shAXL cells (Figure 4.6.F). Next, to confirm the innocuous effect on stemness when depleting AXL, we tested the tumour initiation ability of shAXL cells *in vivo*. As expected from the *in vitro* results, primary tumour initiation was not affected in the absence of AXL (Figure 4.6.G). Collectively, these results indicate that AXL does not affect the intrinsic abilities of tumour cells to self-renew and initiate a tumour mass. Finally, we assessed the ability of shAXL cells to metastasise following intravenous injection. In line with the reduction of THBS2-dependent stromal activation, shAXL cells showed a decrease ability to initiate metastasis (Figure 4.6.H).

We corroborated these results in the highly metastatic 4T1 mouse cell line that homogeneously expresses high levels of AXL (Figure 4.7.A). Either THBS2 or AXL depletion decreased the lung colonisation ability of 4T1 cells in an experimental metastasis setting (Figure 4.7.B).

In summary, we have shown that the AXL-mesenchymal status of cancer cells is functionally required not only to support their invasion at the primary site but also to trigger the niche activation at the target metastatic site. Moreover, these results highlight the crucial requirement of AXL-mesenchymal features as metastatic cells infiltrate the new tissue.

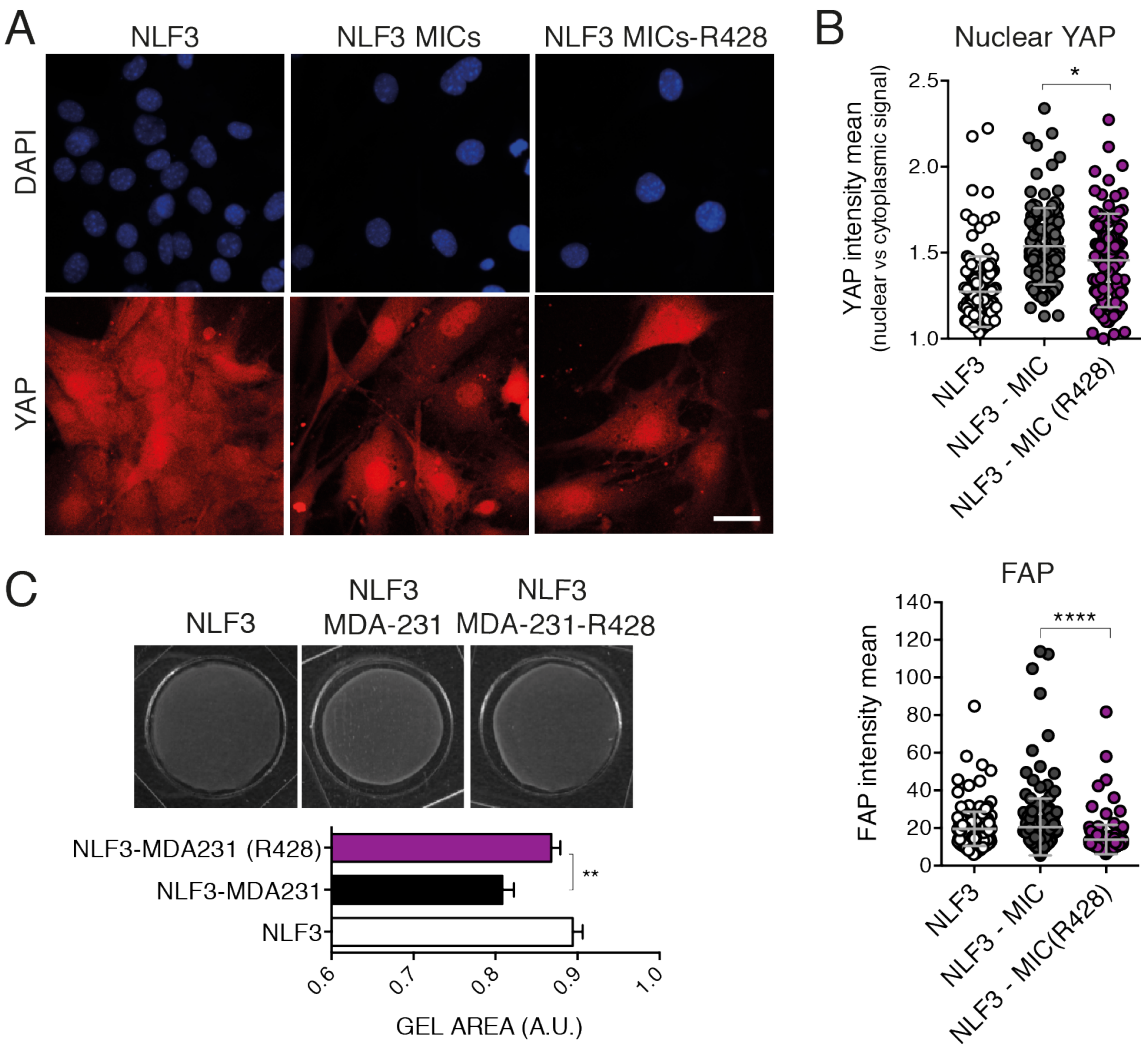


Figure 4.5 AXL inhibition partially blocks the ability of mesenchymal cancer cells to activate fibroblasts

(A) Representative images show YAP nuclear staining in NLF3 cells cultured in the presence of MIC or MICs pre-treated with 1.5 μM R428 for 24h before the coculture. Scale bar, 100 μm. (B) Charts indicate the quantification of the nuclear levels of YAP and total expression of FAP. Data from one representative experiment from a total of 2 are shown. Line and error bar indicate mean ± SD of the populations. (C) (Upper panel) Representative images display the gel contraction assay 3 days after coculture with MDA231 or MDA231 pre-treated with R428 3 μM for 24h. (Lower panel) Histogram displays mean ± sem (n=12-13, number of gels assessed over 3 independent experiments).

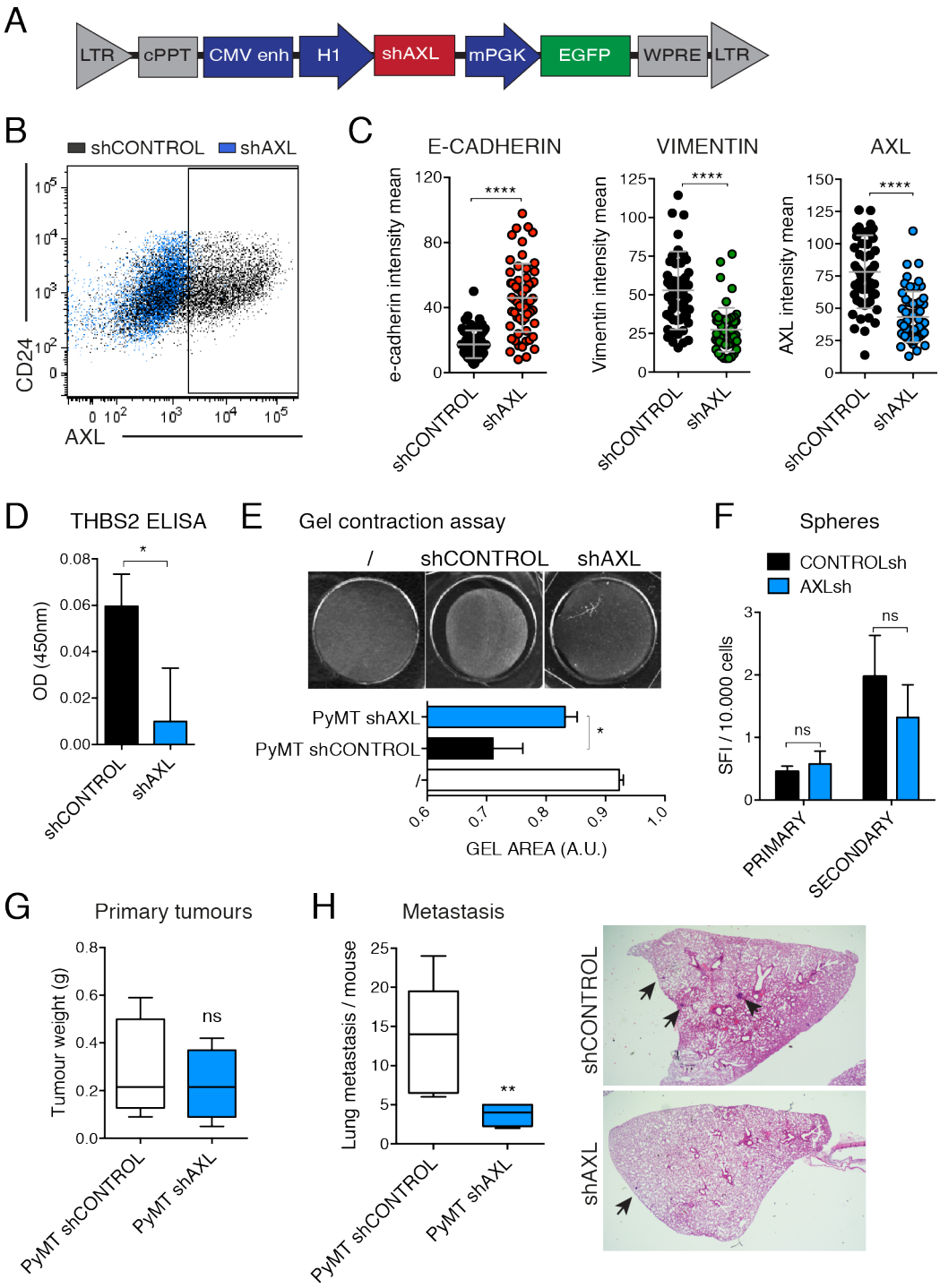


Figure 4.6 AXL knockdown attenuates the mesenchymal phenotype and reduces THBS2 secretion, fibroblast activation and metastatic colonisation
(A) Schematic shows the lentiviral reporter used to express the shRNA sequences. The shRNA sequences were cloned under the expression of a CMV enhancer followed by the H1 promoter. An independent mouse PGK promoter controls EGFP expression. (B) FACS dot plot displays AXL expression levels in PyMT cells expressing shCONTROL or shAXL lentiviral constructs. (C) Charts indicate the expression levels of E-cadherin, Vimentin and AXL in PyMT cells expressing

shCONTROL or shAXL lentiviral constructs. Line and error bars indicate mean ± SD of the population. (D) ELISA assay measures the levels of secreted THBS2 by PyMT cells expressing shCONTROL or shAXL lentiviral constructs. Histogram displays mean ± sem of 2 independent cell preparations. (E) (Upper panel) Representative images of the gel contraction assay 3 days after coculture with PyMT cells expressing shCONTROL or shAXL lentiviral constructs. (Lower panel) Histogram displays mean ± sem (n=7-9, number of gels assessed over multiple experiments). (F) Sphere assay by PyMT cells in suspension expressing shCONTROL and shAXL lentiviral constructs. Histogram shows the mean ± sem of 2 independent experiments. Spheres were quantified 10 days after plating a single PyMT suspension (primary) and 10 days after the first passage (secondary). (G) Primary tumour growth potential. Cells were FACS sorted to select for lentiviral expression and 50,000 shCONTROL or shAXL PyMT cells were injected into the mammary fad pat. Tumour weight was evaluated 20 days post-injection (n=6). (H) Metastatic colonisation ability of shCONTROL and shAXL PyMT cells. Positively infected PyMT cells were sorted as GFP^{+}, and $5x10^{6}$ cells were directly injected into the tail vein of mice. (Left) Micrometastasis were evaluated 30 days post-injection. (Right) Representative images of lung sections with micrometastasis.

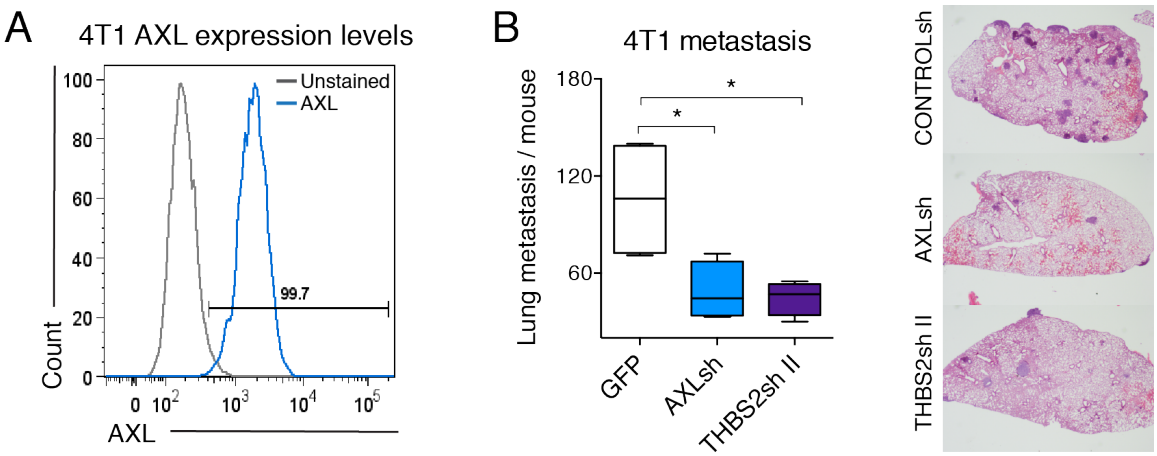


Figure 4.7 Metastatic ability of 4T1 cells upon THBS2 or AXL knockdown
(A) FACS histogram displays the level of AXL expression in 4T1 cell line. (B) Box plot shows the lung experimental metastasis after intravenous injection of 4T1 cell line expressing shCONTROL, shAXL and shTHBS2 lentiviral constructs (n=4 mice per group). Representative images of H&E stain of lung sections displaying 4T1 metastatic nodules.

4.2 The AXL-mesenchymal status of cancer cells is required upon extravasation for effective metastatic colonisation

The previous results show that MICs isolated from primary tumours are in a mesenchymal state characterised by the expression of AXL. Therefore, in order to understand the regulation and dynamics of this mesenchymal phenotype during metastatic colonisation following arrival in the lungs, we monitored the expression of AXL on cancer cells after intravenous injection. Based on the presence of AXL, we could discriminate two distinct phases during lung colonisation: a first phase early after extravasation where MICs still maintain a mesenchymal phenotype; and a second phase where alongside proliferation and enrichment among the total cancer cells in the lungs, MICs lose their AXL-mesenchymal phenotype (Figure 4.8.A).

We have also shown that the ability of metastatic cells to activate the niche is crucial for metastatic establishment (Figure 3.19) and depends on the mesenchymal phenotype of cancer cells (Figure 4.6). Therefore, we next tested whether the AXL-mesenchymal status needs to be maintained specifically during the first colonisation phase when disseminated cancer cells maintain AXL expression. To this end, we used the AXL small molecule inhibitor R428 that allows us to block AXL expression *in vivo* during this specific window of time. We pre-treated cancer cells *in vitro* prior to intravenous injection and then continued the treatment *in vivo* for 7 days (Figure 4.8.B). We checked that *in vivo* administration of R428 could indeed block AXL expression in cancer cells following intravenous injection (Figure 4.8.C), and that extravasation into the lungs 24h post-seeding was not affected by the pre-treatment *in vitro* (Figure 4.8.D). Remarkably, the *in vitro* pre-treatment was sufficient to decrease the metastatic initiating potential of PyMT cells 30 days post-injection, and this effect was further exacerbated by the continuous 1 week treatment *in vivo* (Figure 4.8.E). Notably, these results reaffirm that the highly AXL-expressing MICs (Figure 3.10.A) are the main population driving lung colonisation in this model (Malanchi et al., 2012).

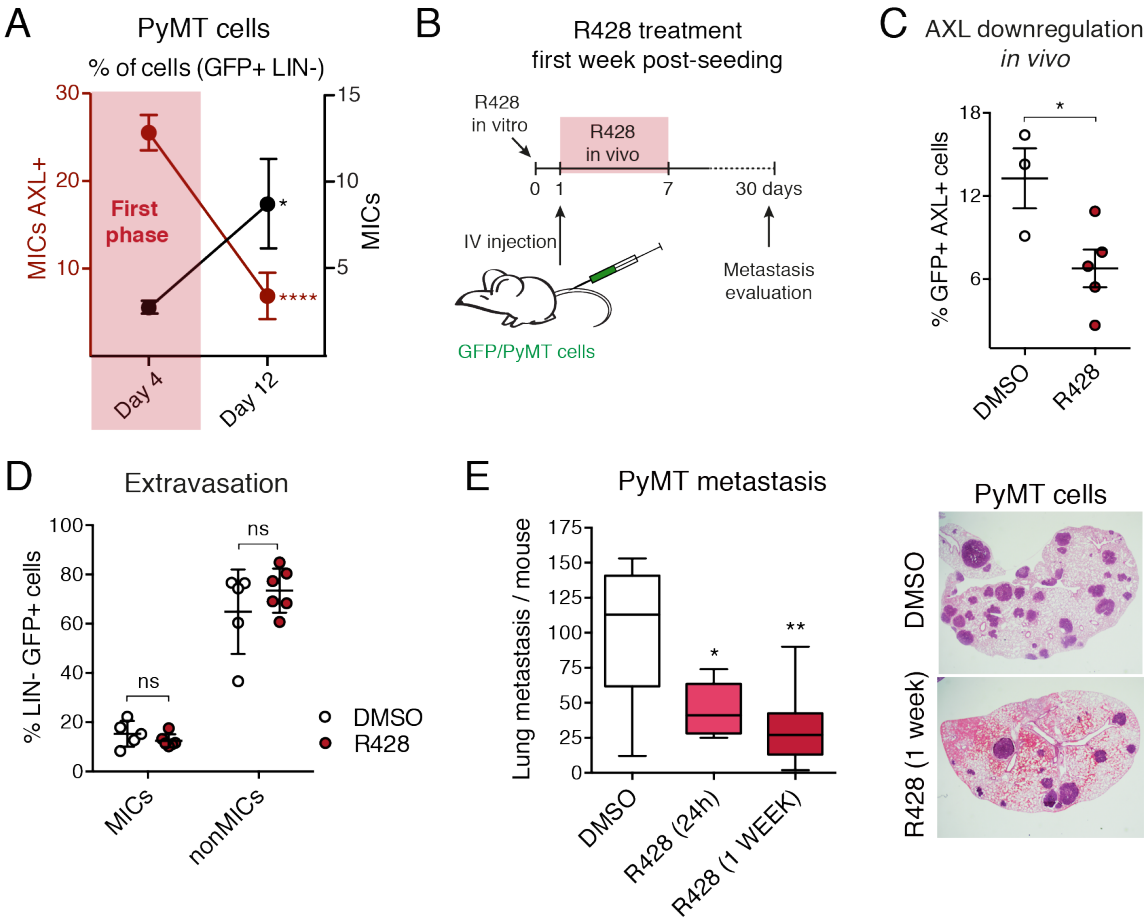


Figure 4.8 AXL–mesenchymal features are required during the early colonisation phase of metastasis

(A) FACS analysis of the percentage of AXL$^+$ MICs (red-left Y axis) and total MICs (black-right Y axis) among total GFP$^+$ cancer cells in the lung at 4 and 12 days after tail vein injection. Graph represents mean ± sem per time point (n=8-9 mice per group of 2 independent experiments). (B) Experimental setup: PyMT cells were pre-treated *in vitro* with vehicle (DMSO) or 1.5 µM R428, 24h hours before intravenous injection. *In vivo*, mice were treated daily with 12.5 mg/kg of R428 during the indicated days, and metastases were evaluated 30 days post-injection.
(C) FACS analysis of AXL$^+$ cancer cells after R428 treatment *in vivo*. Mice were intravenously injected with primary PyMT GFP$^+$ cells, and treated daily with DMSO (vehicle) or R428 (12.5 mg/kg) for 3 days, when lungs were analysed by FACS. Chart shows mean ± sem (n=3-5 per group). (D) FACS analysis of MIC and nonMIC cells presence in the lung. PyMT GFP$^+$ cells were treated *in vitro* with DMSO or 1.5 µM R428 24 hours prior to intravenous injection into mice. FACS analysis was performed 24h post-injection after lung perfusion with PBS. Chart shows mean ± sem (n=5-6 per group). (E) Box plot displays the metastatic burden of PyMT cells intravenously injected into mice as specified in (B). PyMT-DMSO, n= 10 mice per group; PyMT R428 (24h), n= 5 mice per group; PyMT-R428 (1 week), n= 9 mice per group. Superficial lung metastasis number was evaluated. (Right) Representative H&E staining of metastatic lung sections from mice of indicated groups.

Similar results were obtained using the human cell line MDA-MB-231. Following intravenous injection, this AXL highly positive cell line spontaneously downregulates AXL as the cells grow in the lung transitioning from the first to the second phase of metastatic colonisation (Figure 4.9.A). We could confirm the requirement of AXL during the first phase of colonisation applying the same *in vivo* treatment scheme that depletes AXL expression during the first week of colonisation (Figure 4.8.B) obtaining reduced metastatic burden 35 days post-injection (Figure 4.9.B).

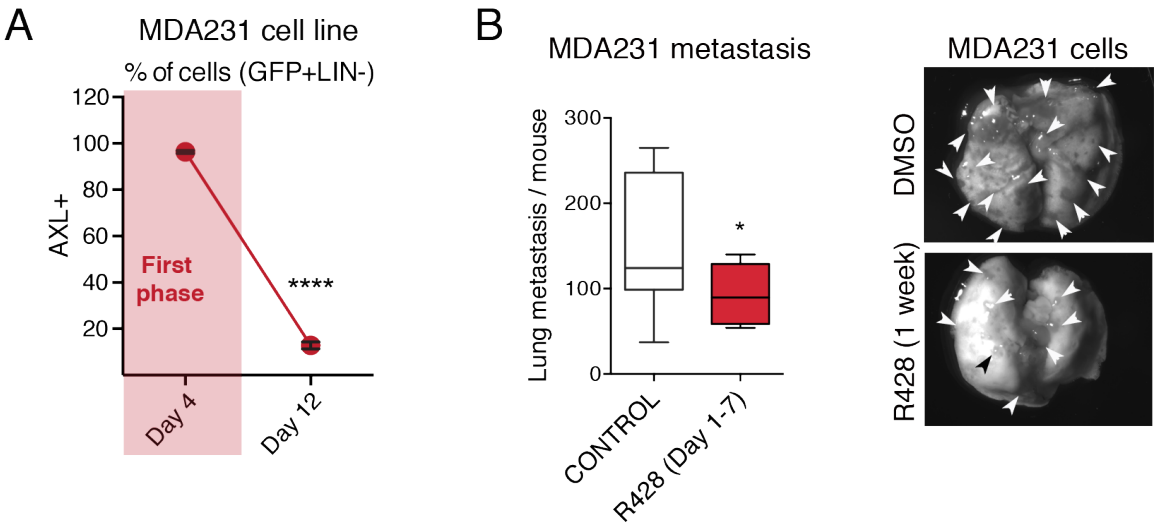


Figure 4.9 The MDA-MB-231 human cell line requires an active AXL–mesenchymal programme during early metastatic colonisation
(A) FACS analysis of the percentage of AXL^+ MDA-MB-231 (MDA231) among total GFP^+ cancer cells in the lung at 4 and 12 days after tail vein injection. Graph represents mean ± sem per time point (n=3 mice per group). (B) Box plot shows the metastatic burden of MDA231 cells after pre-treatment *in vitro* with vehicle (DMSO) or 3 µM R428, 24h hours before intravenous injection. *In vivo*, mice were treated daily with 12.5 mg/kg of R428 from day 1 to 7 post-injection, and metastases were evaluated 25 days post-injection. MDA231-DMSO, n= 13 per group; MDA231-R428 (Day1-7), n= 8 mice per group. Superficial lung metastasis number was evaluated.

These results corroborate the crucial requirement of AXL-mesenchymal features precisely during the first phase of colonisation, when the niche induction ability of metastatic cells provides metastatic advantage.

We next tested whether the used short-term AXL inhibition would have an effect in the intrinsic abilities of cancer cells to grow and self-renew. *In vitro*, AXL inhibition did not have an effect on proliferation as measured in a BrdU incorporation assay in both PyMT cells and the MDA231 cell line (Figure 4.10.A-B). Also, we excluded a negative effect of AXL inhibition on the self-renewal capacity of primary PyMT cells as measured by their sphere formation capacity (Figure 4.10.C) and their ability to grow 3D organoids when embedded into Matrigel (Figure 4.10.D). Last, we checked the effect of a single AXL inhibitor treatment on cancer cells tumorigenic potential. R428 pre-treatment leading to a three-day AXL down-regulation (Figure 4.3.C) is enough to reduce PyMT cells metastatic potential (Figure 4.8.E). However, primary tumour development was not affected by the reversible early inhibition of AXL in both PyMT primary cells and MDA231 cells (Figure 4.10.E). This is in line with the observation that in the PyMT model permanent downregulation of AXL by shRNA does not affect primary tumour initiation (Figure 4.6.G). Altogether, these results suggest that while AXL is required by breast cancer cells to re-initiate tumour growth at distant sites (Figure 4.8.E and 4.9.B), it is dispensable to start a tumour when cells are locally transplanted into the mammary fad pat (Figure 4.6.G and 4.10.E).

Our data indicate that AXL is required during early colonisation to allow efficient metastasis. We have also shown that AXL-mesenchymal features provide an advantage in niche activation via THBS2 secretion. Thus, we next tested whether THBS2 overexpression could partially rescue the effects of blocking AXL *in vivo*. In order to test this, we generated a MDA231 cell line clone that expresses THBS2ORF lentiviral construct (Figure 3.21.A) and confirmed the increase in THBS2 expression (30-fold increment) (Figure 4.11.A). Then, we assessed the metastatic potential of R428 treated MDA231 cells expressing the control VENUS or THBS2ORF lentiviral constructs. Cells were pre-treated with R428 *in vitro*, and then continued treating the mice *in vivo* for the first week of metastatic colonisation after intravenous injection (as in Figure 4.8.B). We showed that despite the broader effects that AXL inhibition might have on tumour cells, THBS2 overexpression could partially rescue the deleterious effects of blocking AXL in this early stage of metastasis (Figure 4.11.B). However, in contrast to what we observed in the PyMT

cells (Figure 3.21.C), THBS2 overexpression was not enough by itself to significantly enhance the overall metastatic efficiency of the MDA231 cell line. This is likely due to the fact that all cells within the MDA231 cell line already express AXL and produce THBS2. These results strengthen the functional link previously described between AXL and THBS2 (Figure 4.4).

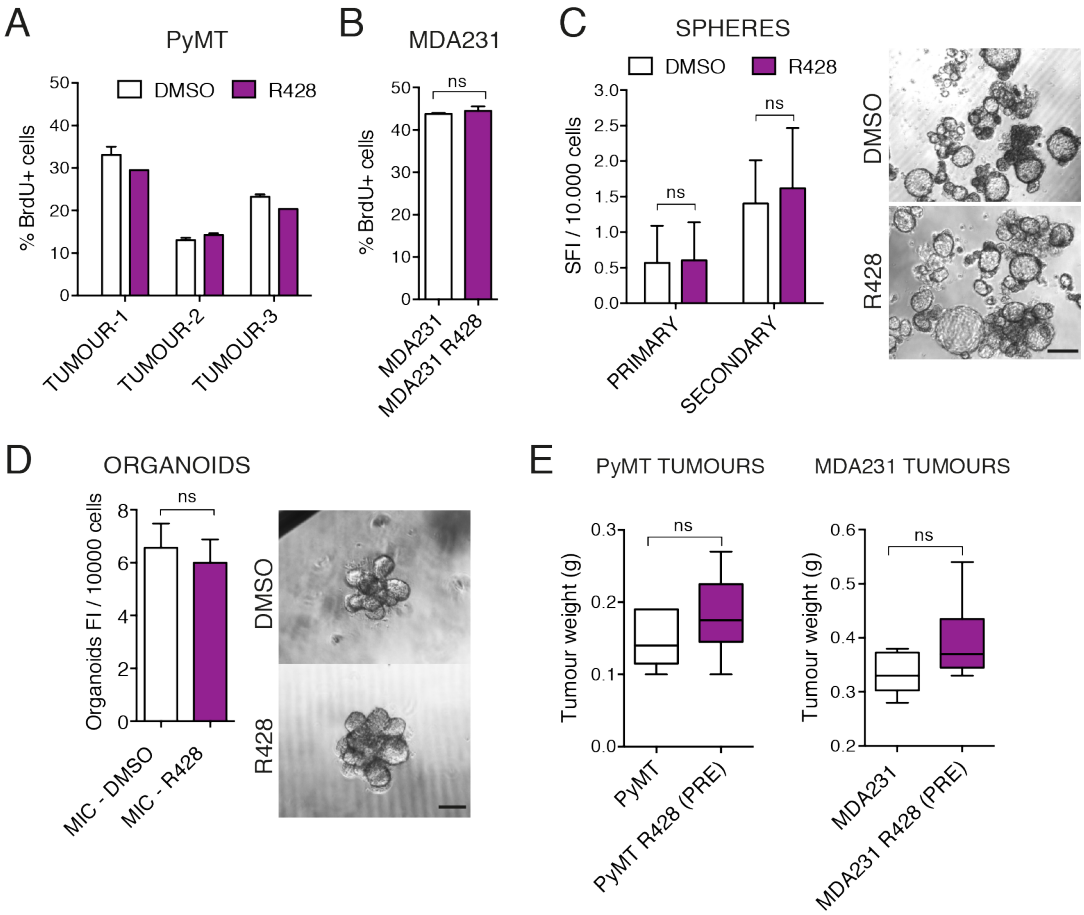


Figure 4.10 AXL inhibition effects on cancer cell proliferation, self-renewal and tumour initiation
(A) FACS analysis shows BrdU incorporation over a 3 hours pulse by primary PyMT cells and PyMT cells treated with 1.5 μM R428. Data from 4 different tumours performed in triplicate. (B) FACS analysis shows BrdU incorporation over a 3 hours pulse by MDA231 and MDA231 cells treated with 3 μM R428. Histogram shows mean ± sem of 3 independent experiments. (C) Sphere culture assay. (Left) Histogram shows mean ± sem of 3 independent experiments. Spheres were quantified 7 days after plating a single PyMT suspension (primary) and 7 days after the first passage (secondary). (Right) Images show representative pictures of spheres formed by primary PyMT cells and PyMT cells treated with 1.5 μM R428 in suspension. Scale bar, 200 μm. (D) Organoid assay. (Upper panel) Images show

representative pictures of organoids formed by primary PyMT cells and PyMT cells treated with 1.5 µM R428, and seeded in Matrigel. Scale bar, 200 µm. (Lower panel) Histogram shows the mean ± sem of 2 independent experiments. (E) Box plots display tumour burden assessed 3 weeks after engraftment of PyMT cells pre-treated with 1.5 µM R428, and MDA231 cells pre-treated with 3 µM R428 (n=6 per group).

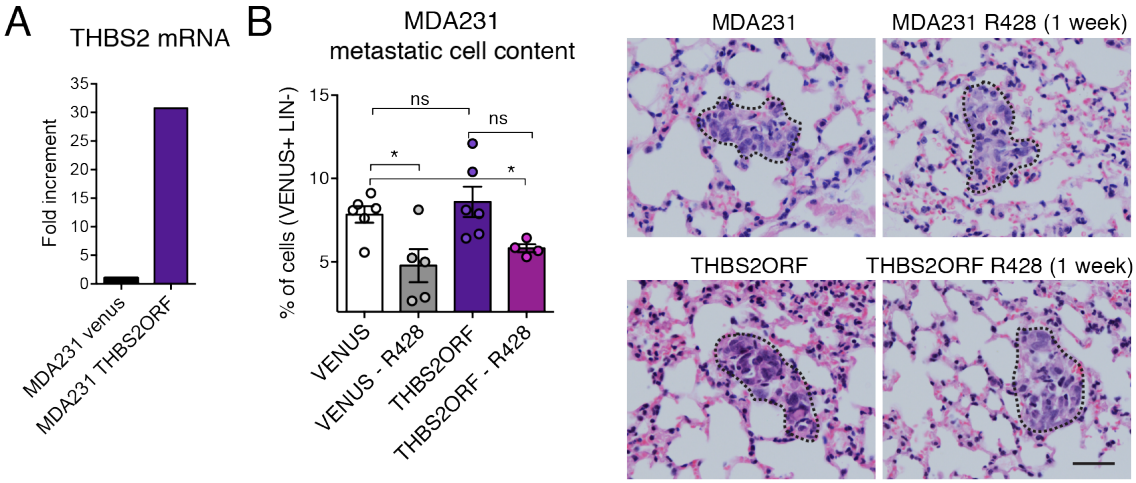


Figure 4.11 THBS2 overexpression partially rescues AXL inhibition in MDA231 cells

(A) Quantitative real time PCR showing a 30-fold THBS2 mRNA expression in MDA231 cell line after infection with the THBS2ORF construct and selection for VENUS positive cells. (B) Box plot shows the flow cytometry quantification of the percentage of $VENUS^{+}$ cells in the lungs of mice injected with MDA231-VENUS or MDA231-THBS2ORF treated with/without R428. Mice were injected with $5x10^{5}$ cells pre-treated with DMSO or 3 µm R428 *in vitro* 24h before. After intravenous injection mice were treated with DMSO or 12.5 mg/kg for 1 week. MDA231 cell content in the lungs was evaluated 20 days post-seeding by flow cytometry of 4 lung lobes (n=4-6 mice per group). (Left) Representative images show lung micrometastasis for the indicated groups in the lobe that was not used for the FACS analysis.

The results presented in this chapter suggest that the niche activation ability provided by the AXL-mesenchymal status and THBS2 is crucial for the first phase of metastatic colonisation, before cancer cells begin to actively grow.

Chapter 5. The activated niche tunes the cancer cell mesenchymal status in a temporally controlled manner during the second phase of metastasis

5.1 The newly activated niche attenuates the mesenchymal features of cancer cells

In the previous chapters, our results highlight the metastatic advantage provided by a mesenchymal status as single cells infiltrate into the target site. Subsequently, we observed that this AXL-mesenchymal programme is spontaneously downregulated *in vivo* as tumour cells start growing in the lungs (Figure 4.8.A and 4.9.A), suggesting that the requirement of the previously reported mesenchymal-to-epithelial transition (MET) at the secondary site (Ocana et al., 2012, Tsai et al., 2012) might be temporally controlled during this second phase of colonisation. Therefore, we next set out to investigate the epithelial modulations occurring in cancer cells during metastatic colonisation. In order to do this, we used ImageStream technology to examine the dynamics of different intracellular factors in disseminated tumour cells in the lungs during colonisation. ImageStream is an imaging flow cytometer that allows the acquisition of high-resolution images of each cell directly in the flow. These capabilities allow quantitating the intensity and location of fluorescent protein stains in rare cell populations. As previously discussed colonisation is the bottleneck of the metastatic process (section 1.2.3), where the total number of tumour cells found in the lungs is lower than 0.1% of the originally disseminated cells (Luzzi et al., 1998). Therefore ImageStream allows the acquisition of quantitative imaging data from each whole lung cell preparation and the analysis of rare populations, which would be very challenging to localise by standard histological sectioning of the tissue and microscopy.

The experimental setting for subsequent ImageStream analysis consisted of seeding GFP labelled $CD24^{+}AXL^{+}$ cancer cells directly into the lungs via tail vein injection two days (first phase) or twelve days (second phase) prior to analysis (Figure 5.1.A). We observed that alongside the previously described AXL down-regulation (Figure 4.8.A and 5.1.B) the EMT-TF Twist1 was concomitantly loss

indicating a reversion of the mesenchymal state. The total percentage of Twist1^{+} cells decreased in the transition from early to late colonisation (Figure 5.1.C). Moreover, in those cells that remained Twist1^{+}, the overall Twist1 protein levels were dramatically reduced as cancer cells expanded in the lungs as shown by the representative images and the contour FACS plot displaying the overall Twist1 intensity decrease in metastatic cells at day 12 in the lungs (Figure 5.1.D-E). These data are in line with the previously reported mesenchymal inhibition at the secondary site (Ocana et al., 2012, Tsai et al., 2012), and reaffirm the use of AXL as a bona-fide marker to track the mesenchymal status of breast metastatic cells *in vivo*.

Next, given the solid evidence for a mesenchymal reversion during late metastatic colonisation we wanted to investigate the causes of these epithelial modulations within the target site. Our previous results show that the induction of a metastatic niche is a pre-requisite for colonisation (Figure 3.20) and its activation temporally coincides AXL is loss (Figure 4.8.A). Therefore, we set out to test the potential regulatory role of the activated niche fibroblasts on cancer cells epithelial plasticity during late metastatic colonisation. First, we analysed *ex vivo* the effects of fibroblasts activation on cancer cells. We repeated the previous coculture assays used to monitor fibroblast activation where freshly isolated MICs are seeded on top of a thin layer of Matrigel:collagen-I gels with GFP labelled normal fibroblasts over 3-5 days (Figure 3.14.A and 5.2.A). We found that as fibroblast become activated, MICs downregulated AXL expression in these cocultures (Figure 5.2.B). Moreover, alongside AXL, MICs downregulated the expression of the core EMT transcription factors as monitored by their RNA expression levels (Figure 5.2.C). Notably, in line with our *in vivo* data Twist1 was the EMT-TF showing a higher decrease upon fibroblast activation in the coculture assays. Accordingly, we specifically checked whether Twist1 protein levels could be negatively regulated by activated fibroblasts *ex vivo*. We monitored Twist1 expression in MICs using ImageStream analysis and found that as fibroblast become activated in the coculture MICs downregulate Twist1 expression (Figure 5.2.D). This result suggests that activated fibroblasts can trigger the re-epithelialisation of mesenchymal cancer cells.

