UNIVERSITAT DE BARCELONA

Análisis de los perfiles de expresión de microRNAs en cáncer de mama y de la implicación de la familia miR-200 en metástasis

Lourdes Sánchez-Cid Pérez

Universitat de Barcelona

Facultat de Biologia

Expression profile of microRNAs in breast cancer and role of the miR-200 family in the metastatic process

Lourdes Sánchez-Cid Pérez

2016

PROGRAMA DE DOCTORAT EN BIOMEDICINA

Universitat de Barcelona

Análisis de los perfiles de expresión de microRNAs en cáncer de mama y de la implicación de la familia miR-200s en metástasis

---

Expression profile of microRNAs in breast cancer and role of the miR-200 family in the metastatic process

Memoria presentada por Lourdes Sánchez-Cid Pérez para optar al título de Doctor por la Universitat de Barcelona

Este trabajo ha sido realizado en el departamento de Biología Celular del Instituto de Biología Molecular de Barcelona (IBMB) del Consejo Superior de Investigaciones Científicas (CSIC) y en el Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS) .

Director: Timothy Thomson Okatsu

Director: Pedro Luis Fernández Ruiz

Tutora: Marta Cascante Serratosa

Doctorando: Lourdes Sánchez-Cid Pérez

Introduction

1. The Mammary Gland

1.1 Anatomy

The mammary gland is a complex tubulolobular gland with apocrine secretion that ensures milk production under hormonal control.

The adult female breast parenchyma contains 15 to 20 lobes of glandular tissue arranged in a radial manner around the nipple (1).The lobes are separated by connective tissue that supports the breast and adipose tissue that determines the size and shape of the organ. Each lobe consists of 20 to 40 lobules that contain the glandular units and a lactiferous duct that collects the milk from the lobules and transports it to the nipple via the lactiferous pore. Immediately deep to the nipple, each lactiferous duct dilates to form the lactiferous sinus, which serves as a reservoir for milk during lactation. Beyond the sinus and deeper into the parenchyma, the large lactiferous ducts subdivide into successive branches of diminishing size until finally subdividing into the terminal ducts that drain the lobule. The lobule is composed of groups of small glandular structures, the acini or alveoli. Although the number of acini in each lobule is extremely variable, it ranges from 10 to 100 acini. The functional units of the breast are the terminal duct lobular units (TDLU), comprising the lobule and its paired terminal duct (Figure 1).

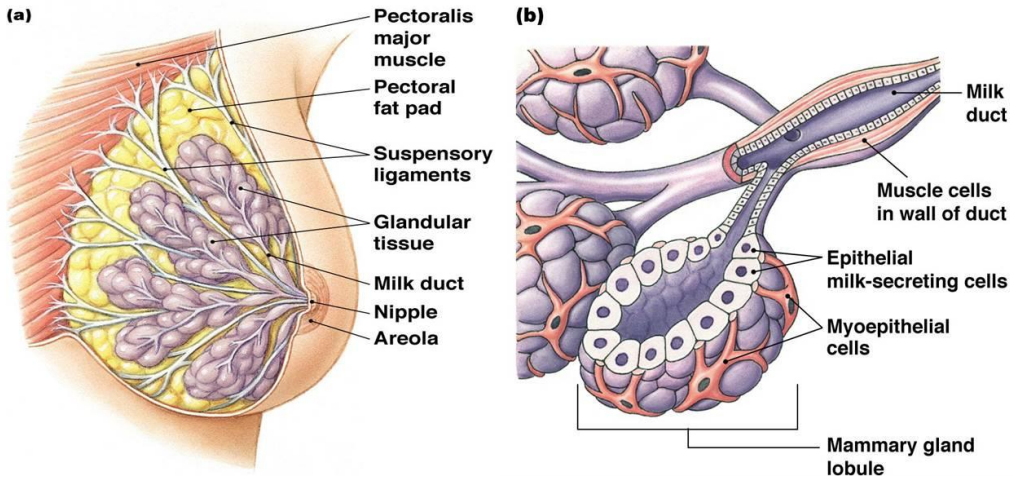

Figure 1. The mammary gland. a) Anatomical structure of the mammary gland. **b)** Representation of the organization of a mammary gland lobule.

1.2 Histology

In the normal mammary gland, ducts and alveoli are lined by an inner layer of secretory luminal epithelial cells, surrounded by transversely oriented myoepithelial cells (ME) (Figure 1b). During lactation, luminal cells produce and secrete milk, whereas basal myoepithelial cells contract to facilitate milk secretion. In addition, myoepithelial cells influence the differentiation, polarity, proliferation, migration and invasion of adjacent luminal epithelial cells. They also produce the basement membrane (BM), which lies immediately outside the myoepithelial cell layer and forms a physical barrier separating the epithelial and stromal compartments (2). The connective tissue or stroma is composed of an extracellular matrix and various cell types including fibroblasts, endothelial cells, and infiltrating leukocytes. Not only does it provide a scaffold and supplies nutrition for the organ, but it also regulates mammary epithelial cell function *via* paracrine, physical, and hormonal interactions.(3)

1.3 Physiological changes

The mammary gland undergoes cycles of development, differentiation and apoptosis during the adult life. These cyclic changes, regulated by hormones and growth factors, affect the tissue structure and function (4).

During childhood and before puberty, the female breast is composed of a branching ductal system that lacks lobular units. At puberty, increasing levels of estrogen stimulate the development of glandular tissue in the female breast and also causes the breast to increase in size through the accumulation of adipose tissue. Progesterone stimulates the development of the duct system. Once formed, the lactiferous ducts and interlobular duct system are stable and unaffected by fluctuating hormone levels during the menstrual cycle, pregnancy, and lactation. The TDLUs, however, are dynamic and undergo changes with alterations in hormone levels, involving both the epithelium and the intralobular stroma (1, 4, 5).

During the menstrual cycle, after ovulation, the terminal duct epithelium proliferates, and the number of terminal ducts within a lobule increases while the intralobular stroma is edematous and loose. As the levels of estrogen and progesterone fall with the onset of menstruation, there is an increase in apoptosis in the TDLU and lymphocytes infiltrate the intralobular stroma, which becomes dense. The TDLU finally regresses to its resting appearance (5).

During pregnancy and lactation, there is a striking increase in the number of terminal ducts and the TDLUs are enlarged as acini increase in number in response to the rising sex hormone levels. Prolactin stimulates the production of milk within the glandular tissue and oxytocin causes myoepithelial cell contraction and the ejection of milk from the glands. With the cessation of lactation, the alveoli collapse and the mammary gland involutes through apoptosis and proteinase-driven tissue remodeling, resulting in a gland that resembles the pre-pregnant state (5, 6).

After menopause, the low hormone levels causes atrophy of TDLUs so that only small residual foci remain. The lactiferous ducts and interlobular duct system remain, but the interlobular stroma is reduced in amount accompanied by a relative increase in adipose tissue (5).

1.4 The mammary stem cell hierarchy

The mammary epithelium is organized into two primary lineages, the luminal and myoepithelial. The luminal lineage comprises the ductal cells lining the lumina and alveolar cells that form the alveolar units that produce milk. Myoepithelial cells are located in a basal position adjacent to the basement membrane and have contractile properties. Each cell lineage is characterized by distinct markers: luminal cells express steroid hormone receptors and “luminal” cytokeratin proteins (CK8, CK18, and CK19), among other markers, while myoepithelial cells express smooth muscle actin and “basal” cytokeratins (CK5 and CK14).

The dynamic nature and regenerative capacity of the mammary epithelium suggests the presence of long-lived and renewable mammary stem cells (MaSCs). The first studies to support de existence of MaSCs were done in mouse. In 1959, De Ome and colleagues developed the fat pad transplantation assay, which involves the implantation of mammary epithelial cells or tissue fragments into cleared mammary fat pads of recipient mice, while the stromal environment remains intact to support growth of transplanted cells (7). It is one of the most used and useful methods to demonstrate repopulating capacity and self-renewing potential of mammary cells. These engrafting cells are termed mammary repopulating units, or MRUs. By transplantation of small epithelium fragments, they obtained morphologically similar outgrowths in recipient mice demonstrating the existence of cells with characteristics of stem cells. Subsequent studies used either mammary explants or cell suspensions to reconstitute the ductal tree that, in addition, could be serially transplanted for several generations (8, 9). It was in 1998 when Kordon and

Smith first suggested that a single cell could give rise to an entire mammary epithelium by tracking epithelial cells with retrovirus (10). In 2006, by isolating cellular subpopulations using specific combinations of antibodies against cell surface proteins for fluorescence activated cell sorting (FACS) of single cell suspensions followed by the "gold standard" transplantation assay, Shackleton and Stingl identified mouse MaSCs (11, 12). These MaSCs are nearly exclusively contained within the CD49f^{hi}CD29^{hi}CD24$^{+/mod}$Sca1^{low} population of basal/myoepithelial (ME) cells. The term basal/ME refers to a heterogeneous, compartment comprising MaSCs (approximately 0.02% of basal cells), mature myoepithelial cells and presumptive basal progenitor intermediates. MaSCs can be preferentially enriched over other basal cells based on their higher expression levels of CD24 and EpCAM (13, 14).