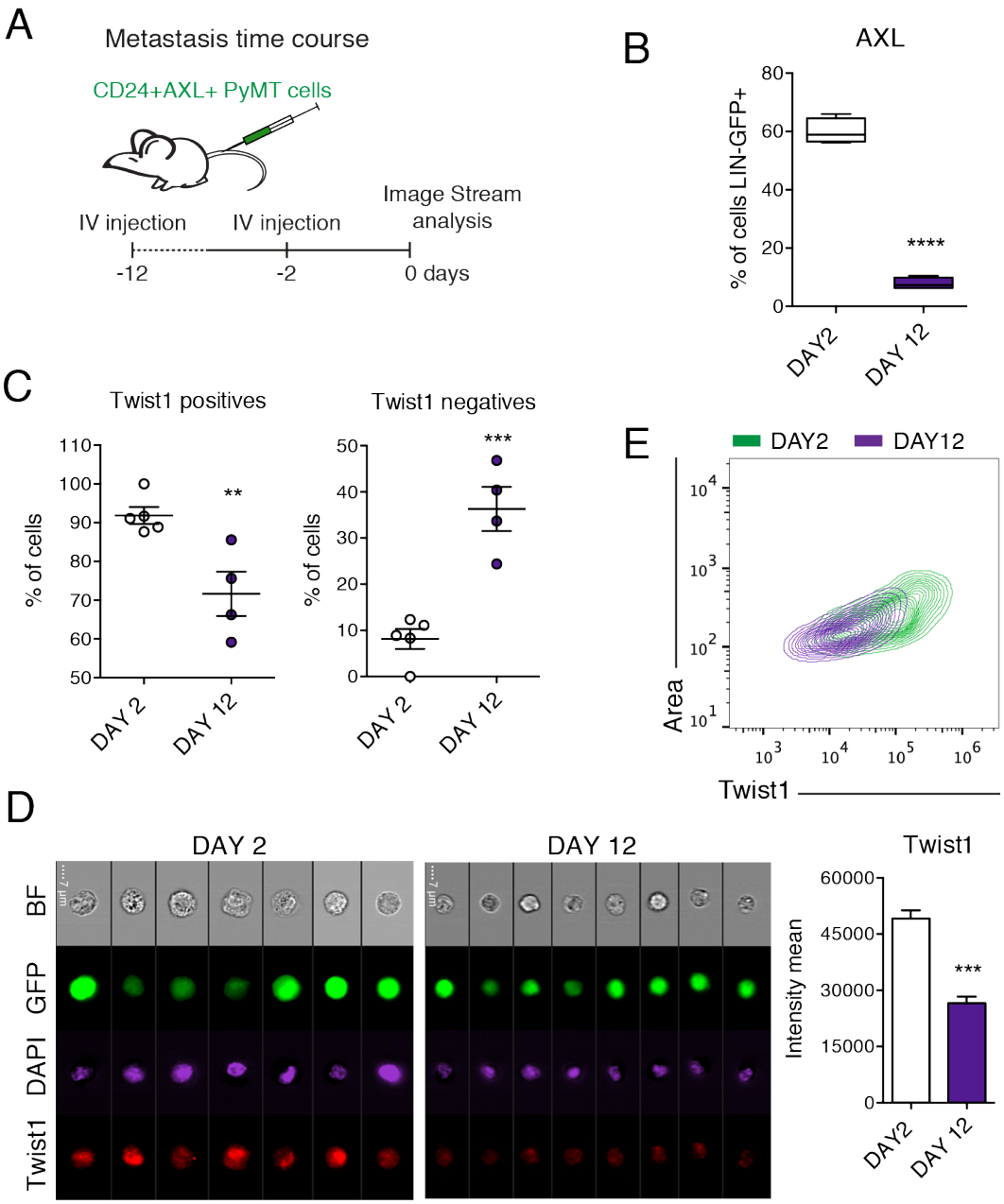


Figure 5.1 AXL and Twist1 expression is downregulated in metastasising cells during late colonisation

(A) Schematic shows the experimental set up for the metastasis time course. Mice were intravenously injected with primary actinGFP-PyMT $CD24^{+}AXL^{+}$ cells 12 or 2 days before analysis. At day 0, lungs were depleted from lineage cells ($CD45^{+}$ $CD31^{+}$ $Ter119^{+}$ cells) and then half of the lung was freshly analysed for AXL expression by FACS, and the other half fixed for intracellular stain and ImageStream analysis. (B) Box plot shows the *in vivo* evaluation of the percentage of AXL^{+} cells in the lung in a metastasis time course. The percentage of $LIN^{-}GFP^{+}$ $CD24^{+}AXL^{+}$ cells was evaluated (n=8-9 mice). One representative experiment out of five is shown. (C) Chart shows the *in vivo* evaluation of the percentage of $Twist1^{+}$ cells in the lung in a metastasis time course. The percentage of $LIN^{-}GFP^{+}$ $Twist1^{+}$ cells was evaluated (n=4-5 mice). Data from one representative experiment out of

three are shown. (D) Representative images show the reduction in Twist1 in cells isolated from 2 or 12 days metastasis, quantified in (C). (E) FACS contour plot shows the intensity of Twist1, and (Left) bar graph the quantification of its intensity mean in $CD24^{+}AXL^{+}$ tumour cells in the lung at 2 and 12 days of metastasis. Data from one representative experiment out of three are shown.

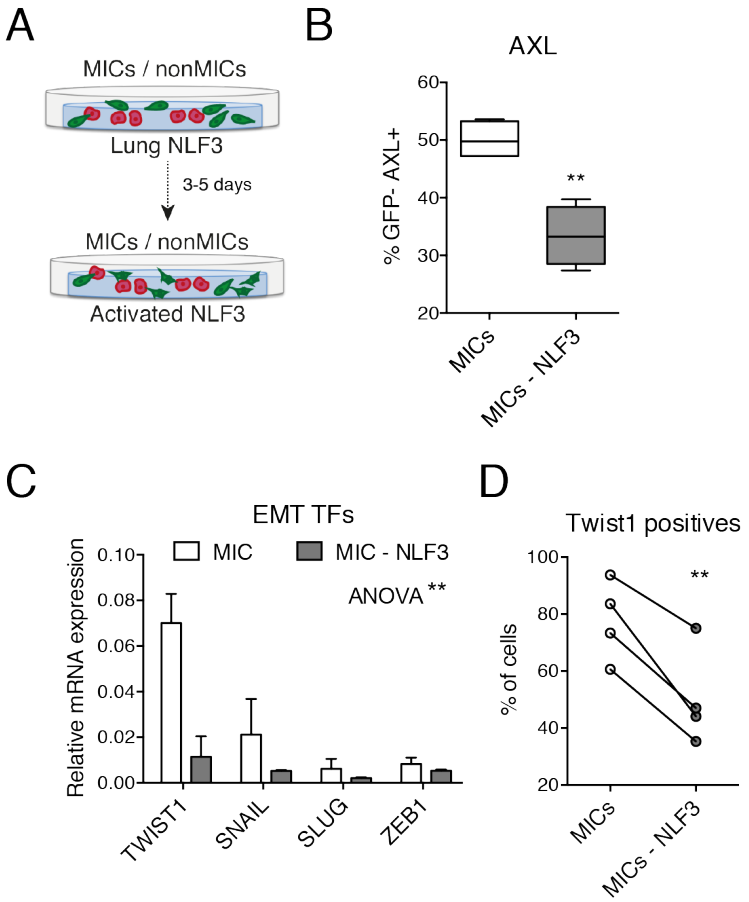


Figure 5.2 Activated fibroblasts induce AXL and Twist1 downregulation in MICs

(A) Schematic shows the coculture setting where nonMICs or MICs are seeded on top of a thin layer of Matrigel-Collagen with GFP^{+} labelled normal lung fibroblast cell line (NLF3) for 3-5 days. (B) FACS quantification of AXL expression levels in MICs and MICs cocultured with NLF3 as in (A). Histogram shows the mean ± sem (n= 2 number of cocultures). (C) Quantitative real time PCR analysis compares the gene expression levels of different EMT transcription factors in MICS and MICS cocultured with NLF3 for 5 days as in (A). Data from 3 different experiments performed in triplicate (normalised to Gapdh). Bar represent mean ± sem. p= 0.0018 by ANOVA comparing the gene set in MICs *versus* MICs-NLF3. (D) ImageStream evaluation of the percentage of $Twist1^{+}$ in MICs and MICs cocultured with NLF3 as in (A). Chart shows data from 4 independent experiments.

To further confirm the phenotypic changes following Twist1 decrease in cancer cells upon contact with activated fibroblasts, we used fully activated CAFs isolated from lung metastases in a coculture with mesenchymal cancer cells for 24h (Figure 5.3.A). Consistent with our previous observations, MICs in the presence of CAFs reduced Vimentin, AXL expression and increased E-cadherin levels (Figure 5.3.B). These results suggest that once fibroblasts change to an activated state they trigger a crosstalk of signals that can modulate cancer cells epithelial plasticity.

As the mesenchymal phenotype confers cancer cells with high invasion ability (section 1.2.1.1), to functionally validate these phenotypic changes upon coculture with CAFs we analysed the three-dimensional invasion modality of cancer cells as they first come into contact with CAFs. In collaboration with Danielle Park from the Sahai Lab we previously showed that in Matrigel:collage-I matrices MICs invade as single cells forming elongated protrusions (Figure 3.7.B), a hallmark of mesenchymal cells. Therefore, we repeated this assay in the presence or absence of CAFs to evaluate their effect on the mesenchymal-like motility of cancer cells. We freshly isolated MICs and nonMICs from primary tumours and made spheroids overnight in suspension. Next day, pure MICs and nonMICs spheroids were embedded into Matrigel:collagen-I gels with or without CAFs and their invasion was evaluated 48h after. As previously reported (Gaggioli et al., 2007), breast cancer cells in a more epithelial non-invasive state (nonMICs) activated both single and collective cell invasion when surrounded by CAFs (Figure 5.3.C – black bars). Strikingly, in line with the observed inhibition of mesenchymal markers MICs showed reduced single cell invasion in the presence of CAFs whereas collective invasion was still enhanced (Figure 5.3.C - blue bars). Consequently, in the presence of CAFs the overall invasion capacity of MICs and nonMICs spheroids results similar (Figure 5.3.D). These results confirm that activated fibroblasts can mitigate the mesenchymal features of cancer cells as they first come into contact with them.

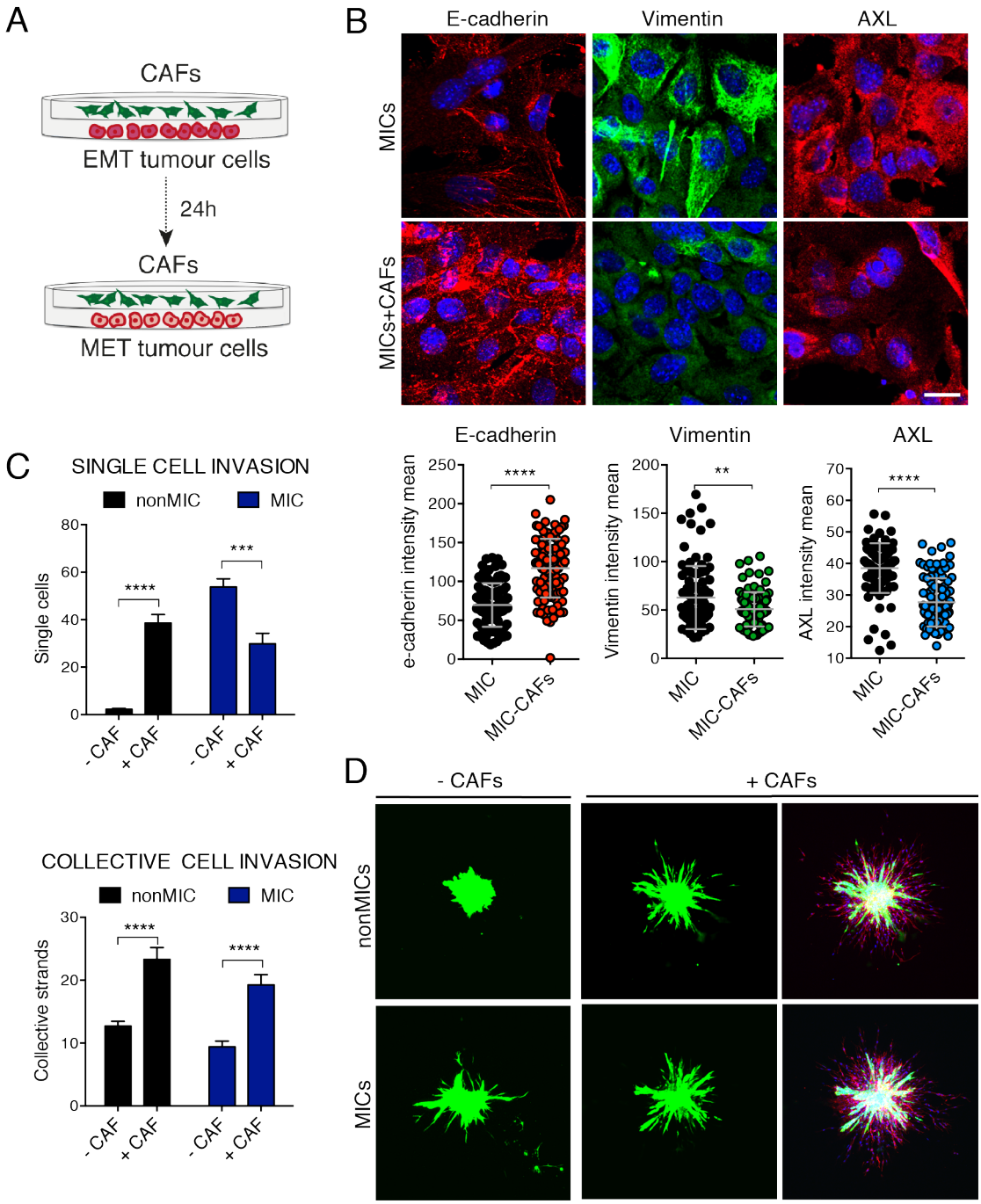


Figure 5.3 CAFs attenuate the mesenchymal status of MICs
(A) Schematic shows the coculture setting where tumours cells are seeded on collagen coated dishes, and CAFs are placed in the upper chamber of the coculture. After 24 hours tumour cells were monitored for changes in the expression of EMT-associated markers. (B) Representative images show the expression of E-cadherin, Vimentin and Axl in MICs and MICs cocultured with CAFs for 24 hours. Scale bar, 20 µm. (Lower panel) Charts indicate the expression levels of E-cadherin, AXL and Vimentin in MICs and MICs cocultured with CAFs for 24 hours. Data from one representative experiment of 2 are shown. Line and error bar indicate mean ± SD of the population. (C) Histograms show the quantification of

single (upper histogram) or collective (lower histogram) cell migration on spheroids made from FACS sorted PyMT/aGFP MICs or nonMICs, embedded in a Matrigel-Collagen matrices with or without CAFs. Data from one representative experiment of two are shown. (D) Representative fluorescent images of spheroids quantified in (C), PyMT/aGFP MICs or nonMICs spheroids (green), embedded in a Matrigel:Collagen-I matrices with or without mCherry labelled CAFs (red). Spheroids were allowed to invade for 48 hours before imaging.

In light of these last results, we next investigated the potential mechanism underlying these CAF-driven changes in cancer cells. As we discussed in the introduction, TGFβ signalling has been extensively implicated in metastasis driving the initial epithelial-to-mesenchymal switch at the invasive front (Heldin et al., 2012), and the maintenance of the EMT state while cells travel in the blood stream (Labelle et al., 2011) (Figure 1.6.B). Particularly during metastatic colonisation, high TGFβ signalling favours extravasation and early colonisation but a subsequent reduction is needed for metastatic outgrowth (Giampieri et al., 2009). Indeed, MICs undergo TGFβ signalling when isolated from late-stage tumours (Figure 3.3) coinciding with the display of mesenchymal traits (Figure 3.4). Also, it was previously reported that breast cancer cells upon infiltration to the distant site negatively modulate TGFβ signalling (Giampieri et al., 2009). Therefore, we hypothesised that the re-acquisition of epithelial characteristics could depend on this reverse modulation of TGFβ signalling upon arrival at the target site. In order to test this idea, we monitored the dynamics of the TGFβ effector SMAD2-3 in metastasising cells *in vivo*. We performed an ImageStream analysis of actinGFP-expressing $CD24^{+}AXL^{+}$ PyMT cells in the lungs two days (first phase) or twelve days (second phase) after intravenous injection (Figure 5.4.A). We found that as cells in an AXL-mesenchymal status transit from the first to the second phase of colonisation, the total number and staining intensity of $pSMAD2\text{-}3^{+}$ cells decreases (Figure 5.4.B-C), following the same overall dynamics than Twist1 (Figure 5.1.C-D). This result is in line with our hypothesis that an attenuation of TGFβ signalling could trigger mesenchymal inhibition during metastatic colonisation.

I have previously discussed in the introduction how the alternative branch of the TGFβ signalling pathway, i.e. canonical BMP signalling can antagonise TGFβ (section 1.2.1.1) (Wakefield and Hill, 2013). Therefore, we next investigated changes in BMP signalling as TGFβ is downregulated in metastasising cells. Indeed, we observed an increase in BMP activity during the second phase of colonisation. We monitored the expression of the immediate BMP effector pSMAD1-5 detecting an increase in the total number of pSMAD1-5^{+} cells in the second phase of colonisation (Figures 5.4.E). In contrast to pSMAD2-3 levels, pSMAD1-5 was maintained during colonisation in the overall tumour cell population in the lungs, inversely correlating with Twist1 dynamics (Figure 5.4.F-G). In line with this increase in BMP activity, we observed the maintenance of the canonical BMP/pSMAD1-5 target gene ID1 (Figure 5.4.H). Moreover, further analysis revealed an increase in the percentage of ID1^{+} cells specifically among those that lost Twist1 expression in the second phase of colonisation (Figure 5.4.I).

Collectively, these data show that the phenotypic changes towards a more epithelial state during the second phase of colonisation correlate with an inhibition of TGFβ signalling, which favours BMP-dependent metastatic outgrowth. Importantly, we also observe the maintenance of ID1 throughout colonisation, a downstream BMP effector previously reported to be required during metastatic outgrowth (Gupta et al., 2007).

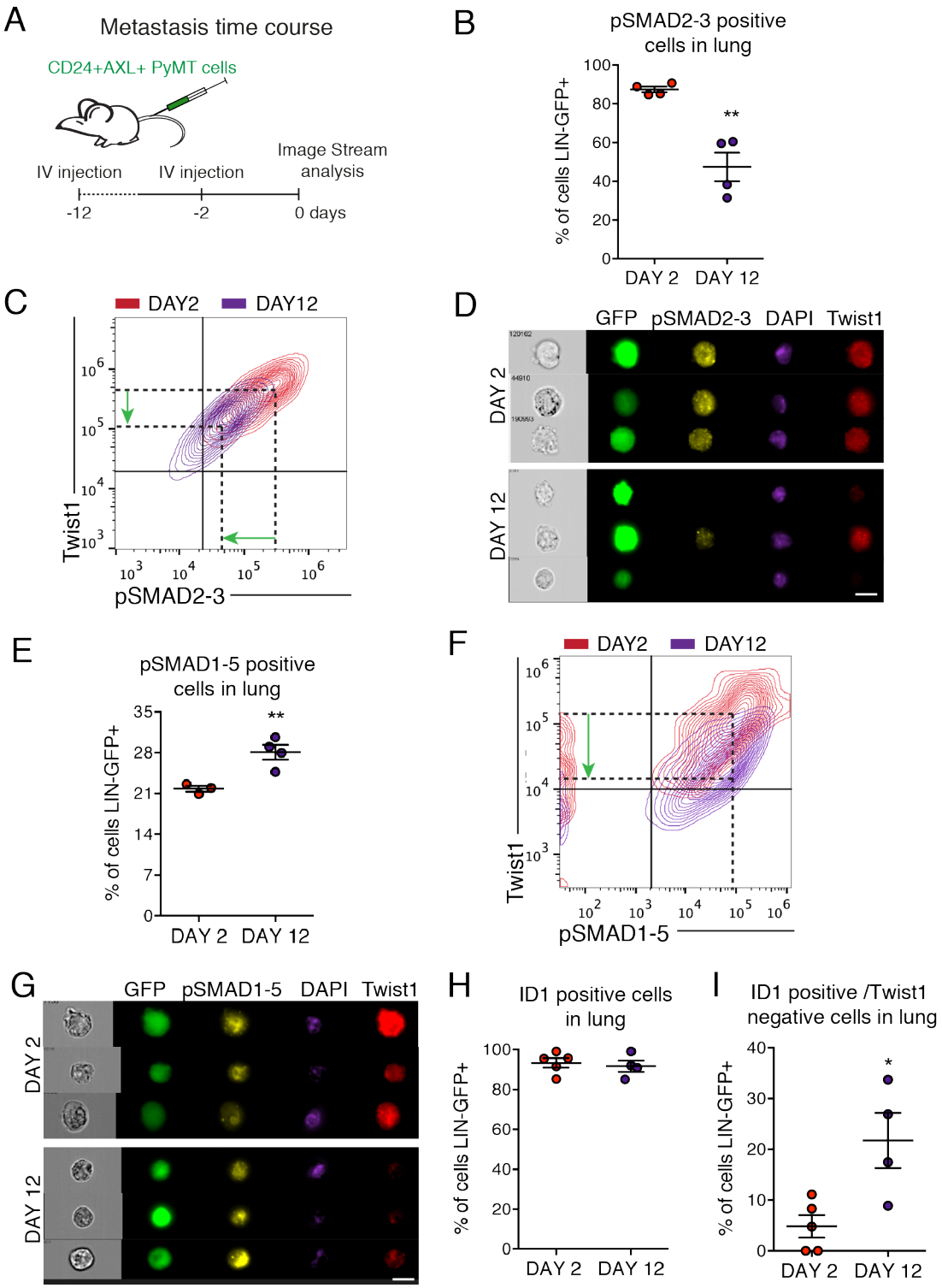


Figure 5.4 *In vivo* dynamics of pSMAD1-5 and pSMAD2-3 during metastatic colonisation

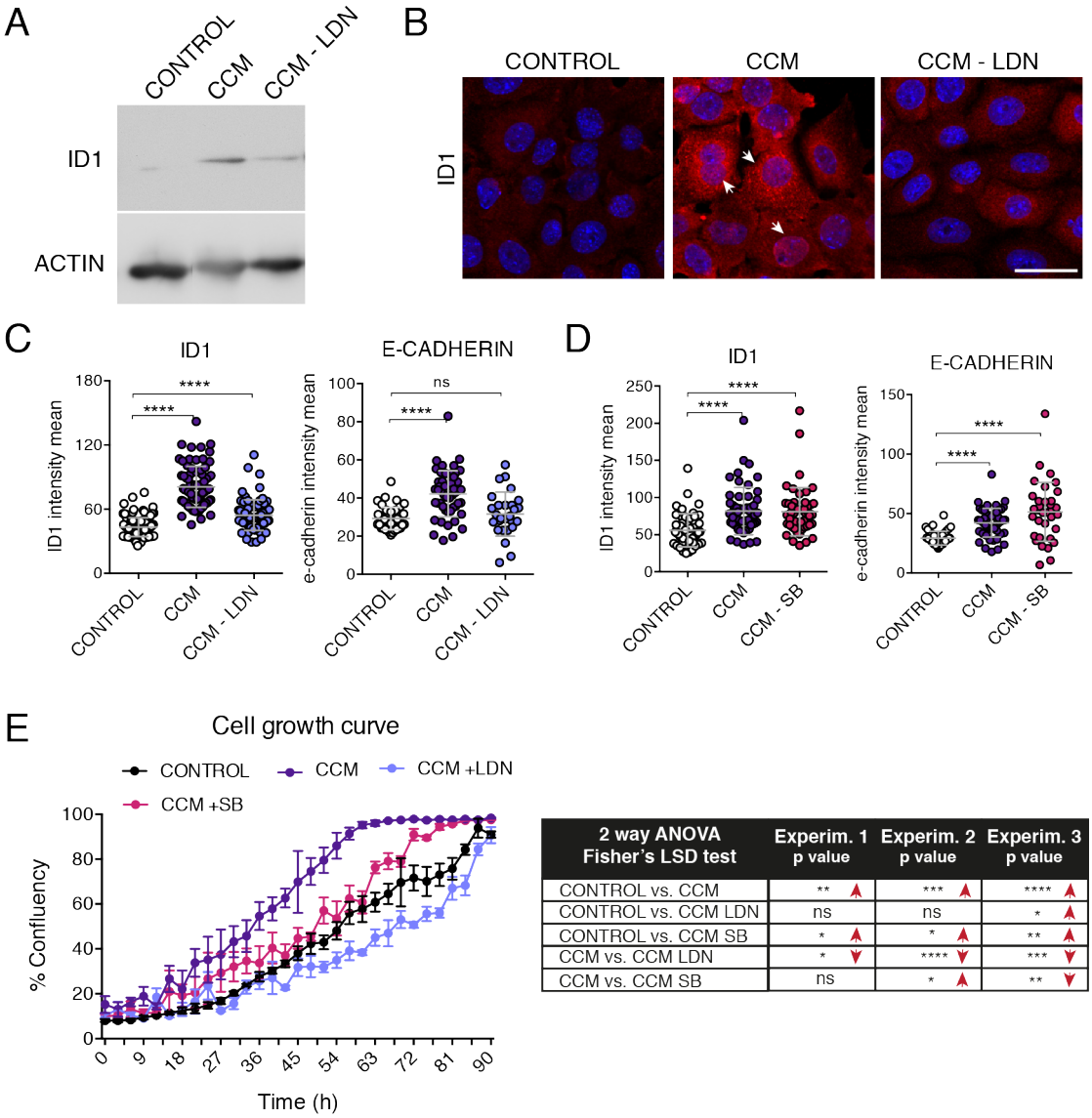
(A) Schematic shows the experimental set up for the metastasis time course. Mice were intravenously injected with primary actinGFP-PyMT $CD24^{+}AXL^{+}$ cells 12 or 2 days before analysis. At day 0, lungs were depleted from lineage cells ($CD45^{+}$ $CD31^{+}$ $Ter119^{+}$ cells) and then half of the lung was freshly analysed for AXL expression, and the other half fixed for intracellular stain and ImageStream

analysis. (B) Chart shows the *in vivo* evaluation of the percentage of pSMAD2-3$^+$ cells in the lung in the metastasis time course in (A). The percentage of LIN$^-$ GFP$^+$/pSMAD2-3$^+$ cells was evaluated (n=4 mice). (C) FACS contour plot show the intensity of Twist1 *versus* pSMAD2-3 in CD24$^+$AXL$^+$ tumour cells in the lung at 2 (red) and 12 (purple) days post metastatic seeding. Green arrows indicate the decrease in Twist1 and pSMAD2-3 expression. (D) Representative images show pSMAD2-3 and Twist1 dynamics in cells isolated from 2 or 12 days metastasis. (E) Chart shows the *in vivo* evaluation of the percentage of pSMAD1-5$^+$ cells in the lung in the metastasis time course in (A). The percentage of LIN$^-$GFP$^+$/pSMAD1-5$^+$ cells was evaluated (n=3-4 mice). (F) FACS contour plot show the intensity of Twist1 *versus* pSMAD1-5 in CD24$^+$AXL$^+$ tumour cells in the lung at 2 (red) and 12 (purple) days post metastatic seeding. The green arrow indicates the decrease in Twist1 expression. (G) Representative images show pSMAD1-5 and Twist1 dynamics in cells isolated from 2 or 12 days metastasis. (H-I) Charts show the *in vivo* evaluation of the percentage of ID1$^+$ cells in the lung in the metastasis time course in (A). The percentage of LIN$^-$GFP$^+$/ID1$^+$ cells (H) and LIN$^-$GFP$^+$Twist1$^-$/ID1$^+$ cells (I) was evaluated (n=4-5 mice).

We have shown that one of the main microenvironmental differences between the first and second phase of colonisation is the presence of the newly activated fibroblasts, which coincides with these TGFβ/BMP modulations observed in tumour cells as colonisation progresses (Figure 5.1.A-B and 5.4.B-I). Therefore, we next analysed the changes that CAFs induce in cancer cells along with the AXL/Twist1 mesenchymal modulation previously shown and how they relate to the *in vivo* situation (Figure 5.2.A-D). CAF-conditioned media (CCM) on primary CD24$^+$AXL$^+$ mesenchymal cells for 24h was enough to induce ID1 expression in cancer cells as measured by western blot (Figure 5.5.A). We confirmed this increase in ID1 expression upon CCM exposure by immunofluorescence, where we could also observed its nuclear localisation (Figure 5.5.B). To assess whether ID1 increase depends on a canonical BMP/pSMAD1-5 signalling as suggested by the *in vivo* data we used the specific BMP inhibitor LDN193189 (LDN) in the CCM. LDN blocks the BMP specific receptors ALK2 and ALK3 preventing SMAD1-5 phosphorylation. Accordingly, when added to the CCM LDN prevented the increased ID1 expression in primary PyMT CD24$^+$AXL$^+$ cells (Figure 5.5.A-B). This result suggests that ID1 induction can be partially explained via BMP activation.

CAFs are known to secrete both TGFβ and BMP ligands (Kalluri and Zeisberg, 2006), and ID1 activity can be transiently induced as well via TGFβ signalling (Kang et al., 2003). Therefore, we next tested the potential contribution of TGFβ to ID1 activation. For this purpose, we used the specific TGFβ inhibitor SB431542 (SB) that blocks the TGFβ receptor ALK5 preventing SMAD2-3 phosphorylation. We tested in parallel both the BMP and the TGFβ inhibitors added to the CCM. Blocking BMP signalling prevented the increase in ID1 and E-cadherin levels induced by the CCM alone (Figure 5.5.C), whereas the TGFβ inhibitor did not have any significant effect on ID1 or E-cadherin levels (Figure 5.5.D). These results suggest that the observed CAF-mediated ID1 induction mainly relies on BMP/pSMAD1-5 activity.

Since ID1 is required for breast cancer cell proliferation during metastatic outgrowth (Gupta et al., 2007), we next tested whether this CAF-mediated ID1 induction would functionally correlate with cell proliferation *ex vivo*. We used the CCM with or without the BMP and TGFβ inhibitors on primary $CD24^{+}AXL^{+}$ mesenchymal cells and monitored cell growth every 3h over 3-4 days. CCM dramatically boosted the proliferation rate of cancer cells (Figure 5.5.E). Note that although there is variability in the growth displayed by the different tumour cell preparations (table in Figure 5.5.E), all experiments showed a consistent increase in tumour cell proliferation upon CCM treatment. Furthermore, in line with ID1 expression, inhibiting BMP/pSMAD1-5 activity reliably decreased the proliferative effects of the CCM, while inhibiting TGFβ/pSMAD2-3 activity did not show a consistent read-out in cell proliferation (Figure 5.5.E). The results shown in the table also indirectly indicate, when comparing the un-stimulated CONTROL *versus* CCM with inhibitors that the negative effect on the proliferation of inhibiting BMP/pSMAD1-5 activity is greater than when inhibiting TGFβ/pSMAD2-3. This suggests an endogenous BMP signalling activity in PyMT cells in basal state sustaining proliferation, as recently reported (Owens et al., 2015).



2 way ANOVA Fisher's LSD test	Experim. 1 p value	Experim. 2 p value	Experim. 3 p value
CONTROL vs. CCM	** ↑	*** ↑	**** ↑
CONTROL vs. CCM LDN	ns	ns	* ↑
CONTROL vs. CCM SB	* ↑	* ↑	** ↑
CCM vs. CCM LDN	* ↓	**** ↓	*** ↓
CCM vs. CCM SB	ns	* ↑	** ↓

Figure 5.5 CAFs induce ID1 expression and cell proliferation in mesenchymal cancer cells *in vitro*

(A) Western blot shows the expression levels of ID1 in primary $CD24^{+}AXL^{+}$ tumour cells plated 2D on collagen coated dishes for 24h with normal media, CAF-conditioned media (CCM) or CCM with 1μm LDN. Actin used as loading control. Data from one representative experiment of two are shown. (B) Representative images show ID1 expression levels in primary $CD24^{+}AXL^{+}$ cells cultured 2D on collagen coated dishes for 24h with normal media, CAF-conditioned media (CCM) or CCM with 1μm LDN. Scale bar, 40 μm. (C) Charts show the immunofluorescence quantification of E-cadherin and ID1 expression levels in primary $CD24^{+}AXL^{+}$ cells cultured 2D on collagen coated dishes for 24h with normal media, CAF-conditioned media (CCM) or CCM with 1μm LDN. Data from one representative experiment out of two are shown. (D) Charts show the immunofluorescence quantification of E-cadherin and ID1 expression levels in primary $CD24^{+}AXL^{+}$ cells cultured 2D on collagen coated dishes for 24h with normal media, CAF-conditioned media (CCM) or CCM with 10μm SB. (E) Graph

shows the growth curve generated by primary $CD24^{+}AXL^{+}$ tumour cells monitor every 3 hours for 90h in normal media (CONTROL), CAF-conditioned media (CCM) and CCM with 1µm LDN or 10 µm SB. Table summarises the effects on cell proliferation in 3 independent experiments using 3 different PyMT primary tumours. The symbols represent the p value generated by ANOVA for each comparison, and the red arrows indicate a significant increase (up) or decrease (down).

In summary, these results suggest that secreted factors from activated fibroblasts increase the active BMP/pSMAD1-5 signalling in mesenchymal cancer cells, inducing *de novo* ID1 expression and a more epithelial/proliferative phenotype.

Finally, we wanted to functionally validate the requirement of pSMAD1-5/ID1 activity during the second phase of metastatic colonisation in our model. To this end, we used the specific BMP inhibitor *in vivo* to block pSMAD1-5/ID1 activity while cells transition from the first to the second phase of colonisation. We treated mice from day 5 to 12 post tail vein injection (Figure 5.6.A), during the time window when Twist1 and AXL are downregulated and BMP becomes the dominant signalling observed in metastasising cancer cells (Figure 5.1.B-E). Remarkably, pSMAD1-5 activity is functionally relevant during the second phase of colonisation as its inhibition reduces metastatic outcome (Figure 5.6.B).

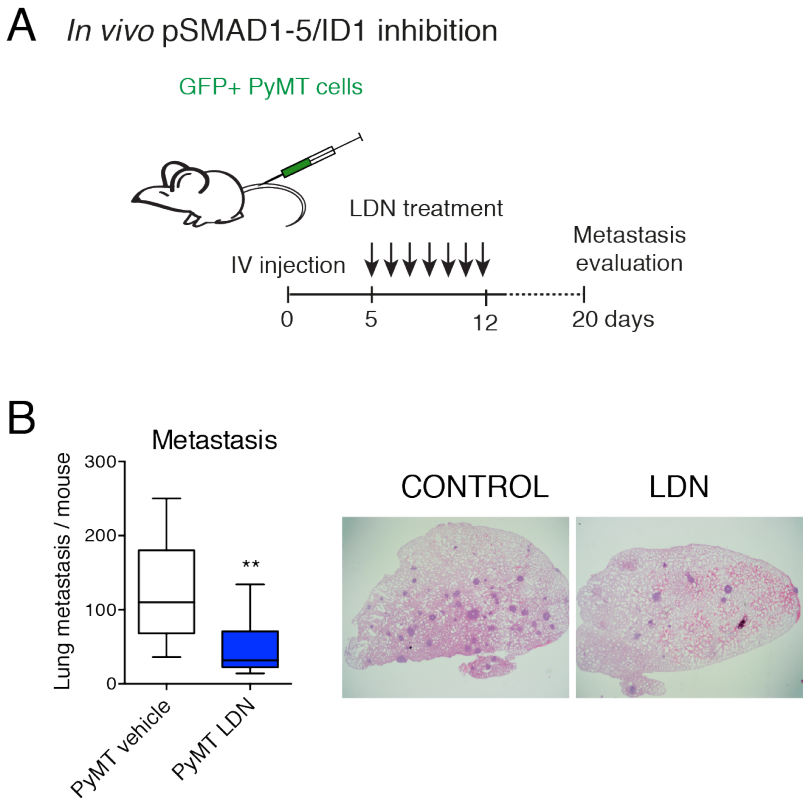


Figure 5.6 BMP signalling is required for metastatic outgrowth
(A) Scheme shows the *in vivo* LDN treatment during the second phase of colonisation. Mice were injected with 10^6 PyMT cells, and treated with LDN193189 (35 mg/kg) twice a day from day 5 to 12 post intravenous injection. Metastases were evaluated 20 days post-injection. (B) (Left) Box plot displays the metastatic burden of PyMT cells in mice treated with vehicle or LDN as specified in (A). Superficial lung metastases were evaluated. (Right) Representative metastatic lung sections of control *versus* LDN treated mice (n=9-12 mice per group).

Collectively, these data support our hypothesis that the phenotypic changes in mesenchymal cancer cells observed *in vivo* during the second colonisation phase, where they switch towards a more epithelial AXL$^-$/Twist1$^-$ state, are the consequence of an inhibition of TGFβ signalling that favours the activation of the BMP branch of the pathway. Subsequently, the increased BMP/pSMAD1-5 activity sustains metastatic outgrowth. Importantly, the secreted factors from activated fibroblasts can induce *ex vivo* the ID1-driven phenotypic changes in cancer cells, switching them towards a more epithelial/proliferative phenotype.

5.2 The second phase of colonisation needs to be temporally controlled for efficient metastasis

We have previously shown that metastatic cells initially require the mesenchymal features to induce their niche as they extravasate to the distant site. However, these mesenchymal features are lost during the second phase of colonisation. These data together with the previous evidence of a mesenchymal-to-epithelial transition during metastatic colonisation (Ocana et al., 2012, Tsai et al., 2012) suggest that these epithelial modulations might be temporally controlled during colonisation. According to this model, a temporal regulation of epithelial plasticity would be critical for metastatic outgrowth and the two events of niche activation and epithelial reversion would be dependent on each other. To validate this hypothesis, we tested whether interfering with the AXL-mesenchymal status of cancer cells would impact on their metastatic capacity.

First, we used the AXL inhibitor R428 *in vivo* during the second phase of colonisation when AXL is transitioning from high to low levels (Figure 4.8.A and 4.9.A) to exogenously trigger a faster epithelial reversion. We injected primary PyMT cells and MDA231 cells into mice and treated them with R428 from day 10 to 15 post-injection (Figure 5.7.A). We previously showed that AXL inhibition during the first week of metastasis reduces the metastatic potential of both PyMT and MDA231 cells (Figure 4.8.E and 4.9.B). Interestingly, in this second phase of colonisation the inhibitor no longer blocked the metastatic establishment of MDA231 cells (Figure 5.7.B) and significantly increased the metastatic burden of primary PyMT cells (Figure 5.7.C). Notably, no effect in the number of metastatic nodules was observed upon AXL inhibition at later stages of metastasis (day 25 – 30 post-injection) when metastatic nodules are fully developed (Figure 5.7.C). These results, in line with the previous studies by Ocana et al. and Tsai et al., suggest that facilitating mesenchymal reversion at the target site enhances metastatic outgrowth.

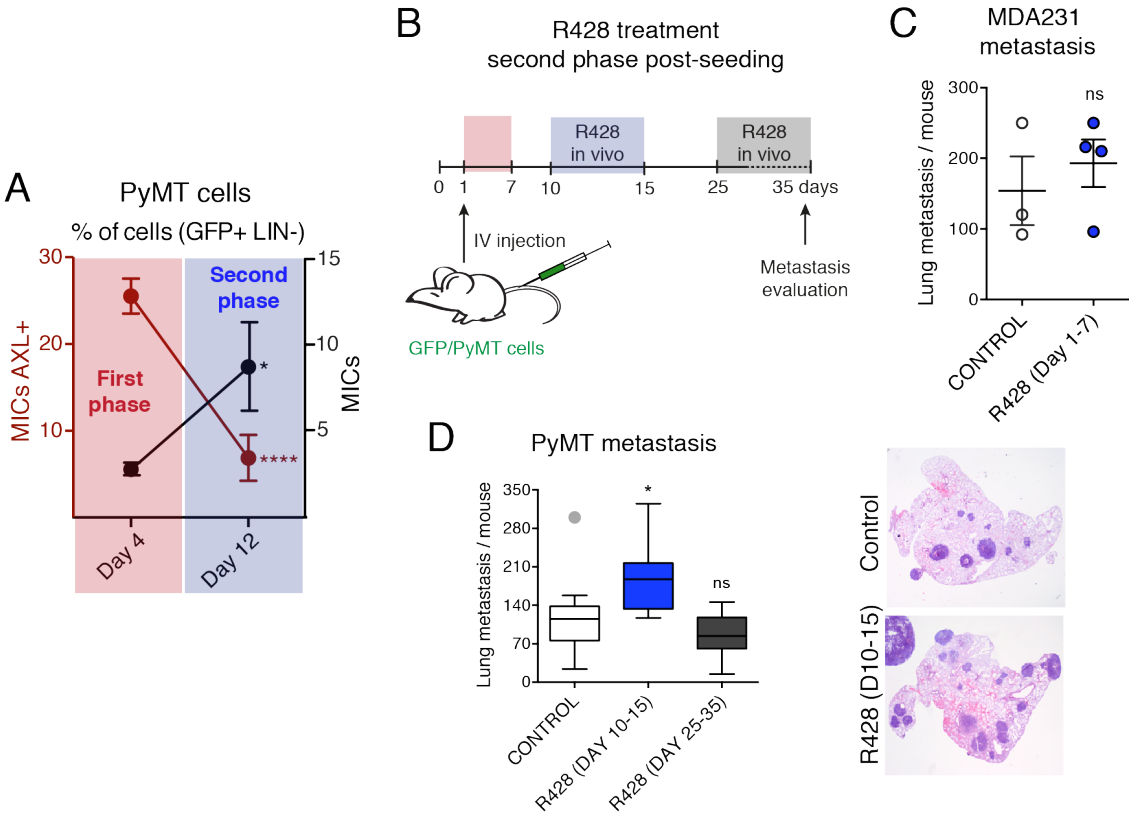


Figure 5.7 AXL inhibition during late metastasis
(A) FACS analysis of the percentage of AXL$^+$ MICs (red-left Y axis) and total MICs (black-right Y axis) among total GFP$^+$ cancer cells in the lung at 4 and 12 days after tail vein injection. Graph represents mean ± sem per time point (n=8-9 mice per group of 2 independent experiments). (B) Experimental setup: Mice injected with 10^6 PyMT or MDA231 cells were treated daily with 12.5 mg/kg of R428 from day 10 to 15 (blue) or from day 25 to 35 (grey), and metastases were evaluated around day 35 post-injection. (C) Box plot displays the metastatic burden of MDA231 cells treated from day 10 to 15 post injection as indicated in (B). MDA231-DMSO, n= 3 mice per group; MDA231-R428 (10-15 days), n= 4 mice per group. Superficial lung metastasis number was evaluated. (D) Box plot displays the metastatic burden of PyMT cells intravenously injected into mice as specified in (B). PyMT-DMSO, n= 11 mice per group; PyMT-R428 (10-15 days), n= 10 mice per group; PyMT-R428 (25-35 days), n= 7 mice per group. Superficial lung metastasis number was evaluated. (Right) Representative H&E staining of metastatic lung sections from mice of indicated groups.

Second, we engineered a lentiviral reporter construct containing the open reading frame (ORF) of the human AXL gene in order to exogenously express AXL in cancer cells delaying their mesenchymal reversion. To avoid the potential confounding effects on cell proliferation that high levels of AXL are reported to cause in other cancer types (Ben-Batalla et al., 2013, Ammoun et al., 2014, Paccez et al., 2013) we used a strategy where the LTR regions of the lentiviral construct controls AXL expression (LTR-AXL) (Figure 5.8.A). This leads to a modest overexpression of the receptor in both PyMT and MDA231 cells (Figure 5.8.B). As expected, AXL overexpression led to an exacerbated mesenchymal phenotype with increased Vimentin expression in the MDA231 cell line (Figure 5.8.C), and increased Vimentin and reduced E-cadherin levels in primary PyMT cells (Figure 5.8.D-E).

Next, we addressed the effect of overexpressing AXL *in vivo* as cells extravasated into the lungs. We used the MDA231 cell line expressing GFP or the LTRAXL constructs monitored their AXL expression levels as they grew in the lungs for a 24h period early after extravasation (from day 3 to 4 post-injection) (Figure 5.9.A). We could observe that AXL overexpression leads to a delay in the ability of the tumour cells to induce its early downregulation preventing their accumulation in the lungs. While the control MDA231-GFP cells reduced to half AXL expression and doubled in number, the MDA231-LTRAXL cells maintained AXL expression levels and did not grow in the lungs (Figure 5.9.B). This result indicates that exogenous AXL expression delays tumour cell growth in the early phase of colonisation. We next excluded that this delayed growth *in vivo* was due to a defect in proliferation in the AXL-overexpressing cells. We performed a BrdU chase experiment *in vitro* and did not detect any difference in the BrdU incorporation rate of MDA231-GFP and MDA231-LTRAXL over 3 hours (Figure 5.9.C) excluding a direct negative impact on tumour cell proliferation. Thus, we next injected both PyMT and MDA231 cells expressing GFP or LTRAXL into the lungs and assessed metastatic outcome 20-30 days post-seeding. In both models the exogenous expression of AXL impaired metastatic outgrowth (Figure 5.9.D-E). This result suggests that for efficient metastasis epithelial reversion needs to occur in a temporally controlled manner during the second phase of metastatic colonisation.

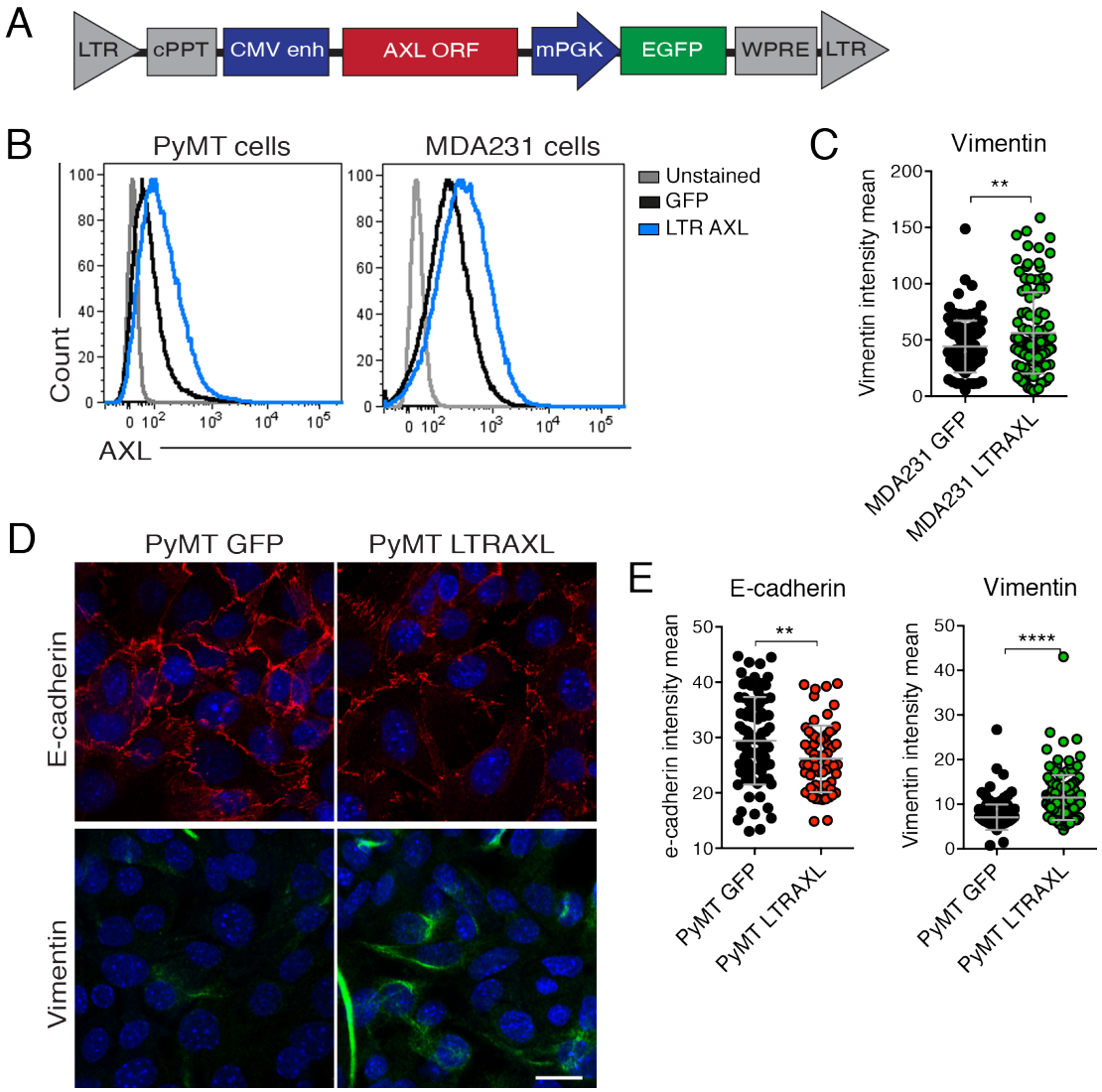


Figure 5.8 LTR-driven AXL overexpression exacerbates the mesenchymal phenotype of cancer cells
(A) Lentiviral construct expressing the human AXL coding sequence under the control of the viral LTR regions (LTRAXL). An independent human PGK promoter controls EGFP reporter gene expression. (B) FACS histogram displays AXL expression levels in PyMT and MDA231 cells expressing GFP (black) or LTRAXL (blue) lentiviral constructs. (C) Chart indicates the expression levels of Vimentin in primary PyMT cells expressing GFP or LTRAXL lentiviral construct assessed by immunofluorescence. (D) Representative images and (E) charts indicate the expression levels of E-cadherin and Vimentin in primary PyMT cells expressing GFP or LTRAXL lentiviral construct. Chart line and error bars indicate mean ± SD of the population. Scale bar, 20 µm.

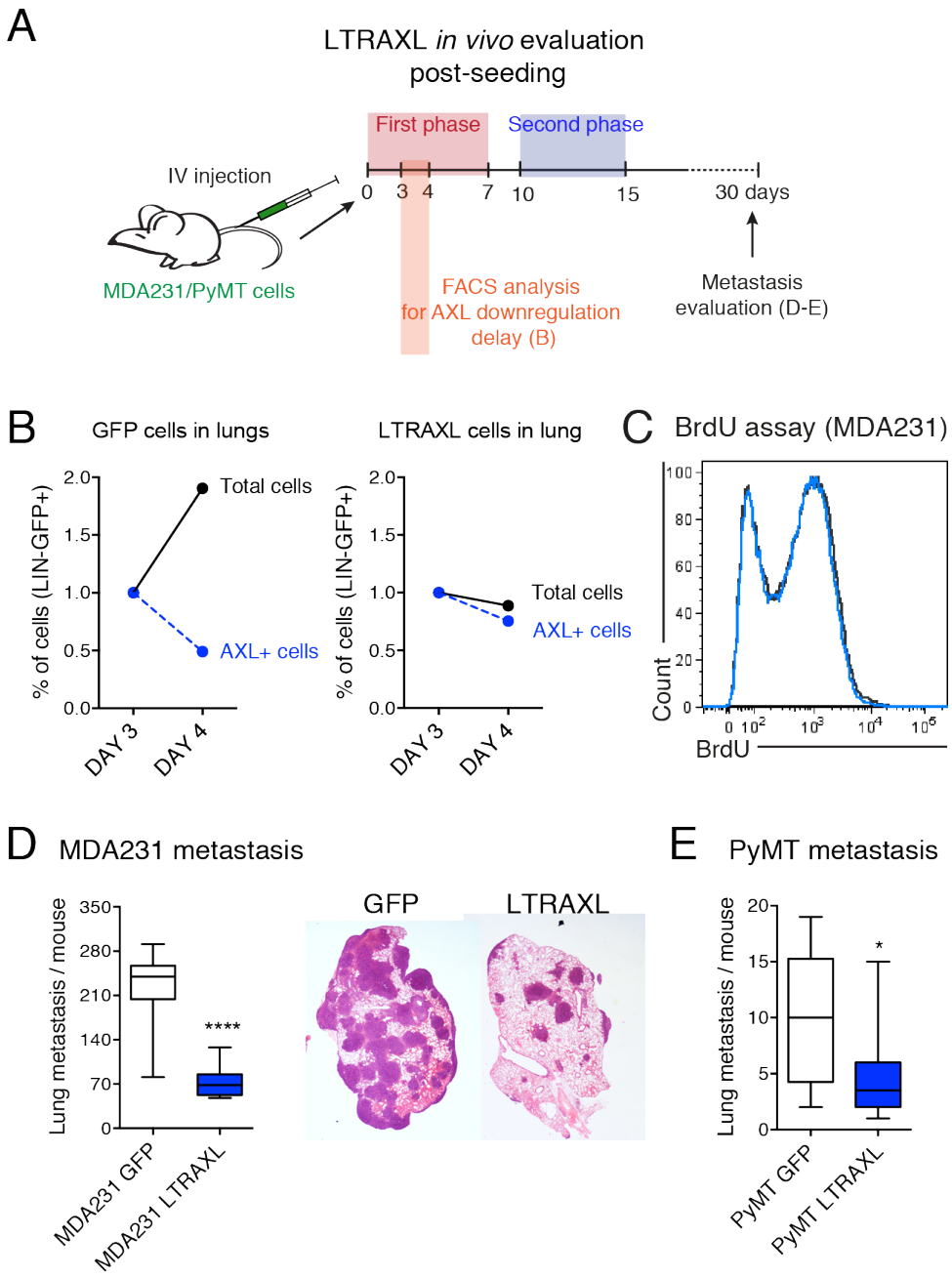


Figure 5.9 AXL exogenous expression effects on cancer cell proliferation and metastatic colonisation

(A) Experimental setup: Mice were injected with 10^6 PyMT or MDA231 expressing GFP or LTRAXL lentiviral constructs. We evaluated the effects of AXL overexpression during the early phase of metastatic colonisation over 24h, from day 3 to 4 (brown window). Metastatic outcome was evaluated 30 days post-injection. (B) FACs analysis of the total cell growth and *in vivo* dynamics of AXL in MDA231 cells expressing GFP (left panel) or LTRAXL (right panel), from day 3 to 4 of the metastatic time course shown in (A). Dot plot represents the average of the percentage of positive cells in the lung: Total cells (measure as GFP^+ cells) and AXL-expressing cells (measure as AXL^+ cells) (n=3 mice per group). (C) FACS histogram plot shows BrdU incorporation over a 3 hours pulse by MDA231 cells

expressing GFP (black) or LTRAXL (blue) lentiviral constructs. (D) Box plot displays the metastatic potential of MDA231 cells expressing GFP or LTRAXL lentiviral constructs, as specified in (A) (n=8 mice per group). (Left) Representative images of H&E stain show MDA231 metastatic nodules in the lung of the indicated groups. (E) Box plot displays the metastatic potential of PyMT cells expressing GFP or LTRAXL lentiviral constructs, as specified in (A) (n=10-12 mice per group).

To further characterise the role of AXL in tumorigenesis, we evaluated the effects of overexpressing AXL on the intrinsic cancer cells abilities. First, we tested the ability of PyMT primary LTRAXL cells to form spheres to assess their self-renewal capacity *in vitro*. We observed that AXL overexpression negatively affects the ability of PyMT cells to grow in suspension (Figure 5.10.A). Second, we performed the gold standard stemness assay *in vivo* analysing their ability to initiate primary tumours. Interestingly, LTRAXL cells displayed an enhanced ability to initiate primary tumours in both models, PyMT and MDA231 (Figure 5.10.B-C) in contrast to their reduced ability to form spheres (Figure 5.10.A). Overall, the enhanced ability to grow primary tumours *in vivo* excludes a direct negative impact on the intrinsic stemness abilities of AXL overexpressing cells. Next, we performed some histological analysis on these fully established tumours and found no difference in proliferation as measured by Ki67 expression (Figure 5.10.D-E). Also, the overall tumour morphology revealed by the H&E stain and the tumour vasculature (endomucin stain) was unaltered in LTRAXL tumours (Figure 5.10.D-E).

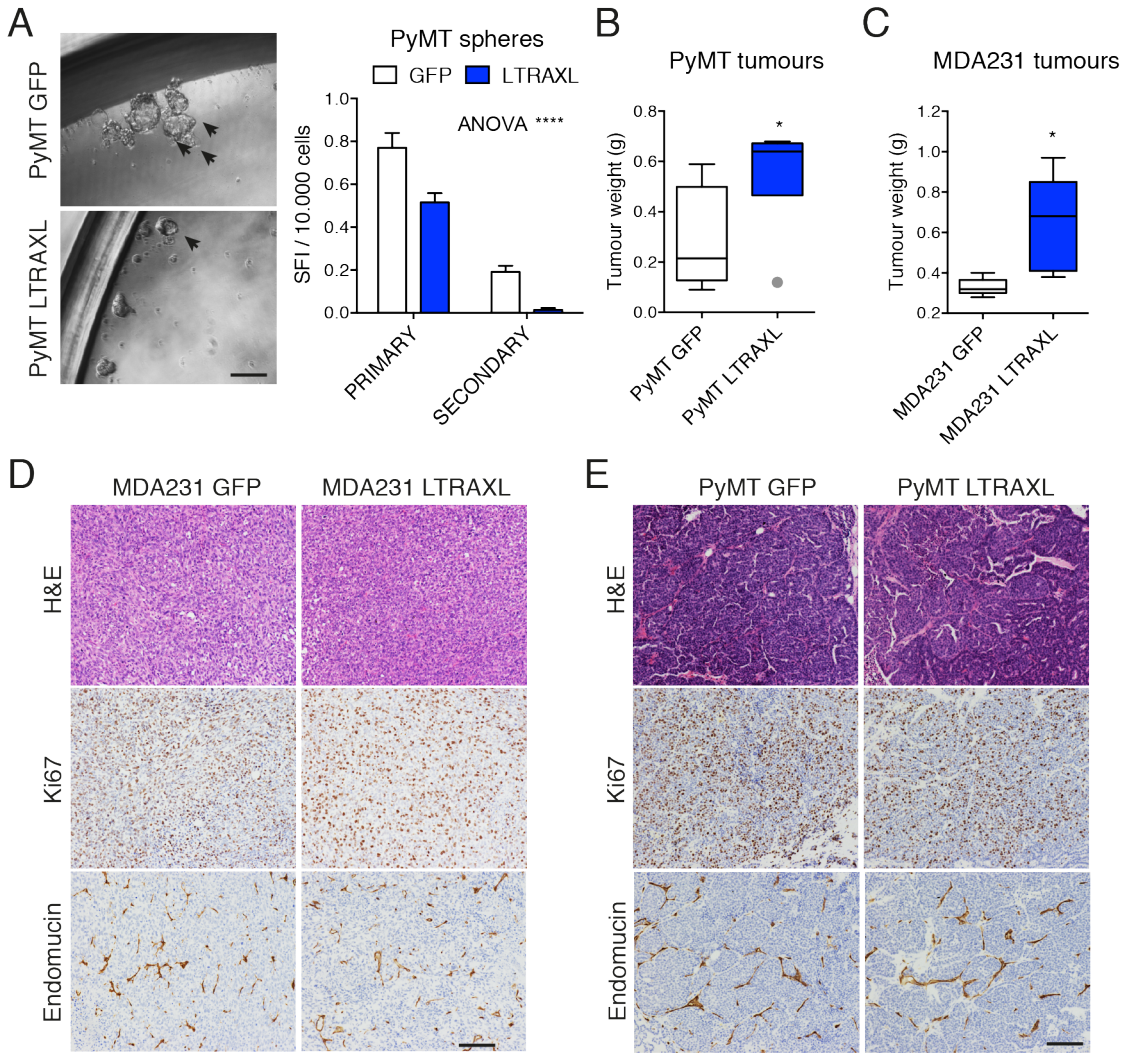


Figure 5.10 Effects of overexpressing AXL on self-renewal and tumour growth

(A) (Left) Images show representative pictures of spheres formed by PyMT cells expressing GFP and LTRAXL lentiviral constructs in suspension. Scale bar, 200 µm. (Right) Histogram shows the mean ± sem of 2 independent experiments. Spheres were quantified 7 days after plating a single PyMT suspension (primary) and 7 days after the first passage (secondary). (B) Box plots display tumour burden assessed 2 weeks after engraftment of PyMT cells expressing GFP or LTR-AXL (n=6 per group). (C) Box plots display tumour burden assessed 2 weeks after engraftment of MDA231 cells expressing GFP or LTR-AXL (n=6 per group). (D) Histological analysis of MDA231 tumours, control and LTRAXL, stain for H&E, Ki67 (brown signal) and Endomucin (brown signal). Scale bar, 100 µm. (E) Histological analysis of PyMT tumours, control and LTRAXL, stain for H&E, Ki67 (brown signal) and Endomucin (brown signal). Scale bar, 100 µm.

These results suggest that AXL displays differential roles during primary tumour initiation and development, and metastatic colonisation, as previously observed with the secreted ECM glycoprotein THBS2.

5.3 AXL loss is a transient effect required during colonisation

We have shown in the previous sections evidence for a biphasic model governing early metastatic colonisation of breast cancer cells to the lungs. Our results support the idea that after extravasation into the target site, tumour cells maintain their mesenchymal status displaying high levels of AXL, Twist1 and the niche activation effector THBS2 (Figures 5.1 and 4.3). Subsequently, following the appearance of the induced niche, activated fibroblasts will trigger the inhibition of cancer cells mesenchymal features favouring cell growth (Figures 5.3 and 5.4). Once metastatic colonisation is completed, tumour cells are re-wired to grow a tumour cell mass that will recapitulate a similar structure to the original primary tumour at a distant location. During this outgrowth phase, new stroma recruitment and stromal-cancer cell interactions will take place in the context of a continuously evolving microenvironment (section 1.3.2). Therefore, we next asked whether in late-stage metastases, when metastatic nodules have successfully outgrown, the original AXL-mesenchymal features displayed by MICs at the primary site would be re-acquired. In order to test this idea, we performed a longer metastasis time course experiment where we monitored AXL expression on total cancer cells and MICs in the lungs after intravenous injection during early colonisation (4 days post-seeding), late colonisation (12 days post-seeding) and metastasis (30 days post-seeding). As previously observed, AXL expression in MICs decreases as the pool expands in the lungs in the transition from early to late colonisation (Figure 5.11.A). Remarkably, in late-stage metastases (30 days post injection) the subpool regains expression of AXL as the number of MICs goes down to their normal frequency in the tumour mass (around 2-3%) (Figure 5.11.A). In contrast, the overall levels of AXL in total tumour cells in the lung did not change significantly throughout the time course, being the percentage of total AXL^+ PyMT cells in the lungs around 10-15% (Figure 5.11.A). These results indicate that MICs in late-stage metastases regain the mesenchymal characteristics displayed at the primary site when they disseminate to colonise distant tissues with a highly secretory mesenchymal phenotype.

Additionally, the analysis of AXL expression in total tumour cells in the lungs shows that AXL dynamics during metastasis are confined to the MIC subpool, highlighting the plasticity of this population of metastasis-initiating cells (Figure 5.11.A).

We previously showed that inhibiting AXL with R428 during this late metastatic stage (day 25-35 post intravenous injection) did not affect the total number of lung metastases (Figure 5.7.D). Further histological analysis revealed a reduction in the size of the metastatic nodules after a 10 days treatment of fully established metastases with the inhibitor (Figure 5.11.B).

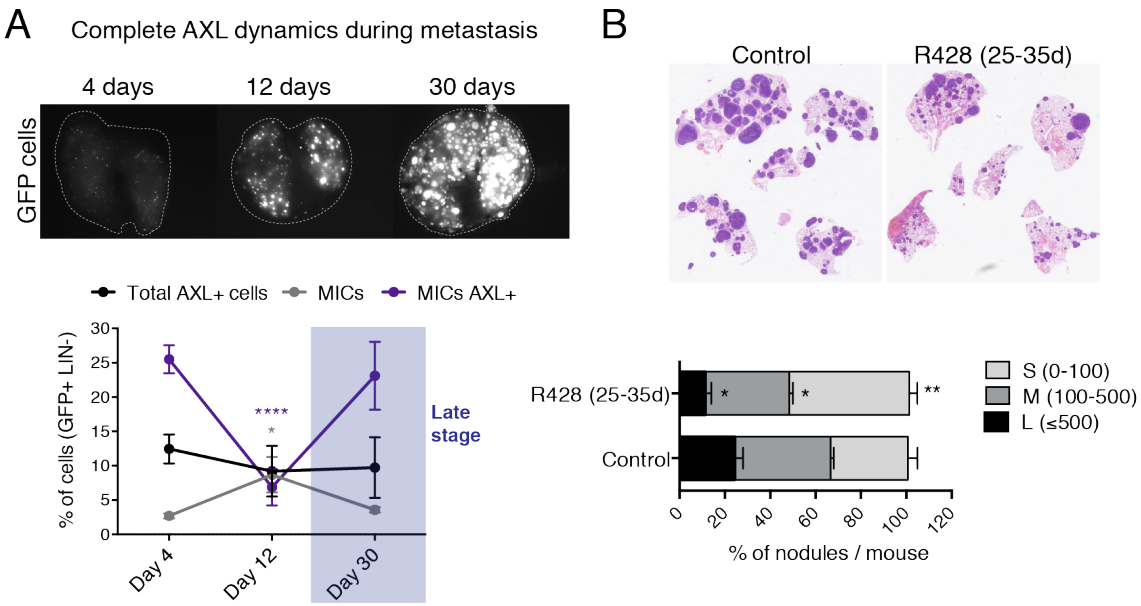


Figure 5.11 Complete AXL dynamics during metastasis and effects of AXL inhibition in fully established nodules
(A) Complete AXL dynamics time course during metastatic progression. Representative images show the amount and size of metastasis in the lung 4 days (first colonisation phase), 12 days (second colonisation phase) and 30 days (late stage metastasis) post-intravenous injection of 10^6 actinGFP/PyMT cells. Graph shows the dynamics in the percentage of total AXL^+ tumour cells (black), the percentage of total MICs (grey) and the percentage of AXL^+ MICs (purple) in the lungs at 4, 12 and 30 days port-seeding measured by FACS analysis (n=8-9 mice per group). (B) AXL inhibition during late stage metastasis. Mice were injected intravenously with 10^6 actinGFP/PyMT cells, and treated with DMSO or 12.5 mg/kg of R428 from day 25 to 35 post-seeding. Metastasis were analysed the last day of treatment. Box plot shows the evaluation of superficial lung metastasis in mice treated with DMSO or R428 (25-35 days). Representative pictures show an H&E staining in whole lung scans of the indicated groups. Bar graph displays the size distribution of the metastatic nodules in mice treated with DMSO or R428 (25-35 days). The area of each metastatic nodule in the lung sections was evaluated using ImageJ and classified attending to their size into the indicated 3 categories.

This result indicates that in late-stage metastasis, when AXL is re-expressed in the MIC subpool, its inhibition can limit metastatic cell growth.

Chapter 6. Discussion

The major goal of this work was to determine what intrinsic features of the previously described metastasis-initiating cells (MICs) in the MMTV-PyMT model (Malanchi et al., 2012) drive breast cancer metastatic colonisation to the lungs.

Collectively our results provide a refined model of distant metastatic colonisation when metastatic cells display a mesenchymal state. We found MICs to display a highly secretory mesenchymal state characterised by the expression of AXL when isolated from invasive carcinomas. Notably, this AXL-mesenchymal status is crucially maintained upon extravasation to the lungs in an experimental-metastasis setting. Dependent on the mesenchymal status of cancer cells, we propose a biphasic model of distant colonisation (Figure 6.1). In the first phase, the AXL-mesenchymal status of cancer cells provides them the crucial advantage of inducing niche activation. Particularly, MICs trigger a novel cancer cell-stromal crosstalk mediated by THBS2 secretion that enhances fibroblast activation at the target site contributing to MICs metastasis initiating advantage. Moreover, our work provides evidence of this cancer cell-stroma crosstalk ultimately governing epithelial plasticity during the second phase of colonisation, when the activated niche feeds back to tune cancer cells epithelial phenotype in a temporally controlled manner. We uncover that the fibroblasts of the newly activated niche trigger a crosstalk of signals that attenuates TGFβ signalling and the mesenchymal status (AXL/Twist1 loss) of cancer cells, driving a BMP-dependent re-epithelialisation.

Importantly, using a combination of *in vivo* primary tumour initiation assays and *ex vivo* sphere formation assays we show that MICs maintain their intrinsic stem-like features throughout epithelial plasticity. These results suggest that the activated niche modulates the epithelial plasticity of metastasis-initiating cancer cells leaving unaltered their intrinsic stemness properties.

Finally, the temporally controlled AXL downregulation that allows metastatic outgrowth in the lungs during late colonisation is a transient event. In fully

developed late-stage metastases the initially high AXL expression in MICs is restored.

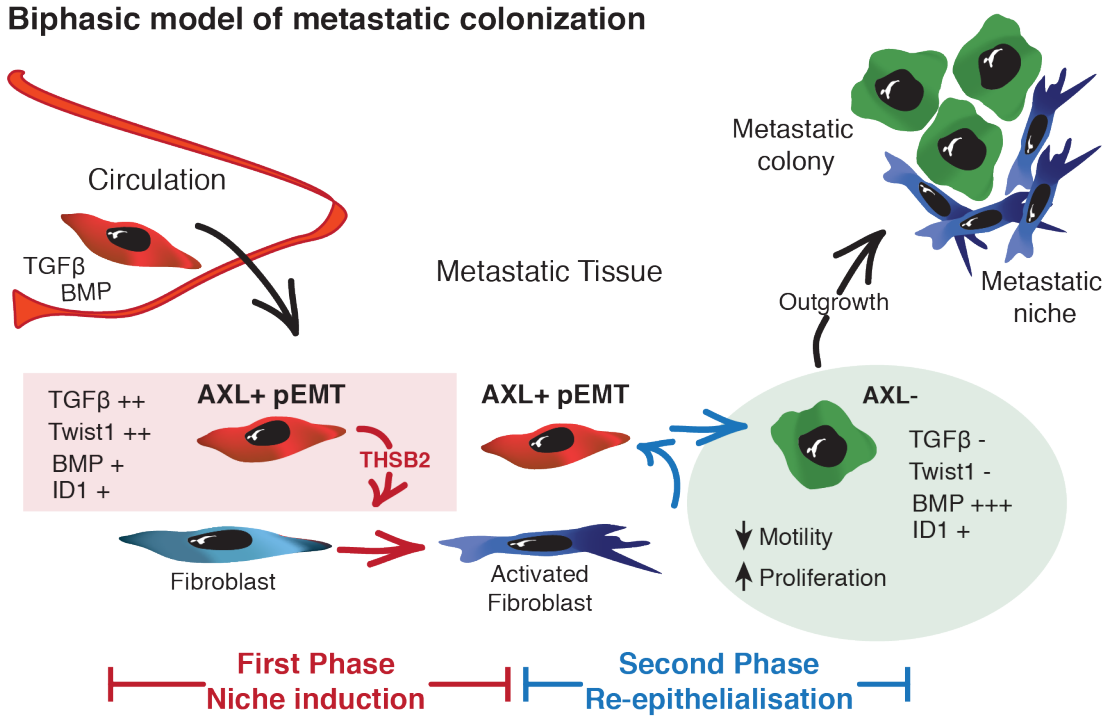


Figure 6.1 Biphasic model of breast to lung metastatic colonisation
In the MMTV-PyMT model metastatic colonisation is divided into two temporally distinct phases defined by the AXL-mesenchymal status of cancer cells. During the first phase upon extravasation MICs display a partial mesenchymal state characterised by high Twist1 and AXL expression, TGFβ signalling (pSMAD2-3) and BMP signalling (pSMAD1-5). Dependent on this mesenchymal status, MICs secrete THBS2 that triggers lung fibroblast activation. In the second phase, the newly activated fibroblasts trigger a crosstalk of signals that mediate cancer cell re-epithelialisation via the attenuation of TGFβ (pSMAD2-3) signalling that favours BMP activation (pSMAD1-5). Alongside TGFβ, AXL and Twist1 expression are also downregulated in this second phase of colonisation. Functionally tumour cells switch from a mesenchymal motile state to a proliferative phenotype sustained by BMP/ID1 that is compatible with metastatic outgrowth.

6.1 AXL-mesenchymal features of MICs promote metastatic colonisation beyond invasion

6.1.1 Metastasis-initiating cells: stemness throughout epithelial plasticity

As previously described, successful establishment of metastases requires cancer cells with self-renewal properties to disseminate from their tissue of origin and re-initiate tumour growth at favourable secondary sites (section 1.3.1). The fact that inducing EMT in non-malignant epithelial cells can lead to the generation of cancer stem cells (Morel et al., 2008, Mani et al., 2008) suggests a unified molecular programme underlying metastasis. In line with this idea we found that the MICs of the MMTV-PyMT tumour model (Malanchi et al., 2012) display a mesenchymal phenotype characterised by high AXL expression (Figure 3.3 and 3.8). However, as previously mentioned, the reversible nature of the EMT programme combined with the crucial requirement of stemness during the metastatic process makes the linear relationship between EMT and stemness controversial (section 1.4). Although we did not performed lineage tracing to determine the epithelial origin of MICs ($CD24^+CD90^+$), the isolation strategy with CD24 allows the isolation of mammary tumour cells of epithelial origin with high purity. Indeed, the isolation of $CD24^-CD90^+$ or $CD24^-CD90^-$ and examination in pure 3D Matrigel cultures showed that these populations are phenotypically and functionally fully mesenchymal cells that form a reticular network invading into the gels (data not shown); in contrast, the $CD24^+CD90^+$ MIC subpool in the same matrices formed mammary branching organoids (Figure 4.10.D).

We used the $CD24^+AXL^+$ more mesenchymal population isolated from the epithelial compartment of PyMT tumours that displays a similar EMT phenotype to MICs (Figure 3.10.A), and compared their tumour initiation ability using orthotopic transplantations as the gold standard test for stemness *in vivo* (Figure 3.10.B). These *in vivo* assays revealed that despite their similar mesenchymal phenotype the tumour-initiating capacity of MICs was higher when compared to the AXL^+ subpool (Figure 3.12). This result suggests that additional stem-like features of MICs such as high Wnt signalling (Figures 3.2 and 3.3) rather than their mesenchymal phenotype confers enhanced tumour-initiating properties to MICs. In

addition, the predominantly epithelial population of $CD24^{+}Sca1^{+}$ cells (Figure 3.11), previously described as a tumour-initiating pool in the MMTV-Neu model (Liu et al., 2007b), displayed a tumorigenic potential similar to that of the more mesenchymal MICs (Figure 3.12). This result indicates that in the MMTV-PyMT model the mesenchymal status *per se* does not correlate with the stemness needed to initiate tumour growth. Notably, the fact that more mesenchymal characteristics did not provide an advantage in tumour initiation ability indicates that both subpools retain similar levels of stemness.

We further analysed the reversible nature of the EMT programme and the crucial requirement of stemness during the metastatic process. We tested the $CD24^{+}AXL^{+}$ mesenchymal PyMT population *versus* the more epithelial AXL^{-} tumour fraction in parallel in tumour initiation and metastatic colonisation assays (Figure 3.10.B-C). Both assays require cancer cell intrinsic stemness properties to self-renew and grow either in a collective challenge when locally injected in its tissue of origin, or in a single cell challenge when directly seeded into a foreign microenvironment. Therefore a comparison of cells behaviour in these two settings allows understanding whether a particular EMT status would provide additional advantage to metastasise. Indeed, the $CD24^{+}AXL^{+}$ mesenchymal PyMT population displayed enhanced metastatic colonisation ability whereas its tumour-initiating capabilities were unaltered compared to the AXL^{-} fraction (Figure 3.10.D-E). This result suggests that the AXL-mesenchymal status provides an advantage in metastatic colonisation independent from cancer cells intrinsic stemness.

Typically, studies analysing the relationship between epithelial plasticity and stemness alter the expression of the EMT core transcription factors to change the EMT status of the cells and assess the impact on stemness (Mani et al., 2008, Beck et al., 2015, De Craene et al., 2014, Chaffer et al., 2013). In this study we chose to manipulate the mesenchymal phenotype of cancer cells via the downstream effector of the EMT programme AXL (section 1.2.1.1). The idea behind this approach was to revert the mesenchymal features of cancer cells without inducing broader changes in cellular functions caused by the alteration of EMT-TFs. Indeed, it is known that manipulating the expression of EMT-TFs can induce extensive epigenetic changes directly affecting survival and self-renewal

(Chaffer et al., 2013, Wellner et al., 2009, Mani et al., 2008). In contrast, by blocking AXL in MICs we confirmed the extensively reported phenotypic inhibition of the AXL-mesenchymal status, and also achieved a modest decrease in the expression levels of the EMT-TFs and a functional reduction in cell motility (Figure 4.2). Similar effects were observed in the MDA-MB-231 basal cell line that shares the AXL-mesenchymal status of the MIC subpool of MMTV-PyMT cells (Figure 4.3). In agreement with our previous data comparing the $CD24^{+}AXL^{+}$ PyMT population to the AXL^{-} tumour fraction, the exogenous depletion of AXL from total PyMT cells using a shRNA against AXL or the specific inhibitor R428 blocked their metastatic ability (Figure 4.6.H and 4.8.E) but did not have any effect on tumour initiation (Figure 4.6.G and 4.10.D). Importantly, we further confirmed that the intrinsic stemness abilities required to initiate tumour growth were not altered upon AXL depletion in PyMT cells performing sphere formation assays *in vitro* (Figure 4.6.F and 4.10.G). These results reiterate the crucial role of AXL expression specifically during metastatic colonisation.

The requirement of the AXL-mesenchymal status during early metastatic colonisation was confirmed using the AXL-expressing human MDA-MB-231 cell line and the mouse 4T1 cell line in experimental metastasis assays (Figure 4.9.B and 4.7.B). In sharp contrast to the expendable role of AXL during primary tumour initiation in the primary MMTV-PyMT model, the tumour growth ability of these cell lines upon AXL depletion have been reported to be impaired (Gjerdrum et al., 2010). This study excluded deleterious effects in cell survival and proliferation upon orthotopic transplantation, two cellular functions that AXL depletion can negatively affect in other cancer types when functioning as a tyrosine kinase receptor upon Gas6 stimulation (Paccez et al., 2014). These results suggest that these basal-like cell lines are highly dependent on AXL, a protein that all cells homogeneously express (Figure 4.3.C and 4.7.A), having AXL additional functions that could impact on cell intrinsic properties like stemness. This possibility could represent a limitation when using highly homogeneous cell lines to study stem-like properties of tumour cells.

The role of the AXL-mesenchymal status and its relationship with stemness was further analysed by exogenously expressing AXL. This resulted in an increase in

the number of AXL-expressing cells among the PyMT population (Figure 5.8.B), whereas in MDA-MB-231 cells AXL exogenous expression increased the overall AXL levels and exacerbated their mesenchymal phenotype (Figure 5.8.B-C). This exogenous AXL expression led to increased tumour growth in both models, but neither the proliferation nor the microenvironmental composition of these fully established tumours were altered (Figure 5.10.B-E). However, we did not exclude that a boost in proliferation due to enhanced AXL signalling occurred earlier during tumour initiation. A possible explanation for this enhanced tumour growth *in vivo* could be a direct AXL stimulation by its ligand Gas6 (Korshunov, 2012, Dormady et al., 2000). Indeed, several components of the tumour microenvironment such as monocytes and fibroblasts secrete Gas6 (Ben-Batalla et al., 2013, Dormady et al., 2000). This could activate AXL tyrosine kinase function triggering PI3K-AKT-NFkB mediated survival (Figure 1.4.B – purple arrows) making tumour cells more resistant to apoptosis just after the post-orthotopic transplantation period. Although it has been reported that AXL does not enhance survival or proliferation in breast cancer cells (Paccez et al., 2014), its overexpression might lower the threshold to respond to Gas6 levels. In agreement with this hypothesis, when AXL-overexpressing cells were assayed in sphere conditions *in vitro* where they lack microenvironmental Gas6-AXL stimulation, their sphere formation ability was indeed impaired (Figure 5.10.A). The efficient primary tumour formation achieved by AXL-overexpressing cells excludes that their reduced *in vitro* sphere formation ability depends on a direct negative effect on self-renewal. Indeed, when culturing total primary PyMT cells in non-adherent conditions the endogenous AXL levels are downregulated (data not shown). This result together with the evidence that AXL depletion does not alter the sphere formation ability of PyMT cells (Figure 4.6.F and 4.10.G) suggest that to adapt to the non-adherent sphere culture conditions AXL loss is required. Collectively, these data suggest that AXL-mesenchymal features *per se* do not alter self-renewal.