1.4.1 Human mammary stem and progenitor cells

It has proven difficult to study MaSCs in the human mammary gland due to the lack of a comparable assay to the mouse fat pad transplantation. Initial studies for prospective isolation of human MaSCs involved sorting of human mammary epithelial cells (HMECs) from human breast epithelium to evaluate their *in vitro* colony-forming ability and subsequent characterization of their progeny. These studies led to the identification of bipotent human mammary stem cells that are able to generate both luminal and basal progeny (15, 16). In this pioneering work, Stingl and colleagues used antibodies against the luminal antigens EpCAM and MUC1 and the basal antigen CALLA to fractionate the epithelium and assess their ability to form colonies *in vitro* (17). They proposed the existence of three types of epithelial progenitors: a luminal-restricted progenitor (MUC-1$^+$/CALLA$^-$/EpCAM$^+$) that could give rise to alveolar colonies, a myoepithelial progenitor (MUC-1$^-$/CALLA$^+$/EpCAM$^-$) and a bipotent progenitor (MUC-1$^+$/CALLA$^+$/EpCAM$^+$) capable of generating both luminal and myoepithelial cells. As EpCAM was found within the luminal layer, they suggested that the bipotent progenitor had a luminal origin. Later on, they added CD49f to further fractionate the human progenitor population and suggested that bipotent cells, characterized by a CD49f$^+$/EpCAM^{low} immunophenotype, had a basal position in the mammary epithelium *in vivo* (15). By also using EpCAM and CD49f to fractionate the lineage-negative (Lin-) epithelium, that is depleted of hematopoietic CD45$^+$ and endothelial CD31$^+$ cells, the Eaves and Lindeman groups obtained similar results (18-20). Eirew and colleagues from the Eaves laboratory developed an analogous method to the mouse fat pad transplantation that allowed the detection of human mammary epithelial stem cells with *in vivo*

regenerative ability. They dissociated cells from reduction mammoplasties and seeded them with fibroblasts in collagen gels which were then implanted under the kidney capsule of hormone-treated immunodeficient mice. The gels contained bilayered epithelial structures with similar spatial organization and marker expression to normal human mammary tissue. They also tested their self-renewal activity by their ability to form secondary transplants. Only $CD49f^{+}/EpCAM^{-/low}$ cells, which showed a basal phenotype, were bipotent while $CD49f^{+}/EpCAM^{+}$ cells were luminal-restricted progenitors (18). Lim and colleagues confirmed that the $CD49f^{+}/EpCAM^{-}$ population of human basal/ME cells had repopulating ability in humanized mouse stroma (termed human-in-mouse or HIM) and together with the luminal $CD49f^{+}/EpCAM^{+}$ population, both exhibited progenitor activity based on CFC assays in Matrigel (19). It was clear from these studies that the mammary epithelium is organized in a hierarchical manner with bipotent MaSC/basal ($CD49f^{hi}/EpCAM^{-/low}$) cells residing at the apex of the hierarchy differentiating to myoepithelial-restricted, contributing to the myoepithelial lineage, and luminal-restricted ($CD49f^{+}/EpCAM^{+}$) progenitors that further differentiate into mature luminal cells ($CD49f^{-}/EpCAM^{+}$) (Figure 2).

It was not until 2012 when a clarifying work on progenitor cells was published (Figure 2). Shehata, Stingl and colleagues demonstrated the existence of heterogeneity within the mammary luminal compartment (13). They subfractionated the luminal compartment, containing luminal-restricted progenitors ($EpCAM^{+}CD49f^{+}$) and mature ($EpCAM^{+}CD49f^{-}$) cells, into three subpopulations of cells with varying degrees of luminal differentiation and proliferative capacity. The relatively differentiated ALDH negative population ($EpCAM^{+}CD49f^{+}ALDH^{-}$) could be ductal progenitors while the less differentiated ALDH positive fraction ($EpCAM^{+}CD49f^{+}ALDH^{+}$) may represent a pool of progenitors that are primed to generate alveoli during pregnancy. One quarter of the samples had an additional population characterized by low expression of ERBB3 with an intermediate phenotype between luminal and basal that they suggest to be also involved in alveologenesis (13).

Current models of the mammary epithelial hierarchy derive mainly from fat pad transplantation studies, which may not reflect the behavior of these cells *in situ*. As a result, several groups have implemented the use of *in vivo* lineage tracing to study the behavior of MaSCs and luminal progenitors in their native tissue as not only stem cells but also their progeny are marked permanently. While transplantation studies suggest that multipotent stem cells maintain the two lineages of the mammary gland, in 2011 van Keymeulen and colleagues used lineage tracing and concluded that two separate populations of unipotent stem cells, rather than bipotent, are

responsible for driving the homeostasis of the mouse mammary gland after birth and during pregnancy (21). They monitored the activity of stem or progenitor cells at different developmental stages by inducible expression of lineage-specific cytokeratin promoters in either basal/ME or luminal cells with doxycycline or tamoxifen, which activated YFP expression. It was the first *in situ* evidence that luminal and basal/ME progenitors exist and that together, they control the regenerative potential of the adult mammary gland. While these evidences do not rule out the presence of multipotent stem cells, they suggest that they do not contribute significantly to adult mammary gland homeostasis (21). In line with their results, the Stingl laboratory observed that myoepithelial cells sorted from myoepithelial-reporter transgenic mice had MRU capacity. Using an inducible lineage model, they also demonstrated that, in the virgin state and during pregnancy, cells of the myoepithelial lineage expressing smooth muscle actin function as unipotent long-lived stem cells and contribute only to the basal cell layer (22). A subsequent lineage-tracing study added a further layer of complexity through the discovery that Wnt-responsive Axin2$^+$ cells in the mouse mammary gland can switch fate according to the developmental stage. When labeled postpuberty they only gave rise to basal cells while in pregnancy they contributed to luminal and myoepithelial cells forming the alveolar structures. Furthermore, the basal-restricted Axin2$^+$ cells defined *in vivo* behaved as bipotent stem cells when transplanted, suggesting that transplantation enhances the regenerative potential and, therefore, may not be physiological (23).

In 2014, the Visvader laboratory used a novel 3D imaging technique to track bipotent MaSCs at a clonal level *in situ*. Combination of a multicolor Confetti reporter 3D imaging techniques allowed them to see portions of the ductal tree at single-cell resolution (24). They marked the basal cell lineage using the K5 promoter and observed clones formed by myoepithelial and luminal cells. They confirmed the existence of bipotency in two additional models driven by K14 and Lgr5 promoters. In fact, they were multipotent as they also contributed to alveologenesis during pregnancy. Curiously, they suggested that during alveolar expansion, myoepithelial cells migrate from adjacent ductal regions. In contradiction with the previous model and in agreement with the results form transplantation studies, they proposed a model whereby bipotent MaSCs with a basal phenotype drive the process of expansion, which is then continued by unipotent progenitors generated from stem cells that have undergone asymmetric division (Figure 2). They suggested that not only do bipotent MaSCs contribute to morphogenesis in the postnatal gland, but that they also have an active role in the homeostasis during the adult phase (24).

The identification of luminal progenitors *via* lineage tracing also changed the paradigm regarding the involution process. It was generally accepted that during mammary involution, differentiated alveolar cells undergo apoptosis and the alveolar compartment is reconstituted in subsequent pregnancies from undifferentiated mammary stem cells or alveolar precursors. In 2002, by mapping the progeny of cells expressing whey acidic protein (WAP), largely activated in late pregnancy and lactation, a subset of differentiated alveolar cells located at the extremity of ducts were able to bypass apoptosis, surviving involution, and functioned as self-renewing alveolar precursors in subsequent pregnancies. In transplants, these cells expressed certain features of multipotent stem cells, displaying the ability to clonally expand and give rise to luminal and alveolar cells (Figure 2). However, because they lacked the capacity to produce myoepithelial cells, they were not considered mammary stem cells. They were termed parity-induced mammary epithelial cells (PI-MECs) (25).

Luminal progenitor cells have been recently tracked in the mouse mammary gland using a number of inducible lineage tracing models. By monitoring cells expressing the luminal progenitor gene Elf5, it has been seen that they are long-lived luminal restricted progenitors that contribute to morphogenesis during puberty and ductal maintenance in adulthood. Furthermore, the Elf5-labeled pool was diminished at the end of each involution cycle, indicating that each round of alveologenesis is driven by a new pool of luminal progenitor cells (24). Similarly, others showed that some lobules generated during the second pregnancy were from previously labeled cells, whereas others derived from unlabeled cells that likely represent new alveolar progenitors (26).

van Keymeulen *et al* traced luminal cells using the K8 promoter. The $K8^{+}$ population contains luminal stem cells with self-renewal capacities that are able to differentiate into both luminal and milk-producing cells during lactation. They even survived three rounds of pregnancy-lactation-involution, making it evident that they are long-term self-renewing cells. On the other hand, when they used an inducible model to track luminal K18-expressing cells, these cells did not clonally expand even during pregnancy and lactation, suggesting that they are more committed luminal cells (21). It is curious, however, that the promoters from K8 and K18 mark different luminal cells when they are coexpressed in the luminal lineage.