Similarly to the requirement of AXL downregulation for cancer cells to adapt to the specific *in vitro* non-adherent conditions in sphere assays, we observed that AXL is lost during metastatic colonisation as cancer cell grow in the lungs (Figure 4.8.B and 4.9.B). Accordingly, maintaining AXL expression with an exogenous promoter

and thereby retarding its downregulation during metastatic colonisation (Figure 5.9.B) reduces metastatic outcome (Figure 5.9.D-E).

As briefly mentioned above, AXL apart from a key EMT downstream effector of the EMT programme (Paccez et al., 2014) functions as a tyrosine kinase receptor activated by its ligand Gas6 (Korshunov, 2012). It has been shown in leukemia that AXL mediates a paracrine crosstalk with the stroma establishing a cancer cell niche at the primary site (Ben-Batalla et al., 2013). Leukemic cells educate bone-marrow stromal cells to secrete Gas6, which in turn stimulates AXL^+ leukemic cells enhancing their proliferation (Ben-Batalla et al., 2013). A similar mechanism in breast cancer metastasis could lead to decreased metastatic potential when inhibiting AXL independently of the epithelial/mesenchymal modulations. Nevertheless, we observe AXL to be downregulated in disseminated cancer cells when they start proliferating in the lungs (Figure 4.8.A and 4.9.A). Therefore, the stromal derived signals promoting metastatic establishment in breast cancer metastasis are likely to be independent from this AXL-Gas6 axis observed in leukemia that promotes proliferation. Moreover, AXL overexpression reduces early cancer cell accumulation in the lungs and impairs metastatic outgrowth (Figure 5.9.B and 5.9.D-E) inversely correlating with the proliferative promoting-effects described in leukemia (Ben-Batalla et al., 2013). Importantly, our data are in line with several studies in breast cancer highlighting the role of AXL as an EMT effector independently of its tyrosine kinase receptor function regulating proliferation and survival (Gjerdrum et al., 2010, Holland et al., 2010, Paccez et al., 2014). Furthermore, we observed that AXL is expendable for cell proliferation *in vitro* and tumour initiation *in vivo* (Figure 4.10.A-E).

In summary, our data illustrates the maintenance of stemness and tumour initiating ability throughout AXL-mesenchymal modulations rather than a direct correlation between EMT and stemness (Figure 6.2). A harmonising model to explain this complex relationships between EMT and stemness in cancer cells is the gradient EMT model (Figure 1.6.A) (Ombrato and Malanchi, 2014). In line with our results, this model proposes that cancer cells with stem-like properties can transit between more mesenchymal or more epithelial phenotypes in a context dependent manner within the limits of a ‘stemness window’ where the maintenance of their intrinsic stemness abilities is preserved (Figure 1.6.A). In line with this model, exacerbating

the mesenchymal state of cancer cells through AXL exogenous expression (Figure 5.9.D-E) or as previously reported inducing the expression of an additional EMT-TF such as Prrx1 (Ocana et al., 2012, Brabletz, 2012a) pushes cells out of the 'stemness window' proposed by the gradient EMT model to a very mesenchymal state that reduces cancer cells plasticity. This impairs the ability of cancer cells to adapt to a new environment such as the one encounter during metastatic colonisation resulting deleterious for the process.

On the other hand, we have described that AXL positive cells show enhanced metastatic activity, and that the AXL-mesenchymal status despite the requirement of its later downregulation, provides an initial advantage in metastasis (Figure 3.10.E and 4.6.H). Therefore, we propose the requirement of a plastic AXL-mesenchymal state that can be modulated throughout metastatic colonisation. Indeed, confirming the previous conclusion reached by Ocana et al., we observed that the mesenchymal phenotype induced via EMT and stemness are uncoupled, as required for a context dependent modulation. In line with this suggested context-dependent modulations of the AXL-mesenchymal status we showed how during early metastatic colonisation AXL downregulation depends on specific features, like the secretion of the niche inducing factor THBS2, linked to its presence (Figure 3.13 and 3.14); whereas AXL absence correlates with active BMP signalling required for metastatic outgrowth (Figure 5.4.E-G).

AXL modulations in MMTV-PyMT cells

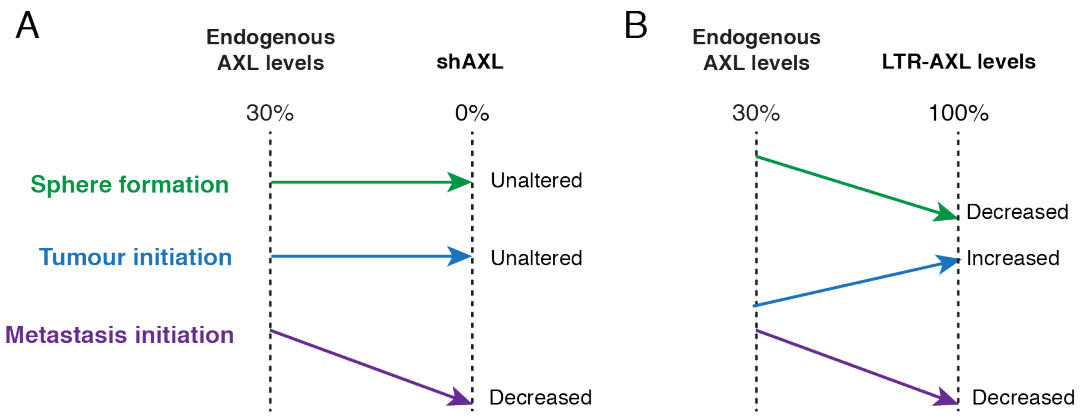


Figure 6.2 Functional readouts of AXL modulations in PyMT cells
Scheme displays the relative comparison of the sphere formation/self-renewal ability *in vitro* (green line), tumour initiation (blue line) and metastatic ability (purple line) of PyMT primary cells according to their AXL expression levels. (A) Effects of AXL depletion. In basal state (endogenous AXL levels) around 30% of PyMT cells express AXL; by infecting the cells with a lentiviral construct expressing shAXL we completely deplete AXL expression (0% of cells). (B) Effects of AXL overexpression. Exogenous expression of AXL (LTR-AXL) generates a homogeneous cell preparation were all PyMT cells express AXL (100%).

6.1.2 THBS2, a potential targetable niche component to prevent metastatic outgrowth

The crucial induction of a favourable metastatic niche upon arrival to the target site is a pre-requisite for metastasis (section 1.3.2). We have previously discussed how fibroblast-derived ECM components, such as POSTN and TNC are crucially required for breast to lung metastatic colonisation (Malanchi et al., 2012, Oskarsson et al., 2011). Our work highlights the requirement of a highly secretory mesenchymal phenotype linked to AXL expression mediating fibroblast activation at the naïve target site (Figure 3.13), providing a potential mechanism to explain how the fibroblastic metastatic niche is induced upon cancer cell arrival to the lungs. Moreover, we show that MICs displaying an AXL-mesenchymal state show an enhanced ability to activate normal lung fibroblasts *ex vivo* compared to the nonMIC subpool of the tumours (Figures 3.14, 3.15).

All cancer cells can trigger stromal activation by secreting factors such as TGFβ. Here we show that cells in an AXL-mesenchymal state show an overall enhanced secretory activity that can potentially perturb their surrounding microenvironment (Figure 3.2 and 3.3). Particularly, we identified the ECM glycoprotein THBS2 to have a novel role as key effector enhancing fibroblast activation during metastatic colonisation. MICs displaying an AXL-mesenchymal phenotype secrete high levels of THBS2 (Figure 3.8 and 3.16), which enhances fibroblast activation through integrin β1 signalling (Figure 3.17 and 3.18). The crucial consequence of this activity is that a single cell expressing THBS2 will be more efficient in activating its surrounding microenvironment compared to a cell that do not secrete this stroma co-activator. Indeed, using experimental metastasis assays we confirmed that the THBS2-mediated MIC-fibroblast crosstalk is required for efficient metastatic colonisation, (Figure 3.19.C and 3.20.C). Conversely, broadening THBS2 expression to all PyMT cells by exogenously expressing it in the nonMIC fraction of the tumour led to an increase in actively metastasising cells in the lungs (Figure 3.21.C), suggesting that expressing THBS2 helps a larger number of tumour cells to successfully induce their niche and thereby enhancing their early metastatic colonisation ability.

In the context of previous studies showing the key role of recruited fibroblasts to form a niche for disseminated breast cancer cells in the lung (Malanchi et al., 2012, Oskarsson et al., 2011), we now provide a molecular mechanism explaining how this cancer cell-fibroblast alliance is established in the first place. THBS2 is one crucial mediator of this fibroblast activation, and thereby promotes metastatic colonisation. Importantly, given its direct constant requirement during early metastatic colonisation (Figure 3.20.C and 3.21.C), THBS2 represents a potential therapeutically targetable molecule to prevent the outgrowth of disseminated tumour cells at distant sites in patients. The therapeutic relevance of THBS2 is supported by computational analyses defining it as one of the three genes associated with stromal desmoplastic reaction and high metastatic risk (Kim et al., 2010), as well as the clinical evidences of its correlation with poor distant metastasis free survival (DMFS) in high grade human breast carcinomas (Figure 3.22).

The decisive pro-metastatic role of THBS2 we observed during early breast to lung metastatic colonisation opposes to the described pro-dormancy function of endothelial cell derived THBS1 in the same context (Ghajar et al., 2013). Ghajar et al. showed that upon extravasation to the lungs, breast cancer cells remain in a perivascular niche were endothelial cell derived THBS1 maintains tumour cells in a dormancy state. Subsequently, upon tumour cell induced vascular sprouting dormancy is overcome dependent on TGFβ signalling stimulation (Ghajar et al., 2013). Although both THBS1 and THBS2 belong to the same glycoprotein family, THBS2 is a smaller protein lacking the TGFβ interacting domain that THBS1 harbours. This could explain the observed TGFβ mediated effect on releasing THBS1 sustained dormancy at perivascular lung niches. Concomitantly, in the very same temporal window of the metastatic process early after extravasation, we have also found AXL-mesenchymal cells undergoing TGFβ signalling (Figure 5.3.B), and as consequence of their TGFβ sustained AXL-mesenchymal phenotype they secrete THBS2 inducing a fibroblastic niche (Figure 3.13.B-C). Notably, we showed that these fibroblastic niches are exclusively formed around mesenchymal AXL^+ cells metastasis-initiating cells, whereas both mesenchymal AXL^+ and more epithelial AXL^- cells can be found in dormant perivascular niches (Figure 3.13.B-C).

The example explained above with THBS1 and THBS2 illustrates how two similar glycoproteins in the same context of early metastatic colonisation can mediate antagonistic functions. However, in fully developed breast tumours, both THBS1 and THBS2 have been shown to exert antiangiogenic functions (Jimenez et al., 2000). The antiangiogenic effect of Thrombospondins is mediated by the common N-terminal TSR domain of these proteins that binds CD36 in endothelial cells triggering apoptosis (Koch et al., 2011). This was shown in a breast cancer xenograft model using the MDA-MB-435 cell line (Koch et al., 2011). In the MMTV-PyMT model, we confirmed that long-term depletion of THBS2 led to increased tumour vascularisation and consequently higher tumour growth (Figure 3.23.A). Importantly, no effects were observed in early primary tumour initiation when depleting THBS2 (Figure 3.20.B), excluding the role of this factor in regulating intrinsic cancer cells features such as stemness. Conversely, overexpressing THBS2 in PyMT tumours impaired tumour vascularisation and led to decreased tumour growth (Figure 3.23.B). Moreover, as expected when altering the ECM composition throughout tumour progression (section 1.1.2.2) histological analysis of THBS2 overexpressing tumours revealed changes in stromal organisation beyond the expected decrease in vasculature. Epithelial cell organisation was altered in THBS2 overexpressing compared to normal PyMT tumours, and the number of infiltrated stroma areas was dramatically increased (Figure 3.23.B).

Altogether, these results depict THBS2 as a key tumour microenvironment component with distinctive roles during tumour initiation, progression and metastatic colonisation (Figure 6.3). Beyond its effects in vascularisation, given that we observed that: 1) THBS2 enhances lung fibroblast activation, and 2) THBS2 overexpressing tumours show a distinct stromal composition, it would be interesting to further characterise THBS2 mediated effects on the fibroblastic stroma at the primary site.

Overall, our work describes a novel role of THBS2 as a key mediator of metastatic colonisation. Importantly, THBS2 represents a potential therapeutically targetable molecule to prevent the outgrowth of disseminated tumour cells at distant sites in patients after primary tumour excision. Given the pro-angiogenic effects following THBS2 depletion from the tumour microenvironment, primary tumour excision

would be a pre-requisite to anti-THBS2 therapies. As commented throughout the introduction (section 1.2), current efforts to develop new anti-metastatic therapies go in this direction, trying to identify new targets that could prevent the outgrowth of disseminated tumour cells. Therefore, THBS2 blocking strategies, such as blocking peptides or antibodies represent a promising approach to test in preclinical spontaneous mouse models upon tumour resection. In this study, the differences in tumour growth observed when genetically manipulating THBS2 expression levels (Figure 3.23) precluded the examination of metastatic outcome in a spontaneous metastasis setting, as differences in tumour size do not necessarily correlate in a linear manner with dissemination and metastatic outcome.

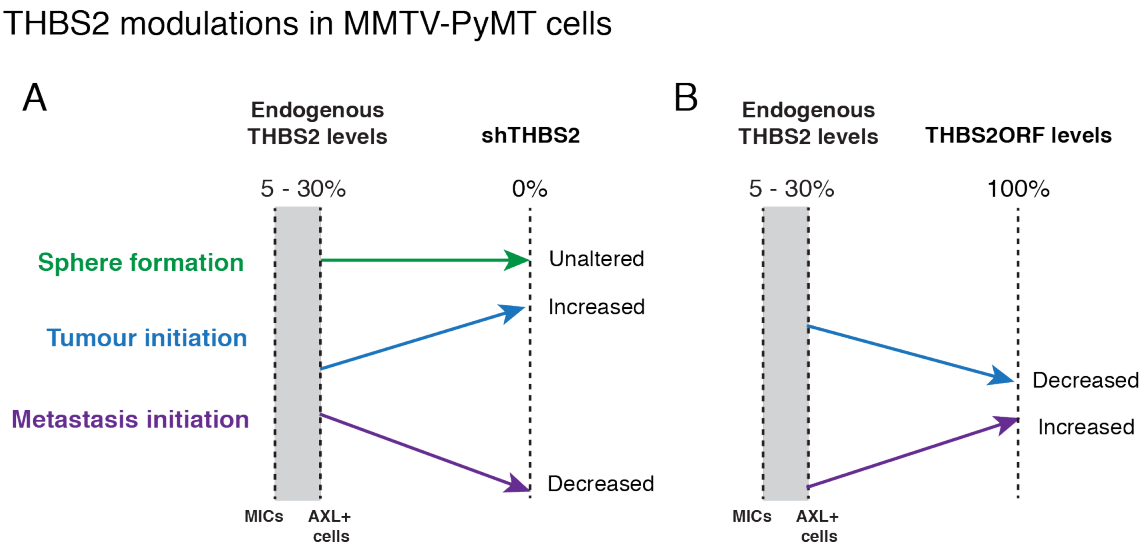


Figure 6.3 Functional readouts of THBS2 modulations in PyMT cells
Scheme displays the relative comparison of the sphere formation/self-renewal ability *in vitro* (green line), tumour initiation (blue line) and metastatic ability (purple line) of PyMT primary cells according to their THBS2 expression levels. (A) Effects of THBS2 depletion. In basal state (endogenous THBS2 levels) THBS2 is highly expressed by MICs that constitute around 5% of the tumour cells, and AXL^+ cells that conform around 30%; by infecting the cells with a lentiviral construct expressing shTHBS2 we completely deplete THBS2 expression (0% of cells). (B) Effects of THBS2 overexpression. Exogenous expression of THBS2 (THBS2ORF) generates a homogeneous cell preparation were all PyMT cells express AXL (100%).

6.1.3 The mesenchymal status of MICs defines their niche induction ability

We have described that AXL-mesenchymal MICs are in a highly secretory state defined by the expression of different previously characterised secreted factors that to contribute to their metastatic potential, such as TNC, SPARC, MMPs (Figure 3.2 and 3.3). Importantly, we observed that THBS2 is dependent on the AXL mesenchymal traits in metastatic cells (Figure 4.4.A-B), which are indeed maintained upon extravasation at the distant tissue (Figure 4.8.A). These results establish a direct link between the niche induction ability via THBS2 and the AXL-mesenchymal status of MICs (AXL). Remarkably, both the MIC subpool of MMTV-PyMT cancer cells as well as the metastatic human MDA-MB-231 cell line show higher niche activation ability dependent on their mesenchymal phenotype (Figure 4.5). Furthermore, in microarray data from human carcinoma biopsies THBS2 expression correlates with AXL, Vimentin, Twist1 and Zeb1 expression supporting this link (Table 4.1 and Figure 4.1). Indeed, the dependency of the niche induction on AXL-mesenchymal status during metastatic colonisation was confirmed *in vivo* when overexpression of THBS2 in the MDA-MB-231 cell line partially recued the deleterious effects of inhibiting AXL (Figure 4.11).

These results show for the first time that EMT is not only required during cancer cell dissemination (Figure 1.6.B) but also early upon arrival to the target site. We describe that the mesenchymal status of cancer cells is required during early metastatic colonisation as it mediates the expression of crucial niche co-activators such as THBS2. These niche co-activators trigger a faster ‘education’ of the surrounding stroma that in turn creates a favourable microenvironment supporting metastatic cell growth.

Whereas we believe that the mesenchymal status of MICs confers an advantage when metastasising as single cells, certainly there might be other mechanisms by which tumour cells metastasise. As previously referred, epithelial cells clusters can locally invade and can be found in patients blood samples (section 1.2.1.2 and 1.2.2). Although it is still unclear how tumour cell clusters sometimes composed by up to 10 cells can extravasate at distant sites (Krebs et al., 2014), a recent study by Aceto et al. showed that rare CTC-clusters are 50-fold more efficiently

metastasising than single CTCs in a mouse model of experimental metastasis (Aceto et al., 2014). It seems evident that being part of a cell cluster has several advantages to promote metastatic colonisation: 1) More efficient niche induction ability - cancer cells within a cluster collectively secrete a higher concentration of any niche-activation factor compared to a single disseminated cell, therefore they can more effectively induce changes in the naïve microenvironment upon extravasation; and 2) Existence of an in-house niche before extravasation - cell clusters can serve as their own niche: the first possibility is that the tumour cell heterogeneity within the oligoclonal cluster might be sufficient to self provide a favourable niche; the second possibility is that different stromal cells and ECM niche components are part of these epithelial clusters, which so far can not be excluded (Yu et al., 2013). With either possibility, epithelial tumour cells in the cluster would circumvent the requirement of early niche induction within the metastatic tissue, being the mesenchymal features described in this work redundant for metastasising cell clusters.

Moreover, as previously mentioned (section 1.2.2), circulating tumour cell analyses in breast, prostate and lung cancer patient samples have revealed that within the clusters there are cells displaying partial EMT states (Armstrong et al., 2011, Hou et al., 2011). The fact that cells within the cluster show cell-cell junctions does not exclude EMT heterogeneity. Therefore, cells displaying a partial EMT state, such as the MMTV-PyMT MICs (Figure 3.4), could form part of these clusters. This possibility could explain why ‘epithelial looking’ clusters display enhanced survival and metastatic ability (Aceto et al., 2014).

It still remains to be determined whether the main route towards metastatic lung colonisation is mediated by more abundant mesenchymal-like single MICs or rare but more potent epithelial cell clusters. Our work using single cell experimental metastatic assays describes a novel targetable mechanism that prevents metastatic outgrowth when cells in a more mesenchymal state are driving the metastatic process.

6.2 The metastatic niche: key determinant for metastatic progression

We have discussed in the previous section how the niche activation in the early phase of colonisation is a crucial step for efficient metastasis and it depends on the AXL-mesenchymal status of MICs. It is now accepted that mesenchymal activation within the primary tumour is needed for dissemination while its inhibition within the distant site is required for metastatic outgrowth (Ocaña et al., 2012; Tsai et al., 2012). In line with the requirement of mesenchymal reversion at the metastatic site, we observed that AXL as well as the highly expressed EMT-TF Twist1 are spontaneously downregulated in the lungs as metastatic colonisation progresses (Figure 4.8.A, 4.9.A and 5.1.B-E). The fact that the attenuation of the AXL-mesenchymal features temporally coincides with the niche activation, led us to test whether the newly activated fibroblast at the metastatic niche could modulate the epithelial status of cancer cells. Indeed, co-culturing MICs with fully activated fibroblasts (CAFs) mimicked the observed *in vivo* modulations mediating both AXL and Twist1 downregulation in MICs (Figure 5.2.A-D). Moreover, CAFs phenotypically and functionally reverted MICs mesenchymal state (Figure 5.3).

We have previously introduced TGFβ as one of the main EMT inducers providing cancer cells with mesenchymal features at the primary site (Figure 1.4.B) (section 1.2.1.1). Indeed, TGFβ positively regulates AXL expression alongside the EMT-TFs (Li et al., 2014b). In line with these observations we showed that during the second phase of colonisation TGFβ activity, as monitored by pSMAD2-3 expression, is downregulated along with the AXL-mesenchymal features and Twist1 in metastasising cells in the lung (Figure 5.4.B-D). Importantly, it has been widely characterised that the two different branches of the TGFβ pathway, TGFβ and BMP interfere each others activity (Wakefield and Hill, 2013). Accordingly we detected alongside TGFβ downregulation that the overall number of cells undergoing BMP signalling increased in metastasising cells being BMP-dependent pSMAD1-5 activity maintained as cells transit from early to late colonisation (Figure 5.4.E-G). This was further corroborated by the increased in the number of cells that gained ID1 expression, a canonical BMP downstream effector, as they lost Twist1 expression (Figure 5.4.H-I). These data suggest that the attenuation of TGFβ

signalling during the second phase of colonisation as a potential cause for the reversion of the AXL-mesenchymal state in cancer cells transitioning towards a more epithelial phenotype.

Certainly, in contrast to the pro-invasive abilities triggered by TGFβ, BMP usually controls cell proliferation (Wakefield and Hill, 2013) and finely tunes the balance between self-renewal and differentiation in normal and cancer stem cells (Lonardo et al., 2011, He et al., 2004). Therefore, the TGFβ-BMP switch could mediate the functionally antagonistic 'go or grow' cell states observed in primary tumours (Patsialou et al., 2015) at distant secondary sites, reversing the initial mesenchymal migratory state to allow proliferation. Indeed, it was previously described that the downstream BMP effectors ID proteins are required for breast cancer cell proliferation and are critically expressed during metastatic outgrowth (Gupta et al., 2007). Accordingly, we confirmed that blocking BMP signalling and thereby its downstream effector ID1 in PyMT cells specifically during the second phase of colonisation, prevents metastatic outgrowth (Figure 5.6.A-B).

Importantly, CAF-secreted factors were capable of mimicking *in vitro* the BMP-dependent ID1 induction on AXL-mesenchymal cells (Figure 5.5.A-B) suggesting that the stromal activation at the metastatic site *in vivo* would be capable of inducing the TGFβ-BMP switch. The TGFβ superfamily of ligands is one of the defining signatures of activated fibroblasts (Calvo et al., 2013). Thus, activated fibroblasts can secrete ligands from the two different branches of the pathway, TGFβ and BMP. By using inhibitors to disrupt the downstream signalling of each branch of the pathway specifically, we found that CAF-mediated ID1 induction is primarily sustained by BMP signalling (Figure 5.5.C-D). Additionally, we observed that CAF-mediated BMP-dependent ID1 induction correlates with increased cell proliferation and E-cadherin expression (Figure 5.5.C-E) suggesting a switch in AXL-mesenchymal cells towards a more epithelial proliferative state *in vitro*.

Altogether, our findings suggest that the newly activated fibroblasts during the second phase of colonisation trigger a crosstalk of signals that attenuates TGFβ activity, leading to the re-epithelialisation of cancer cells. Based on these data, it is tempting to think that once activated, lung fibroblasts could secrete TGFβ

antagonists such as Decorin (Zhang et al., 2007), THBS1 (Rojas et al., 2008) or LTBP1 (Lorda-Diez et al., 2010) that would decrease TGFβ signalling favouring BMP activation, however further experimental evidence would be required to prove this.

The idea of a balance between TGFβ and BMP signalling regulating EMT modulations at the target site is in line with the fact that TGFβ signalling is finely controlled throughout metastasis (Figure 1.6.B). TGFβ is the master EMT inducer at the invasive front of primary tumours (Giampieri et al., 2009, Xu et al., 2009); subsequently in the circulation platelet-derived TGFβ have been shown to maintain the mesenchymal characteristics of CTCs (Labelle et al., 2011). Our data complete this picture by describing a bi-phasic colonisation model correlating with TGFβ dynamics: in the early phase extravasated mesenchymal tumour cells display active TGFβ signalling (pSMAD2-3 expression) along with AXL and Twist1, whereas in the second phase TGFβ signalling is attenuated and BMP becomes the predominant signal of the pathway driving a phenotypic switch that mediates metastatic outgrowth.

Additionally, we observed that BMP signalling rather than being induced *de novo* during the second phase colonisation when TGFβ activity is lost, is active throughout the process as indicated by the high pSMAD1-5 activity already present upon extravasation (Figure 5.4.E-G). This is line with a recently published report describing that inhibiting BMP signalling during primary tumour development in the MMTV-PyMT model impairs metastatic outcome in this preclinical therapeutic setting (Owens et al., 2015). The authors highlight that beyond BMP inhibition affecting primary tumour development at the primary site, BMP signalling is required through metastatic invasion (Owens et al., 2015). Based on these observations, it is reasonable to speculate that BMP activity in PyMT tumour cells is a fundamental signalling that is always active, and modulates its activation level in the presence or absence of TGFβ signalling, exerting a more recessive or dominant effect on cell functions. Epithelial plasticity could be the functional outcome of these TGFβ-BMP modulations shifting cancer cells purposes (i.e. high niche induction *versus* cancer cell outgrowth) throughout the dynamic metastatic progression.

Interestingly, although the effector ID1 is a canonical BMP target and its expression is sustained after BMP stimulation (Hollnagel et al., 1999), TGFβ signalling has been shown to transiently activate ID1 coupled to stress signalling in epithelial cells (Kang et al., 2003). Indeed, we observed that in the early phase of colonisation breast cancer cells undergoing TGFβ signalling and displaying an AXL/Twist1 mesenchymal state also express ID1 (Figure 5.1.C and 5.4.H). Moreover, ID1 expression is maintained in the second phase of colonisation when TGFβ is downregulated and BMP/pSMAD1-5 signalling increases (Figure 5.4.H), suggesting that ID1 expression in metastasising cancer cells is sustained by TGFβ attenuation rather than its stimulation (Figure 5.4). In contrast to our results showing the concomitant expression of Twist1 and ID1 during early colonisation, a recent report from Stankic et al. suggested that TGFβ driven ID1 expression could antagonise Twist1 during metastatic colonisation. This correlative antagonistic expression of ID1 and Twist1 was examined *in vitro* using the human epithelial HMLE exogenously expressing Twist1. When stimulated with TGFβ, HMLE-Twist1 mesenchymal cells induced ID1 expression while normal HMLE epithelial cells cannot induce ID1 expression (Stankic et al., 2013). The discrepancy between our results and the observed function of ID1 as a negative regulator of Twist1 observed by Stankic et al. could have many explanations. An important difference to note is the fact that Stankic et al. used normal epithelial cells overexpressing Twist1 and not cancer cells to define this antagonistic activity *in vitro*, whereas we are monitoring ID1 and Twist1 expression in cancer cells metastasising *in vivo*. This fundamental difference in the approach used to observe ID1 and Twist1 dynamics makes it difficult to relate their findings to ours in the context of metastatic colonisation *in vivo*.

Adding to the proposed complex TGFβ-BMP modulations described in our work, it has been shown that during the first colonisation phase, when we observed tumour cells to display AXL-mesenchymal features and active TGFβ signalling, breast cancer cells secrete the TGFβ ligands antagonist Coco to exit dormancy and progress towards colonisation (Gao et al., 2012). In this study Gao et al. propose that Coco secretion enables a fraction of metastatic cells to overcome BMP-driven dormancy in the lung parenchyma. Although Coco has been shown to function as a

ligand trap for BMPs, it can also antagonise TGFβ ligands (Walsh et al., 2010, Bates et al., 2013). Therefore, as discussed by the authors, the possibility that in the lung metastatic niche Coco induces metastatic reactivation through BMP independent mechanisms cannot be excluded (Gao et al., 2012). Indeed, in line with our data, Coco expression during early colonisation could mediate an autocrine mechanism that reinforces the initial requirement to maintain the TGFβ-BMP balance in favour of TGFβ. Upon extravasation, Coco mediated TGFβ signalling maintenance would ensure the mesenchymal-dependent niche activation ability of MICs, which likely reduces their entrance into a dormant state.

6.3 Biphasic model of lung metastatic colonisation

We have observed that based on AXL expression metastatic colonisation can be divided in two temporal phases: during the early phase after extravasation, single disseminated cells maintain AXL-expression, along with Twist1 expression and TGFβ signalling, which provides an enhanced niche induction activity; in a later stage, temporally coinciding with the appearance of the activated niche, AXL is downregulated alongside Twist1 and TGFβ shifting cancer cell functions towards a BMP-dependent proliferation (Figure 4.8.A and 5.4). We also showed that AXL loss during the second phase of metastasis is functionally required for efficient metastatic growth (Figure 5.7) confirming that in order to efficiently generate a cancer cell mass at the distant site cells regain a more epithelial phenotype (Tsai et al., 2012). Indeed, inducing AXL-mesenchymal reversion using AXL chemical inhibition during this second phase of metastasis promoted metastatic outgrowth (Figure 5.7.D).

Previous reports proposed that Twist1 driven mesenchymal features need to be spatiotemporally controlled during colonisation (Tsai et al., 2012). Our results refine the suggested temporal regulation according to cancer cells AXL-mesenchymal status, and show that cancer cell epithelial modulations during the second phase of metastatic colonisation depend on the metastatic niche induced during the first phase. In this view, the mesenchymal status represents the trigger of its own inhibition via stromal activation. Indeed, stabilisation of the mesenchymal status via exogenous AXL expression in cancer cells delays AXL downregulation and strongly impairs metastatic ability *in vivo* (Figure 5.9.C-E). Supporting this model, Giampieri et al. showed that breast cancer cells overexpressing TGFβ1, which likely stabilises the AXL-mesenchymal status, are impaired in their colonisation capacity (Giampieri et al., 2009). Altogether these findings suggest epithelial cancer cell plasticity to be temporally regulated by a unified mechanism involving the inhibition of TGFβ signalling, and subsequently AXL and Twist1 expression. Importantly, exogenously expressing any of these EMT related components halts epithelial reversion and metastatic outgrowth.

Altogether, our findings show that the reversion of the mesenchymal status of cancer cells during the second phase of metastasis is a pre-requisite for metastatic outgrowth. Importantly, the fact that chemical inhibition of AXL during this second phase of metastasis favours epithelial reversion and enhances metastatic growth supports the previously raised concerns about anti-EMT therapies as double-edged swords in cancer therapeutics (Franco-Chuaire et al., 2013, Nieto, 2013). Indeed, the AXL inhibitor used in this work (R428 or BGB324) was the first AXL inhibitor to enter clinical trials two year ago (Sheridan, 2013). According to our results, AXL treatment should prevent dissemination from the primary site in patients; at the target site, we have shown that upon extravasation AXL is crucially maintained mediating niche initiation (Figure 4.8), therefore its inhibition could halt metastatic outgrowth. However, due to the asynchronous nature of the metastatic process, disseminated cells in different EMT states coexist at the target site in patients making it difficult to predict treatment outcome. Moreover, AXL inhibition could affect different microenvironmental components at the target site making it even more difficult to predict the global effect of this treatment.

Therefore, trying to shed some light into the potential beneficial effects of AXL inhibition during metastasis, we further examined AXL dynamics at the target site not only during colonisation, but also in fully established metastasis. We showed that after the previously reported AXL loss during the late colonisation phase, AXL expression in MICs is re-acquired in fully developed metastatic nodules (Figure 5.11.A). As expected, inhibiting AXL from tumour cells specifically during this late metastatic stage did not affect the number of metastases found in the lungs (Figure 5.7.D). However, we observed a dramatic reduction in the size of the metastatic nodules, indicating that tumour cell growth is being directly or indirectly sustained by AXL (Figure 5.11). This effect on metastatic outgrowth can be explained through two complementary hypotheses: 1) A direct effect on tumour cells; AXL at this late metastatic stage is highly expressed in MICs (Figure 5.11.A). The role of AXL expression on MIC in metastasis is not defined, therefore AXL-mesenchymal MICs could be required either to maintain functional heterogeneity among cancer cells or as the sub-set of cells responsible to sustain metastatic growth. In any case, AXL inhibition will mainly target MICs within the tumour cell compartment impairing

metastatic growth. 2) An indirect effect in tumour growth exerted by AXL-expressing microenvironmental components that sustain metastatic outgrowth. There are two key cellular components in the tumour microenvironment that highly express AXL and could be affected by the inhibitor treatment. TAMs usually express AXL in the tumour microenvironment (Sheridan, 2013) and play a pro-tumorigenic role during metastasis supporting tumour cell growth (Qian and Pollard, 2010); indeed, relevant for breast to lung metastasis, resident lung macrophages have been shown to upregulate AXL in the inflamed lung environment (Fujimori et al., 2015) being a potential target of the inhibitor during this late stage of metastasis. Also dendritic cells are known to express AXL (Scutera et al., 2009), although their role at metastatic sites is less characterised.
Overall, our results using an experimental synchronised metastatic setting support a global beneficial effect of AXL inhibition; early-disseminated cells would be prevented from growing and late metastatic nodules would partially regress, being only those cells transitioning from the first to the second colonisation phase positively supported by AXL inhibition.

Mechanistically, the fact that MICs re-gain AXL expression in late-stage metastasis poises intriguing questions. Based on this observation it could be hypothesised that at this stage, with entirely reconstituted tumours formed at the distant site, metastatic cells could become invasive again re-initiating the metastatic cascade. However, whether established metastases can metastasise have been long debate and still remains unclear (Tait et al., 2004, Holzel et al., 2010, Goldstein et al., 2005). Additionally, the overall AXL dynamics in metastasising cells in the lungs indicate that AXL modulations are mainly occurring within the MIC compartment while total tumour cells maintain AXL expression constant throughout the process (Figure 5.11.A). This is suggestive of the high plasticity displayed by MICs compared to the bulk of the tumour cells, correlating with their stem-like features. Also, this high plasticity illustrated by the changes in AXL expression supports the experimental evidence that MICs drive the metastatic process (Malanchi et al., 2012), and their presence is crucial throughout.

6.4 Concluding remarks

Mesenchymal features enhance stemness and metastatic spreading (Mani et al., 2008, Rhim et al., 2012), while growth at the distant site requires its inhibition (Ocaña et al., 2012; Tsai et al., 2012). Our findings indicate that a mesenchymal state characterised by AXL expression expeditiously triggers a cancer cell-stroma interaction upon infiltrating the secondary site. This crosstalk critically governs metastatic colonisation, and ultimately leads cancer cells to regain a more epithelial phenotype compatible with proliferation. Importantly, we identify THBS2 as a novel cofactor in the AXL-mesenchymal cell–stromal interplay. This study highlights the importance of a dynamic modulation of the epithelial-mesenchymal status of cancer cells during metastatic progression and supports the previously reported concerns about anti-EMT therapies as double-edged swords in cancer therapeutics (Tsai et al., 2012).

Chapter 7. Appendix

This appendix table contains the list of genes that show ≥3 fold increment expression in MICs ($CD24^+CD90^+$) compared to nonMICs ($CD24^+CD90^-$). The complete microarray raw data are deposited in the Gene Expression Omnibus (GEO) database and are publicly available at http://www.ncbi.nlm.nih.gov/geo/ with accession number GEO:GSE63558.

Red – genes described in this thesis to play a crucial role in the metastatic potential of MICs.

Blue – EMT-related genes.

Bold – genes previously shown to contribute to the metastatic potential of breast cancer cells.

POSITION	GENE SYMBOL	FOLD INCREMENT	DEFINITION
1	**Fn1**	**73.60964635**	**Mus musculus fibronectin 1 (Fn1), mRNA.**
2	Serping1	60.46920116	Mus musculus serine (or cysteine) peptidase inhibitor, clade G, member 1 (Serping1), mRNA.
3	Lum	60.06918146	Mus musculus lumican (Lum), mRNA.
4	**Thy1**	**59.32025165**	**Mus musculus thymus cell antigen 1, theta (Thy1), mRNA. CD90.**
5	**Mmp3**	**56.8142886**	**Mus musculus matrix metallopeptidase 3 (Mmp3), mRNA.**
6	Serping1	54.48133702	
7	Mmp13	50.80097994	Mus musculus matrix metallopeptidase 13 (Mmp13), mRNA.
8	Loxl1	49.87156687	Mus musculus lysyl oxidase-like 1 (Loxl1), mRNA.
9	Col1a1	47.68421086	
10	Ctsk	44.40448725	Mus musculus cathepsin K (Ctsk), mRNA.
11	Dcn	43.61584256	Mus musculus decorin (Dcn), mRNA.
12	Col6a1	43.02501894	Mus musculus procollagen, type VI, alpha 1 (Col6a1), mRNA.
13	**Mmp2**	**42.58322206**	**Mus musculus matrix metallopeptidase 2 (Mmp2), mRNA.**
14	Adamts2	41.75490813	Mus musculus a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 2 (Adamts2), mRNA.
15	Lrrc15	41.68029907	
16	Mfap5	41.64301515	Mus musculus microfibrillar associated protein 5 (Mfap5), mRNA.
17	Col5a1	41.04235801	Mus musculus procollagen, type V, alpha 1 (Col5a1), mRNA.
18	Rnase4	40.18861336	Mus musculus ribonuclease, RNase A family 4 (Rnase4), transcript variant 1, mRNA.
19	Pi16	39.30356586	Mus musculus peptidase inhibitor 16 (Pi16), mRNA.
20	Rarres2	38.82640932	Mus musculus retinoic acid receptor responder (tazarotene induced) 2 (Rarres2), mRNA.
21	Adamts2	37.03161082	Mus musculus a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 2 (Adamts2), mRNA.
22	Pla1a	35.33017006	Mus musculus phospholipase A1 member A (Pla1a), mRNA.

23	Cpxm1	34.56071134	Mus musculus carboxypeptidase X 1 (M14 family) (Cpxm1), mRNA.
24	Cygb	34.15923397	Mus musculus cytoglobin (Cygb), mRNA.
25	Ccdc80	34.05706562	Mus musculus coiled-coil domain containing 80 (Ccdc80), mRNA.
26	Ccl11	33.36436321	Mus musculus small chemokine (C-C motif) ligand 11 (Ccl11), mRNA.
27	Fstl1	32.97793885	Mus musculus follistatin-like 1 (Fstl1), mRNA.
28	**Lox**	**32.66522761**	**Mus musculus lysyl oxidase (Lox), mRNA.**
29	Col6a1	32.32838235	Mus musculus procollagen, type VI, alpha 1 (Col6a1), mRNA.
30	Col6a2	32.24366031	Mus musculus procollagen, type VI, alpha 2 (Col6a2), mRNA.
31	Serpinf1	32.22446558	Mus musculus serine (or cysteine) peptidase inhibitor, clade F, member 1 (Serpinf1), mRNA.
32	Pcolce	32.16013333	Mus musculus procollagen C-endopeptidase enhancer protein (Pcolce), mRNA.
33	Scara3	31.79005897	Mus musculus scavenger receptor class A, member 3 (Scara3), mRNA.
34	Aebp1	31.55275609	Mus musculus AE binding protein 1 (Aebp1), mRNA.
35	Fbln1	31.30597931	Mus musculus fibulin 1 (Fbln1), mRNA.
36	**Mmp3**	**31.29927231**	**Mus musculus matrix metallopeptidase 3 (Mmp3), mRNA.**
37	Fbn1	30.84942594	
38	Ogn	30.62286362	Mus musculus osteoglycin (Ogn), mRNA.
39	**Mmp3**	**30.22038636**	**Mus musculus matrix metallopeptidase 3 (Mmp3), mRNA.**
40	Tnxb	29.18681299	Mus musculus tenascin XB (Tnxb), mRNA.
41	Dpt	28.81349432	Mus musculus dermatopontin (Dpt), mRNA.
42	**Thbs2**	**27.92699494**	**Mus musculus thrombospondin 2 (Thbs2), mRNA.**
43	LOC100047583	27.8890865	PREDICTED: Mus musculus similar to apolipoprotein D (LOC100047583), mRNA.
44	Serpina3n	27.67160631	Mus musculus serine (or cysteine) peptidase inhibitor, clade A, member 3N (Serpina3n), mRNA.
45	Col12a1	27.39104496	Mus musculus collagen, type XII, alpha 1 (Col12a1), mRNA.
46	Knsl5	26.95470532	
47	Svep1	26.90604697	Mus musculus sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1 (Svep1), mRNA.
48	Spon2	26.83853041	Mus musculus spondin 2, extracellular matrix protein (Spon2), mRNA.
49	Plat	26.34455797	
50	Cd248	26.15887428	Mus musculus CD248 antigen, endosialin (Cd248), mRNA.
51	Col14a1	25.92587138	
52	LOC638301	25.55375508	PREDICTED: Mus musculus similar to interferon activated gene 204 (LOC638301), mRNA.
53	Bicc1	25.1432103	Mus musculus bicaudal C homolog 1 (Drosophila) (Bicc1), mRNA.
54	Srpx2	24.90728486	Mus musculus sushi-repeat-containing protein, X-linked 2 (Srpx2), transcript variant 2, mRNA.
55	LOC100044430	24.55997339	PREDICTED: Mus musculus similar to Interferon activated gene 205 (LOC100044430), mRNA.
56	Tnxb	24.27757116	Mus musculus tenascin XB (Tnxb), mRNA.
57	**Sparc**	**24.11332411**	**Mus musculus secreted acidic cysteine rich glycoprotein (Sparc), mRNA.**
58	Igfbp4	23.58079355	
59	scl0001849.1_2273	23.17580957	
60	6330406I15Rik	22.6360887	Mus musculus RIKEN cDNA 6330406I15 gene (6330406I15Rik), mRNA.

61	Mfap2	22.34155105	Mus musculus microfibrillar-associated protein 2 (Mfap2), mRNA.
62	Pdgfra	22.29305127	Mus musculus platelet derived growth factor receptor, alpha polypeptide (Pdgfra), transcript variant 1, mRNA.
63	Itga11	22.19115202	Mus musculus integrin alpha 11 (Itga11), mRNA.
64	Fbln1	22.09933361	
65	Htra3	21.72598387	
66	Cxcl14	21.70535391	Mus musculus chemokine (C-X-C motif) ligand 14 (Cxcl14), mRNA.
67	Nid1	21.42034706	Mus musculus nidogen 1 (Nid1), mRNA.
68	Prelp	21.29436558	Mus musculus proline arginine-rich end leucine-rich repeat (Prelp), mRNA.
69	6330406I15Rik	20.90904317	Mus musculus RIKEN cDNA 6330406I15 gene (6330406I15Rik), mRNA.
70	Gpx8	20.8173571	Mus musculus glutathione peroxidase 8 (putative) (Gpx8), mRNA.
71	Scarf2	20.81035285	Mus musculus scavenger receptor class F, member 2 (Scarf2), mRNA.
72	**Sparc**	**20.49187503**	**Mus musculus secreted acidic cysteine rich glycoprotein (Sparc), mRNA.**
73	Ccl7	20.40200837	
74	Lum	20.27088078	
75	Bicc1	20.12116588	Mus musculus bicaudal C homolog 1 (Drosophila) (Bicc1), mRNA.
76	Osr2	19.93992721	Mus musculus odd-skipped related 2 (Drosophila) (Osr2), mRNA.
77	Olfml2b	19.85429131	Mus musculus olfactomedin-like 2B (Olfml2b), mRNA.
78	Serpine2	19.81592693	
79	Itgbl1	19.72979134	Mus musculus integrin, beta-like 1 (Itgbl1), mRNA.
80	Col8a1	19.66821469	
81	ENSMUSG00000043795	19.60575329	PREDICTED: Mus musculus predicted gene, ENSMUSG00000043795 (ENSMUSG00000043795), mRNA.
82	Timp1	19.04713677	Mus musculus tissue inhibitor of metalloproteinase 1 (Timp1), transcript variant 2, mRNA.
83	Dcn	18.84139087	Mus musculus decorin (Dcn), mRNA.
84	Timp1	18.41986061	
85	C1qtnf3	18.41977383	Mus musculus C1q and tumor necrosis factor related protein 3 (C1qtnf3), mRNA.
86	Igfbp4	18.36908763	Mus musculus insulin-like growth factor binding protein 4 (Igfbp4), mRNA.
87	Ly6c1	18.30177212	Mus musculus lymphocyte antigen 6 complex, locus C1 (Ly6c1), mRNA.
88	Serpina3h	18.14573816	Mus musculus serine (or cysteine) peptidase inhibitor, clade A, member 3H (Serpina3h), mRNA.
89	Gp38	17.90928912	
90	scl0002507.1_236	17.80258841	
91	Ppic	17.68622567	Mus musculus peptidylprolyl isomerase C (Ppic), mRNA.
92	Prg4	17.59252918	
93	Col3a1	17.5644643	Mus musculus collagen, type III, alpha 1 (Col3a1), mRNA.
94	C2	17.55475936	Mus musculus complement component 2 (within H-2S) (C2), mRNA.
95	2310016C16Rik	17.26052088	Mus musculus RIKEN cDNA 2310016C16 gene (2310016C16Rik), mRNA.
96	Has1	17.09402607	Mus musculus hyaluronan synthase1 (Has1), mRNA.
97	Gpx3	17.07563602	Mus musculus glutathione peroxidase 3 (Gpx3), transcript variant 2, mRNA.
98	Olfml3	16.73507945	Mus musculus olfactomedin-like 3 (Olfml3), mRNA.

99	Rarres2	16.7212608	Mus musculus retinoic acid receptor responder (tazarotene induced) 2 (Rarres2), mRNA.
100	**Axl**	**16.71865497**	**Mus musculus AXL receptor tyrosine kinase (Axl), mRNA.**
101	Nbl1	16.51841458	Mus musculus neuroblastoma, suppression of tumorigenicity 1 (Nbl1), mRNA.
102	Tnxb	16.48466591	Mus musculus tenascin XB (Tnxb), mRNA.
103	Cthrc1	16.29152784	Mus musculus collagen triple helix repeat containing 1 (Cthrc1), mRNA.
104	Igfbp6	16.24186696	Mus musculus insulin-like growth factor binding protein 6 (Igfbp6), mRNA.
105	C1qtnf3	16.18213201	Mus musculus C1q and tumor necrosis factor related protein 3 (C1qtnf3), mRNA.
106	Ndn	16.00439486	Mus musculus necdin (Ndn), mRNA.
107	Pla1a	15.98069561	Mus musculus phospholipase A1 member A (Pla1a), mRNA.
108	Htra1	15.89958642	Mus musculus HtrA serine peptidase 1 (Htra1), mRNA.
109	Fbln2	15.59466235	Mus musculus fibulin 2 (Fbln2), transcript variant 2, mRNA.
110	Apbh	15.5632681	Mus musculus androgen-binding protein eta (Apbh), mRNA.
111	Emp3	15.4061342	Mus musculus epithelial membrane protein 3 (Emp3), mRNA.
112	Eln	15.3439034	Mus musculus elastin (Eln), mRNA.
113	Adamts12	15.29473237	
114	Srpx2	15.20458847	Mus musculus sushi-repeat-containing protein, X-linked 2 (Srpx2), mRNA.
115	Gpc3	15.12293152	Mus musculus glypican 3 (Gpc3), mRNA.
116	Gpc3	15.06075292	Mus musculus glypican 3 (Gpc3), mRNA.
117	Hsd11b1	15.05374854	Mus musculus hydroxysteroid 11-beta dehydrogenase 1 (Hsd11b1), transcript variant 1, mRNA.
118	Gstm2	14.98089895	Mus musculus glutathione S-transferase, mu 2 (Gstm2), mRNA.
119	Ppic	14.92166286	Mus musculus peptidylprolyl isomerase C (Ppic), mRNA.
120	Htra1	14.87834633	Mus musculus HtrA serine peptidase 1 (Htra1), mRNA.
121	4632401N01Rik	14.87217679	
122	Mxra8	14.60561279	Mus musculus matrix-remodelling associated 8 (Mxra8), mRNA.
123	Ptx3	14.47368284	
124	Mgp	14.41053925	Mus musculus matrix Gla protein (Mgp), mRNA.
125	Hsd11b1	14.34060143	
126	Gpx8	14.23827364	Mus musculus glutathione peroxidase 8 (putative) (Gpx8), mRNA.
127	Mmp14	14.21967435	Mus musculus matrix metallopeptidase 14 (membrane-inserted) (Mmp14), mRNA.
128	Col4a2	14.15140118	Mus musculus collagen, type IV, alpha 2 (Col4a2), mRNA.
129	**Tnc**	**13.99480871**	**Mus musculus tenascin C (Tnc), mRNA.**
130	Aoc3	13.97009189	Mus musculus amine oxidase, copper containing 3 (Aoc3), mRNA.
131	**Vim**	**13.96666284**	**Mus musculus vimentin (Vim), mRNA.**
132	Slit2	13.93949435	Mus musculus slit homolog 2 (Drosophila) (Slit2), mRNA.
133	Pla1a	13.86746308	
134	Bgn	13.84886971	Mus musculus biglycan (Bgn), mRNA.
135	Gstm2	13.74789223	Mus musculus glutathione S-transferase, mu 2 (Gstm2), mRNA.
136	Abi3bp	13.73997695	Mus musculus ABI gene family, member 3 (NESH) binding protein (Abi3bp), transcript variant 1, mRNA.
137	LOC100044411	13.5902578	PREDICTED: Mus musculus similar to Epidermal growth factor-containing fibulin-like extracellular matrix protein 1 (LOC100044411), mRNA.

138	Reg1	13.53237984	Mus musculus regenerating islet-derived 1 (Reg1), mRNA.
139	Col12a1	13.35145322	Mus musculus collagen, type XII, alpha 1 (Col12a1), mRNA.
140	Fhl1	13.34526116	Mus musculus four and a half LIM domains 1 (Fhl1), transcript variant 1, mRNA.
141	Hsd11b1	13.25327208	
142	Mdk	13.2393174	Mus musculus midkine (Mdk), transcript variant 3, mRNA.
143	Col6a3	13.20660539	Mus musculus collagen, type VI, alpha 3 (Col6a3), mRNA.
144	Rgs5	13.16965062	Mus musculus regulator of G-protein signaling 5 (Rgs5), mRNA.
145	Osr2	13.07743619	Mus musculus odd-skipped related 2 (Drosophila) (Osr2), mRNA.
146	Srpx	13.03833319	Mus musculus sushi-repeat-containing protein (Srpx), mRNA.
147	Dpep1	12.98310729	Mus musculus dipeptidase 1 (renal) (Dpep1), mRNA.
148	Col15a1	12.95927171	Mus musculus collagen, type XV, alpha 1 (Col15a1), mRNA.
149	Gpx7	12.94201412	Mus musculus glutathione peroxidase 7 (Gpx7), mRNA.
150	**Tnc**	**12.67230295**	
151	A730054J21Rik	12.6517808	
152	Igf1	12.60755528	Mus musculus insulin-like growth factor 1 (Igf1), transcript variant 1, mRNA.
153	LOC98434	12.50090703	
154	Islr	12.47561738	
155	Cnrip1	12.45654948	Mus musculus cannabinoid receptor interacting protein 1 (Cnrip1), mRNA.
156	Fbn1	12.41247655	
157	F2r	12.26877546	Mus musculus coagulation factor II (thrombin) receptor (F2r), mRNA.
158	C1qtnf2	12.26538819	Mus musculus C1q and tumor necrosis factor related protein 2 (C1qtnf2), mRNA.
159	Mfap4	12.20664523	Mus musculus microfibrillar-associated protein 4 (Mfap4), mRNA.
160	Cxcl14	12.20266731	
161	Pcolce2	12.18776088	Mus musculus procollagen C-endopeptidase enhancer 2 (Pcolce2), mRNA.
162	Gsn	12.17223337	Mus musculus gelsolin (Gsn), mRNA.
163	Gfpt2	12.17179246	Mus musculus glutamine fructose-6-phosphate transaminase 2 (Gfpt2), mRNA.
164	Col4a1	12.09414721	Mus musculus procollagen, type IV, alpha 1 (Col4a1), mRNA.
165	Mmp23	11.92813464	Mus musculus matrix metallopeptidase 23 (Mmp23), mRNA.
166	Col8a1	11.77557179	Mus musculus collagen, type VIII, alpha 1 (Col8a1), mRNA.
167	Nfatc4	11.72374054	Mus musculus nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 4 (Nfatc4), mRNA.
168	Fhl1	11.66187462	Mus musculus four and a half LIM domains 1 (Fhl1), transcript variant 2, mRNA.
169	Fstl	11.5832149	
170	scl0003799.1_2	11.57461662	
171	4930533K18Rik	11.47648384	
172	Mmp11	11.293549	Mus musculus matrix metallopeptidase 11 (Mmp11), mRNA.
173	Srpx	11.22391775	Mus musculus sushi-repeat-containing protein (Srpx), mRNA.
174	1110032E23Rik	11.22230787	Mus musculus RIKEN cDNA 1110032E23 gene (1110032E23Rik), mRNA.
175	Nuak1	11.10986255	Mus musculus NUAK family, SNF1-like kinase, 1 (Nuak1),

			mRNA.
176	Itm2a	11.09825089	Mus musculus integral membrane protein 2A (Itm2a), mRNA.
177	Emp3	10.99672096	Mus musculus epithelial membrane protein 3 (Emp3), mRNA.
178	Nav1	10.9777688	Mus musculus neuron navigator 1 (Nav1), mRNA.
179	Rbms3	10.96545835	Mus musculus RNA binding motif, single stranded interacting protein (Rbms3), mRNA.
180	Nrp1	10.89297674	Mus musculus neuropilin 1 (Nrp1), mRNA.
181	Mfap2	10.85579253	Mus musculus microfibrillar-associated protein 2 (Mfap2), mRNA.
182	Tpst1	10.84192048	Mus musculus protein-tyrosine sulfotransferase 1 (Tpst1), transcript variant 2, mRNA.
183	Col14a1	10.79839664	Mus musculus collagen, type XIV, alpha 1 (Col14a1), mRNA.
184	Slc10a6	10.70926301	Mus musculus solute carrier family 10 (sodium/bile acid cotransporter family), member 6 (Slc10a6), mRNA.
185	Sulf1	10.69460088	Mus musculus sulfatase 1 (Sulf1), mRNA.
186	Pdgfrl	10.56613483	Mus musculus platelet-derived growth factor receptor-like (Pdgfrl), mRNA.
187	Hoxc6	10.55822625	Mus musculus homeo box C6 (Hoxc6), mRNA.
188	Cxcl12	10.55634498	Mus musculus chemokine (C-X-C motif) ligand 12 (Cxcl12), transcript variant 1, mRNA.
189	Leprel2	10.53419854	Mus musculus leprecan-like 2 (Leprel2), mRNA.
190	Pmp22	10.53268759	Mus musculus peripheral myelin protein 22 (Pmp22), mRNA.
191	Ltbp4	10.52977486	Mus musculus latent transforming growth factor beta binding protein 4 (Ltbp4), mRNA.
192	Gstm2	10.49749813	Mus musculus glutathione S-transferase, mu 2 (Gstm2), mRNA.
193	Adam23	10.47713536	Mus musculus a disintegrin and metallopeptidase domain 23 (Adam23), mRNA.
194	Fmo1	10.45863328	Mus musculus flavin containing monooxygenase 1 (Fmo1), mRNA.
195	Pdlim2	10.4544022	Mus musculus PDZ and LIM domain 2 (Pdlim2), mRNA.
196	Plac8	10.41172549	Mus musculus placenta-specific 8 (Plac8), mRNA.
197	Ccl7	10.40747058	Mus musculus chemokine (C-C motif) ligand 7 (Ccl7), mRNA.
198	Igf1	10.38754511	Mus musculus insulin-like growth factor 1 (Igf1), transcript variant 1, mRNA.
199	Adamts4	10.38026328	
200	Emilin2	10.37389995	Mus musculus elastin microfibril interfacer 2 (Emilin2), mRNA.
201	Dcn	10.31850318	
202	Mxra8	10.26065219	Mus musculus matrix-remodelling associated 8 (Mxra8), mRNA.
203	Dkk3	10.08815324	Mus musculus dickkopf homolog 3 (Xenopus laevis) (Dkk3), mRNA.
204	Antxr1	10.08575451	Mus musculus anthrax toxin receptor 1 (Antxr1), mRNA.
205	Figf	10.07256421	Mus musculus c-fos induced growth factor (Figf), mRNA.
206	6720469N11Rik	10.066945	
207	Igsf10	10.06111353	
208	Rab31	10.02736051	Mus musculus RAB31, member RAS oncogene family (Rab31), mRNA.
209	Tmem204	9.920325228	Mus musculus transmembrane protein 204 (Tmem204), mRNA.
210	Dpysl3	9.889673519	Mus musculus dihydropyrimidinase-like 3 (Dpysl3), mRNA.
211	1200002N14Rik	9.88003959	Mus musculus RIKEN cDNA 1200002N14 gene (1200002N14Rik), mRNA.
212	Cyp7b1	9.868252777	Mus musculus cytochrome P450, family 7, subfamily b, polypeptide 1 (Cyp7b1), mRNA.

213	Gtl2	9.848498444	
214	Nrp1	9.759641328	Mus musculus neuropilin 1 (Nrp1), mRNA.
215	Lama2	9.75878158	Mus musculus laminin, alpha 2 (Lama2), mRNA.
216	Galntl1	9.726280213	Mus musculus UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-like 1 (Galntl1), mRNA.
217	Ugt1a10	9.589501087	Mus musculus UDP glycosyltransferase 1 family, polypeptide A10 (Ugt1a10), mRNA.
218	Bmp1	9.566890084	Mus musculus bone morphogenetic protein 1 (Bmp1), mRNA.
219	Ptx3	9.539933898	Mus musculus pentraxin related gene (Ptx3), mRNA.
220	E430002G05Rik	9.486767531	
221	Wisp1	9.358076979	Mus musculus WNT1 inducible signaling pathway protein 1 (Wisp1), mRNA.
222	Fcgrt	9.293550364	Mus musculus Fc receptor, IgG, alpha chain transporter (Fcgrt), mRNA.
223	Igfbp2	9.250518105	Mus musculus insulin-like growth factor binding protein 2 (Igfbp2), mRNA.
224	Adamts2	9.153997096	Mus musculus a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 2 (Adamts2), mRNA.
225	Sulf1	9.133300368	
226	Vgll3	9.126427489	PREDICTED: Mus musculus vestigial like 3 (Drosophila) (Vgll3), mRNA.
227	Lgals1	9.096955479	
228	Lpl	9.073909345	Mus musculus lipoprotein lipase (Lpl), mRNA.
229	0610007N19Rik	9.019834825	
230	Nav1	9.009249262	Mus musculus neuron navigator 1 (Nav1), mRNA.
231	Dkk3	9.003944861	Mus musculus dickkopf homolog 3 (Xenopus laevis) (Dkk3), mRNA.
232	H19	8.956570966	Mus musculus H19 fetal liver mRNA (H19), non-coding RNA.
233	Sema7a	8.950043087	Mus musculus sema domain, immunoglobulin domain (Ig), and GPI membrane anchor, (semaphorin) 7A (Sema7a), mRNA.
234	Hspg2	8.881396465	PREDICTED: Mus musculus perlecan (heparan sulfate proteoglycan 2) (Hspg2), mRNA.
235	Gdf10	8.860918267	Mus musculus growth differentiation factor 10 (Gdf10), mRNA.
236	Fkbp10	8.761584854	Mus musculus FK506 binding protein 10 (Fkbp10), mRNA.
237	B230343A10Rik	8.753546693	
238	5430433G21Rik	8.738872938	PREDICTED: Mus musculus RIKEN cDNA 5430433G21 gene (5430433G21Rik), mRNA.
239	6330404C01Rik	8.676706937	
240	Nnmt	8.569832134	Mus musculus nicotinamide N-methyltransferase (Nnmt), mRNA.
241	Cxcl1	8.566788776	Mus musculus chemokine (C-X-C motif) ligand 1 (Cxcl1), mRNA.
242	Il33	8.558051769	Mus musculus interleukin 33 (Il33), mRNA.
243	Ltbp4	8.540455004	Mus musculus latent transforming growth factor beta binding protein 4 (Ltbp4), mRNA.
244	Mfap2	8.535282497	Mus musculus microfibrillar-associated protein 2 (Mfap2), mRNA.
245	Col8a2	8.513234912	
246	Ybx3	8.476330854	
247	Ugt1a10	8.456354016	Mus musculus UDP glycosyltransferase 1 family, polypeptide A10 (Ugt1a10), mRNA.

248	Msc	8.443293072	Mus musculus musculin (Msc), mRNA.
249	LOC381283	8.365092809	
250	Igfbp2	8.267617353	Mus musculus insulin-like growth factor binding protein 2 (Igfbp2), mRNA.
251	C1qtnf3	8.170119331	Mus musculus C1q and tumor necrosis factor related protein 3 (C1qtnf3), mRNA.
252	Spsb1	8.153919249	Mus musculus splA/ryanodine receptor domain and SOCS box containing 1 (Spsb1), mRNA.
253	Kdelr3	8.152718986	Mus musculus KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3 (Kdelr3), mRNA.
254	4732458O05Rik	8.088876917	
255	Lrrn4cl	8.047674754	Mus musculus LRRN4 C-terminal like (Lrrn4cl), mRNA.
256	Bmp1	8.032494117	Mus musculus bone morphogenetic protein 1 (Bmp1), mRNA.
257	Angpt2	8.026455493	Mus musculus angiopoietin 2 (Angpt2), mRNA.
258	C3	8.018252266	
259	LOC100046044	7.961090315	PREDICTED: Mus musculus similar to COUP-TFI (LOC100046044), mRNA.
260	Abca8a	7.857631269	Mus musculus ATP-binding cassette, sub-family A (ABC1), member 8a (Abca8a), mRNA.
261	BC034076	7.827990125	Mus musculus cDNA sequence BC034076 (BC034076), mRNA.
262	C1s	7.821812507	Mus musculus complement component 1, s subcomponent (C1s), transcript variant 1, mRNA.
263	2310047A01Rik	7.805485724	PREDICTED: Mus musculus RIKEN cDNA 2310047A01 gene (2310047A01Rik), mRNA.
264	Antxr1	7.802778109	Mus musculus anthrax toxin receptor 1 (Antxr1), mRNA.
265	Lamc1	7.763953761	Mus musculus laminin, gamma 1 (Lamc1), mRNA.
266	Serpine2	7.75617188	Mus musculus serine (or cysteine) peptidase inhibitor, clade E, member 2 (Serpine2), mRNA.
267	Ccnd2	7.749395349	Mus musculus cyclin D2 (Ccnd2), mRNA.
268	Srpx2	7.72744666	Mus musculus sushi-repeat-containing protein, X-linked 2 (Srpx2), transcript variant 2, mRNA.
269	Tnfrsf11b	7.676307058	Mus musculus tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin) (Tnfrsf11b), mRNA.
270	Flnc	7.65353693	Mus musculus filamin C, gamma (Flnc), mRNA.
271	Ahnak2	7.629407811	Mus musculus AHNAK nucleoprotein 2 (Ahnak2), mRNA.
272	Ace	7.623792453	Mus musculus angiotensin I converting enzyme (peptidyl-dipeptidase A) 1 (Ace), transcript variant 1, mRNA.
273	Ly6a	7.61962991	Mus musculus lymphocyte antigen 6 complex, locus A (Ly6a), mRNA.
274	LOC100048332	7.600835989	PREDICTED: Mus musculus similar to a disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motif, 5 (aggrecanase-2) (LOC100048332), mRNA.
275	Lgmn	7.586219868	Mus musculus legumain (Lgmn), mRNA.
276	Fkbp9	7.577225363	Mus musculus FK506 binding protein 9 (Fkbp9), mRNA.
277	Bdh2	7.573490344	Mus musculus 3-hydroxybutyrate dehydrogenase, type 2 (Bdh2), mRNA.
278	Zeb1	7.56712924	Mus musculus zinc finger E-box binding homeobox 1 (Zeb1), mRNA.
279	Fcgrt	7.542821239	
280	Clec11a	7.537562454	Mus musculus C-type lectin domain family 11, member a (Clec11a), mRNA.
281	Cxcl12	7.524150068	Mus musculus chemokine (C-X-C motif) ligand 12 (Cxcl12), transcript variant 3, mRNA.
282	Col3a1	7.520186864	Mus musculus collagen, type III, alpha 1 (Col3a1), mRNA.
283	C4b	7.502715142	Mus musculus complement component 4B (Childo blood group) (C4b), mRNA. XM_921663 XM_921673 XM_921676 XM_921678
284	Lhfp	7.49817863	Mus musculus lipoma HMGIC fusion partner (Lhfp), mRNA.

285	E430002G05Rik	7.443727172	Mus musculus RIKEN cDNA E430002G05 gene (E430002G05Rik), mRNA.
286	Nkd2	7.400644254	Mus musculus naked cuticle 2 homolog (Drosophila) (Nkd2), mRNA.
287	Bmp1	7.361966014	Mus musculus bone morphogenetic protein 1 (Bmp1), mRNA.
288	Vcam1	7.287574407	Mus musculus vascular cell adhesion molecule 1 (Vcam1), mRNA.
289	Tshz3	7.278179942	Mus musculus teashirt zinc finger family member 3 (Tshz3), mRNA.
290	Emilin2	7.254095849	
291	Ifitm1	7.189531058	Mus musculus interferon induced transmembrane protein 1 (Ifitm1), mRNA.
292	Hmgcs2	7.181414572	Mus musculus 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (Hmgcs2), nuclear gene encoding mitochondrial protein, mRNA.
293	Hspg2	7.173981521	
294	Csrp2	7.112671625	Mus musculus cysteine and glycine-rich protein 2 (Csrp2), mRNA.
295	Hbb-b1	7.098136446	Mus musculus hemoglobin, beta adult major chain (Hbb-b1), mRNA.
296	Lbh	6.990659631	Mus musculus limb-bud and heart (Lbh), mRNA.
297	Ecm1	6.97615217	Mus musculus extracellular matrix protein 1 (Ecm1), mRNA.
298	Itga11	6.973263024	Mus musculus integrin alpha 11 (Itga11), mRNA.
299	Saa3	6.95296686	Mus musculus serum amyloid A 3 (Saa3), mRNA.
300	C4a	6.937980951	Mus musculus complement component 4A (Rodgers blood group) (C4a), mRNA.
301	Pdlim2	6.833377099	Mus musculus PDZ and LIM domain 2 (Pdlim2), mRNA.
302	Olfml1	6.806360733	Mus musculus olfactomedin-like 1 (Olfml1), mRNA.
303	2310014H01Rik	6.789759138	PREDICTED: Mus musculus RIKEN cDNA 2310014H01 gene, transcript variant 3 (2310014H01Rik), mRNA.
304	9530064J02	6.7670163	
305	Pdgfrb	6.764712002	Mus musculus platelet derived growth factor receptor, beta polypeptide (Pdgfrb), mRNA.
306	Fxyd5	6.741233889	Mus musculus FXYD domain-containing ion transport regulator 5 (Fxyd5), mRNA.
307	Rgs4	6.70527763	Mus musculus regulator of G-protein signaling 4 (Rgs4), mRNA.
308	B3gnt9	6.694206763	Mus musculus UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 9 (B3gnt9), mRNA.
309	Crlf1	6.691824294	Mus musculus cytokine receptor-like factor 1 (Crlf1), mRNA.
310	Adra2a	6.690351503	Mus musculus adrenergic receptor, alpha 2a (Adra2a), mRNA.
311	Hbb-b1	6.663288727	
312	Gpr176	6.656021427	Mus musculus G protein-coupled receptor 176 (Gpr176), mRNA.
313	Fbln1	6.6170484	Mus musculus fibulin 1 (Fbln1), mRNA.
314	Crtap	6.570610638	Mus musculus cartilage associated protein (Crtap), mRNA.
315	Avpr1a	6.554498008	Mus musculus arginine vasopressin receptor 1A (Avpr1a), mRNA.
316	H6pd	6.502122685	Mus musculus hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase) (H6pd), mRNA.
317	Adap1	6.496166082	Mus musculus ArfGAP with dual PH domains 1 (Adap1), mRNA.
318	Hba-a1	6.495490043	Mus musculus hemoglobin alpha, adult chain 1 (Hba-a1), mRNA.
319	D1Ertd471e	6.487376378	
320	Crtap	6.462248258	Mus musculus cartilage associated protein (Crtap), mRNA.
321	Mrgprf	6.459751937	Mus musculus MAS-related GPR, member F (Mrgprf), mRNA.

322	C4a	6.452434731	Mus musculus complement component 4A (Rodgers blood group) (C4a), mRNA.
323	Clec3b	6.451191499	Mus musculus C-type lectin domain family 3, member b (Clec3b), mRNA.
324	Cyp1b1	6.435526321	Mus musculus cytochrome P450, family 1, subfamily b, polypeptide 1 (Cyp1b1), mRNA.
325	Emb	6.430169501	Mus musculus embigin (Emb), mRNA.
326	Ccl11	6.419525408	Mus musculus small chemokine (C-C motif) ligand 11 (Ccl11), mRNA.
327	Crim2	6.403401936	Mus musculus cysteine rich BMP regulator 2 (chordin like) (Crim2), mRNA.
328	Msx1	6.384307352	Mus musculus homeobox, msh-like 1 (Msx1), mRNA.
329	H19	6.371270015	Mus musculus H19 fetal liver mRNA (H19), non-coding RNA.
330	Ppap2b	6.348022187	Mus musculus phosphatidic acid phosphatase type 2B (Ppap2b), mRNA.
331	Tppp3	6.328415276	Mus musculus tubulin polymerization-promoting protein family member 3 (Tppp3), mRNA.
332	Lamc1	6.253883258	Mus musculus laminin, gamma 1 (Lamc1), mRNA.
333	Akr1b8	6.25130073	Mus musculus aldo-keto reductase family 1, member B8 (Akr1b8), mRNA.
334	Lhfp	6.238413758	Mus musculus lipoma HMGIC fusion partner (Lhfp), mRNA.
335	Twist2	6.170337225	
336	P4ha2	6.155396133	
337	Hbb-b1	6.125085656	Mus musculus hemoglobin, beta adult major chain (Hbb-b1), mRNA.
338	Twist1	6.120875056	Mus musculus twist gene homolog 1 (Drosophila) (Twist1), mRNA.
339	Antxr2	6.095402066	Mus musculus anthrax toxin receptor 2 (Antxr2), mRNA.
340	LOC214403	6.087443937	
341	Lhfp	6.070946877	Mus musculus lipoma HMGIC fusion partner (Lhfp), mRNA.
342	Igfbp3	6.069970871	Mus musculus insulin-like growth factor binding protein 3 (Igfbp3), mRNA.
343	Tbc1d2b	6.068722696	Mus musculus TBC1 domain family, member 2B (Tbc1d2b), mRNA.
344	Rasa3	6.045737317	Mus musculus RAS p21 protein activator 3 (Rasa3), mRNA.
345	Epha3	6.038110421	Mus musculus Eph receptor A3 (Epha3), mRNA.
346	Tgfbr2	6.032781351	Mus musculus transforming growth factor, beta receptor II (Tgfbr2), transcript variant 1, mRNA.
347	Col1a2	5.969973792	Mus musculus collagen, type I, alpha 2 (Col1a2), mRNA.
348	Gpc1	5.965503473	Mus musculus glypican 1 (Gpc1), mRNA.
349	Bmper	5.959845423	Mus musculus BMP-binding endothelial regulator (Bmper), mRNA.
350	Inmt	5.954676497	Mus musculus indolethylamine N-methyltransferase (Inmt), mRNA.
351	2210023G05Rik	5.941521573	Mus musculus RIKEN cDNA 2210023G05 gene (2210023G05Rik), mRNA.
352	Hbb-b1	5.913740116	
353	Glipr2	5.909285785	Mus musculus GLI pathogenesis-related 2 (Glipr2), mRNA.
354	Ugt1a10	5.885704487	Mus musculus UDP glycosyltransferase 1 family, polypeptide A10 (Ugt1a10), mRNA.
355	Tmem100	5.864433868	Mus musculus transmembrane protein 100 (Tmem100), mRNA.
356	Dlk1	5.855761782	Mus musculus delta-like 1 homolog (Drosophila) (Dlk1), mRNA.
357	Dnm3os	5.84507897	Mus musculus dynamin 3, opposite strand (Dnm3os), non-coding RNA.
358	Hexa	5.840892392	Mus musculus hexosaminidase A (Hexa), mRNA.
359	Sfrp2	5.833282582	Mus musculus secreted frizzled-related protein 2 (Sfrp2), mRNA.
360	Apcdd1	5.82763171	Mus musculus adenomatosis polyposis coli down-regulated

			1 (Apcdd1), mRNA.
361	Rhoj	5.810529961	Mus musculus ras homolog gene family, member J (Rhoj), mRNA.
362	Agtr1a	5.773748093	Mus musculus angiotensin II receptor, type 1a (Agtr1a), mRNA.
363	AI607873	5.759651473	PREDICTED: Mus musculus expressed sequence AI607873 (AI607873), mRNA.
364	B3gnt9	5.752546721	Mus musculus UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 9 (B3gnt9), mRNA.
365	Sfrp1	5.722724878	Mus musculus secreted frizzled-related protein 1 (Sfrp1), mRNA.
366	Gadd45g	5.690641453	Mus musculus growth arrest and DNA-damage-inducible 45 gamma (Gadd45g), mRNA.
367	LOC100045680	5.65258775	PREDICTED: Mus musculus similar to complement C4 (LOC100045680), mRNA.
368	Hbb-b1	5.645586056	Mus musculus hemoglobin, beta adult major chain (Hbb-b1), mRNA.
369	Bcat1	5.635819657	Mus musculus branched chain aminotransferase 1, cytosolic (Bcat1), transcript variant 2, mRNA.
370	Prnp	5.629650665	Mus musculus prion protein (Prnp), mRNA.
371	C1s	5.622593016	Mus musculus complement component 1, s subcomponent (C1s), mRNA.
372	Fap	5.610018287	Mus musculus fibroblast activation protein (Fap), mRNA.
373	Tnfaip2	5.593331022	Mus musculus tumor necrosis factor, alpha-induced protein 2 (Tnfaip2), mRNA.
374	Vegfc	5.586212784	Mus musculus vascular endothelial growth factor C (Vegfc), mRNA.
375	Angptl4	5.552567496	Mus musculus angiopoietin-like 4 (Angptl4), mRNA.
376	Gucy1a3	5.535833951	Mus musculus guanylate cyclase 1, soluble, alpha 3 (Gucy1a3), mRNA.
377	Olfm1	5.530388783	Mus musculus olfactomedin 1 (Olfm1), transcript variant 1, mRNA. XM_923993 XM_923996 XM_924001
378	Dpp7	5.510989633	Mus musculus dipeptidylpeptidase 7 (Dpp7), mRNA.
379	E130203B14Rik	5.507898256	Mus musculus RIKEN cDNA E130203B14 gene (E130203B14Rik), mRNA.
380	Igsf10	5.507376589	PREDICTED: Mus musculus immunoglobulin superfamily, member 10, transcript variant 4 (Igsf10), mRNA.
381	Mmp13	5.505782672	Mus musculus matrix metallopeptidase 13 (Mmp13), mRNA.
382	Ptgis	5.500260194	Mus musculus prostaglandin I2 (prostacyclin) synthase (Ptgis), mRNA.
383	Col18a1	5.492024252	Mus musculus procollagen, type XVIII, alpha 1 (Col18a1), mRNA.
384	LOC100048554	5.490793963	PREDICTED: Mus musculus similar to monocyte chemoattractant protein-2 (MCP-2) (LOC100048554), mRNA.
385	LOC631037	5.467126093	PREDICTED: Mus musculus similar to CYP4B1 (LOC631037), mRNA.
386	Tmem119	5.456973331	Mus musculus transmembrane protein 119 (Tmem119), mRNA.
387	Syde1	5.440898397	PREDICTED: Mus musculus synapse defective 1, Rho GTPase, homolog 1 (C. elegans) (Syde1), mRNA.
388	Prkcdbp	5.429395253	Mus musculus protein kinase C, delta binding protein (Prkcdbp), mRNA.
389	Nav1	5.395931352	
390	Hbb-b1	5.392454602	
391	Mmp10	5.381211415	Mus musculus matrix metallopeptidase 10 (Mmp10), mRNA.
392	BC020108	5.38074711	
393	LOC100045780	5.377562508	PREDICTED: Mus musculus similar to metalloprotease-disintegrin meltrin beta (LOC100045780), mRNA.
394	Pcolce2	5.377187401	Mus musculus procollagen C-endopeptidase enhancer 2 (Pcolce2), mRNA.