Although these studies highlight the great heterogeneity and complexity of the breast differentiation hierarchy, it is true that the different techniques and models used have added further confusion (27). On the one hand, some lineage-tracing studies have revealed that the differentiation potential of epithelial cells in transplantation assays does not necessarily reflect

their fate under physiological conditions (21, 23) as tissue disaggregation, niche removal and its random replacement during transplantation may alter the stem and progenitor cell phenotype and activity. *In vivo* cell tracking has been regarded as a potential solution to avoid this problem. However, their disadvantages should also be taken into account. It could be possible that quiescent cells such as multipotent stem cells are more difficult to label than progenitors. Also, promoter infidelity or low transcriptional activity could determine differences in labeling of identical cells. In transgenic models, the level and timing of expression of a promoter could be conditioned by its insertion site in the genome. In addition, since ER is essential in the mammary development, the dose of tamoxifen in tamoxifen-inducible models is critical when using ER-driven vectors. In fact, it has been seen that the MaSC pool is very sensitive to hormone deprivation (28). Lastly, the results from these studies rely on the selection of highly specific promoters that accurately mimic expression of the endogenous gene.

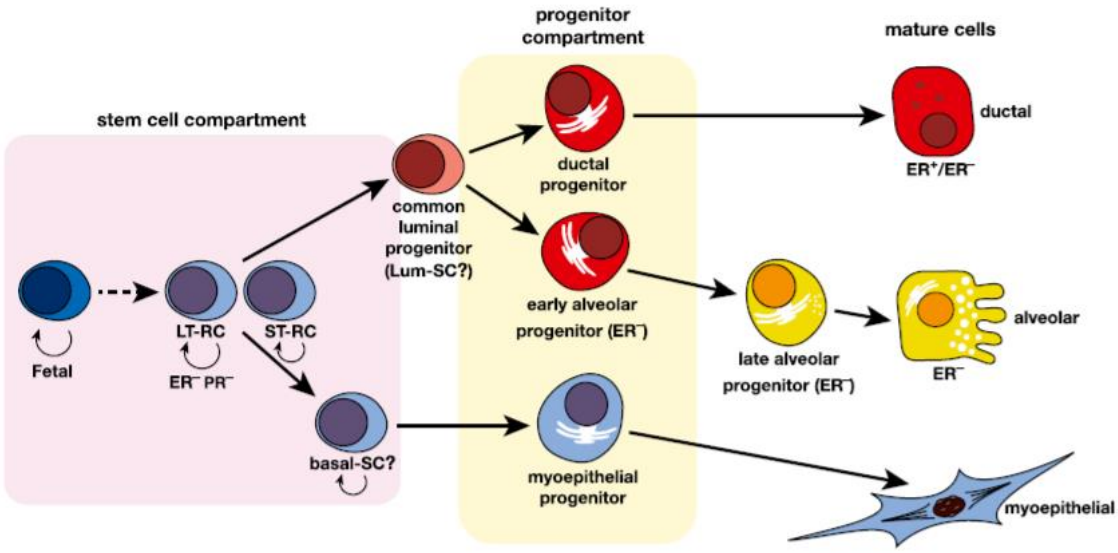

Figure 2. Hypothetical model of the mammary epithelial hierarchy. The mammary stem cell compartment comprise a long-term (LT-RC) and short-term (ST-RC) repopulating population of cells that are multipotent and can give rise to commited progenitors for the myoepithelial and luminal (ductal and alveolar) epithelial lineages. Luminal progenitors are restricted to either a ductal or alveolar cell fate while myoepithelial progenitors are restricted to myoepithelial cells (29).

1.4.2 Molecular regulators of the MaSC function

In a similar way to the differentiation of embryonic stem cell into different lineages, organized transcriptional networks governed by master transcription factors (TFs) play a key role in the differentiation of tissue-specific stem cells. However, the master TFs that control adult SC programs remain poorly defined, particularly in epithelial tissues.

Slug and Sox9 have been identified by the Weinberg laboratory as TFs that regulate the MaSC state, acting cooperatively to determine the MaSC state (30). It was observed that Slug alone is sufficient to induce MaSCs from luminal progenitors but not from differentiated luminal cells, while expression of Sox9 in differentiated luminal cells can led by itself to the formation of luminal cells with progenitor activity. Transient coexpression of both TFs is sufficient to convert differentiated luminal cells into bipotent MaSCs with long-term repopulating activity. In addition, Sox9 collaborates with endogenously expressed Slug to induce MaSC activity in the population of basal cells. However, as MaSCs and myoepithelial cells in the basal population could not be separated with flow cytometry markers, it was not possible to determine whether these basal cells were progenitors or differentiated myoepithelial cells or Sox9 had expanded a preexisting MaSC population. What Guo et *al*. did reveal was that Slug and Sox9 regulate the basal and luminal lineage programs, respectively, and aggregation of both was required for MECs to enter and remain in the MaSC state. Previous work from the Weinberg laboratory showed that induction of an epithelia-to-mesenchymal transition (EMT) program through EMT-inducing TFs could induce stem-like cells in immortalized human MECs (31). In line with their results, they suggest that other potent EMT-TFs, such as Snail, could also cooperate with Sox9 to induce SC formation. Other studies have also highlighted the importance of Slug in maintaining the MaSC/basal cell population (32).

The Wnt/β-catenin pathway is another important pathway that controls self-renewal of stem cells in the mammary gland and other organs. Studies in mouse using the Wnt-responsive $Axin2^{lacZ}$ reporter showed that adult mammary glands contain a Wnt-responsive cell population that is enriched for stem cells. MaSC sensitized to Wnt signals have a competitive advantage in mammary gland reconstitution assays. In addition, treatment with the Wnt3a ligand allows long-term culture of MaSCs that retain their multipotency following many generations in culture (33). Knockout of Lrp5 (34) and loss of function mutations of Lrp6 (35), both Wnt co-receptors, caused significantly reduced Wnt activity in the mammary basal cell compartment and impaired gland branching, suggesting their implication in maintaining stem cell activity. There is also evidence that the Wnt pathway has a role in regulating the human MaSC subset isolated from reduction

mammoplasties. In addition to transforming growth factor (TGF)-β, canonical and non-canonical Wnt signaling collaborate to first induce activation of the EMT program via paracrine signals and thereafter maintaining the mesenchymal/stem cell state in an autocrine manner. These signals also control the interconversion of primary stem- and progenitor cell-containing (basal) to lineage-restricted (luminal) MECs (36). Activation of Wnt/β-catenin signaling via Pygo2, a histone methylation reader, suppresses luminal and alveolar differentiation of the MaSC-enriched basal cells. Pygo2 facilitates binding of β-catenin to the Notch3 gene in MaSC/basal cells to maintain this gene in a ‘‘poised’’ bivalent chromatin state (37).

Other positive epigenetic regulators are the polycomb-repressive complex-1 (PRC1) component Bmi1 and the PRC2 component Ezh2. Bmi-1 regulates the self-renewal of human mammary stem cells and controls premature alveolar differentiation induced by pregnancy hormones (38-40). Ezh2 controls the activity of mammary stem cells or the descendant progenitors as targeted deletion reduces the repopulating frequency *in vivo* and it decreases clonogenic activity of both basal and luminal progenitors. It also impairs alveologenesis during pregnancy and lactation (40).

On the contrary, the Notch signaling pathway restricts the expansion of the MaSC pool. Downregulation of the Notch effector Cbf-1 in the MaSCs-enriched population leads to increased repopulating activity *in vivo* and aberrant ductal morphogenesis (41). Likewise, loss of p53 leads to the expansion of stem cells both *in vitro* and *in vivo* (42). p53 regulates self-renewal of mammary SCs, preferentially through asymmetric divisions, to maintain a constant number of SCs in the mammary gland.

1.4.3 Molecular regulators of the luminal lineage

Elf5 is a master regulator in alveolar development (43). In mammary glands, Elf5 is highly induced during pregnancy and is primarily restricted to the luminal epithelial cell (43, 44). Studies performed with conditional knockout mice (KO) showed that ablation of Elf5 caused lactation failure and blocked alveolar morphogenesis due in part to impaired Stat5 activity (43, 45). Conversely, overexpression of Elf5 in an inducible transgenic model caused alveolar differentiation and milk secretion in virgin mice, disrupting ductal morphogenesis (45). In addition to regulating terminal differentiation of alveolar cells, Elf5 also plays a critical role in determining cell fate and regulating the stem/progenitor function of the mammary

epithelium. Targeted deletion of Elf5 in the mammary glands leads to accumulation of cell types with dual luminal/basal properties such as coexpression of cytokeratins 8 (K8) and 14 (K14) and an increase in $CD61^+$ luminal progenitor population during pregnancy (46). The abnormal increase in $K14^+K8^+$ cells may represent the $CD61^+$ luminal progenitors blocked in the differentiation towards mature luminal cells that had already been reported in previous studies (45). In addition, it was also reported that loss of Elf5 triggers an increase in MaSCs during pregnancy and in virgin mice that could be explained in part by hyperactivation of the Notch signaling pathway (46). Also, Elf5 was found to directly repress the transcription of Slug, a known EMT inducer and master regulator of mammary stem cells, suggesting that it suppresses a basal program while promoting a luminal lineage by preventing the dedifferentiation of luminal progenitors to stem-like cells (47).