395	Fhl2	5.371558153	Mus musculus four and a half LIM domains 2 (Fhl2), mRNA.
396	Cdkn2b	5.36666227	Mus musculus cyclin-dependent kinase inhibitor 2B (p15, inhibits CDK4) (Cdkn2b), mRNA.
397	Nox4	5.362865481	Mus musculus NADPH oxidase 4 (Nox4), mRNA.
398	C1qtnf5	5.353976382	Mus musculus C1q and tumor necrosis factor related protein 5 (C1qtnf5), transcript variant 1, mRNA.
399	Lgi2	5.338651634	Mus musculus leucine-rich repeat LGI family, member 2 (Lgi2), mRNA.
400	Arhgap24	5.298224244	Mus musculus Rho GTPase activating protein 24 (Arhgap24), transcript variant 1, mRNA.
401	Snx24	5.26805474	Mus musculus sorting nexing 24 (Snx24), mRNA.
402	Flrt2	5.255317218	Mus musculus fibronectin leucine rich transmembrane protein 2 (Flrt2), mRNA.
403	Lrrc17	5.245065145	PREDICTED: Mus musculus leucine rich repeat containing 17, transcript variant 1 (Lrrc17), mRNA.
404	Rgl1	5.222704127	Mus musculus ral guanine nucleotide dissociation stimulator,-like 1 (Rgl1), mRNA. XM_924428
405	Serpinh1	5.220445571	Mus musculus serine (or cysteine) peptidase inhibitor, clade H, member 1 (Serpinh1), mRNA.
406	Cercam	5.214353606	Mus musculus cerebral endothelial cell adhesion molecule (Cercam), mRNA.
407	Olfm1	5.205603895	Mus musculus olfactomedin 1 (Olfm1), transcript variant 4, mRNA.
408	C4b	5.204238332	Mus musculus complement component 4B (Childo blood group) (C4b), mRNA. XM_921663 XM_921673 XM_921676 XM_921678
409	Epb4.1l3	5.18599776	Mus musculus erythrocyte protein band 4.1-like 3 (Epb4.1l3), mRNA.
410	Palld	5.157538801	Mus musculus palladin, cytoskeletal associated protein (Palld), mRNA.
411	Thbs3	5.144567551	Mus musculus thrombospondin 3 (Thbs3), mRNA.
412	Plekha4	5.140839916	Mus musculus pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 4 (Plekha4), mRNA.
413	Cpz	5.13657056	Mus musculus carboxypeptidase Z (Cpz), mRNA.
414	Nope	5.126967854	Mus musculus neighbor of Punc E11 (Nope), mRNA.
415	Gstm1	5.12583511	Mus musculus glutathione S-transferase, mu 1 (Gstm1), mRNA.
416	Prkcdbp	5.11828371	Mus musculus protein kinase C, delta binding protein (Prkcdbp), mRNA.
417	Col6a2	5.092910982	
418	LOC100047339	5.058052108	PREDICTED: Mus musculus similar to lysyl oxidase-like 2 (LOC100047339), mRNA.
419	Ednra	5.054522861	Mus musculus endothelin receptor type A (Ednra), mRNA.
420	Enc1	5.049121692	Mus musculus ectodermal-neural cortex 1 (Enc1), mRNA.
421	1500015O10Rik	5.03376504	Mus musculus RIKEN cDNA 1500015O10 gene (1500015O10Rik), mRNA.
422	Tns1	4.998380128	PREDICTED: Mus musculus tensin 1 (Tns1), mRNA.
423	Fbln2	4.988346562	Mus musculus fibulin 2 (Fbln2), transcript variant 1, mRNA.
424	Efemp2	4.988084206	Mus musculus epidermal growth factor-containing fibulin-like extracellular matrix protein 2 (Efemp2), mRNA.
425	Fbln2	4.98218623	Mus musculus fibulin 2 (Fbln2), transcript variant 2, mRNA.
426	Ntrk2	4.962973058	Mus musculus neurotrophic tyrosine kinase, receptor, type 2 (Ntrk2), transcript variant 1, mRNA.
427	Sfrp1	4.960135914	
428	Tgfbi	4.94676836	Mus musculus transforming growth factor, beta induced (Tgfbi), mRNA.
429	Hbb-b1	4.945302652	
430	Pmp22	4.94469321	Mus musculus peripheral myelin protein 22 (Pmp22), mRNA.
431	1200009O22R	4.928538517	Mus musculus RIKEN cDNA 1200009O22 gene

	ik		(1200009O22Rik), mRNA.
432	Adora2b	4.919606764	Mus musculus adenosine A2b receptor (Adora2b), mRNA.
433	Hspb8	4.900864319	Mus musculus heat shock protein 8 (Hspb8), mRNA.
434	Slc10a6	4.835781344	Mus musculus solute carrier family 10 (sodium/bile acid cotransporter family), member 6 (Slc10a6), mRNA.
435	Pscd3	4.814836213	Mus musculus pleckstrin homology, Sec7 and coiled-coil domains 3 (Pscd3), mRNA.
436	Ugcg	4.808652342	
437	Rgs16	4.797272207	Mus musculus regulator of G-protein signaling 16 (Rgs16), mRNA.
438	Plau	4.77643973	Mus musculus plasminogen activator, urokinase (Plau), mRNA.
439	Ccnd2	4.767402436	
440	Pdlim2	4.762010824	Mus musculus PDZ and LIM domain 2 (Pdlim2), mRNA.
441	Prrx1	4.754127085	Mus musculus paired related homeobox 1 (Prrx1), transcript variant 2, mRNA.
442	Gli3	4.752647745	Mus musculus GLI-Kruppel family member GLI3 (Gli3), mRNA.
443	Adamts1	4.747836998	Mus musculus a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 1 (Adamts1), mRNA.
444	Timp2	4.747585368	Mus musculus tissue inhibitor of metalloproteinase 2 (Timp2), mRNA.
445	Emilin1	4.734763949	
446	B430201A12Rik	4.721831546	PREDICTED: Mus musculus RIKEN cDNA B430201A12 gene (B430201A12Rik), mRNA.
447	Frmd6	4.690317255	Mus musculus FERM domain containing 6 (Frmd6), mRNA.
448	Sema5a	4.670822095	Mus musculus sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5A (Sema5a), mRNA.
449	C4b	4.656553995	Mus musculus complement component 4B (Childo blood group) (C4b), mRNA. XM_921663 XM_921673 XM_921676 XM_921678
450	Pftk1	4.653916766	Mus musculus PFTAIRE protein kinase 1 (Pftk1), mRNA.
451	Entpd2	4.640691497	Mus musculus ectonucleoside triphosphate diphosphohydrolase 2 (Entpd2), mRNA.
452	Wnt2	4.624019829	Mus musculus wingless-related MMTV integration site 2 (Wnt2), mRNA.
453	Mxra7	4.604726233	Mus musculus matrix-remodelling associated 7 (Mxra7), mRNA.
454	Prrx2	4.59972339	
455	Arhgdib	4.597389683	Mus musculus Rho, GDP dissociation inhibitor (GDI) beta (Arhgdib), mRNA.
456	Igfbp3	4.582770632	Mus musculus insulin-like growth factor binding protein 3 (Igfbp3), mRNA.
457	LOC100046883	4.58146411	PREDICTED: Mus musculus similar to CKLF-like MARVEL transmembrane domain containing 3 (LOC100046883), mRNA.
458	Angptl1	4.572594928	Mus musculus angiopoietin-like 1 (Angptl1), mRNA.
459	Palld	4.569097527	PREDICTED: Mus musculus palladin, cytoskeletal associated protein, transcript variant 6 (Palld), mRNA.
460	Snai2	4.565399611	Mus musculus snail homolog 2 (Drosophila) (Snai2), mRNA.
461	Igfbp7	4.542608254	Mus musculus insulin-like growth factor binding protein 7 (Igfbp7), mRNA.
462	Ccdc3	4.534040135	
463	Klf4	4.527074658	Mus musculus Kruppel-like factor 4 (gut) (Klf4), mRNA.
464	LOC100047419	4.509440587	PREDICTED: Mus musculus similar to c-Maf long form (LOC100047419), mRNA.
465	Tgfb1i1	4.509108841	Mus musculus transforming growth factor beta 1 induced transcript 1 (Tgfb1i1), mRNA.

466	Podn	4.498653098	Mus musculus podocan (Podn), mRNA.
467	1200015N20Rik	4.478369979	
468	Fst	4.471524522	Mus musculus follistatin (Fst), mRNA.
469	Ugcg	4.47030687	Mus musculus UDP-glucose ceramide glucosyltransferase (Ugcg), mRNA.
470	Dpysl2	4.457439995	Mus musculus dihydropyrimidinase-like 2 (Dpysl2), mRNA.
471	Hic1	4.450206366	Mus musculus hypermethylated in cancer 1 (Hic1), transcript variant 1, mRNA.
472	2010004A03Rik	4.4419843	Mus musculus RIKEN cDNA 2010004A03 gene (2010004A03Rik), mRNA.
473	Crabp1	4.408758003	Mus musculus cellular retinoic acid binding protein I (Crabp1), mRNA.
474	Fcgr2b	4.38846322	Mus musculus Fc receptor, IgG, low affinity IIb (Fcgr2b), transcript variant 2, mRNA.
475	Lhfpl2	4.386253632	Mus musculus lipoma HMGIC fusion partner-like 2 (Lhfpl2), mRNA.
476	Mdk	4.38506791	Mus musculus midkine (Mdk), transcript variant 1, mRNA.
477	Fgf7	4.377519277	Mus musculus fibroblast growth factor 7 (Fgf7), mRNA.
478	Angptl2	4.355187382	Mus musculus angiopoietin-like 2 (Angptl2), mRNA.
479	Fam13c	4.342874705	Mus musculus family with sequence similarity 13, member C (Fam13c), mRNA.
480	Epdr1	4.340659051	Mus musculus ependymin related protein 1 (zebrafish) (Epdr1), mRNA.
481	Ifi27	4.34043673	Mus musculus interferon, alpha-inducible protein 27 (Ifi27), mRNA.
482	Larp6	4.337924773	Mus musculus La ribonucleoprotein domain family, member 6 (Larp6), mRNA.
483	Hoxc9	4.333790914	Mus musculus homeo box C9 (Hoxc9), mRNA.
484	Hdac7	4.331227078	Mus musculus histone deacetylase 7 (Hdac7), mRNA. XM_001003311 XM_001003316 XM_001004168 XM_001004172 XM_001004178 XM_001004180 XM_001004182 XM_001004187 XM_001004194 XM_001004198
485	Eml1	4.330343362	Mus musculus echinoderm microtubule associated protein like 1 (Eml1), transcript variant 1, mRNA. XM_901775 XM_901779 XM_978946 XM_978978 XM_979015 XM_979053 XM_979124 XM_979166 XM_979204
486	Mylk	4.322904976	Mus musculus myosin, light polypeptide kinase (Mylk), mRNA.
487	Lepre1	4.319777906	Mus musculus leprecan 1 (Lepre1), transcript variant 1, mRNA.
488	Id3	4.307567608	Mus musculus inhibitor of DNA binding 3 (Id3), mRNA.
489	Papss2	4.305862379	Mus musculus 3'-phosphoadenosine 5'-phosphosulfate synthase 2 (Papss2), mRNA.
490	Fkbp14	4.295855041	Mus musculus FK506 binding protein 14 (Fkbp14), mRNA.
491	Fscn1	4.294279192	Mus musculus fascin homolog 1, actin bundling protein (Strongylocentrotus purpuratus) (Fscn1), mRNA.
492	Inhba	4.284512396	Mus musculus inhibin beta-A (Inhba), mRNA.
493	Emid2	4.284184633	Mus musculus EMI domain containing 2 (Emid2), mRNA.
494	Gpr23	4.270372953	
495	Timp3	4.244932016	Mus musculus tissue inhibitor of metalloproteinase 3 (Timp3), mRNA.
496	Serpina1b	4.243623953	Mus musculus serine (or cysteine) preptidase inhibitor, clade A, member 1b (Serpina1b), mRNA.
497	Slit3	4.239530102	
498	Tbc1d2b	4.228200931	Mus musculus TBC1 domain family, member 2B (Tbc1d2b), mRNA.
499	Ebf2	4.20390548	Mus musculus early B-cell factor 2 (Ebf2), mRNA.
500	Ramp2	4.199349443	Mus musculus receptor (calcitonin) activity modifying protein 2 (Ramp2), mRNA.

501	Acta2	4.199110926	
502	Meox1	4.197568662	Mus musculus mesenchyme homeobox 1 (Meox1), mRNA.
503	Moxd1	4.187369732	Mus musculus monooxygenase, DBH-like 1 (Moxd1), mRNA.
504	Ppap2b	4.186768068	Mus musculus phosphatidic acid phosphatase type 2B (Ppap2b), mRNA.
505	Gadd45b	4.18350809	Mus musculus growth arrest and DNA-damage-inducible 45 beta (Gadd45b), mRNA.
506	Ncam1	4.168593966	Mus musculus neural cell adhesion molecule 1 (Ncam1), transcript variant 1, mRNA.
507	Pvrl3	4.159275865	Mus musculus poliovirus receptor-related 3 (Pvrl3), transcript variant alpha, mRNA.
508	Ifi202b	4.156378254	Mus musculus interferon activated gene 202B (Ifi202b), mRNA.
509	Ccl2	4.138499247	Mus musculus chemokine (C-C motif) ligand 2 (Ccl2), mRNA.
510	Tgfb1i1	4.138408835	Mus musculus transforming growth factor beta 1 induced transcript 1 (Tgfb1i1), mRNA.
511	Cd248	4.132471638	Mus musculus CD248 antigen, endosialin (Cd248), mRNA.
512	Serpinh1	4.125075992	
513	Fn1	4.121693703	Mus musculus fibronectin 1 (Fn1), mRNA.
514	Ror2	4.118040162	Mus musculus receptor tyrosine kinase-like orphan receptor 2 (Ror2), mRNA.
515	Hspb8	4.115789016	Mus musculus heat shock protein 8 (Hspb8), mRNA.
516	Edg2	4.114688284	
517	Il11ra1	4.112155465	Mus musculus interleukin 11 receptor, alpha chain 1 (Il11ra1), mRNA.
518	Prrx1	4.109263218	
519	Calu	4.108320913	Mus musculus calumenin (Calu), transcript variant 1, mRNA.
520	Rgs10	4.105957554	Mus musculus regulator of G-protein signalling 10 (Rgs10), mRNA.
521	Pdia5	4.096982649	Mus musculus protein disulfide isomerase associated 5 (Pdia5), mRNA.
522	Gprc5b	4.079881579	Mus musculus G protein-coupled receptor, family C, group 5, member B (Gprc5b), mRNA.
523	Adh1	4.061195457	Mus musculus alcohol dehydrogenase 1 (class I) (Adh1), mRNA.
524	B430305P08Rik	4.05116071	
525	Cd47	4.040192953	Mus musculus CD47 antigen (Rh-related antigen, integrin-associated signal transducer) (Cd47), mRNA.
526	Kremen	4.035825323	
527	Tnc	3.984531463	Mus musculus tenascin C (Tnc), mRNA.
528	Ak1	3.977036046	Mus musculus adenylate kinase 1 (Ak1), mRNA.
529	Cep170	3.975884225	Mus musculus centrosomal protein 170 (Cep170), mRNA.
530	LOC100047200	3.974096024	PREDICTED: Mus musculus similar to T-box 3 protein (LOC100047200), mRNA.
531	Meis2	3.965613813	Mus musculus Meis homeobox 2 (Meis2), transcript variant 2, mRNA.
532	Acta2	3.958913088	Mus musculus actin, alpha 2, smooth muscle, aorta (Acta2), mRNA.
533	Lgals3bp	3.952935763	Mus musculus lectin, galactoside-binding, soluble, 3 binding protein (Lgals3bp), mRNA.
534	Mylk	3.951876438	Mus musculus myosin, light polypeptide kinase (Mylk), mRNA.
535	Tmem45a	3.951013258	Mus musculus transmembrane protein 45a (Tmem45a), mRNA.
536	Rassf2	3.944571743	Mus musculus Ras association (RalGDS/AF-6) domain family member 2 (Rassf2), mRNA.
537	Dpp4	3.941721695	Mus musculus dipeptidylpeptidase 4 (Dpp4), mRNA.
538	LOC10004556	3.941014782	PREDICTED: Mus musculus similar to purine nucleoside

	7		phosphorylase (LOC100045567), mRNA.
539	6430511F03	3.932307008	
540	MGC41689	3.926167661	
541	Emp1	3.925984664	Mus musculus epithelial membrane protein 1 (Emp1), mRNA.
542	Sema5a	3.922217117	
543	Icam1	3.915216655	Mus musculus intercellular adhesion molecule 1 (Icam1), mRNA.
544	Sdk1	3.910325094	Mus musculus sidekick homolog 1 (chicken) (Sdk1), mRNA.
545	Wipf1	3.900819437	Mus musculus WAS/WASL interacting protein family, member 1 (Wipf1), mRNA.
546	Pygl	3.89256788	Mus musculus liver glycogen phosphorylase (Pygl), mRNA.
547	Gli2	3.88701841	Mus musculus GLI-Kruppel family member GLI2 (Gli2), mRNA.
548	Klf4	3.873314389	Mus musculus Kruppel-like factor 4 (gut) (Klf4), mRNA.
549	Rftn1	3.866958425	Mus musculus raftlin lipid raft linker 1 (Rftn1), mRNA.
550	Tnfsf13b	3.865518993	
551	Zeb2	3.862980214	Mus musculus zinc finger E-box binding homeobox 2 (Zeb2), transcript variant 2, mRNA.
552	Rcn3	3.855784511	Mus musculus reticulocalbin 3, EF-hand calcium binding domain (Rcn3), mRNA.
553	Pip4k2a	3.836679409	Mus musculus phosphatidylinositol-5-phosphate 4-kinase, type II, alpha (Pip4k2a), mRNA.
554	Tek	3.835260998	Mus musculus endothelial-specific receptor tyrosine kinase (Tek), mRNA.
555	AI115600	3.834594177	PREDICTED: Mus musculus expressed sequence AI115600 (AI115600), mRNA.
556	Ptrf	3.831666768	Mus musculus polymerase I and transcript release factor (Ptrf), mRNA.
557	Magee1	3.786953719	Mus musculus melanoma antigen, family E, 1 (Magee1), mRNA.
558	Plxdc1	3.78297656	Mus musculus plexin domain containing 1 (Plxdc1), mRNA.
559	Nnmt	3.782062493	Mus musculus nicotinamide N-methyltransferase (Nnmt), mRNA.
560	En1	3.779463664	
561	Pard6g	3.77164835	Mus musculus par-6 partitioning defective 6 homolog gamma (C. elegans) (Pard6g), mRNA.
562	Adm	3.753670118	Mus musculus adrenomedullin (Adm), mRNA.
563	Klf4	3.743955282	Mus musculus Kruppel-like factor 4 (gut) (Klf4), mRNA.
564	Vcan	3.736829115	Mus musculus versican (Vcan), transcript variant 1, mRNA.
565	Cbr3	3.726992773	Mus musculus carbonyl reductase 3 (Cbr3), mRNA.
566	LOC100043822	3.725877207	PREDICTED: Mus musculus hypothetical protein LOC100043822 (LOC100043822), mRNA.
567	Gm106	3.724318701	Mus musculus gene model 106, (NCBI) (Gm106), mRNA.
568	Rasl11b	3.723342891	Mus musculus RAS-like, family 11, member B (Rasl11b), mRNA.
569	Palm	3.721679806	Mus musculus paralemmin (Palm), mRNA.
570	D130032J17Rik	3.717827164	
571	LOC100045981	3.709831039	PREDICTED: Mus musculus similar to synaptotagmin XI (LOC100045981), mRNA.
572	Clip4	3.708089268	Mus musculus CAP-GLY domain containing linker protein family, member 4 (Clip4), mRNA.
573	3830612M24	3.683409541	
574	Gpnmb	3.680596775	Mus musculus glycoprotein (transmembrane) nmb (Gpnmb), mRNA.
575	Cyp4b1	3.67698453	Mus musculus cytochrome P450, family 4, subfamily b, polypeptide 1 (Cyp4b1), mRNA.
576	Tgfbr3	3.668251385	Mus musculus transforming growth factor, beta receptor III

			(Tgfbr3), mRNA.
577	Egr2	3.666314795	
578	Ifitm3	3.664312778	Mus musculus interferon induced transmembrane protein 3 (Ifitm3), mRNA.
579	Gadd45g	3.660280398	Mus musculus growth arrest and DNA-damage-inducible 45 gamma (Gadd45g), mRNA.
580	Hoxd8	3.659365193	Mus musculus homeo box D8 (Hoxd8), mRNA.
581	Qpct	3.658413624	Mus musculus glutaminyl-peptide cyclotransferase (glutaminyl cyclase) (Qpct), mRNA.
582	Ebf2	3.654617266	Mus musculus early B-cell factor 2 (Ebf2), mRNA.
583	D10Ertd610e	3.650436698	Mus musculus DNA segment, Chr 10, ERATO Doi 610, expressed (D10Ertd610e), mRNA.
584	Pcsk6	3.647645426	PREDICTED: Mus musculus proprotein convertase subtilisin/kexin type 6, transcript variant 4 (Pcsk6), mRNA.
585	Lrrk1	3.64510029	Mus musculus leucine-rich repeat kinase 1 (Lrrk1), mRNA.
586	Gas7	3.629626594	Mus musculus growth arrest specific 7 (Gas7), mRNA.
587	Larp6	3.623947602	Mus musculus La ribonucleoprotein domain family, member 6 (Larp6), mRNA.
588	Enpp2	3.62100446	Mus musculus ectonucleotide pyrophosphatase/phosphodiesterase 2 (Enpp2), mRNA.
589	Itgb5	3.614958138	Mus musculus integrin beta 5 (Itgb5), mRNA.
590	Marcks	3.612535024	Mus musculus myristoylated alanine rich protein kinase C substrate (Marcks), mRNA.
591	Serpina1b	3.611669653	
592	Myl9	3.607691967	PREDICTED: Mus musculus myosin, light polypeptide 9, regulatory (Myl9), mRNA.
593	Actg2	3.599723104	Mus musculus actin, gamma 2, smooth muscle, enteric (Actg2), mRNA.
594	Acta2	3.596437311	Mus musculus actin, alpha 2, smooth muscle, aorta (Acta2), mRNA.
595	Gm22	3.595226124	PREDICTED: Mus musculus gene model 22, (NCBI) (Gm22), mRNA.
596	2810022L02Rik	3.594467067	Mus musculus RIKEN cDNA 2810022L02 gene (2810022L02Rik), mRNA.
597	Glipr2	3.592952285	
598	Kng1	3.58600218	Mus musculus kininogen 1 (Kng1), mRNA.
599	Acta2	3.583407209	Mus musculus actin, alpha 2, smooth muscle, aorta (Acta2), mRNA.
600	Ebf3	3.579098377	Mus musculus early B-cell factor 3 (Ebf3), mRNA.
601	Edg2	3.567951381	Mus musculus endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 2 (Edg2), mRNA.
602	A830080H07Rik	3.566599725	
603	Fzd2	3.565383173	Mus musculus frizzled homolog 2 (Drosophila) (Fzd2), mRNA.
604	Reck	3.558061288	Mus musculus reversion-inducing-cysteine-rich protein with kazal motifs (Reck), mRNA.
605	Agpat4	3.549790501	Mus musculus 1-acylglycerol-3-phosphate O-acyltransferase 4 (lysophosphatidic acid acyltransferase, delta) (Agpat4), mRNA.
606	Ccbp2	3.544428803	Mus musculus chemokine binding protein 2 (Ccbp2), mRNA.
607	Calu	3.535731204	Mus musculus calumenin (Calu), transcript variant 1, mRNA.
608	Cacna2d1	3.529198204	Mus musculus calcium channel, voltage-dependent, alpha2/delta subunit 1 (Cacna2d1), mRNA.
609	Rasl11a	3.520396698	Mus musculus RAS-like, family 11, member A (Rasl11a), mRNA.
610	Ifitm1	3.518466255	Mus musculus interferon induced transmembrane protein 1 (Ifitm1), mRNA.
611	P4ha2	3.517754806	Mus musculus procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha II polypeptide (P4ha2), transcript variant 2, mRNA.

612	BC028528	3.51213461	Mus musculus cDNA sequence BC028528 (BC028528), mRNA.
613	Hmcn1	3.503830688	Mus musculus hemicentin 1 (Hmcn1), mRNA.
614	Palld	3.500867359	Mus musculus palladin, cytoskeletal associated protein (Palld), mRNA.
615	Cd34	3.496261377	Mus musculus CD34 antigen (Cd34), mRNA.
616	Slc8a1	3.495638719	
617	H2-M3	3.494915791	Mus musculus histocompatibility 2, M region locus 3 (H2-M3), mRNA.
618	Mylk	3.483111196	Mus musculus myosin, light polypeptide kinase (Mylk), mRNA.
619	Nlgn2	3.476929054	Mus musculus neuroligin 2 (Nlgn2), mRNA.
620	Robo1	3.473528285	Mus musculus roundabout homolog 1 (Drosophila) (Robo1), mRNA.
621	Gpr133	3.468579242	Mus musculus G protein-coupled receptor 133 (Gpr133), mRNA.
622	LOC100038993	3.467787997	Mus musculus similar to interleukin 11 receptor, alpha chain 2 (LOC100038993), mRNA.
623	Olfm1	3.439957859	Mus musculus olfactomedin 1 (Olfm1), transcript variant 2, mRNA.
624	Fxyd6	3.431361652	Mus musculus FXYD domain-containing ion transport regulator 6 (Fxyd6), mRNA.
625	Evc	3.429880756	Mus musculus Ellis van Creveld gene homolog (human) (Evc), mRNA.
626	6720477C19Rik	3.422870973	
627	Ugt1a6b	3.419591896	Mus musculus UDP glucuronosyltransferase 1 family, polypeptide A6B (Ugt1a6b), mRNA.
628	Ckb	3.40549571	Mus musculus creatine kinase, brain (Ckb), mRNA.
629	Hoxb7	3.389518398	Mus musculus homeo box B7 (Hoxb7), mRNA.
630	Rnf144a	3.381684332	Mus musculus ring finger protein 144A (Rnf144a), transcript variant 2, mRNA.
631	Gpr153	3.376826404	Mus musculus G protein-coupled receptor 153 (Gpr153), mRNA.
632	Txndc5	3.375589741	Mus musculus thioredoxin domain containing 5 (Txndc5), mRNA.
633	A730040I05Rik	3.367608382	
634	Pdlim7	3.364379593	
635	Higd1b	3.363995028	Mus musculus HIG1 domain family, member 1B (Higd1b), mRNA.
636	Ugt1a6a	3.357359355	Mus musculus UDP glucuronosyltransferase 1 family, polypeptide A6A (Ugt1a6a), mRNA.
637	BC023892	3.350637106	PREDICTED: Mus musculus cDNA sequence BC023892, transcript variant 2 (BC023892), mRNA.
638	Ttc28	3.343255499	PREDICTED: Mus musculus tetratricopeptide repeat domain 28 (Ttc28), mRNA.
639	Mapre2	3.332880818	Mus musculus microtubule-associated protein, RP/EB family, member 2 (Mapre2), mRNA.
640	Nid2	3.310204895	Mus musculus nidogen 2 (Nid2), mRNA.
641	Evc2	3.298378115	Mus musculus Ellis van Creveld syndrome 2 homolog (human) (Evc2), mRNA.
642	Vasn	3.296063147	Mus musculus vasorin (Vasn), mRNA.
643	Col5a2	3.292749309	Mus musculus collagen, type V, alpha 2 (Col5a2), mRNA.
644	Thbd	3.283632522	Mus musculus thrombomodulin (Thbd), mRNA.
645	1700023M03Rik	3.282949054	
646	LOC385644	3.280234247	
647	Sphk1	3.263337115	Mus musculus sphingosine kinase 1 (Sphk1), transcript variant 1, mRNA.
648	Serpina1b	3.261771015	Mus musculus serine (or cysteine) preptidase inhibitor, clade A, member 1b (Serpina1b), mRNA.

649	Myh10	3.253924087	Mus musculus myosin, heavy polypeptide 10, non-muscle (Myh10), mRNA.
650	Abpe	3.24865393	Mus musculus androgen binding protein epsilon (Abpe), mRNA.
651	Cdon	3.242064552	
652	Ifi202b	3.239265413	Mus musculus interferon activated gene 202B (Ifi202b), mRNA.
653	Col7a1	3.233973377	Mus musculus collagen, type VII, alpha 1 (Col7a1), mRNA.
654	9130415E20Rik	3.218585116	
655	Plod2	3.217801918	
656	Vkorc1	3.208534942	Mus musculus vitamin K epoxide reductase complex, subunit 1 (Vkorc1), mRNA.
657	Adh1	3.208316914	Mus musculus alcohol dehydrogenase 1 (class I) (Adh1), mRNA.
658	Sdc2	3.205144247	Mus musculus syndecan 2 (Sdc2), mRNA.
659	Tpbg	3.203659409	Mus musculus trophoblast glycoprotein (Tpbg), mRNA.
660	Camk2n1	3.203230182	Mus musculus calcium/calmodulin-dependent protein kinase II inhibitor 1 (Camk2n1), mRNA.
661	Tnfaip3	3.199697623	Mus musculus tumor necrosis factor, alpha-induced protein 3 (Tnfaip3), mRNA.
662	Serpina1d	3.196721631	Mus musculus serine (or cysteine) peptidase inhibitor, clade A, member 1d (Serpina1d), mRNA.
663	Tnfsf13b	3.19091891	Mus musculus tumor necrosis factor (ligand) superfamily, member 13b (Tnfsf13b), mRNA.
664	Il11ra1	3.190578862	Mus musculus interleukin 11 receptor, alpha chain 1 (Il11ra1), mRNA.
665	Fkbp7	3.189973451	Mus musculus FK506 binding protein 7 (Fkbp7), mRNA.
666	Eno2	3.184671905	Mus musculus enolase 2, gamma neuronal (Eno2), mRNA.
667	Efnb1	3.18393579	Mus musculus ephrin B1 (Efnb1), mRNA.
668	Spnb2	3.174541699	Mus musculus spectrin beta 2 (Spnb2), transcript variant 2, mRNA.
669	Lynx1	3.173694319	Mus musculus Ly6/neurotoxin 1 (Lynx1), mRNA.
670	Tcea3	3.17086412	
671	Ext1	3.166493427	Mus musculus exostoses (multiple) 1 (Ext1), mRNA.
672	Islr2	3.151379122	Mus musculus immunoglobulin superfamily containing leucine-rich repeat 2 (Islr2), mRNA.
673	Nid1	3.140692241	
674	Rbms1	3.139106411	Mus musculus RNA binding motif, single stranded interacting protein 1 (Rbms1), mRNA.
675	Vtn	3.131049002	Mus musculus vitronectin (Vtn), mRNA.
676	Gngt2	3.117441027	Mus musculus guanine nucleotide binding protein (G protein), gamma transducing activity polypeptide 2 (Gngt2), transcript variant 2, mRNA.
677	Mras	3.116815796	Mus musculus muscle and microspikes RAS (Mras), mRNA.
678	Xpnpep2	3.113105435	Mus musculus X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound (Xpnpep2), transcript variant 1, mRNA.
679	Gria3	3.112921797	Mus musculus glutamate receptor, ionotropic, AMPA3 (alpha 3) (Gria3), mRNA.
680	Col6a3	3.104032826	Mus musculus procollagen, type VI, alpha 3 (Col6a3), mRNA.
681	Pkig	3.102270657	Mus musculus protein kinase inhibitor, gamma (Pkig), transcript variant 1, mRNA.
682	Gata6	3.094015265	Mus musculus GATA binding protein 6 (Gata6), mRNA.
683	Adprh	3.084452344	Mus musculus ADP-ribosylarginine hydrolase (Adprh), mRNA.
684	Gcnt2	3.083116715	Mus musculus glucosaminyl (N-acetyl) transferase 2, I-branching enzyme (Gcnt2), transcript variant 3, mRNA.
685	9030024J15Rik	3.076528325	

686	Vkorc1	3.074806113	Mus musculus vitamin K epoxide reductase complex, subunit 1 (Vkorc1), mRNA.
687	Sod3	3.070193872	Mus musculus superoxide dismutase 3, extracellular (Sod3), mRNA.
688	Ddah1	3.057856839	
689	Mtap1b	3.053114628	Mus musculus microtubule-associated protein 1 B (Mtap1b), mRNA.
690	Synpo	3.045770447	
691	Pcdh19	3.034510441	PREDICTED: Mus musculus protocadherin 19 (Pcdh19), mRNA.
692	Cacna1g	3.031008926	Mus musculus calcium channel, voltage-dependent, T type, alpha 1G subunit (Cacna1g), mRNA.
693	LOC676420	3.026731005	PREDICTED: Mus musculus similar to ceramide kinases (LOC676420), misc RNA.
694	Abcc9	3.023704495	Mus musculus ATP-binding cassette, sub-family C (CFTR/MRP), member 9 (Abcc9), transcript variant 4, mRNA.
695	Gpsm1	3.017429681	Mus musculus G-protein signalling modulator 1 (AGS3-like, C. elegans) (Gpsm1), mRNA.
696	Cdc42ep2	3.015979174	Mus musculus CDC42 effector protein (Rho GTPase binding) 2 (Cdc42ep2), mRNA.
697	Ppap2a	3.003742226	Mus musculus phosphatidic acid phosphatase 2a (Ppap2a), transcript variant 1, mRNA.
698	Ctgf	3.003457243	Mus musculus connective tissue growth factor (Ctgf), mRNA.
699	H2-Q5	3.001312804	Mus musculus histocompatibility 2, Q region locus 5 (H2-Q5), mRNA.

Reference List

ACETO, N., BARDIA, A., MIYAMOTO, D. T., DONALDSON, M. C., WITTNER, B. S., SPENCER, J. A., YU, M., PELY, A., ENGSTROM, A., ZHU, H., BRANNIGAN, B. W., KAPUR, R., STOTT, S. L., SHIODA, T., RAMASWAMY, S., TING, D. T., LIN, C. P., TONER, M., HABER, D. A. & MAHESWARAN, S. 2014. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell,* 158**,** 1110-22.

AKKARI, L., GOCHEVA, V., KESTER, J. C., HUNTER, K. E., QUICK, M. L., SEVENICH, L., WANG, H. W., PETERS, C., TANG, L. H., KLIMSTRA, D. S., REINHECKEL, T. & JOYCE, J. A. 2014. Distinct functions of macrophage-derived and cancer cell-derived cathepsin Z combine to promote tumor malignancy via interactions with the extracellular matrix. *Genes Dev,* 28**,** 2134-50.

AL-HAJJ, M., WICHA, M. S., BENITO-HERNANDEZ, A., MORRISON, S. J. & CLARKE, M. F. 2003. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A,* 100**,** 3983-8.

AMMOUN, S., PROVENZANO, L., ZHOU, L., BARCZYK, M., EVANS, K., HILTON, D. A., HAFIZI, S. & HANEMANN, C. O. 2014. Axl/Gas6/NFkappaB signalling in schwannoma pathological proliferation, adhesion and survival. *Oncogene,* 33**,** 336-46.

ARMSTRONG, A. J., MARENGO, M. S., OLTEAN, S., KEMENY, G., BITTING, R. L., TURNBULL, J. D., HEROLD, C. I., MARCOM, P. K., GEORGE, D. J. & GARCIA-BLANCO, M. A. 2011. Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers. *Mol Cancer Res,* 9**,** 997-1007.

ARNOUX, V., NASSOUR, M., L'HELGOUALC'H, A., HIPSKIND, R. A. & SAVAGNER, P. 2008. Erk5 controls Slug expression and keratinocyte activation during wound healing. *Mol Biol Cell,* 19**,** 4738-49.

ARTANDI, S. E. & DEPINHO, R. A. 2010. Telomeres and telomerase in cancer. *Carcinogenesis,* 31**,** 9-18.

ASIEDU, M. K., BEAUCHAMP-PEREZ, F. D., INGLE, J. N., BEHRENS, M. D., RADISKY, D. C. & KNUTSON, K. L. 2013. AXL induces epithelial-to-mesenchymal transition and regulates the function of breast cancer stem cells. *Oncogene*.

BACCELLI, I., SCHNEEWEISS, A., RIETHDORF, S., STENZINGER, A., SCHILLERT, A., VOGEL, V., KLEIN, C., SAINI, M., BAUERLE, T., WALLWIENER, M., HOLLAND-LETZ, T., HOFNER, T., SPRICK, M., SCHARPFF, M., MARME, F., SINN, H. P., PANTEL, K., WEICHERT, W. & TRUMPP, A. 2013. Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nat Biotechnol,* 31**,** 539-44.

BADVE, S. & NAKSHATRI, H. 2012. Breast-cancer stem cells-beyond semantics. *Lancet Oncol,* 13**,** e43-8.

BAKHOUM, S. F. & SWANTON, C. 2014. Chromosomal instability, aneuploidy, and cancer. *Front Oncol,* 4**,** 161.

BAO, S., WU, Q., MCLENDON, R. E., HAO, Y., SHI, Q., HJELMELAND, A. B., DEWHIRST, M. W., BIGNER, D. D. & RICH, J. N. 2006. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature,* 444**,** 756-60.

BARKER, N., VAN ES, J. H., KUIPERS, J., KUJALA, P., VAN DEN BORN, M., COZIJNSEN, M., HAEGEBARTH, A., KORVING, J., BEGTHEL, H., PETERS, P.

J. & CLEVERS, H. 2007. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature,* 449**,** 1003-7.

BARRETT, L. E., GRANOT, Z., COKER, C., IAVARONE, A., HAMBARDZUMYAN, D., HOLLAND, E. C., NAM, H. S. & BENEZRA, R. 2012. Self-renewal does not predict tumor growth potential in mouse models of high-grade glioma. *Cancer Cell,* 21**,** 11-24.

BATES, T. J., VONICA, A., HEASMAN, J., BRIVANLOU, A. H. & BELL, E. 2013. Coco regulates dorsoventral specification of germ layers via inhibition of TGFbeta signalling. *Development,* 140**,** 4177-81.

BECK, B., DRIESSENS, G., GOOSSENS, S., YOUSSEF, K. K., KUCHNIO, A., CAAUWE, A., SOTIROPOULOU, P. A., LOGES, S., LAPOUGE, G., CANDI, A., MASCRE, G., DROGAT, B., DEKONINCK, S., HAIGH, J. J., CARMELIET, P. & BLANPAIN, C. 2011. A vascular niche and a VEGF-Nrp1 loop regulate the initiation and stemness of skin tumours. *Nature,* 478**,** 399-403.

BECK, B., LAPOUGE, G., RORIVE, S., DROGAT, B., DESAEDELAERE, K., DELAFAILLE, S., DUBOIS, C., SALMON, I., WILLEKENS, K., MARINE, J. C. & BLANPAIN, C. 2015. Different levels of Twist1 regulate skin tumor initiation, stemness, and progression. *Cell Stem Cell,* 16**,** 67-79.

BEN-BATALLA, I., SCHULTZE, A., WROBLEWSKI, M., ERDMANN, R., HEUSER, M., WAIZENEGGER, J. S., RIECKEN, K., BINDER, M., SCHEWE, D., SAWALL, S., WITZKE, V., CUBAS-CORDOVA, M., JANNING, M., WELLBROCK, J., FEHSE, B., HAGEL, C., KRAUTER, J., GANSER, A., LORENS, J. B., FIEDLER, W., CARMELIET, P., PANTEL, K., BOKEMEYER, C. & LOGES, S. 2013. Axl, a prognostic and therapeutic target in acute myeloid leukemia mediates paracrine cross-talk of leukemia cells with bone marrow stroma. *Blood*.

BENNETT, D. C., PEACHEY, L. A., DURBIN, H. & RUDLAND, P. S. 1978. A possible mammary stem cell line. *Cell,* 15**,** 283-98.

BERCLAZ, G., ALTERMATT, H. J., ROHRBACH, V., KIEFFER, I., DREHER, E. & ANDRES, A. C. 2001. Estrogen dependent expression of the receptor tyrosine kinase axl in normal and malignant human breast. *Ann Oncol,* 12**,** 819-24.

BERENDSEN, A. D. & OLSEN, B. R. 2014. How vascular endothelial growth factor-A (VEGF) regulates differentiation of mesenchymal stem cells. *J Histochem Cytochem,* 62**,** 103-8.

BERGAMASCHI, A., TAGLIABUE, E., SORLIE, T., NAUME, B., TRIULZI, T., ORLANDI, R., RUSSNES, H. G., NESLAND, J. M., TAMMI, R., AUVINEN, P., KOSMA, V. M., MENARD, S. & BORRESEN-DALE, A. L. 2008. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. *J Pathol,* 214**,** 357-67.

BERNSTEIN, B. E., MIKKELSEN, T. S., XIE, X., KAMAL, M., HUEBERT, D. J., CUFF, J., FRY, B., MEISSNER, A., WERNIG, M., PLATH, K., JAENISCH, R., WAGSCHAL, A., FEIL, R., SCHREIBER, S. L. & LANDER, E. S. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell,* 125**,** 315-26.

BONNET, D. & DICK, J. E. 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med,* 3**,** 730-7.

BRABLETZ, T. 2012a. EMT and MET in metastasis: where are the cancer stem cells? *Cancer Cell,* 22**,** 699-701.

BRABLETZ, T. 2012b. To differentiate or not--routes towards metastasis. *Nat Rev Cancer,* 12**,** 425-36.

BRAGADO, P., ESTRADA, Y., PARIKH, F., KRAUSE, S., CAPOBIANCO, C., FARINA, H. G., SCHEWE, D. M. & AGUIRRE-GHISO, J. A. 2013. TGF-beta2 dictates disseminated tumour cell fate in target organs through TGF-beta-RIII and p38alpha/beta signalling. *Nat Cell Biol,* 15**,** 1351-61.

BREKKEN, R. A., PUOLAKKAINEN, P., GRAVES, D. C., WORKMAN, G., LUBKIN, S. R. & SAGE, E. H. 2003. Enhanced growth of tumors in SPARC null mice is associated with changes in the ECM. *J Clin Invest,* 111**,** 487-95.

BRUCE ALBERTS, A. J., JULIAN LEWIS, MARTIN RAFF, KEITH ROBERTS, AND PETER WALTER. 2002. *Molecular Biology of the Cell, 4th edition,* New York, Garland Science.

BRUZZESE, F., HAGGLOF, C., LEONE, A., SJOBERG, E., ROCA, M. S., KIFLEMARIAM, S., SJOBLOM, T., HAMMARSTEN, P., EGEVAD, L., BERGH, A., OSTMAN, A., BUDILLON, A. & AUGSTEN, M. 2014. Local and systemic protumorigenic effects of cancer-associated fibroblast-derived GDF15. *Cancer Res,* 74**,** 3408-17.

BUDAY, L. & DOWNWARD, J. 2008. Many faces of Ras activation. *Biochim Biophys Acta,* 1786**,** 178-87.

BUIJS, J. T., RENTSCH, C. A., VAN DER HORST, G., VAN OVERVELD, P. G., WETTERWALD, A., SCHWANINGER, R., HENRIQUEZ, N. V., TEN DIJKE, P., BOROVECKI, F., MARKWALDER, R., THALMANN, G. N., PAPAPOULOS, S. E., PELGER, R. C., VUKICEVIC, S., CECCHINI, M. G., LOWIK, C. W. & VAN DER PLUIJM, G. 2007. BMP7, a putative regulator of epithelial homeostasis in the human prostate, is a potent inhibitor of prostate cancer bone metastasis in vivo. *Am J Pathol,* 171**,** 1047-57.

BURK, U., SCHUBERT, J., WELLNER, U., SCHMALHOFER, O., VINCAN, E., SPADERNA, S. & BRABLETZ, T. 2008. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep,* 9**,** 582-9.

BURKHART, D. L. & SAGE, J. 2008. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat Rev Cancer,* 8**,** 671-82.

CALABRESE, C., POPPLETON, H., KOCAK, M., HOGG, T. L., FULLER, C., HAMNER, B., OH, E. Y., GABER, M. W., FINKLESTEIN, D., ALLEN, M., FRANK, A., BAYAZITOV, I. T., ZAKHARENKO, S. S., GAJJAR, A., DAVIDOFF, A. & GILBERTSON, R. J. 2007. A perivascular niche for brain tumor stem cells. *Cancer Cell,* 11**,** 69-82.

CALVO, F., EGE, N., GRANDE-GARCIA, A., HOOPER, S., JENKINS, R. P., CHAUDHRY, S. I., HARRINGTON, K., WILLIAMSON, P., MOEENDARBARY, E., CHARRAS, G. & SAHAI, E. 2013. Mechanotransduction and YAP-dependent matrix remodelling is required for the generation and maintenance of cancer-associated fibroblasts. *Nat Cell Biol,* 15**,** 637-46.

CANCER GENOME ATLAS, N. 2012. Comprehensive molecular portraits of human breast tumours. *Nature,* 490**,** 61-70.

CANO, A., PEREZ-MORENO, M. A., RODRIGO, I., LOCASCIO, A., BLANCO, M. J., DEL BARRIO, M. G., PORTILLO, F. & NIETO, M. A. 2000. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol,* 2**,** 76-83.

CASTELLANO, E. & DOWNWARD, J. 2011. RAS Interaction with PI3K: More Than Just Another Effector Pathway. *Genes Cancer,* 2**,** 261-74.

CAZES, A., GALAUP, A., CHOMEL, C., BIGNON, M., BRECHOT, N., LE JAN, S., WEBER, H., CORVOL, P., MULLER, L., GERMAIN, S. & MONNOT, C. 2006. Extracellular matrix-bound angiopoietin-like 4 inhibits endothelial cell adhesion, migration, and sprouting and alters actin cytoskeleton. *Circ Res,* 99**,** 1207-15.

CEDAR, H. & BERGMAN, Y. 2008. Epigenetic silencing during early lineage commitment. *StemBook.* Cambridge (MA).

CERAMI, E., GAO, J., DOGRUSOZ, U., GROSS, B. E., SUMER, S. O., AKSOY, B. A., JACOBSEN, A., BYRNE, C. J., HEUER, M. L., LARSSON, E., ANTIPIN, Y., REVA, B., GOLDBERG, A. P., SANDER, C. & SCHULTZ, N. 2012. The cBio

cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov,* 2**,** 401-4.

CHAFFER, C. L., MARJANOVIC, N. D., LEE, T., BELL, G., KLEER, C. G., REINHARDT, F., D'ALESSIO, A. C., YOUNG, R. A. & WEINBERG, R. A. 2013. Poised Chromatin at the ZEB1 Promoter Enables Breast Cancer Cell Plasticity and Enhances Tumorigenicity. *Cell,* 154**,** 61-74.

CHAFFER, C. L. & WEINBERG, R. A. 2011. A Perspective on Cancer Cell Metastasis. *Science,* 331**,** 1559-1564.

CHEN, J., MCKAY, R. M. & PARADA, L. F. 2012. Malignant glioma: lessons from genomics, mouse models, and stem cells. *Cell,* 149**,** 36-47.

CHEUNG, K. J. & EWALD, A. J. 2014. Illuminating breast cancer invasion: diverse roles for cell-cell interactions. *Curr Opin Cell Biol,* 30**,** 99-111.

CHEUNG, K. J., GABRIELSON, E., WERB, Z. & EWALD, A. J. 2013. Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell,* 155**,** 1639-51.

CHI, A. S. & BERNSTEIN, B. E. 2009. Developmental biology. Pluripotent chromatin state. *Science,* 323**,** 220-1.

CHO, R. W., WANG, X., DIEHN, M., SHEDDEN, K., CHEN, G. Y., SHERLOCK, G., GURNEY, A., LEWICKI, J. & CLARKE, M. F. 2008. Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors. *Stem Cells,* 26**,** 364-71.

CHONG, H. C., TAN, C. K., HUANG, R. L. & TAN, N. S. 2012. Matricellular proteins: a sticky affair with cancers. *J Oncol,* 2012**,** 351089.

CLARKE, M. F., DICK, J. E., DIRKS, P. B., EAVES, C. J., JAMIESON, C. H., JONES, D. L., VISVADER, J., WEISSMAN, I. L. & WAHL, G. M. 2006. Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. *Cancer Res,* 66**,** 9339-44.

CLARKSON, B. D. 1969. Review of recent studies of cellular proliferation in acute leukemia. *Natl Cancer Inst Monogr,* 30**,** 81-120.

COFFELT, S. B., KERSTEN, K., DOORNEBAL, C. W., WEIDEN, J., VRIJLAND, K., HAU, C. S., VERSTEGEN, N. J., CIAMPRICOTTI, M., HAWINKELS, L. J., JONKERS, J. & DE VISSER, K. E. 2015. IL-17-producing gammadelta T cells and neutrophils conspire to promote breast cancer metastasis. *Nature,* 522**,** 345-8.

DALERBA, P., DYLLA, S. J., PARK, I. K., LIU, R., WANG, X., CHO, R. W., HOEY, T., GURNEY, A., HUANG, E. H., SIMEONE, D. M., SHELTON, A. A., PARMIANI, G., CASTELLI, C. & CLARKE, M. F. 2007. Phenotypic characterization of human colorectal cancer stem cells. *Proc Natl Acad Sci U S A,* 104**,** 10158-63.

DALERBA, P., KALISKY, T., SAHOO, D., RAJENDRAN, P. S., ROTHENBERG, M. E., LEYRAT, A. A., SIM, S., OKAMOTO, J., JOHNSTON, D. M., QIAN, D., ZABALA, M., BUENO, J., NEFF, N. F., WANG, J., SHELTON, A. A., VISSER, B., HISAMORI, S., SHIMONO, Y., VAN DE WETERING, M., CLEVERS, H., CLARKE, M. F. & QUAKE, S. R. 2011. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol,* 29**,** 1120-7.

DAVIES, M. A. & SAMUELS, Y. 2010. Analysis of the genome to personalize therapy for melanoma. *Oncogene,* 29**,** 5545-55.

DE CRAENE, B. & BERX, G. 2013. Regulatory networks defining EMT during cancer initiation and progression. *Nat Rev Cancer,* 13**,** 97-110.

DE CRAENE, B., DENECKER, G., VERMASSEN, P., TAMINAU, J., MAUCH, C., DERORE, A., JONKERS, J., FUCHS, E. & BERX, G. 2014. Epidermal Snail expression drives skin cancer initiation and progression through enhanced cytoprotection, epidermal stem/progenitor cell expansion and enhanced metastatic potential. *Cell Death Differ,* 21**,** 310-20.

DEATON, A. M. & BIRD, A. 2011. CpG islands and the regulation of transcription. *Genes Dev,* 25**,** 1010-22.

DESPRAT, N., SUPATTO, W., POUILLE, P. A., BEAUREPAIRE, E. & FARGE, E. 2008. Tissue deformation modulates twist expression to determine anterior midgut differentiation in Drosophila embryos. *Dev Cell,* 15**,** 470-7.

DIEHN, M., CHO, R. W., LOBO, N. A., KALISKY, T., DORIE, M. J., KULP, A. N., QIAN, D., LAM, J. S., AILLES, L. E., WONG, M., JOSHUA, B., KAPLAN, M. J., WAPNIR, I., DIRBAS, F. M., SOMLO, G., GARBEROGLIO, C., PAZ, B., SHEN, J., LAU, S. K., QUAKE, S. R., BROWN, J. M., WEISSMAN, I. L. & CLARKE, M. F. 2009. Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature,* 458**,** 780-3.

DONG, C., YUAN, T., WU, Y., WANG, Y., FAN, T. W., MIRIYALA, S., LIN, Y., YAO, J., SHI, J., KANG, T., LORKIEWICZ, P., ST CLAIR, D., HUNG, M. C., EVERS, B. M. & ZHOU, B. P. 2013. Loss of FBP1 by Snail-mediated repression provides metabolic advantages in basal-like breast cancer. *Cancer Cell,* 23**,** 316-31.

DORMADY, S. P., ZHANG, X. M. & BASCH, R. S. 2000. Hematopoietic progenitor cells grow on 3T3 fibroblast monolayers that overexpress growth arrest-specific gene-6 (GAS6). *Proc Natl Acad Sci U S A,* 97**,** 12260-5.

DRISKELL, R. R. & WATT, F. M. 2015. Understanding fibroblast heterogeneity in the skin. *Trends Cell Biol,* 25**,** 92-9.

DU, P., KIBBE, W. A. & LIN, S. M. 2008. lumi: a pipeline for processing Illumina microarray. *Bioinformatics,* 24**,** 1547-8.

DUDLEY, A. T., LYONS, K. M. & ROBERTSON, E. J. 1995. A requirement for bone morphogenetic protein-7 during development of the mammalian kidney and eye. *Genes Dev,* 9**,** 2795-807.

DUMONT, N., LIU, B., DEFILIPPIS, R. A., CHANG, H., RABBAN, J. T., KARNEZIS, A. N., TJOE, J. A., MARX, J., PARVIN, B. & TLSTY, T. D. 2013. Breast fibroblasts modulate early dissemination, tumorigenesis, and metastasis through alteration of extracellular matrix characteristics. *Neoplasia,* 15**,** 249-62.

DUPONT, S., MORSUT, L., ARAGONA, M., ENZO, E., GIULITTI, S., CORDENONSI, M., ZANCONATO, F., LE DIGABEL, J., FORCATO, M., BICCIATO, S., ELVASSORE, N. & PICCOLO, S. 2011. Role of YAP/TAZ in mechanotransduction. *Nature,* 474**,** 179-83.

DYKXHOORN, D. M., WU, Y., XIE, H., YU, F., LAL, A., PETROCCA, F., MARTINVALET, D., SONG, E., LIM, B. & LIEBERMAN, J. 2009. miR-200 enhances mouse breast cancer cell colonization to form distant metastases. *PLoS One,* 4**,** e7181.

ERLER, J. T. & GIACCIA, A. J. 2006. Lysyl oxidase mediates hypoxic control of metastasis. *Cancer Res,* 66**,** 10238-41.

ESTELLER, M., GARCIA-FONCILLAS, J., ANDION, E., GOODMAN, S. N., HIDALGO, O. F., VANACLOCHA, V., BAYLIN, S. B. & HERMAN, J. G. 2000. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med,* 343**,** 1350-4.

FARMER, P., BONNEFOI, H., ANDERLE, P., CAMERON, D., WIRAPATI, P., BECETTE, V., ANDRE, S., PICCART, M., CAMPONE, M., BRAIN, E., MACGROGAN, G., PETIT, T., JASSEM, J., BIBEAU, F., BLOT, E., BOGAERTS, J., AGUET, M., BERGH, J., IGGO, R. & DELORENZI, M. 2009. A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer. *Nat Med,* 15**,** 68-74.

FLANAGAN, S. P. 1966. 'Nude', a new hairless gene with pleiotropic effects in the mouse. *Genet Res,* 8**,** 295-309.

FORRESTER, E., CHYTIL, A., BIERIE, B., AAKRE, M., GORSKA, A. E., SHARIF-AFSHAR, A. R., MULLER, W. J. & MOSES, H. L. 2005. Effect of conditional

knockout of the type II TGF-beta receptor gene in mammary epithelia on mammary gland development and polyomavirus middle T antigen induced tumor formation and metastasis. *Cancer Res,* 65**,** 2296-302.

FRANCIA, G., CRUZ-MUNOZ, W., MAN, S., XU, P. & KERBEL, R. S. 2011. Mouse models of advanced spontaneous metastasis for experimental therapeutics. *Nat Rev Cancer,* 11**,** 135-41.

FRANCISCHETTI, I. M., KOTSYFAKIS, M., ANDERSEN, J. F. & LUKSZO, J. 2010. Cyr61/CCN1 displays high-affinity binding to the somatomedin B(1-44) domain of vitronectin. *PLoS One,* 5**,** e9356.

FRANCO-CHUAIRE, M. L., MAGDA CAROLINA, S. C. & CHUAIRE-NOACK, L. 2013. Epithelial-mesenchymal transition (EMT): principles and clinical impact in cancer therapy. *Invest Clin,* 54**,** 186-205.

FRIEDL, P. & ALEXANDER, S. 2011. Cancer invasion and the microenvironment: plasticity and reciprocity. *Cell,* 147**,** 992-1009.

FRIEDL, P., LOCKER, J., SAHAI, E. & SEGALL, J. E. 2012. Classifying collective cancer cell invasion. *Nat Cell Biol,* 14**,** 777-83.

FRIEDL, P., WOLF, K. & LAMMERDING, J. 2011. Nuclear mechanics during cell migration. *Curr Opin Cell Biol,* 23**,** 55-64.

FU, J., XU, D., LIU, Z., SHI, M., ZHAO, P., FU, B., ZHANG, Z., YANG, H., ZHANG, H., ZHOU, C., YAO, J., JIN, L., WANG, H., YANG, Y., FU, Y. X. & WANG, F. S. 2007. Increased regulatory T cells correlate with CD8 T-cell impairment and poor survival in hepatocellular carcinoma patients. *Gastroenterology,* 132**,** 2328-39.

FUJIMORI, T., GRABIEC, A. M., KAUR, M., BELL, T. J., FUJINO, N., COOK, P. C., SVEDBERG, F. R., MACDONALD, A. S., MACIEWICZ, R. A., SINGH, D. & HUSSELL, T. 2015. The Axl receptor tyrosine kinase is a discriminator of macrophage function in the inflamed lung. *Mucosal Immunol*.

GABRILOVICH, D. I., OSTRAND-ROSENBERG, S. & BRONTE, V. 2012. Coordinated regulation of myeloid cells by tumours. *Nat Rev Immunol,* 12**,** 253-68.

GAGGIOLI, C., HOOPER, S., HIDALGO-CARCEDO, C., GROSSE, R., MARSHALL, J. F., HARRINGTON, K. & SAHAI, E. 2007. Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nat Cell Biol,* 9**,** 1392-400.

GAO, H., CHAKRABORTY, G., LEE-LIM, A. P., MO, Q., DECKER, M., VONICA, A., SHEN, R., BROGI, E., BRIVANLOU, A. H. & GIANCOTTI, F. G. 2012. The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. *Cell,* 150**,** 764-79.

GAO, J., AKSOY, B. A., DOGRUSOZ, U., DRESDNER, G., GROSS, B., SUMER, S. O., SUN, Y., JACOBSEN, A., SINHA, R., LARSSON, E., CERAMI, E., SANDER, C. & SCHULTZ, N. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal,* 6**,** pl1.

GARBER, K. 2008. Epithelial-to-mesenchymal transition is important to metastasis, but questions remain. *J Natl Cancer Inst,* 100**,** 232-3, 239.

GARCION, E., HALILAGIC, A., FAISSNER, A. & FFRENCH-CONSTANT, C. 2004. Generation of an environmental niche for neural stem cell development by the extracellular matrix molecule tenascin C. *Development,* 131**,** 3423-32.

GASTEIGER, G., HEMMERS, S., FIRTH, M. A., LE FLOC'H, A., HUSE, M., SUN, J. C. & RUDENSKY, A. Y. 2013. IL-2-dependent tuning of NK cell sensitivity for target cells is controlled by regulatory T cells. *J Exp Med,* 210**,** 1167-78.

GHAJAR, C. M., PEINADO, H., MORI, H., MATEI, I. R., EVASON, K. J., BRAZIER, H., ALMEIDA, D., KOLLER, A., HAJJAR, K. A., STAINIER, D. Y., CHEN, E. I., LYDEN, D. & BISSELL, M. J. 2013. The perivascular niche regulates breast tumour dormancy. *Nat Cell Biol,* 15**,** 807-17.

GIAMPIERI, S., MANNING, C., HOOPER, S., JONES, L., HILL, C. S. & SAHAI, E. 2009. Localized and reversible TGFbeta signalling switches breast cancer cells from cohesive to single cell motility. *Nat Cell Biol,* 11**,** 1287-96.

GJERDRUM, C., TIRON, C., HOIBY, T., STEFANSSON, I., HAUGEN, H., SANDAL, T., COLLETT, K., LI, S., MCCORMACK, E., GJERTSEN, B. T., MICKLEM, D. R., AKSLEN, L. A., GLACKIN, C. & LORENS, J. B. 2010. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proc Natl Acad Sci U S A,* 107**,** 1124-9.

GLOGAUER, J., SUN, C. X., BRADLEY, G. & MAGALHAES, M. A. 2015. Neutrophils Increase Oral Squamous Cell Carcinoma Invasion Through An Invadopodia-Dependent Pathway. *Cancer Immunol Res*.

GOLDSTEIN, N. S., VICINI, F. A., HUNTER, S., ODISH, E., FORBES, S. & KESTIN, L. L. 2005. Molecular clonality relationships in initial carcinomas, ipsilateral breast failures, and distant metastases in patients treated with breast-conserving therapy: evidence suggesting that some distant metastases are derived from ipsilateral breast failures and that metastases can metastasize. *Am J Clin Pathol,* 124**,** 49-57.

GOUWY, M., STRUYF, S., NOPPEN, S., SCHUTYSER, E., SPRINGAEL, J. Y., PARMENTIER, M., PROOST, P. & VAN DAMME, J. 2008. Synergy between coproduced CC and CXC chemokines in monocyte chemotaxis through receptor-mediated events. *Mol Pharmacol,* 74**,** 485-95.

GRANGE, C., LANZARDO, S., CAVALLO, F., CAMUSSI, G. & BUSSOLATI, B. 2008. Sca-1 identifies the tumor-initiating cells in mammary tumors of BALB-neuT transgenic mice. *Neoplasia,* 10**,** 1433-43.

GREGORY, A. D. & HOUGHTON, A. M. 2011. Tumor-associated neutrophils: new targets for cancer therapy. *Cancer Res,* 71**,** 2411-6.

GRIVENNIKOV, S. I. & KARIN, M. 2010. Inflammation and oncogenesis: a vicious connection. *Curr Opin Genet Dev,* 20**,** 65-71.

GRONROOS, E., KINGSTON, I. J., RAMACHANDRAN, A., RANDALL, R. A., VIZAN, P. & HILL, C. S. 2012. Transforming growth factor beta inhibits bone morphogenetic protein-induced transcription through novel phosphorylated Smad1/5-Smad3 complexes. *Mol Cell Biol,* 32**,** 2904-16.

GUO, W. & GIANCOTTI, F. G. 2004. Integrin signalling during tumour progression. *Nat Rev Mol Cell Biol,* 5**,** 816-26.

GUPTA, G. P., PERK, J., ACHARYYA, S., DE CANDIA, P., MITTAL, V., TODOROVA-MANOVA, K., GERALD, W. L., BROGI, E., BENEZRA, R. & MASSAGUE, J. 2007. ID genes mediate tumor reinitiation during breast cancer lung metastasis. *Proc Natl Acad Sci U S A,* 104**,** 19506-11.

GUPTA, P. B., CHAFFER, C. L. & WEINBERG, R. A. 2009. Cancer stem cells: mirage or reality? *Nat Med,* 15**,** 1010-2.

GUY, C. T., CARDIFF, R. D. & MULLER, W. J. 1992. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol,* 12**,** 954-61.

HAERYFAR, S. M. & HOSKIN, D. W. 2004. Thy-1: more than a mouse pan-T cell marker. *J Immunol,* 173**,** 3581-8.

HAGER, J. C., RUSSO, J., CERIANI, R. L., PETERSON, J. A., FLIGIEL, S., JOLLY, G. & HEPPNER, G. H. 1981. Epithelial characteristics of five subpopulations of a heterogeneous strain BALB/cfC3H mouse mammary tumor. *Cancer Res,* 41**,** 1720-30.

HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.

HANAHAN, D. & WEINBERG, ROBERT A. 2011. Hallmarks of Cancer: The Next Generation. *Cell,* 144**,** 646-674.

HANKENSON, K. D. & BORNSTEIN, P. 2002. The secreted protein thrombospondin 2 is an autocrine inhibitor of marrow stromal cell proliferation. *J Bone Miner Res,* 17**,** 415-25.

HE, X. C., ZHANG, J., TONG, W. G., TAWFIK, O., ROSS, J., SCOVILLE, D. H., TIAN, Q., ZENG, X., HE, X., WIEDEMANN, L. M., MISHINA, Y. & LI, L. 2004. BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt-beta-catenin signaling. *Nat Genet,* 36**,** 1117-21.

HEGERFELDT, Y., TUSCH, M., BROCKER, E. B. & FRIEDL, P. 2002. Collective cell movement in primary melanoma explants: plasticity of cell-cell interaction, beta1-integrin function, and migration strategies. *Cancer Res,* 62**,** 2125-30.

HELDIN, C. H., VANLANDEWIJCK, M. & MOUSTAKAS, A. 2012. Regulation of EMT by TGFbeta in cancer. *FEBS Lett,* 586**,** 1959-70.

HENNESSY, B. T., GONZALEZ-ANGULO, A. M., STEMKE-HALE, K., GILCREASE, M. Z., KRISHNAMURTHY, S., LEE, J. S., FRIDLYAND, J., SAHIN, A., AGARWAL, R., JOY, C., LIU, W., STIVERS, D., BAGGERLY, K., CAREY, M., LLUCH, A., MONTEAGUDO, C., HE, X., WEIGMAN, V., FAN, C., PALAZZO, J., HORTOBAGYI, G. N., NOLDEN, L. K., WANG, N. J., VALERO, V., GRAY, J. W., PEROU, C. M. & MILLS, G. B. 2009. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res,* 69**,** 4116-24.

HERMANN, P. C., HUBER, S. L., HERRLER, T., AICHER, A., ELLWART, J. W., GUBA, M., BRUNS, C. J. & HEESCHEN, C. 2007. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell,* 1**,** 313-23.

HERZIG, M. & CHRISTOFORI, G. 2002. Recent advances in cancer research: mouse models of tumorigenesis. *Biochim Biophys Acta,* 1602**,** 97-113.

HIDALGO-CARCEDO, C., HOOPER, S., CHAUDHRY, S. I., WILLIAMSON, P., HARRINGTON, K., LEITINGER, B. & SAHAI, E. 2011. Collective cell migration requires suppression of actomyosin at cell-cell contacts mediated by DDR1 and the cell polarity regulators Par3 and Par6. *Nat Cell Biol,* 13**,** 49-58.

HIRATA, E., GIROTTI, M. R., VIROS, A., HOOPER, S., SPENCER-DENE, B., MATSUDA, M., LARKIN, J., MARAIS, R. & SAHAI, E. 2015. Intravital imaging reveals how BRAF inhibition generates drug-tolerant microenvironments with high integrin beta1/FAK signaling. *Cancer Cell,* 27**,** 574-88.

HIROSE, Y., CHIBA, K., KARASUGI, T., NAKAJIMA, M., KAWAGUCHI, Y., MIKAMI, Y., FURUICHI, T., MIO, F., MIYAKE, A., MIYAMOTO, T., OZAKI, K., TAKAHASHI, A., MIZUTA, H., KUBO, T., KIMURA, T., TANAKA, T., TOYAMA, Y. & IKEGAWA, S. 2008. A functional polymorphism in THBS2 that affects alternative splicing and MMP binding is associated with lumbar-disc herniation. *Am J Hum Genet,* 82**,** 1122-9.

HOLLAND, S. J., PAN, A., FRANCI, C., HU, Y., CHANG, B., LI, W., DUAN, M., TORNEROS, A., YU, J., HECKRODT, T. J., ZHANG, J., DING, P., APATIRA, A., CHUA, J., BRANDT, R., PINE, P., GOFF, D., SINGH, R., PAYAN, D. G. & HITOSHI, Y. 2010. R428, a selective small molecule inhibitor of Axl kinase, blocks tumor spread and prolongs survival in models of metastatic breast cancer. *Cancer Res,* 70**,** 1544-54.

HOLLNAGEL, A., OEHLMANN, V., HEYMER, J., RUTHER, U. & NORDHEIM, A. 1999. Id genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. *J Biol Chem,* 274**,** 19838-45.

HOLZEL, D., ECKEL, R., EMENY, R. T. & ENGEL, J. 2010. Distant metastases do not metastasize. *Cancer Metastasis Rev,* 29**,** 737-50.

HOSTE, E., ARWERT, E. N., LAL, R., SOUTH, A. P., SALAS-ALANIS, J. C., MURRELL, D. F., DONATI, G. & WATT, F. M. 2015. Innate sensing of microbial products promotes wound-induced skin cancer. *Nat Commun,* 6**,** 5932.

HOU, J. M., KREBS, M., WARD, T., SLOANE, R., PRIEST, L., HUGHES, A., CLACK, G., RANSON, M., BLACKHALL, F. & DIVE, C. 2011. Circulating tumor cells as a window on metastasis biology in lung cancer. *Am J Pathol,* 178**,** 989-96.

HOU, J. M., KREBS, M. G., LANCASHIRE, L., SLOANE, R., BACKEN, A., SWAIN, R. K., PRIEST, L. J., GREYSTOKE, A., ZHOU, C., MORRIS, K., WARD, T., BLACKHALL, F. H. & DIVE, C. 2012. Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer. *J Clin Oncol,* 30**,** 525-32.

HYNES, R. O. 2009. The extracellular matrix: not just pretty fibrils. *Science,* 326**,** 1216-9.

IRMISCH, A. & HUELSKEN, J. 2013. Metastasis: New insights into organ-specific extravasation and metastatic niches. *Exp Cell Res*.

ISHIKAWA, F., YOSHIDA, S., SAITO, Y., HIJIKATA, A., KITAMURA, H., TANAKA, S., NAKAMURA, R., TANAKA, T., TOMIYAMA, H., SAITO, N., FUKATA, M., MIYAMOTO, T., LYONS, B., OHSHIMA, K., UCHIDA, N., TANIGUCHI, S., OHARA, O., AKASHI, K., HARADA, M. & SHULTZ, L. D. 2007. Chemotherapy-resistant human AML stem cells home to and engraft within the bone-marrow endosteal region. *Nat Biotechnol,* 25**,** 1315-21.

ITO, T., KWON, H. Y., ZIMDAHL, B., CONGDON, K. L., BLUM, J., LENTO, W. E., ZHAO, C., LAGOO, A., GERRARD, G., FORONI, L., GOLDMAN, J., GOH, H., KIM, S. H., KIM, D. W., CHUAH, C., OEHLER, V. G., RADICH, J. P., JORDAN, C. T. & REYA, T. 2010. Regulation of myeloid leukaemia by the cell-fate determinant Musashi. *Nature,* 466**,** 765-8.

JIANG, B. H. & LIU, L. Z. 2009. PI3K/PTEN signaling in angiogenesis and tumorigenesis. *Adv Cancer Res,* 102**,** 19-65.

JIMENEZ, B., VOLPERT, O. V., CRAWFORD, S. E., FEBBRAIO, M., SILVERSTEIN, R. L. & BOUCK, N. 2000. Signals leading to apoptosis-dependent inhibition of neovascularization by thrombospondin-1. *Nat Med,* 6**,** 41-8.

JOHANN, P. D. & MULLER, I. 2015. Multipotent Mesenchymal Stromal Cells: Possible Culprits in Solid Tumors? *Stem Cells Int,* 2015**,** 914632.

JUNTTILA, M. R. & DE SAUVAGE, F. J. 2013. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature,* 501**,** 346-54.

KALLERGI, G., PAPADAKI, M. A., POLITAKI, E., MAVROUDIS, D., GEORGOULIAS, V. & AGELAKI, S. 2011. Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients. *Breast Cancer Res,* 13**,** R59.

KALLURI, R. & WEINBERG, R. A. 2009. The basics of epithelial-mesenchymal transition. *J Clin Invest,* 119**,** 1420-8.

KALLURI, R. & ZEISBERG, M. 2006. Fibroblasts in cancer. *Nat Rev Cancer,* 6**,** 392-401.

KANG, Y., CHEN, C. R. & MASSAGUE, J. 2003. A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. *Mol Cell,* 11**,** 915-26.

KIM, H., WATKINSON, J., VARADAN, V. & ANASTASSIOU, D. 2010. Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. *BMC Med Genomics,* 3**,** 51.

KOCH, M., HUSSEIN, F., WOESTE, A., GRUNDKER, C., FRONTZEK, K., EMONS, G. & HAWIGHORST, T. 2011. CD36-mediated activation of endothelial cell apoptosis by an N-terminal recombinant fragment of thrombospondin-2 inhibits

breast cancer growth and metastasis in vivo. *Breast Cancer Res Treat,* 128**,** 337-46.

KORSHUNOV, V. A. 2012. Axl-dependent signalling: a clinical update. *Clin Sci (Lond),* 122**,** 361-8.

KREBS, M. G., METCALF, R. L., CARTER, L., BRADY, G., BLACKHALL, F. H. & DIVE, C. 2014. Molecular analysis of circulating tumour cells-biology and biomarkers. *Nat Rev Clin Oncol,* 11**,** 129-44.

KRESO, A. & DICK, J. E. 2014. Evolution of the Cancer Stem Cell Model. *Cell Stem Cell,* 14**,** 275-291.

KUHN, A., BRACHTENDORF, G., KURTH, F., SONNTAG, M., SAMULOWITZ, U., METZE, D. & VESTWEBER, D. 2002. Expression of endomucin, a novel endothelial sialomucin, in normal and diseased human skin. *J Invest Dermatol,* 119**,** 1388-93.

LABELLE, M., BEGUM, S. & HYNES, R. O. 2011. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell,* 20**,** 576-90.

LAPIDOT, T., SIRARD, C., VORMOOR, J., MURDOCH, B., HOANG, T., CACERES-CORTES, J., MINDEN, M., PATERSON, B., CALIGIURI, M. A. & DICK, J. E. 1994. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature,* 367**,** 645-8.

LEDFORD, H. 2011. Cancer theory faces doubts. *Nature,* 472**,** 273.

LEONG, H. S., ROBERTSON, A. E., STOLETOV, K., LEITH, S. J., CHIN, C. A., CHIEN, A. E., HAGUE, M. N., ABLACK, A., CARMINE-SIMMEN, K., MCPHERSON, V. A., POSTENKA, C. O., TURLEY, E. A., COURTNEIDGE, S. A., CHAMBERS, A. F. & LEWIS, J. D. 2014. Invadopodia are required for cancer cell extravasation and are a therapeutic target for metastasis. *Cell Rep,* 8**,** 1558-70.

LEVY, L. & HILL, C. S. 2006. Alterations in components of the TGF-beta superfamily signaling pathways in human cancer. *Cytokine Growth Factor Rev,* 17**,** 41-58.

LI, B., SHI, Y., SHU, J., GAO, J., WU, P. & TANG, S. J. 2013. Wingless-type mammary tumor virus integration site family, member 5A (Wnt5a) regulates human immunodeficiency virus type 1 (HIV-1) envelope glycoprotein 120 (gp120)-induced expression of pro-inflammatory cytokines via the Ca2+/calmodulin-dependent protein kinase II (CaMKII) and c-Jun N-terminal kinase (JNK) signaling pathways. *J Biol Chem,* 288**,** 13610-9.

LI, M., LU, J., ZHANG, F., LI, H., ZHANG, B., WU, X., TAN, Z., ZHANG, L., GAO, G., MU, J., SHU, Y., BAO, R., DING, Q., WU, W., DONG, P., GU, J. & LIU, Y. 2014a. Yes-associated protein 1 (YAP1) promotes human gallbladder tumor growth via activation of the AXL/MAPK pathway. *Cancer Lett*.

LI, Y., JIA, L., LIU, C., GONG, Y., REN, D., WANG, N., ZHANG, X. & ZHAO, Y. 2014b. Axl as a downstream effector of TGF-beta1 via PI3K/Akt-PAK1 signaling pathway promotes tumor invasion and chemoresistance in breast carcinoma. *Tumour Biol*.

LIM, D. A., HUANG, Y. C., SWIGUT, T., MIRICK, A. L., GARCIA-VERDUGO, J. M., WYSOCKA, J., ERNST, P. & ALVAREZ-BUYLLA, A. 2009. Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. *Nature,* 458**,** 529-33.

LIN, E. Y., JONES, J. G., LI, P., ZHU, L., WHITNEY, K. D., MULLER, W. J. & POLLARD, J. W. 2003. Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am J Pathol,* 163**,** 2113-26.

LIU, C., YU, S., KAPPES, J., WANG, J., GRIZZLE, W. E., ZINN, K. R. & ZHANG, H. G. 2007a. Expansion of spleen myeloid suppressor cells represses NK cell cytotoxicity in tumor-bearing host. *Blood,* 109**,** 4336-42.

LIU, J. C., DENG, T., LEHAL, R. S., KIM, J. & ZACKSENHAUS, E. 2007b. Identification of tumorsphere- and tumor-initiating cells in HER2/Neu-induced mammary tumors. *Cancer Res,* 67**,** 8671-81.

LO, C. M., WANG, H. B., DEMBO, M. & WANG, Y. L. 2000. Cell movement is guided by the rigidity of the substrate. *Biophys J,* 79**,** 144-52.

LOHELA, M., CASBON, A. J., OLOW, A., BONHAM, L., BRANSTETTER, D., WENG, N., SMITH, J. & WERB, Z. 2014. Intravital imaging reveals distinct responses of depleting dynamic tumor-associated macrophage and dendritic cell subpopulations. *Proc Natl Acad Sci U S A,* 111**,** E5086-95.

LONARDO, E., HERMANN, P. C., MUELLER, M. T., HUBER, S., BALIC, A., MIRANDA-LORENZO, I., ZAGORAC, S., ALCALA, S., RODRIGUEZ-ARABAOLAZA, I., RAMIREZ, J. C., TORRES-RUIZ, R., GARCIA, E., HIDALGO, M., CEBRIAN, D. A., HEUCHEL, R., LOHR, M., BERGER, F., BARTENSTEIN, P., AICHER, A. & HEESCHEN, C. 2011. Nodal/Activin signaling drives self-renewal and tumorigenicity of pancreatic cancer stem cells and provides a target for combined drug therapy. *Cell Stem Cell,* 9**,** 433-46.

LOPEZ-NOVOA, J. M. & NIETO, M. A. 2009. Inflammation and EMT: an alliance towards organ fibrosis and cancer progression. *EMBO Mol Med,* 1**,** 303-14.

LORDA-DIEZ, C. I., MONTERO, J. A., GARCIA-PORRERO, J. A. & HURLE, J. M. 2010. Tgfbeta2 and 3 are coexpressed with their extracellular regulator Ltbp1 in the early limb bud and modulate mesodermal outgrowth and BMP signaling in chicken embryos. *BMC Dev Biol,* 10**,** 69.

LU, H., CLAUSER, K. R., TAM, W. L., FROSE, J., YE, X., EATON, E. N., REINHARDT, F., DONNENBERG, V. S., BHARGAVA, R., CARR, S. A. & WEINBERG, R. A. 2014. A breast cancer stem cell niche supported by juxtacrine signalling from monocytes and macrophages. *Nat Cell Biol,* 16**,** 1105-17.

LU, P., WEAVER, V. M. & WERB, Z. 2012. The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol,* 196**,** 395-406.