GATA3 is a master regulator of the luminal lineage. In a microarray screening, GATA3 appeared to be the most highly expressed transcription factor in the pubertal mammary epithelium compared to the stroma. Its expression is restricted to luminal cells and absent in myoepithelial cells (48). Using a mammary epithelium-specific Gata3 knockout, it was shown that GATA3 is necessary for mammary gland development. With the onset of puberty, the mammary glands of Gata3 knockout mice fail to develop terminal end buds (TEBs) and the epithelium fails to invade the stroma, suggesting a role in ductal elongation and branching (48, 49). Deletion of Gata3 using a doxycycline-inducible system caused severe cellular defects in the luminal epithelium including de-differentiation of the luminal cells, disorganization of the duct, a decrease in cell–cell adhesion and an increase in cell proliferation. Acute loss of Gata3 leads to an expansion of the de-differentiated luminal cell population prior to cell death while long-term loss led to caspase-mediated luminal cell death and insufficient lactation. Therefore, GATA3 is necessary in the adult mammary gland to maintain the integrity and function of the luminal epithelium (48, 49). Not only is it essential for the differentiation of luminal progenitor cells into alveolar cells during pregnancy (48, 49), but GATA3 also regulates the differentiation of progenitors into mature ductal cells and the activity of MaSCs or their immediate descendants (50). It is necessary and sufficient for luminal cell specification since ectopic expression in the MaSC subset induces milk protein gene synthesis in the absence of a lactogenic stimulus (49). Among the several epithelial transcription factors that cooperate with Gata3, FOXA1, an important regulator of estrogen receptor (ER), was identified as a downstream target that acts in coordination with Gata3 (49). Additional transcription factors that may cooperate with GATA3 include MSX2, FOXP4, TRPS1, ELF5, EHF and RUNX1 (49).

The canonical Notch pathway is more active in the luminal cell compartment with prominent expression of the active form of Notch1 and its target genes in luminal progenitor cells *in vivo*. Constitutive activation of Notch1 in MaSCs promotes luminal cell commitment but not commitment to the myoepithelial lineage (41). A recent lineage tracing study also identified a population of mammary luminal progenitors that express the Notch1 receptor (51). Notch1-expressing cells were found to be multipotent in the embryo stage, when they can give rise to all mammary cell types including myoepithelial and both $ERα^{pos}$ and $ERα^{neg}$ luminal cells, but they progressively restrict their lineage potential. In the postnatal mammary gland, Notch1 expression is restricted to ERα- and PR-negative luminal cells. However, despite lacking hormone receptors, they are highly responsive to hormones. In fact, $Nothch1^{pos}$ cells traced in pregnancy, when high levels of circulating hormones are released, massively expande and give rise to alveoli. They represent alveolar progenitors and may correspond to the previously mentioned PI-MECs (26). Non-canonical Notch signaling, via Notch3, modulates Wnt signaling through the Wnt signaling receptor Frizzled7 (FZD7). FZD7-expressing cells are found more frequently in the luminal progenitor-enriched subpopulation of cells obtained from breast reduction samples compared to the undifferentiated bipotent progenitors (52). Concordantly, Notch3 alone regulates the commitment of the undifferentiated bipotential progenitors to the luminal cell fate (53). While lineage tracing experiments show that all three Notch receptors are exclusively expressed in the luminal compartment (51, 54, 55), they seem to define distinct cell types.

Besides contributing to the maintenance of stem cell identity through epigenetic modifications, EZH2 is required to maintain the luminal cell pool during puberty and may limit differentiation of luminal progenitors into differentiated luminal cells. Consistent with this, knockdown of EZH2 impaired TEB formation and ductal elongation and it also reduced lobuloalveolar expansion during pregnancy (56).

The formation of alveoli during pregnancy is under the influence of progesterone and prolactin. Prolactin mediates its effects through the Prolactin receptor and activation of the downstream Jak2-Stat5 pathway and Elf5 (57). Similar to the phenotypes observed in mammary glands deficient for Gata3 and Elf5, loss of Stat5 impairs alveologenesis. While loss of STAT5A/5B does not affect the stem cell population and its ability to form mammary ducts, luminal progenitors are greatly reduced and unable to form alveoli during pregnancy. Therefore, Stat5 is required for the generation of alveolar progenitors from stem cells (58) as well as for alveolar proliferation and differentiation during pregnancy (59).

1.4.4 Hormonal signaling between mammary epithelial subtypes

There is substantial evidence that estrogen and progesterone mediate their morphogenic effects on the mammary gland through paracrine signals from steroid receptor-positive cells to adjacent hormone receptor-negative cells stimulating their proliferation (Figure 3) (60-62). A population of ER-expressing luminal epithelial cells with little *in vivo* stem cell activity was isolated, that is distinct from the stem/progenitor cells located in the basal compartment, suggesting that the effects of estrogen on MaSCs are indirect (63). It was later observed that mouse MaSCs are highly responsive to hormone signaling, despite lacking the receptors for estrogen or progesterone, and at least one mediator has been identified. The MaSC activity increases upon treatment with steroid hormones and conversely, a reduction is observed after deprivation. A remarkable increase in the number of MaSCs is observed in pregnancy and it was proposed to be mediated in a paracrine manner by Rank ligand (RankL), which is released from neighboring steroid receptor-positive luminal cells where it is induced by progesterone (28, 64). RankL also mediates the induction of Elf5 that occurs in PR^{neg} luminal progenitors in response to progesterone forcing their differentiation towards the alveolar phenotype (65). Forced Elf5 expression also reduces stem cell numbers, suggesting that RankL may also act on stem cells to force their division via induction of Elf5 expression.

In addition to paracrine signals that travel from luminal sensor population to other luminal cells and basally located mammary stem cells, basal cells themselves can send signals back to regulate the fate of the luminal epithelium. Deletion of the transcription factor p63 (Trp63) exclusively within basal cells of the adult gland during pregnancy leads to accumulation of luminal progenitors that faile to differentiate into alveolar cells resulting in lactation failure. NRG1 was identified as the paracrine factor, the transcription of which depends on p63, required to promote maturation of luminal progenitors through the activation of the ERBB4/STAT5A signaling (66).

Progesterone also induces changes in chromatin structure and gene expression, via Ezh2, that determine cellular fate. During pregnancy, activation of Ezh2 upon hormone sensing results in a decrease of repressive H3K27me3 marks in key luminal genes required for luminal differentiation and milk production that are therefore upregulated while the opposite pattern is found in luminal genes expressed in the steady-state gland and downregulated during pregnancy (40). Therefore, hormones can regulate cell-fate switching by changing the histone methylation profile besides activating specific pathways by paracrine effectors.

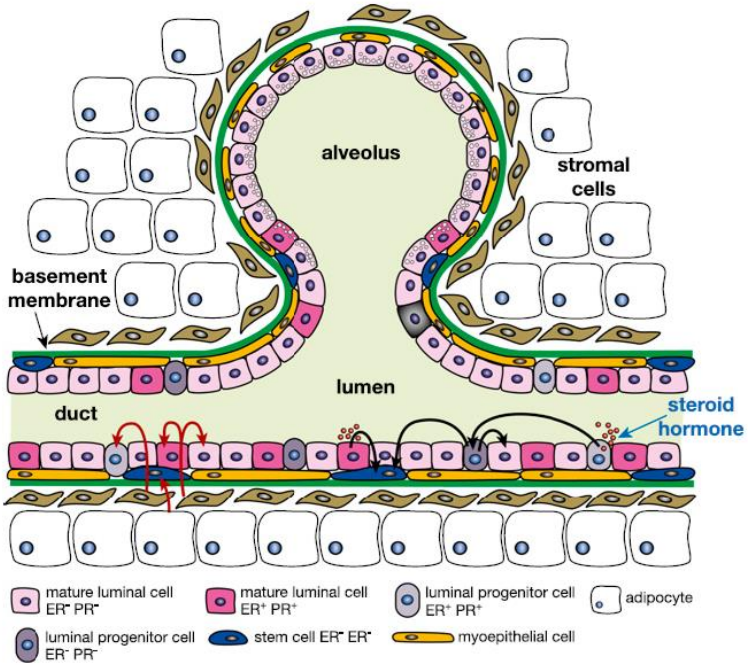

Figure 3. Schematic representation of the potential regulatory crosstalk between the epithelial subtypes. Steroid hormones activate either mature or progenitor ER positive epithelial cells that in turn, activate hormone receptor negative cells (depicted by black arrows). The red arrows depict a further layer of interaction between stromal cells (fibroblasts and adipocytes) and mammary epithelial cells lining the ducts (29).

2. Breast Cancer

2.1 Classification

2.1.1 Histological classification

Human breast cancers are heterogeneous, both in their pathology and in their molecular profiles. Unlike colon cancer, breast tumorigenesis does not necessarily progress in a sequential linear way from well-differentiated to poorly differentiated tumors (67, 68).

According to histological features, they have been categorized into 18 different subtypes. However, they can be broadly divided into *in situ* and invasive/ infiltrating carcinomas (IBC). Breast carcinoma *in situ* is further sub-classified as either ductal (DCIS) or lobular (LCIS). DCIS is more common and heterogeneous, which has given rise to five histological subtypes: comedo, cribiform, micropapillary, papillary and solid. It is defined as a premalignant proliferation of neoplastic epithelial cells contained within the lumen of mammary ducts (intraductal). DCIS are lined by a layer of semi-continuous myoepithelial cells and surrounded by an intact basement membrane (67). The fact that DCIS is found adjacent to invasive disease in the vast majority of IBCs at the time of diagnosis and that IBCs developes in the same anatomical site as DCIS lesions led to the hypotheses that it is a preceding stage.

As tumor proliferation persists by interrupting the basement membrane and entering into the surrounding breast tissue, the cancer is referred to as either invasive lobular carcinoma (ILC), or invasive ductal carcinoma (IDC) (69). The most common type is invasive ductal carcinoma, not otherwise specified (IDC NOS), comprising about ~70-80% of cases while invasive lobular carcinomas (ILC) represent, with 10% of cases, the next most frequent type (70). Together, they account for the majority of IBC while the remainder, known as “special types”, are classified as medullary, neuroendocrine, tubular, apocrine, metaplastic, mucinous, inflammatory, comedo, adenoid cystic and micropapillary type (70). The histological type is linked to prognosis. While IDC NOS, ILC, medullary and apocrine carcinomas have similar 10-year survival rates, adenoid cystic, medullary, mucinous and tubular carcinomas show better overall outcomes (71-73). ILC, characterized by the lack of E-cadherin expression, has a substantially better clinical outcome than does IDC (74).

In the past, breast cancer progression was thought of as a linear pathway with progressive steps starting with normal mammary epithelial cells, followed by hyperplasia, DCIS and LCIS resulting in IDC and ILC, respectively, and finally metastasis (69, 75). While DCIS often presents itself as a precursor to IDC, the progression is frequently nonlinear and breast cancer as a whole is considered to be a complex process that evolves through distinct pathways (75, 76).

2.1.2 Histological grade

Another parameter assessed by pathologists is the grade of differentiation of the tumor. It is a powerful indicator of prognosis based on the combination of the nuclear grade, tubule formation and mitotic rate. Following the Scarff-Bloom-Richardson system, each element is given a score of 1 (best) to 3 (worst) and the scores of all three components are added together to determine the grade. The lowest score is given to well-differentiated tumors that all form tubules and have a low mitotic rate. This allows classification of tumors into grade I, intermediate or grade II and grade III tumors; each of them corresponding to well differentiated, moderately and poorly differentiated tumors (77).

2.1.3 Immunopathological classification

Besides the histological features, specific markers are assessed by immunohistochemistry (IHC) to define subtypes with differential prognosis and identify tumors susceptible to targeted treatments.

The most commonly analyzed markers are estrogen receptor (ER), progesterone receptor (PR) and human epidermal receptor 2 (HER2). A rough measure of the protein expression levels is given by combining the percentage of tumoral cells expressing the protein and the staining intensity to determine the positivity or negativity. When HER2 results are ambiguous by IHC, samples are subjected to fluorescence in situ hybridization (FISH) testing for genomic amplifications of HER2. If the overall ratio of copies of the HER2 gene related to those of its chromosome is greater than 2.2, the case is considered HER2 positive (HER2+) (78). Cases are assigned individually to one of these categories based on the combination of these markers: ER+ (ER+/HER2–), HER2+ (ER–/HER2+), triple negative (TN; ER–/PR–/HER2–), and triple positive (ER+/PR+/HER2+) (79).

From a prognostic viewpoint, ER+ tumors exhibit the best overall outcome. HER2+ tumors had been previously associated with poor outcome but thanks to the use of HER2-targeted therapies, they now exhibit an improved outcome when treated (80, 81). On the other hand, TN tumors are linked to the worst prognosis among these subtypes, while triple positive have an intermediate prognosis between ER+ cases and HER2+ prior to the introduction of the targeted treatment (79).

From the point of view of their usefulness in guiding treatment decisions, ER status is used to identify tumors that may respond to endocrine therapy, including ER antagonists and aromatase inhibitors, which target ER-dependent signaling. PR status correlates in general with ER status but it has less clinical significance because ER+/PR+ cases may not receive additional benefit from endocrine therapy compared with ER+/PR- cases (82, 83). HER2+ cases are treated with the anti-HER2 monoclonal antibody trastuzumab, which binds to HER2 disrupting HER2-dependent signaling and mediates antibody-dependent cytotoxicity (84). The TN group presents significant heterogeneity as they are only defined by the absence of markers. In fact, an investigation of expression profiles from 587 TN breast tumors further subdivided these into 6 groups (85). There is currently no standard targeted therapy for cases assessed as ER– and HER– by IHC, although this represents an intensive area of research.

2.1.4 TNM classification

One of the most important factors in determining therapeutic options and, more importantly, prognosis, is tumor stage. In 1987, the Union International Against Cancer (UICC) and the American Joint Committee of Cancer (AJCC) staging systems were unified into a single TNM staging system to be able to communicate clinical information without ambiguity. It considers the extent of the primary tumor (T), regional lymph node involvement (N) and presence of distant metastasis (M).

The TNM system offers a very precise description of the degree of dissemination but there existed the need to condense all the information into homogeneous groups regarding survival data. As a result, once the TNM categories have been determined, this information is combined in a process called stage grouping, which allows patient stratification into categories with similar prognosis and therapeutic decisions. There are five stages, in which stages 0-II are considered early stage breast cancers. Stage 0 is non-invasive, Stage I displays some indication of local

invasion, while Stage II represents cell invasion into neighboring lymph nodes. By stage III, the breast cancer is considered late-stage due to extensive lymph node involvement. Stage IV is represented by metastasis in which the cancer has spread to distant organs (86).

2.1.5 Molecular subtypes

The advent of high-throughput gene expression profiling technologies has allowed large-scale studies of breast cancer cohorts at the molecular level. The first molecular classification with a major impact on breast cancer research was proposed by Perou and colleagues (87, 88). They characterized variations in gene expression patterns in a set of 65 surgical specimens of human breast tumors from 42 different individuals, using cDNA microarrays representing 8,102 human genes. They selected a subset of 496 genes, termed the "intrinsic" gene signature, which consisted of genes with significantly greater variation in expression between different tumors than between paired samples from the same tumor. Hierarchical clustering of the samples based on the use of the "intrinsic signature" provided a classification into 4 main "intrinsic subtypes" or "molecular subtypes" that have prognostic relevance in terms of survival: Luminal-like, Basal-like, Normal-like and the HER-2-enriched (87). In a subsequent study, they refined the classification by separating the luminal-like into the Luminal A and Luminal B groups, with characteristic gene expression profiles and different prognosis (88). A detailed analysis of genes differentially expressed in ER-negative tumors has demonstrated that basal breast cancers are a heterogeneous group with at least four main subtypes (89). Other recent studies have also identified a new breast cancer intrinsic subtype known as Claudin-low or mesenchymal-like (90). This subtype is also a triple negative and is marked by high expression of mesenchymal genes such as ZEB1 and ZEB2 as well as low expression levels of claudin-3, claudin-4, claudin-7, and E-cadherin (90). It appears to be enriched with cells showing distinct biological properties associated with mammary stem cells and tumor initiating potential (19, 85, 91, 92). It is linked to metaplastic breast cancer and poor outcome (92).

It is important to note that breast cancer heterogeneity makes it difficult to reconcile the clinicopathological features with the molecular data. The classical pathological markers ER, PR and HER2 used in the clinic for tumor classification do not fully recapitulate the intrinsic subtypes (90). These markers, alone or in combination, are not accurate surrogates (93). Thus, tumors with positive expression of hormone receptors and negative for HER2 (HR+/HER2-) contain mainly the luminal A and luminal B subtypes. In the HR+/HER2- group of tumors, the most

frequent subtypes are luminal B and HER2-enriched. Within HR-/HER2+ tumors, about 50-88% correspond to the HER2-enriched subtype followed by the other poor prognostic subtypes. Finally, within the triple negative subtype (HR-/HER2-), the basal-like and claudin-low account for the majority (90).

Luminal tumors are the most common subtype. Both luminal A and B are ER+ as well, as they express the luminal cytokeratins 8 and 18. Luminal A tumors are associated with higher levels of ESR1, ER and ER-regulated genes (88), decreased proliferation (94-96) and improved overall outcome (88, 94, 97), and patients with this tumor subtype exhibit better prognosis than patients with luminal B tumors. Conversely, luminal B tumors often have lower expression levels of hormone receptors, higher grade and index proliferation rates; and can be HER2 positive (98).

The rest of subtypes contain predominantly ER- cases (87). The molecular ERBB2+ subtype overlaps in general with the clinically defined HER2+ tumors and are high grade (88). The normal-like molecular subtype resembles normal epithelial tissue (87) and may comprise cases in which samples contain large amounts of non-tumor tissue (93, 94). The basal-like subtype appears to overlap substantially with TN breast cancers (ER-, PR-, HER2-). These tumors are associated with younger patient age, develop more commonly in African-American women and especially among pre-menopausal women (99). They are associated with high-grade and also characteristic histologic features (100). They also express cytokeratins 5/6 from the basal epithelial layer and epidermal growth factor receptor (EGFR), although clinically validated thresholds for these are still lacking. Furthermore, it revealed an immune response gene expression signature that identifies a good prognosis subtype in ER-negative tumors.