LUHR, I., FRIEDL, A., OVERATH, T., THOLEY, A., KUNZE, T., HILPERT, F., SEBENS, S., ARNOLD, N., ROSEL, F., OBERG, H. H., MAASS, N., MUNDHENKE, C., JONAT, W. & BAUER, M. 2012. Mammary fibroblasts regulate morphogenesis of normal and tumorigenic breast epithelial cells by mechanical and paracrine signals. *Cancer Lett,* 325**,** 175-88.

LUZZI, K. J., MACDONALD, I. C., SCHMIDT, E. E., KERKVLIET, N., MORRIS, V. L., CHAMBERS, A. F. & GROOM, A. C. 1998. Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *Am J Pathol,* 153**,** 865-73.

MALANCHI, I. 2013. Tumour cells coerce host tissue to cancer spread. *Bonekey Rep,* 2**,** 371.

MALANCHI, I., PEINADO, H., KASSEN, D., HUSSENET, T., METZGER, D., CHAMBON, P., HUBER, M., HOHL, D., CANO, A., BIRCHMEIER, W. & HUELSKEN, J. 2008. Cutaneous cancer stem cell maintenance is dependent on beta-catenin signalling. *Nature,* 452**,** 650-3.

MALANCHI, I., SANTAMARIA-MARTINEZ, A., SUSANTO, E., PENG, H., LEHR, H. A., DELALOYE, J. F. & HUELSKEN, J. 2012. Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature,* 481**,** 85-9.

MANI, S. A., GUO, W., LIAO, M. J., EATON, E. N., AYYANAN, A., ZHOU, A. Y., BROOKS, M., REINHARD, F., ZHANG, C. C., SHIPITSIN, M., CAMPBELL, L. L., POLYAK, K., BRISKEN, C., YANG, J. & WEINBERG, R. A. 2008. The

epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell,* 133**,** 704-15.

MASSAGUE, J. 2012. TGF-beta signaling in development and disease. *FEBS Lett,* 586**,** 1833.

MAZZONI, A., BRONTE, V., VISINTIN, A., SPITZER, J. H., APOLLONI, E., SERAFINI, P., ZANOVELLO, P. & SEGAL, D. M. 2002. Myeloid suppressor lines inhibit T cell responses by an NO-dependent mechanism. *J Immunol,* 168**,** 689-95.

MCALLISTER, S. S. & WEINBERG, R. A. 2014. The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. *Nat Cell Biol,* 16**,** 717-27.

MERIC, F., LEE, W. P., SAHIN, A., ZHANG, H., KUNG, H. J. & HUNG, M. C. 2002. Expression profile of tyrosine kinases in breast cancer. *Clin Cancer Res,* 8**,** 361-7.

MOMBAERTS, P., IACOMINI, J., JOHNSON, R. S., HERRUP, K., TONEGAWA, S. & PAPAIOANNOU, V. E. 1992. RAG-1-deficient mice have no mature B and T lymphocytes. *Cell,* 68**,** 869-77.

MOREL, A. P., HINKAL, G. W., THOMAS, C., FAUVET, F., COURTOIS-COX, S., WIERINCKX, A., DEVOUASSOUX-SHISHEBORAN, M., TREILLEUX, I., TISSIER, A., GRAS, B., POURCHET, J., PUISIEUX, I., BROWNE, G. J., SPICER, D. B., LACHUER, J., ANSIEAU, S. & PUISIEUX, A. 2012. EMT inducers catalyze malignant transformation of mammary epithelial cells and drive tumorigenesis towards claudin-low tumors in transgenic mice. *PLoS Genet,* 8**,** e1002723.

MOREL, A. P., LIEVRE, M., THOMAS, C., HINKAL, G., ANSIEAU, S. & PUISIEUX, A. 2008. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS One,* 3**,** e2888.

MORIN, P., WICKMAN, G., MUNRO, J., INMAN, G. J. & OLSON, M. F. 2011. Differing contributions of LIMK and ROCK to TGFbeta-induced transcription, motility and invasion. *Eur J Cell Biol,* 90**,** 13-25.

MORRIS, P. G., HUDIS, C. A., GIRI, D., MORROW, M., FALCONE, D. J., ZHOU, X. K., DU, B., BROGI, E., CRAWFORD, C. B., KOPELOVICH, L., SUBBARAMAIAH, K. & DANNENBERG, A. J. 2011. Inflammation and increased aromatase expression occur in the breast tissue of obese women with breast cancer. *Cancer Prev Res (Phila),* 4**,** 1021-9.

MUDDULURU, G., CEPPI, P., KUMARSWAMY, R., SCAGLIOTTI, G. V., PAPOTTI, M. & ALLGAYER, H. 2011. Regulation of Axl receptor tyrosine kinase expression by miR-34a and miR-199a/b in solid cancer. *Oncogene,* 30**,** 2888-99.

MULLER, V., STAHMANN, N., RIETHDORF, S., RAU, T., ZABEL, T., GOETZ, A., JANICKE, F. & PANTEL, K. 2005. Circulating tumor cells in breast cancer: correlation to bone marrow micrometastases, heterogeneous response to systemic therapy and low proliferative activity. *Clin Cancer Res,* 11**,** 3678-85.

MURAOKA-COOK, R. S., SHIN, I., YI, J. Y., EASTERLY, E., BARCELLOS-HOFF, M. H., YINGLING, J. M., ZENT, R. & ARTEAGA, C. L. 2006. Activated type I TGFbeta receptor kinase enhances the survival of mammary epithelial cells and accelerates tumor progression. *Oncogene,* 25**,** 3408-23.

MYERS, K. A., APPLEGATE, K. T., DANUSER, G., FISCHER, R. S. & WATERMAN, C. M. 2011. Distinct ECM mechanosensing pathways regulate microtubule dynamics to control endothelial cell branching morphogenesis. *J Cell Biol,* 192**,** 321-34.

NABA, A., CLAUSER, K. R., HOERSCH, S., LIU, H., CARR, S. A. & HYNES, R. O. 2012. The matrisome: in silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. *Mol Cell Proteomics,* 11**,** M111 014647.

NEVE, R. M., CHIN, K., FRIDLYAND, J., YEH, J., BAEHNER, F. L., FEVR, T., CLARK, L., BAYANI, N., COPPE, J. P., TONG, F., SPEED, T., SPELLMAN, P. T., DEVRIES, S., LAPUK, A., WANG, N. J., KUO, W. L., STILWELL, J. L., PINKEL, D., ALBERTSON, D. G., WALDMAN, F. M., MCCORMICK, F., DICKSON, R. B., JOHNSON, M. D., LIPPMAN, M., ETHIER, S., GAZDAR, A. & GRAY, J. W. 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell,* 10**,** 515-27.

NEWMAN, A. C., NAKATSU, M. N., CHOU, W., GERSHON, P. D. & HUGHES, C. C. 2011. The requirement for fibroblasts in angiogenesis: fibroblast-derived matrix proteins are essential for endothelial cell lumen formation. *Mol Biol Cell,* 22**,** 3791-800.

NGUYEN, D. X., BOS, P. D. & MASSAGUÉ, J. 2009. Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer,* 9**,** 274-284.

NGUYEN, L. V., VANNER, R., DIRKS, P. & EAVES, C. J. 2012. Cancer stem cells: an evolving concept. *Nat Rev Cancer*.

NGUYEN-NGOC, K. V., CHEUNG, K. J., BRENOT, A., SHAMIR, E. R., GRAY, R. S., HINES, W. C., YASWEN, P., WERB, Z. & EWALD, A. J. 2012. ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium. *Proc Natl Acad Sci U S A,* 109**,** E2595-604.

NIE, S., GURREA, M., ZHU, J., THAKOLWIBOON, S., HETH, J. A., MURASZKO, K. M., FAN, X. & LUBMAN, D. M. 2015. Tenascin-C: a novel candidate marker for cancer stem cells in glioblastoma identified by tissue microarrays. *J Proteome Res,* 14**,** 814-22.

NIEMAN, K. M., KENNY, H. A., PENICKA, C. V., LADANYI, A., BUELL-GUTBROD, R., ZILLHARDT, M. R., ROMERO, I. L., CAREY, M. S., MILLS, G. B., HOTAMISLIGIL, G. S., YAMADA, S. D., PETER, M. E., GWIN, K. & LENGYEL, E. 2011. Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. *Nat Med,* 17**,** 1498-503.

NIETO, M. A. 2011. The ins and outs of the epithelial to mesenchymal transition in health and disease. *Annu Rev Cell Dev Biol,* 27**,** 347-76.

NIETO, M. A. 2013. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science,* 342**,** 1234850.

NIK-ZAINAL, S., ALEXANDROV, L. B., WEDGE, D. C., VAN LOO, P., GREENMAN, C. D., RAINE, K., JONES, D., HINTON, J., MARSHALL, J., STEBBINGS, L. A., MENZIES, A., MARTIN, S., LEUNG, K., CHEN, L., LEROY, C., RAMAKRISHNA, M., RANCE, R., LAU, K. W., MUDIE, L. J., VARELA, I., MCBRIDE, D. J., BIGNELL, G. R., COOKE, S. L., SHLIEN, A., GAMBLE, J., WHITMORE, I., MADDISON, M., TARPEY, P. S., DAVIES, H. R., PAPAEMMANUIL, E., STEPHENS, P. J., MCLAREN, S., BUTLER, A. P., TEAGUE, J. W., JONSSON, G., GARBER, J. E., SILVER, D., MIRON, P., FATIMA, A., BOYAULT, S., LANGEROD, A., TUTT, A., MARTENS, J. W., APARICIO, S. A., BORG, A., SALOMON, A. V., THOMAS, G., BORRESEN-DALE, A. L., RICHARDSON, A. L., NEUBERGER, M. S., FUTREAL, P. A., CAMPBELL, P. J., STRATTON, M. R. & BREAST CANCER WORKING GROUP OF THE INTERNATIONAL CANCER GENOME, C. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell,* 149**,** 979-93.

NIOLA, F., ZHAO, X., SINGH, D., CASTANO, A., SULLIVAN, R., LAURIA, M., NAM, H. S., ZHUANG, Y., BENEZRA, R., DI BERNARDO, D., IAVARONE, A. & LASORELLA, A. 2012. Id proteins synchronize stemness and anchorage to the niche of neural stem cells. *Nat Cell Biol,* 14**,** 477-87.

NIOLA, F., ZHAO, X., SINGH, D., SULLIVAN, R., CASTANO, A., VERRICO, A., ZOPPOLI, P., FRIEDMANN-MORVINSKI, D., SULMAN, E., BARRETT, L.,

ZHUANG, Y., VERMA, I., BENEZRA, R., ALDAPE, K., IAVARONE, A. & LASORELLA, A. 2013. Mesenchymal high-grade glioma is maintained by the ID-RAP1 axis. *J Clin Invest,* 123**,** 405-17.

NISHIMURA, T. & KAIBUCHI, K. 2007. Numb controls integrin endocytosis for directional cell migration with aPKC and PAR-3. *Dev Cell,* 13**,** 15-28.

O'BRIEN, C. A., POLLETT, A., GALLINGER, S. & DICK, J. E. 2007. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature,* 445**,** 106-10.

OCANA, O. H., CORCOLES, R., FABRA, A., MORENO-BUENO, G., ACLOQUE, H., VEGA, S., BARRALLO-GIMENO, A., CANO, A. & NIETO, M. A. 2012. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell,* 22**,** 709-24.

OHM, J. E. & BAYLIN, S. B. 2007. Stem cell chromatin patterns: an instructive mechanism for DNA hypermethylation? *Cell Cycle,* 6**,** 1040-3.

OLUMI, A. F., GROSSFELD, G. D., HAYWARD, S. W., CARROLL, P. R., TLSTY, T. D. & CUNHA, G. R. 1999. Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. *Cancer Res,* 59**,** 5002-11.

OMBRATO, L. & MALANCHI, I. 2014. The EMT universe: space between cancer cell dissemination and metastasis initiation. *Crit Rev Oncog,* 19**,** 349-61.

OSHIMORI, N., ORISTIAN, D. & FUCHS, E. 2015. TGF-beta promotes heterogeneity and drug resistance in squamous cell carcinoma. *Cell,* 160**,** 963-76.

OSKARSSON, T., ACHARYYA, S., ZHANG, X. H.-F., VANHARANTA, S., TAVAZOIE, S. F., MORRIS, P. G., DOWNEY, R. J., MANOVA-TODOROVA, K., BROGI, E. & MASSAGU&EACUTE, J. 2011. Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nature Medicine,* 17**,** 867-874.

OSKARSSON, T., BATLLE, E. & MASSAGUE, J. 2014. Metastatic Stem Cells: Sources, Niches, and Vital Pathways. *Cell Stem Cell,* 14**,** 306-321.

OTTEWELL, P. D., COLEMAN, R. E. & HOLEN, I. 2006. From genetic abnormality to metastases: murine models of breast cancer and their use in the development of anticancer therapies. *Breast Cancer Res Treat,* 96**,** 101-13.

OWENS, P., PICKUP, M. W., NOVITSKIY, S. V., GILTNANE, J. M., GORSKA, A. E., HOPKINS, C. R., HONG, C. C. & MOSES, H. L. 2015. Inhibition of BMP signaling suppresses metastasis in mammary cancer. *Oncogene,* 34**,** 2437-49.

PACCEZ, J. D., VASQUES, G. J., CORREA, R. G., VASCONCELLOS, J. F., DUNCAN, K., GU, X., BHASIN, M., LIBERMANN, T. A. & ZERBINI, L. F. 2013. The receptor tyrosine kinase Axl is an essential regulator of prostate cancer proliferation and tumor growth and represents a new therapeutic target. *Oncogene,* 32**,** 689-98.

PACCEZ, J. D., VOGELSANG, M., PARKER, M. I. & ZERBINI, L. F. 2014. The receptor tyrosine kinase Axl in cancer: biological functions and therapeutic implications. *Int J Cancer,* 134**,** 1024-33.

PAGET, S. 1989. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev,* 8**,** 98-101.

PANG, R., LAW, W. L., CHU, A. C., POON, J. T., LAM, C. S., CHOW, A. K., NG, L., CHEUNG, L. W., LAN, X. R., LAN, H. Y., TAN, V. P., YAU, T. C., POON, R. T. & WONG, B. C. 2010. A subpopulation of CD26+ cancer stem cells with metastatic capacity in human colorectal cancer. *Cell Stem Cell,* 6**,** 603-15.

PANTEL, K., BRAUN, S., SCHLIMOK, G. & RIETHMULLER, G. 1993. Micrometastatic tumour cells in bone marrow in colorectal cancer. *Lancet,* 341**,** 501.

PATSIALOU, A., WANG, Y., PIGNATELLI, J., CHEN, X., ENTENBERG, D., OKTAY, M. & CONDEELIS, J. S. 2015. Autocrine CSF1R signaling mediates switching

between invasion and proliferation downstream of TGFbeta in claudin-low breast tumor cells. *Oncogene,* 34**,** 2721-31.

PATSIALOU, A., WYCKOFF, J., WANG, Y., GOSWAMI, S., STANLEY, E. R. & CONDEELIS, J. S. 2009. Invasion of human breast cancer cells in vivo requires both paracrine and autocrine loops involving the colony-stimulating factor-1 receptor. *Cancer Res,* 69**,** 9498-506.

PAUNESCU, V., BOJIN, F. M., TATU, C. A., GAVRILIUC, O. I., ROSCA, A., GRUIA, A. T., TANASIE, G., BUNU, C., CRISNIC, D., GHERGHICEANU, M., TATU, F. R., TATU, C. S. & VERMESAN, S. 2011. Tumour-associated fibroblasts and mesenchymal stem cells: more similarities than differences. *J Cell Mol Med,* 15**,** 635-46.

PECE, S., TOSONI, D., CONFALONIERI, S., MAZZAROL, G., VECCHI, M., RONZONI, S., BERNARD, L., VIALE, G., PELICCI, P. G. & DI FIORE, P. P. 2010. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell,* 140**,** 62-73.

PEINADO, H., OLMEDA, D. & CANO, A. 2007. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer,* 7**,** 415-28.

PEINADO, H. C., LAVOTSHKIN, S. & LYDEN, D. 2011. The secreted factors responsible for pre-metastatic niche formation: Old sayings and new thoughts. *Seminars in Cancer Biology,* 21**,** 139-146.

PFLAUM, J., SCHLOSSER, S. & MULLER, M. 2014. p53 Family and Cellular Stress Responses in Cancer. *Front Oncol,* 4**,** 285.

PICCOLO, S., CORDENONSI, M. & DUPONT, S. 2013. Molecular pathways: YAP and TAZ take center stage in organ growth and tumorigenesis. *Clin Cancer Res,* 19**,** 4925-30.

PITTENGER, M. F., MACKAY, A. M., BECK, S. C., JAISWAL, R. K., DOUGLAS, R., MOSCA, J. D., MOORMAN, M. A., SIMONETTI, D. W., CRAIG, S. & MARSHAK, D. R. 1999. Multilineage potential of adult human mesenchymal stem cells. *Science,* 284**,** 143-7.

PLAKS, V., KONG, N. & WERB, Z. 2015. The Cancer Stem Cell Niche: How Essential Is the Niche in Regulating Stemness of Tumor Cells? *Cell Stem Cell,* 16**,** 225-238.

QIAN, B. Z. & POLLARD, J. W. 2010. Macrophage diversity enhances tumor progression and metastasis. *Cell,* 141**,** 39-51.

QUAIL, D. F. & JOYCE, J. A. 2013. Microenvironmental regulation of tumor progression and metastasis. *Nat Med,* 19**,** 1423-37.

QUAIL, D. F., TAYLOR, M. J. & POSTOVIT, L. M. 2012. Microenvironmental regulation of cancer stem cell phenotypes. *Curr Stem Cell Res Ther,* 7**,** 197-216.

RAOUF, A., ZHAO, Y., TO, K., STINGL, J., DELANEY, A., BARBARA, M., ISCOVE, N., JONES, S., MCKINNEY, S., EMERMAN, J., APARICIO, S., MARRA, M. & EAVES, C. 2008. Transcriptome analysis of the normal human mammary cell commitment and differentiation process. *Cell Stem Cell,* 3**,** 109-18.

REYA, T. & CLEVERS, H. 2005. Wnt signalling in stem cells and cancer. *Nature,* 434**,** 843-50.

REYA, T., MORRISON, S. J., CLARKE, M. F. & WEISSMAN, I. L. 2001. Stem cells, cancer, and cancer stem cells. *Nature,* 414**,** 105-11.

REYMOND, N., D'AGUA, B. B. & RIDLEY, A. J. 2013. Crossing the endothelial barrier during metastasis. *Nat Rev Cancer,* 13**,** 858-70.

RHIM, A. D., MIREK, E. T., AIELLO, N. M., MAITRA, A., BAILEY, J. M., MCALLISTER, F., REICHERT, M., BEATTY, G. L., RUSTGI, A. K., VONDERHEIDE, R. H., LEACH, S. D. & STANGER, B. Z. 2012. EMT and dissemination precede pancreatic tumor formation. *Cell,* 148**,** 349-61.

RIDLEY, A. J., SCHWARTZ, M. A., BURRIDGE, K., FIRTEL, R. A., GINSBERG, M. H., BORISY, G., PARSONS, J. T. & HORWITZ, A. R. 2003. Cell migration: integrating signals from front to back. *Science,* 302**,** 1704-9.

RINGNER, M., FREDLUND, E., HAKKINEN, J., BORG, A. & STAAF, J. 2011. GOBO: gene expression-based outcome for breast cancer online. *PLoS One,* 6**,** e17911.

ROCHLITZ, C., LOHRI, A., BACCHI, M., SCHMIDT, M., NAGEL, S., FOPP, M., FEY, M. F., HERRMANN, R. & NEUBAUER, A. 1999. Axl expression is associated with adverse prognosis and with expression of Bcl-2 and CD34 in de novo acute myeloid leukemia (AML): results from a multicenter trial of the Swiss Group for Clinical Cancer Research (SAKK). *Leukemia,* 13**,** 1352-8.

ROJAS, A., MEHEREM, S., KIM, Y. H., WASHINGTON, M. K., WILLIS, J. E., MARKOWITZ, S. D. & GRADY, W. M. 2008. The aberrant methylation of TSP1 suppresses TGF-beta1 activation in colorectal cancer. *Int J Cancer,* 123**,** 14-21.

ROUSSOS, E. T., KECKESOVA, Z., HALEY, J. D., EPSTEIN, D. M., WEINBERG, R. A. & CONDEELIS, J. S. 2010. AACR special conference on epithelial-mesenchymal transition and cancer progression and treatment. *Cancer Res,* 70**,** 7360-4.

SAITO, Y., LIANG, G., EGGER, G., FRIEDMAN, J. M., CHUANG, J. C., COETZEE, G. A. & JONES, P. A. 2006. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. *Cancer Cell,* 9**,** 435-43.

SAITO, Y., UCHIDA, N., TANAKA, S., SUZUKI, N., TOMIZAWA-MURASAWA, M., SONE, A., NAJIMA, Y., TAKAGI, S., AOKI, Y., WAKE, A., TANIGUCHI, S., SHULTZ, L. D. & ISHIKAWA, F. 2010. Induction of cell cycle entry eliminates human leukemia stem cells in a mouse model of AML. *Nat Biotechnol,* 28**,** 275-80.

SAKUMA, K., AOKI, M. & KANNAGI, R. 2012. Transcription factors c-Myc and CDX2 mediate E-selectin ligand expression in colon cancer cells undergoing EGF/bFGF-induced epithelial-mesenchymal transition. *Proc Natl Acad Sci U S A,* 109**,** 7776-81.

SAMAVARCHI-TEHRANI, P., GOLIPOUR, A., DAVID, L., SUNG, H. K., BEYER, T. A., DATTI, A., WOLTJEN, K., NAGY, A. & WRANA, J. L. 2010. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell,* 7**,** 64-77.

SAUNDERS, R., SUTCLIFFE, A., WOODMAN, L., KAUR, D., SIDDIQUI, S., OKAYAMA, Y., WARDLAW, A., BRADDING, P. & BRIGHTLING, C. 2008. The airway smooth muscle CCR3/CCL11 axis is inhibited by mast cells. *Allergy,* 63**,** 1148-55.

SCAFFIDI, P. & MISTELI, T. 2011. In vitro generation of human cells with cancer stem cell properties. *Nat Cell Biol,* 13**,** 1051-61.

SCHEEL, C. & WEINBERG, R. A. 2011. Phenotypic plasticity and epithelial-mesenchymal transitions in cancer and normal stem cells? *Int J Cancer,* 129**,** 2310-4.

SCHLEGEL, N. C., VON PLANTA, A., WIDMER, D. S., DUMMER, R. & CHRISTOFORI, G. 2015. PI3K signalling is required for a TGFbeta-induced epithelial-mesenchymal-like transition (EMT-like) in human melanoma cells. *Exp Dermatol,* 24**,** 22-8.

SCHNITT, S. J. 2010. Classification and prognosis of invasive breast cancer: from morphology to molecular taxonomy. *Mod Pathol,* 23 Suppl 2**,** S60-4.

SCUTERA, S., FRAONE, T., MUSSO, T., CAPPELLO, P., ROSSI, S., PIEROBON, D., ORINSKA, Z., PAUS, R., BULFONE-PAUS, S. & GIOVARELLI, M. 2009.

Survival and migration of human dendritic cells are regulated by an IFN-alpha-inducible Axl/Gas6 pathway. *J Immunol,* 183**,** 3004-13.

SEGUIN, L., KATO, S., FRANOVIC, A., CAMARGO, M. F., LESPERANCE, J., ELLIOTT, K. C., YEBRA, M., MIELGO, A., LOWY, A. M., HUSAIN, H., CASCONE, T., DIAO, L., WANG, J., WISTUBA, II, HEYMACH, J. V., LIPPMAN, S. M., DESGROSELLIER, J. S., ANAND, S., WEIS, S. M. & CHERESH, D. A. 2014. An integrin beta(3)-KRAS-RalB complex drives tumour stemness and resistance to EGFR inhibition. *Nat Cell Biol,* 16**,** 457-68.

SEVENICH, L., BOWMAN, R. L., MASON, S. D., QUAIL, D. F., RAPAPORT, F., ELIE, B. T., BROGI, E., BRASTIANOS, P. K., HAHN, W. C., HOLSINGER, L. J., MASSAGUE, J., LESLIE, C. S. & JOYCE, J. A. 2014. Analysis of tumour- and stroma-supplied proteolytic networks reveals a brain-metastasis-promoting role for cathepsin S. *Nat Cell Biol,* 16**,** 876-88.

SHAMIR, E. R., PAPPALARDO, E., JORGENS, D. M., COUTINHO, K., TSAI, W. T., AZIZ, K., AUER, M., TRAN, P. T., BADER, J. S. & EWALD, A. J. 2014. Twist1-induced dissemination preserves epithelial identity and requires E-cadherin. *J Cell Biol,* 204**,** 839-56.

SHERIDAN, C. 2013. First Axl inhibitor enters clinical trials. *Nat Biotechnol,* 31**,** 775-6.

SHIBUE, T., BROOKS, M. W. & WEINBERG, R. A. 2013. An Integrin-Linked Machinery of Cytoskeletal Regulation that Enables Experimental Tumor Initiation and Metastatic Colonization. *Cancer Cell*.

SHLUSH, L. I., ZANDI, S., MITCHELL, A., CHEN, W. C., BRANDWEIN, J. M., GUPTA, V., KENNEDY, J. A., SCHIMMER, A. D., SCHUH, A. C., YEE, K. W., MCLEOD, J. L., DOEDENS, M., MEDEIROS, J. J., MARKE, R., KIM, H. J., LEE, K., MCPHERSON, J. D., HUDSON, T. J., CONSORTIUM, H. P.-L. G. P., BROWN, A. M., TRINH, Q. M., STEIN, L. D., MINDEN, M. D., WANG, J. C. & DICK, J. E. 2014. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature,* 506**,** 328-33.

SHOEMAKER, A. R., HAIGIS, K. M., BAKER, S. M., DUDLEY, S., LISKAY, R. M. & DOVE, W. F. 2000. Mlh1 deficiency enhances several phenotypes of Apc(Min)/+ mice. *Oncogene,* 19**,** 2774-9.

SIEMENS, H., JACKSTADT, R., HUNTEN, S., KALLER, M., MENSSEN, A., GOTZ, U. & HERMEKING, H. 2011. miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. *Cell Cycle,* 10**,** 4256-71.

SINGH, S. K., HAWKINS, C., CLARKE, I. D., SQUIRE, J. A., BAYANI, J., HIDE, T., HENKELMAN, R. M., CUSIMANO, M. D. & DIRKS, P. B. 2004. Identification of human brain tumour initiating cells. *Nature,* 432**,** 396-401.

SLEEMAN, K. E., KENDRICK, H., ASHWORTH, A., ISACKE, C. M. & SMALLEY, M. J. 2006. CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast Cancer Res,* 8**,** R7.

SNEDDON, J. B. & WERB, Z. 2007. Location, location, location: the cancer stem cell niche. *Cell Stem Cell,* 1**,** 607-11.

SNEDDON, J. B., ZHEN, H. H., MONTGOMERY, K., VAN DE RIJN, M., TWARD, A. D., WEST, R., GLADSTONE, H., CHANG, H. Y., MORGANROTH, G. S., ORO, A. E. & BROWN, P. O. 2006. Bone morphogenetic protein antagonist gremlin 1 is widely expressed by cancer-associated stromal cells and can promote tumor cell proliferation. *Proc Natl Acad Sci U S A,* 103**,** 14842-7.

SOSA, M. S., BRAGADO, P. & AGUIRRE-GHISO, J. A. 2014. Mechanisms of disseminated cancer cell dormancy: an awakening field. *Nat Rev Cancer,* 14**,** 611-22.

SPARMANN, A. & VAN LOHUIZEN, M. 2006. Polycomb silencers control cell fate, development and cancer. *Nat Rev Cancer,* 6**,** 846-56.

STANKIC, M., PAVLOVIC, S., CHIN, Y., BROGI, E., PADUA, D., NORTON, L., MASSAGUE, J. & BENEZRA, R. 2013. TGF-beta-Id1 signaling opposes Twist1 and promotes metastatic colonization via a mesenchymal-to-epithelial transition. *Cell Rep,* 5**,** 1228-42.

STRATTON, M. R. 2011. Exploring the genomes of cancer cells: progress and promise. *Science,* 331**,** 1553-8.

SWEET, D. T., CHEN, Z., WILEY, D. M., BAUTCH, V. L. & TZIMA, E. 2012. The adaptor protein Shc integrates growth factor and ECM signaling during postnatal angiogenesis. *Blood,* 119**,** 1946-55.

TAIT, C. R., DODWELL, D. & HORGAN, K. 2004. Do metastases metastasize? *J Pathol,* 203**,** 515-8.

TALMADGE, J. E. & GABRILOVICH, D. I. 2013. History of myeloid-derived suppressor cells. *Nat Rev Cancer,* 13**,** 739-52.

TAM, W. L. & WEINBERG, R. A. 2013. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat Med,* 19**,** 1438-49.

TARIN, D. & CROFT, C. B. 1969. Ultrastructural features of wound healing in mouse skin. *J Anat,* 105**,** 189-90.

TAUBE, J. H., HERSCHKOWITZ, J. I., KOMUROV, K., ZHOU, A. Y., GUPTA, S., YANG, J., HARTWELL, K., ONDER, T. T., GUPTA, P. B., EVANS, K. W., HOLLIER, B. G., RAM, P. T., LANDER, E. S., ROSEN, J. M., WEINBERG, R. A. & MANI, S. A. 2010. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci U S A,* 107**,** 15449-54.

THIERY, J. P. 2009. [Epithelial-mesenchymal transitions in cancer onset and progression]. *Bull Acad Natl Med,* 193**,** 1969-78; discussion 1978-9.

THIERY, J. P., ACLOQUE, H., HUANG, R. Y. & NIETO, M. A. 2009. Epithelial-mesenchymal transitions in development and disease. *Cell,* 139**,** 871-90.

TIWARI, A., JUNG, J. J., INAMDAR, S. M., NIHALANI, D. & CHOUDHURY, A. 2013. The myosin motor Myo1c is required for VEGFR2 delivery to the cell surface and for angiogenic signaling. *Am J Physiol Heart Circ Physiol,* 304**,** H687-96.

TSAI, J. H., DONAHER, J. L., MURPHY, D. A., CHAU, S. & YANG, J. 2012. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell,* 22**,** 725-36.

UBERTI, B., DENTELLI, P., ROSSO, A., DEFILIPPI, P. & BRIZZI, M. F. 2010. Inhibition of beta1 integrin and IL-3Rbeta common subunit interaction hinders tumour angiogenesis. *Oncogene,* 29**,** 6581-90.

VALIENTE, M., OBENAUF, A. C., JIN, X., CHEN, Q., ZHANG, X. H., LEE, D. J., CHAFT, J. E., KRIS, M. G., HUSE, J. T., BROGI, E. & MASSAGUE, J. 2014. Serpins promote cancer cell survival and vascular co-option in brain metastasis. *Cell,* 156**,** 1002-16.

VESUNA, F., VAN DIEST, P., CHEN, J. H. & RAMAN, V. 2008. Twist is a transcriptional repressor of E-cadherin gene expression in breast cancer. *Biochem Biophys Res Commun,* 367**,** 235-41.

VISCO, V., BAVA, F. A., D'ALESSANDRO, F., CAVALLINI, M., ZIPARO, V. & TORRISI, M. R. 2009. Human colon fibroblasts induce differentiation and proliferation of intestinal epithelial cells through the direct paracrine action of keratinocyte growth factor. *J Cell Physiol,* 220**,** 204-13.

VOGELSTEIN, B., FEARON, E. R., HAMILTON, S. R., KERN, S. E., PREISINGER, A. C., LEPPERT, M., NAKAMURA, Y., WHITE, R., SMITS, A. M. & BOS, J. L. 1988. Genetic alterations during colorectal-tumor development. *N Engl J Med,* 319**,** 525-32.

VOGELSTEIN, B., PAPADOPOULOS, N., VELCULESCU, V. E., ZHOU, S., DIAZ, L. A., JR. & KINZLER, K. W. 2013. Cancer genome landscapes. *Science,* 339**,** 1546-58.

VUORILUOTO, K., HAUGEN, H., KIVILUOTO, S., MPINDI, J. P., NEVO, J., GJERDRUM, C., TIRON, C., LORENS, J. B. & IVASKA, J. 2011. Vimentin regulates EMT induction by Slug and oncogenic H-Ras and migration by governing Axl expression in breast cancer. *Oncogene,* 30**,** 1436-48.

WAKEFIELD, L. M. & HILL, C. S. 2013. Beyond TGFbeta: roles of other TGFbeta superfamily members in cancer. *Nat Rev Cancer,* 13**,** 328-41.

WALSH, D. W., GODSON, C., BRAZIL, D. P. & MARTIN, F. 2010. Extracellular BMP-antagonist regulation in development and disease: tied up in knots. *Trends Cell Biol,* 20**,** 244-56.

WANG, W., WYCKOFF, J. B., FROHLICH, V. C., OLEYNIKOV, Y., HUTTELMAIER, S., ZAVADIL, J., CERMAK, L., BOTTINGER, E. P., SINGER, R. H., WHITE, J. G., SEGALL, J. E. & CONDEELIS, J. S. 2002. Single cell behavior in metastatic primary mammary tumors correlated with gene expression patterns revealed by molecular profiling. *Cancer Res,* 62**,** 6278-88.

WANG, X., KRUITHOF-DE JULIO, M., ECONOMIDES, K. D., WALKER, D., YU, H., HALILI, M. V., HU, Y. P., PRICE, S. M., ABATE-SHEN, C. & SHEN, M. M. 2009. A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature,* 461**,** 495-500.

WELLNER, U., SCHUBERT, J., BURK, U. C., SCHMALHOFER, O., ZHU, F., SONNTAG, A., WALDVOGEL, B., VANNIER, C., DARLING, D., ZUR HAUSEN, A., BRUNTON, V. G., MORTON, J., SANSOM, O., SCHULER, J., STEMMLER, M. P., HERZBERGER, C., HOPT, U., KECK, T., BRABLETZ, S. & BRABLETZ, T. 2009. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat Cell Biol,* 11**,** 1487-95.

WELM, B. E., TEPERA, S. B., VENEZIA, T., GRAUBERT, T. A., ROSEN, J. M. & GOODELL, M. A. 2002. Sca-1(pos) cells in the mouse mammary gland represent an enriched progenitor cell population. *Dev Biol,* 245**,** 42-56.

WICKI, A., LEHEMBRE, F., WICK, N., HANTUSCH, B., KERJASCHKI, D. & CHRISTOFORI, G. 2006. Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer Cell,* 9**,** 261-72.

WONG, G. S. & RUSTGI, A. K. 2013. Matricellular proteins: priming the tumour microenvironment for cancer development and metastasis. *Br J Cancer,* 108**,** 755-61.

WOOD, L. D., PARSONS, D. W., JONES, S., LIN, J., SJOBLOM, T., LEARY, R. J., SHEN, D., BOCA, S. M., BARBER, T., PTAK, J., SILLIMAN, N., SZABO, S., DEZSO, Z., USTYANKSKY, V., NIKOLSKAYA, T., NIKOLSKY, Y., KARCHIN, R., WILSON, P. A., KAMINKER, J. S., ZHANG, Z., CROSHAW, R., WILLIS, J., DAWSON, D., SHIPITSIN, M., WILLSON, J. K., SUKUMAR, S., POLYAK, K., PARK, B. H., PETHIYAGODA, C. L., PANT, P. V., BALLINGER, D. G., SPARKS, A. B., HARTIGAN, J., SMITH, D. R., SUH, E., PAPADOPOULOS, N., BUCKHAULTS, P., MARKOWITZ, S. D., PARMIGIANI, G., KINZLER, K. W., VELCULESCU, V. E. & VOGELSTEIN, B. 2007. The genomic landscapes of human breast and colorectal cancers. *Science,* 318**,** 1108-13.

WYCKOFF, J. B., WANG, Y., LIN, E. Y., LI, J. F., GOSWAMI, S., STANLEY, E. R., SEGALL, J. E., POLLARD, J. W. & CONDEELIS, J. 2007. Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res,* 67**,** 2649-56.

XU, J., LAMOUILLE, S. & DERYNCK, R. 2009. TGF-beta-induced epithelial to mesenchymal transition. *Cell Res,* 19**,** 156-72.

YACHIDA, S., JONES, S., BOZIC, I., ANTAL, T., LEARY, R., FU, B., KAMIYAMA, M., HRUBAN, R. H., ESHLEMAN, J. R., NOWAK, M. A., VELCULESCU, V. E., KINZLER, K. W., VOGELSTEIN, B. & IACOBUZIO-DONAHUE, C. A. 2010. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature,* 467**,** 1114-7.

YI, J. M., DHIR, M., VAN NESTE, L., DOWNING, S. R., JESCHKE, J., GLOCKNER, S. C., DE FREITAS CALMON, M., HOOKER, C. M., FUNES, J. M., BOSHOFF, C., SMITS, K. M., VAN ENGELAND, M., WEIJENBERG, M. P., IACOBUZIO-DONAHUE, C. A., HERMAN, J. G., SCHUEBEL, K. E., BAYLIN, S. B. & AHUJA, N. 2011. Genomic and epigenomic integration identifies a prognostic signature in colon cancer. *Clin Cancer Res,* 17**,** 1535-45.

YOSHIDA, T., AKATSUKA, T. & IMANAKA-YOSHIDA, K. 2015. Tenascin-C and integrins in cancer. *Cell Adh Migr,* 9**,** 96-104.

YU, M., BARDIA, A., WITTNER, B. S., STOTT, S. L., SMAS, M. E., TING, D. T., ISAKOFF, S. J., CICILIANO, J. C., WELLS, M. N., SHAH, A. M., CONCANNON, K. F., DONALDSON, M. C., SEQUIST, L. V., BRACHTEL, E., SGROI, D., BASELGA, J., RAMASWAMY, S., TONER, M., HABER, D. A. & MAHESWARAN, S. 2013. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science,* 339**,** 580-4.

YUAN, T. L. & CANTLEY, L. C. 2008. PI3K pathway alterations in cancer: variations on a theme. *Oncogene,* 27**,** 5497-510.

ZACKULAR, J. P., BAXTER, N. T., IVERSON, K. D., SADLER, W. D., PETROSINO, J. F., CHEN, G. Y. & SCHLOSS, P. D. 2013. The gut microbiome modulates colon tumorigenesis. *MBio,* 4**,** e00692-13.

ZAMAN, M. H., MATSUDAIRA, P. & LAUFFENBURGER, D. A. 2007. Understanding effects of matrix protease and matrix organization on directional persistence and translational speed in three-dimensional cell migration. *Ann Biomed Eng,* 35**,** 91-100.

ZHANG, X. H., JIN, X., MALLADI, S., ZOU, Y., WEN, Y. H., BROGI, E., SMID, M., FOEKENS, J. A. & MASSAGUE, J. 2013. Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell,* 154**,** 1060-73.

ZHANG, Z., LI, X. J., LIU, Y., ZHANG, X., LI, Y. Y. & XU, W. S. 2007. Recombinant human decorin inhibits cell proliferation and downregulates TGF-beta1 production in hypertrophic scar fibroblasts. *Burns,* 33**,** 634-41.