In terms of patient outcome, luminal A tumors are the ones exhibiting the best prognosis while there are no differences in survival among the rest of subtypes with poor prognosis; these being luminal B, HER2-enriched, basal-like and claudin-low (90).

Luminal tumors are treated with hormone therapy with or without chemotherapy. While showing a good response to hormone therapy, they respond poorly to conventional chemotherapy (101). According to prognostic predictors, such as OncotypeDx, that strongly consider expression of proliferation related genes, luminal B as well as HER2-enriched and basal-like tumors are classified as associated with a high risk of recurrence. In contrast to luminal B, luminal A tumors are given a low recurrence score showing that treatment responses vary between luminal subtypes (98). Thus, luminal A tumors could be adequately treated with endocrine therapy as they show favorable relapse-free and disease specific survival outcomes

after treatment with tamoxifen (98, 102). Luminal B tumors are not only relatively chemo-insensitive tumors, but they are also associated with poor prognoss and are relatively hormone-resistant. However, in HR+/HER2+ breast tumors, which represent about 20% of Luminal B tumors, combined endocrine and anti-HER2 therapy enhances progression-free survival (103).

Though HER2 overexpression and basal-like tumors have poor prognosis, they are sensitive to neoadjuvant chemotherapy with anthracyclines and taxanes, showing a significantly higher rates of complete response than luminal tumors (104). Other studies evaluating IHC markers as surrogates of the molecular subtypes also found that a complete response to chemotherapy was significantly better among triple negative and ER-/HER2+ than ER+ tumors (105). However, despite the lower rates of response to therapy, disease-free survival was still better for patients with ER+ tumors due to the higher rate of relapse in triple negative and ER-/HER2+ patients when tumor cells have not been completely eradicated, which explains their worse prognosis. Thus, when there is no residual disease, some triple-negative patients can have good long-term survival outcomes (105). Unlike basal tumors, those that overexpress HER2 can be treated with targeted therapy. Despite this success, many patients do not benefit from trastuzumab (106). Basal tumors are the ones with a lower survival and a higher risk of local or regional relapse. Besides, they have a rapid growth and a tendency to metastasize to visceral organs. Their only therapeutic option is chemotherapy as there are no targeted therapies available for these tumors (101). However, studies are focusing on the efficacy of particular chemotherapies and/or targeted therapies such as the PARP inhibitors and anti-CSC therapies to treat the triple-negative population (107).

3. Overview on Cancer Progression

According to Doug Hanahan and Robert Weinberg, despite their phenotypic and genotypic diversity, most if not all human tumor types have in common six biological capabilities that are acquired during their multistep development. These hallmarks of cancer are: sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis (108). Two additional emerging hallmarks were proposed a decade later. One involves the capability to modify, or reprogram, cellular metabolism to support neoplastic proliferation. The second allows cancer cells to evade immunological destruction, in particular by T and B lymphocytes, macrophages, and natural killer cells (109). Additionally, two characteristics allow the acquisition of these hallmarks. On one hand, genomic instability endows cancer cells with genetic alterations that drive tumor progression. On the other hand, inflammation by innate immune cells can promote tumorogenesis (109). In addition, cancer cells recruit normal cells residing in the stroma to create the tumor microenvironment (TME), contributing to the acquisition of cancer hallmarks (110).

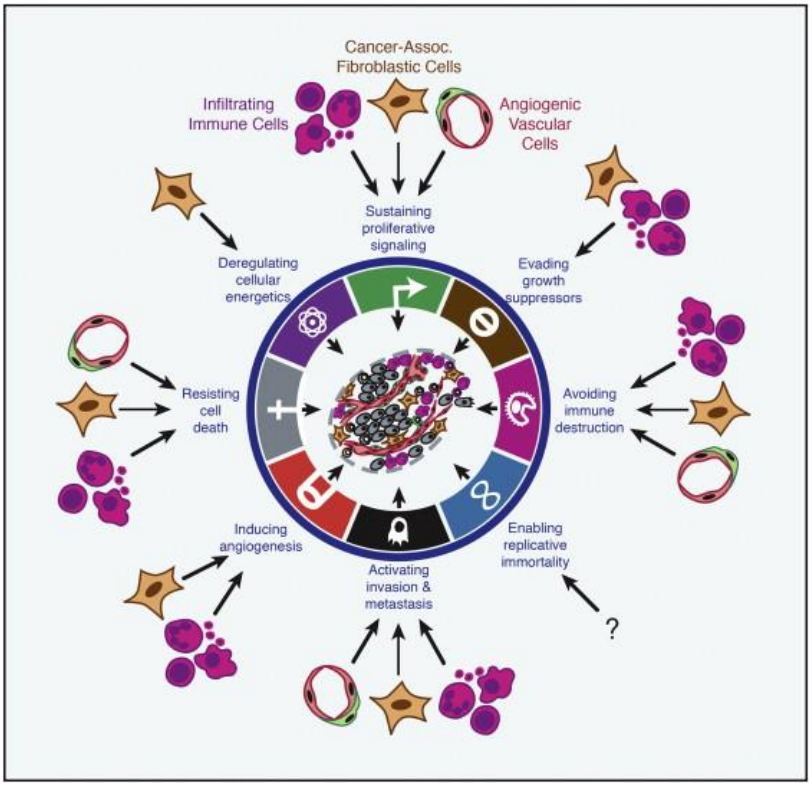

Figure 4. Hallmarks of cancer. Six established and two emerging biological capabilities for cancer progression established in 2000 and 2011, respectively (108, 109).

3.1 Tumor Heterogeneity

Deep sequencing studies have revealed the subclonal composition and thus, the genetic heterogeneity that exists across all types of cancers (111). These molecular differences can occur either between different patients with the same tumour type (intertumour heterogeneity), or within the same patient (intratumour heterogeneity). Intratumoral heterogeneity is the coexistence of different subpopulations of tumor cells in a tumor; each endowed with particular phenotypes, as a result of the acquisition of genetic and epigenetic events that can emerge either sequentially within a single lineage or in parallel in multiple, independent lineages (112). Evidence from tumor sequencing suggests that these genetically different subclones are generated by branching evolution (113-115). Furthermore, different subclones can appear within different regions of a tumor. It is unclear how spatial heterogeneity arises in primary tumours but it could possibly reflect the presence of different micro-environmental niches (116). There also exists a temporal heterogeneity, between the primary tumour and subsequent local or distant recurrence in the same patients which further complicates its clinical management (111).

Evolution of the tumorogenic process is not only determined by genetic alterations in tumor cells but also by the advantage such mutations confer in a given environment. This is determined by the interaction of neoplastic cells with the extracellular matrix, vascular endothelial and immune cells that form the tumor niche and impose selective pressures such as hypoxia, acidity and the presence of growth factors, further contributing to the tumor heterogeneity (117).

3.1.1 Models accounting for heterogeneity in tumors

The first model to explain tumor heterogeneity was proposed by Nowell in 1976, stating that cancer development follows an evolutionary process parallel to the Darwinian natural selection (118) . The **clonal evolution model** postulates that tumor initiation and progression results from the acquisition of multiple and sequential genetic mutations in a random single cell providing it with a survival advantage over adjacent normal cells. The progeny of the cell with increased fitness will expand and produce a clonal dominating population. Over time and following the Darwinian rules, additional advantageous mutations can arise, endowing a further growth

advantage to another cell within the clone. As unique subclones arise, the fittest subclone will dominate while less fit subclones can disappear or remain forming reservoirs from which evolution can continue. In either case, several subpopulations coexist within the tumor and evolve independently resulting in tumor heterogeneity.

The cancer stem cell **(CSC) model** proposes instead that a subset of tumor cells with stem cell-like properties is responsible for driving tumor progression and recurrence. In a similar way to normal stem cells albeit aberrantly, they have unlimited self-renew capacity (by symmetric and asymmetric division) and the ability to differentiate (119). Because of this, normal stem and progenitor cells are considered the most likely targets of transformation although the model does not address the cell of origin (120). However, the cell-of-origin can influence the hierarchical organization of cancers (121). It is proposed that CSCs undergo epigenetic changes differentiating into a phenotypically diverse progeny of nontumorigenic cancer cells that compose the bulk of the tumor and lack both the self-renewal and long-term repopulating abilities. Therefore, this model implies that tumors are hierarchically organized, with a small subset of dedifferentiated CSCs lying at the apex of the hierarchy being responsible for sustaining tumorigenesis and establishing the cellular heterogeneity. Though not all cancers follow the CSC model, evidence of a hierarchical organization has been demonstrated in various human cancers by isolation of tumorigenic from nontumorigenic cancer cells by marker expression and xenotransplantation in immunocompromised mice (122-125). However, the potential of a cell to produce a tumor in a xenotransplantation assay might not reflect its real behaviour in a cancer patient (126).

Notably, the CSC model proposes that CSCs are responsible for metastatic spread (called metastatic CSCs or mCSCs), therapy failure and recurrence (127).

The main limitation of the CSC/hierarchical model is that it views the tumor as genetically homogeneous. However, as CSC can undergo clonal evolution, the two theories can be coupled (128). Both models state that the origin of a tumor is a single cell that has acquired mutations and gained unlimited proliferative potential. In both cases, microenvironmental factors may influence tumor progression. Furthermore, both theories consider that the presence of stem-like properties would confer a selective growth advantage over the rest of the tumor population. Evidence from the convergence of these models has been found in leukemia studies (113) showing that CSCs within individual cancer patients can be genetically heterogeneous.

In addition to the genetic variation, **tumor cell plasticity** may also contribute to phenotypic and functional plasticity. For instance, melanoma cells can reversibly express cell surface markers and recapitulate the tumor heterogeneity of the initial tumor whether arising from maker-positive or marker-negative cells (126). In addition, regulatory genes may also be transiently or stochastically expressed and induce a nontumorigenic cell to reacquire stem cell-like properties (129). Furthermore, cells within breast cancer cell lines were found to transition stochastically between phenotypic states, based on cell surface expression, to maintain equilibrium (130). This progression towards equilibrium was the result of interconversion between different phenotypic states, which can be explained by a Markov model in which cells transition in a stochastic manner. A prediction of this model is that breast cancer stem-like cells arise *de novo* from non-stem-like cells. In fact, this was precisely observed in a subpopulation of basal-like human mammary epithelial cells; normal and CSC-like cells can arise *de novo* from more differentiated cell types (131).

3.1.2 Cancer Stem Cells

The term CSC refers to malignant cells that share properties and gene programs with *bona fide* stem cells but as a result of transformation upon genetic and epigenetic events, they are no longer stem cells. In consequence, they do not have the capacity to give rise to a wide variety of differentiated progeny as embryonic stem cells do (132). In some cancers, they can arise from transformation of normal stem cells (133).whereas in others, they can arise from restricted progenitors or differentiated cells upon mutations that activate self-renewal mechanisms (134). Either case, they are distinguished from other cancer cells by their ability to self-renew, generating more CSCs, their tumorigenic potential and that they display multilineage differentiation potential giving rise to phenotypically different non-tumorigenic cancer cells with more limited proliferative potential (135). Moreover, they can transition in either way between the stem and non-stem cell like states, demonstrating their plasticity (130).

CSCs have been functionally defined on the basis of their immunophenotype and their clonal long-term repopulation ability when transplanted into immunodeficient mice. However, CSC markers are not universal as they differ between different types of cancer and even within tumors of the same subtype. For example, breast CSC have been extensively isolated based on the immunophenotype CD44+/CD24- and also by high ALDH activity (125, 136). In ER-negative and triple negative breast tumors they did not selectively enrich for CSCs as they were

additionally found in the CD44+/CD24+ fraction (137). Furthermore, within the same tumor, the $ALDH^{high}$ subset overlapped little with the CD44+/CD24- subset (136). Given the significant heterogeneity among patients, CSC markers should be validated in a great number of patients. The existence of multiple CSC pools within a tumor has been described in other solid tumors besides breast cancer (137) and in acute myeloid leukemia (138, 139).

There exists also a great variability in the number of CSCs. Initial reports on several types of cancer suggested that there could be found in a ratio of less than 1 CSC per 1000 cells. Conversely, more recent data indicates that the number of cells with tumorogenic capacity might be much higher and can vary widely from 2.5% to 41% (140, 141). The experimental procedure used, particularly the type of immunodeficient mouse strain, has an important influence (141). As human cells transplanted into mice activate a powerful immune response, transplantation into highly immunocompromised mice can significantly increase the number of cells with tumorogenic ability in some cancers (141, 142).

3.1.2.1 CSC gene transcriptional networks

CSC display many features associated to normal stem cells including core regulatory genes and developmental pathways such as the Hedgehog (Hh), Notch and Wnt pathways (143). Importantly, it was shown that embryonic stem cells (ESCs) and cancer stem cells share transcriptional programs (144). However, a consensus stemness signature has not been yet defined maybe reflecting the existence of different types of stem cells using different mechanisms to achieve pluripotency and self-renewal or due to technical variations.

In the study by Wong and colleagues, they used a gene module map tool to evaluate the presence of stem cell programs in tumors of epithelial origin (144). An “ESC-like gene module” was identified containing 335 genes; several of them associated with pluripotency such as SOX2, c-MYC, DNMT1 or HDAC1. This gene expression program is activated in many human epithelial cancers and repressed in normal tissues. Activation of the ESC-like module is positively associated with poor tumor differentiation, increased risk of metastasis and worse outcome in lung and breast adenocarcinomas. Furthermore, it was detected in the CSC-enriched subpopulation ($CD44^{+}/CD24^{low}$) of human breast cancers indicating that CSC may be at least responsible for the presence of this signature in bulk tumors. In addition, they identified c-Myc as a potential regulator of the ESC-like program. c-Myc, an important reprogramming factor that enhances the formation of induced pluripotent stem cells (iPSCs) (144, 145), was found to be

sufficient to activate the program ESC-like program in adult epithelial cells and reprogram them to epithelial cancer stem cells (144). These finding suggest that induction of CSC and the process of reprogramming may display overlapping programs.

Another group established the first link between an ES-like gene signature and histopathological tumor traits (146). Overexpression of genes associated with embryonic stem cell identity was found in histologically poorly differentiated tumors. The targets of key regulators of ES cell identity – NANOG, OCT4, SOX2 and c-MYC- were also more frequently overexpressed in poorly differentiated aggressive tumors than in well-differentiated tumors. In breast cancers, this ES-like signature was associated with high-grade ER-negative tumors, often of the basal-like subtype, and with poor clinical outcome (146).

3.1.2.2 Role of the CSC niche

As is the case for normal stem cells, CSCs are believed to reside in niches with non-CSC tumoral cells. The CSC niche itself is an anatomically distinct microenvironment that is part of the overall tumor microenvironment. Cells within the CSC niche produce factors that stimulate CSC self-renewal, induce angiogenesis, and recruit immune and other stromal cells that secrete additional factors to promote tumor cell invasion and metastasis (147, 148).

By secreting CXCL12, IL6, and IL8, MSCs promote cancer cell stemness through upregulating NF-kB while CSCs secrete IL6 to attract more MSCs (149). Cancer cells secrete chemokines and cytokines to recruit inflammatory cells and induce immunosuppression. TAM produces TNFa and TGF-b to promote epithelial to mesenchymal transition (EMT) and thus enhance CSC plasticity (149, 150). TGF-b can also directly interact with NF-kB signalling pathways to further enhance cancer cell stemness. Hypoxia further promotes cancer cell stemness by promoting an undifferentiated state through TGF-b and the WNT signalling pathway (36). Besides activation of the WNT pathway, CAFs also stimulate stemness via activation of NOTCH pathway . Both pathways are interconnected with other pathways, like bone morphogenetic protein (BMP) and Hedgehog signalling pathways, to determine the state of differentiation (151). For some of these self-renewal pathways to activate, CSCs need to establish cell-cell contacts that anchor them to the niche and locate them physically adjacent to niche factors that specify self-renewal (152).

The microenvironment surrounding CSCs plays multiple roles including as a mechanical anchorage for the stem cells and in cross-talk communication mediated by direct contact and/or

indirect extracellular factors. CSCs protect their niche and, vice versa, the niche contributes to enhance therapy resistance of CSCs by creating a physical barrier blocking therapeutics.

3.1.2.3 CSC and therapy response

Cancer stem cells are more resistant than differentiated cells and survive to most conventional therapies(153)Additionally, an enrichment of residual CSC-like cells was found in patients after chemotherapy, resulting in tumor recurrence and progression. However, it may not be a general rule and depend on the CSC marker analysed (153, 154).

One of the molecular mechanisms that may explain their intrinsic resistance is the aberrant activation of pathways governing self-renewal and differentiation in stem cells. Developmental signalling pathways, such as Wnt, Notch, and Hedgehog signalling, can activate diverse target genes that mediate chemo-resistance and radio-resistance of CSCs (155, 156). Evasion of apoptosis is also associated with increased CSC survival to chemotherapy. Another intrinsic mechanism that CSC display is overexpression of ABC transporters which reduce intracellular drug levels through efflux of chemotherapeutic agents (157). In addition, they have active DNA damage checkpoints and activate the DNA repair response upon radiotherapy (158, 159). Finally, an extrinsic mechanism of therapeutic resistance would be the microenvironment, through formation of specific niches such as perivascular and hypoxic niches (160).

3.1.3 The epithelial- mesenchymal transition program

The epithelial-mesenchymal transition (EMT) is a naturally occurring transdifferentiation process by which epithelial cells change their phenotype to mesenchymal-like, with the concomitant acquisition of a marked migratory and invasive behaviour (Figure 5). It is a highly conserved and essential mechanism during implantation, embryogenesis and organ formation (161).In adult tissues, it is quickly activated in response to wound healing, organ regeneration and organ fibrosis (161). More than a decade of research has recognized its pathological role in the progression and dissemination of cancer.

Under normal conditions, epithelial cells exhibit apical-basal polarity and are tightly connected to one another laterally via adherens and tight junctions ensuring the structural integrity of epithelial cell sheets. Normally, they interact with the basement membrane via its basal surface.

In contrast, mesenchymal cells exhibit front-back polarity, repress cell-cell adhesion molecules and attach loosely to the extracellular matrix via focal adhesions, which favours their motile and invasive phenotype (162). The epithelial and mesenchymal states can be regarded as the two extremes of a transition where intermediate states correspond to partial EMTs. In fact, intermediate states in wound healing have been described, where cells are at equilibrium and can revert to the epithelial phenotype before the wound is healed (Figure5). Breast carcinosarcomas also exhibit an intermediate EMT with cells from the epithelial component switching from a fully epithelial to a partially mesenchymal state while maintaining epithelial characteristics (163). That is the case for metaplastic and claudin-low carcinomas that have a tendency to undergo EMT (164).

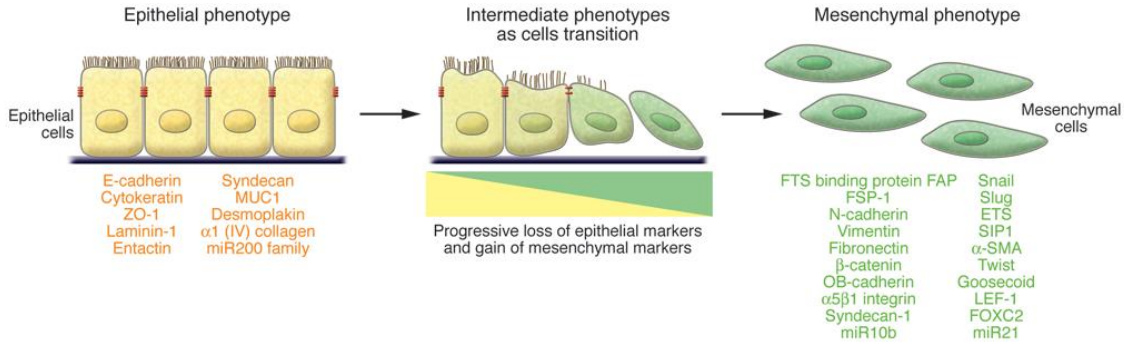

Figure 5. The epithelial to mesenchymal transition. Cells with an epithelial phenotype, characterized by the markers shown in orange, experience a transition to a mesenchymal phenotype, characterized by the expression of markers shown in green. In this transition, a wide array of intermediate phenotypes expressing both types of markers can be generated (161).

Hallmarks of the EMT process are the activation of transcription factors known as EMT-inducers (Snail, Twist, Zeb and others) that repress the transcription of E-cadherin and polarity genes resulting in a loss of polarity and cell-cell junctions, enhancing the migratory and invasive properties (165). Downregulation of E-cadherin is accompanied by up-regulation of mesenchymal-type cadherins, such as N-cadherin, which results in weaker and transient cell-cell adhesions (166, 167).

Once activated, the reactive stroma of tumors can maintain the continued expression of an EMT program through paracrine signals. When these signals cease, mesenchymal-like cells can revert back to the epithelial state in a reverse process known as mesenchymal to epithelial transition (MET), which is required for organ formation once embryonic cells have migrated and also for the establishment of distant metastases (36, 168). Alternatively, carcinoma cells can maintain

activated the EMT program via autocrine signals in self-reinforcing positive feedback loops (36). Among these signals, TGF-b is an important EMT inducer. In addition, morphogens such as Wnt, Notch, Shh, cytokines, prostaglandin E2 (PGE2) and growth factors such as EGF, FGF, HGF, PDGF and VEGF are also inducers (167). The effect of a given EMT-inducer is context-dependent; it can inhibit EMT in other processes (167).

3.1.3.1 Regulatory networks defining EMT in cancer

Several regulatory mechanisms exert a tight control to maintain the epithelial homeostasis. EMT is initiated when the balance in these regulatory networks is disturbed; the degree of EMT is further defined by how far the balance is tipped in one or the other direction.

The variety of signals mentioned above activate an array of EMT transcription factors (EMT-TFs) that are potent enough to drive EMT. Among others, members of the Snail family (SNAI1, SNAI2) and ZEB family (ZEB1, ZEB2) repress E-cadherin by directly binding to the CDH1 promoter resulting in loss of E-cadherin and gain of vimentin (165, 169). Simultaneously, they induce repression of the transcription of several other protein junctions, including claudins and demosomes, facilitating dedifferentiation (170) which may not be complete; the different degrees reflecting a variety of partial EMTs. Other transcription factors, such as TWIST (171) or PRRX1 (172), initiate EMT without direct binding to the CDH1 promoter. Many of this transcription factors often function in the same pathway or in synergy with each other but SNAIL, ZEB and TWIST are considered master EMT-TF. Their potent induction of EMT may be explained by its cooperation with different epigenetic modifiers that modulate methylation of the CDH1 promoter (173, 174). Besides the methylation pattern, histone modifications are also involved. For example, SNAI1 recruits histone deacetylase 1 (HDAC1) and HDAC2 to the CDH1 promoter inducing repressive histone modifications(175, 176). Similarly, ZEB1 represses E-cadherin by recruiting a subunit of the SWI-SNF chromatin-remodelling complex, BRG1, which interacts with several histone-modifying enzymes (177). Modification of the CDH1 promoter is part of the program, as there exists a global epigenetic reprogramming during EMT (178).

MicroRNAs (miRNAs) are also essential regulators of the plasticity of the epithelial phenotype.The miR-200 and miR-34 members establish a negative regulatory loop with both Zeb and Snail factors (179, 180), ensuring maintenance of the epithelial and differentiated phenotype. They also do so by regulating the expression of epigenetic modifiers such as BMI1 and SUZ12 (181). Conversely, epigenetic regulators control the expression of miRNAs associated

the epithelial phenotype. Another miRNA downregulated in several carcinomas and considered to have a tumor suppressor function, miR-205 is epigenetically silenced together with miR-200 resulting in an EMT, acquisition of stem-like features and promoting cancer initiation (182). However, the pattern of expression of many miRNAs depends on the context. For example, miR-9 directly targets CDH1 but does not induce EMT (183).

An additional layer of gene regulation is pre-mRNA splicing. Specific splicing factors of either the epithelial or mesenchymal phenotypes determine specific epithelial or mesenchymal splice variants. One such factor are the epithelial splicing regulatory proteins (ESRPs), which maintain epithelial homeostasis by promoting epithelial specific isoforms of EMT-associated genes including CTNND1 (which encodes p120-catenin), FGFR2, CD44 or MENA (184). They are downregulated by the EMT-inducing transcription factors SNAI1, ZEB1 and ZEB2 (185, 186). The absence of ESRPs has a strong impact on cancer invasion. In human mammary epithelial cells, it was observed that knockdown of ESRPs increased expression of the mesenchymal markers vimentin, fibronectin and N-cadherin, whereas no changes of epithelial markers were detected, and cells became more motile.). In its absence the mesenchymal isoforms CD44s and IIIc of FGFRs, associated to EMT and metastasis in carcinomas, are generated (187, 188).

3.1.3.2 Role of EMT in cancer progression

Progression from an early-stage carcinoma in situ to a high-grade invasive carcinoma involves epithelial cells losing their polarity and detaching from the basement membrane (108). Activation of EMT was proposed as an ideal mechanism for epithelial cells to become motile, invasive cells that can break the basal lamina and open paths through the extracellular matrix to reach the bloodstream and intravasate into it facilitating dissemination (Figure 6) (189). It has been observed that the most aggressive breast cancer subtypes display mesenchymal features and human breast cancer cell lines with such features are the most invasive and have metastatic potential *in vivo* (190, 191).

Despite the large experimental evidence in support of its relevance in tumor progression, many clinical pathologists are sceptical about its role in generating high-grade carcinomas given that, with the EMT-associated markers used in clinical tissue sections, it is difficult to distinguish bona fide carcinoma cells that have undergone an EMT from the stromal cells recruited by the tumor (192). Since carcinoma cells often undergo a partial EMT, cells with co-expression of mesenchymal and epithelial markers should be detected. Indeed, rare tumor cells with epithelial

morphology that stained for both epithelial and mesenchymal markers were found in primary tumors from the ER+/PR+, HER2+ and TN breast subtypes. In addition, circulating tumor cells (CTCs) exhibited a mesenchymal phenotype, expressing known EMT regulators. Clinical response to therapy was accompanied by a reduction of CTCs and a switch to the epithelial phenotype while disease progression and recurrence was associated with an increase in mesenchymal CTCs (193).

These sceptical positions regarding EMT also come from the observation at the histopathological level that metastatic samples do not always exhibit the mesenchymal phenotype assigned to metastatic carcinoma cells and indeed, they resemble the primary tumor. To reconcile with this observations, it was proposed that EMT is required for epithelial cells to intravasate, survive in the bloodstream and extravasate in secondary sites but that colonization of distant organs requires a mesenchymal-epithelial transition (MET), which would afford them to recover epithelial identity and regain their original proliferative ability to form tumors at secondary sites (Figure 6) (168). It seems that EMT is a necessary step for intravasation regardless of the type of invasion as breast tumor cells that undergo collective migration are able to invade lymphatic vessels but only individual cells seem to disseminate through the bloodstream in addition to the lymphatic system (194). Attenuation of cell proliferation favours invasion at the expense of tumor growth, and resistance to cell death enhances survival, which confers a selective advantage invasive cells to populate distant organs.

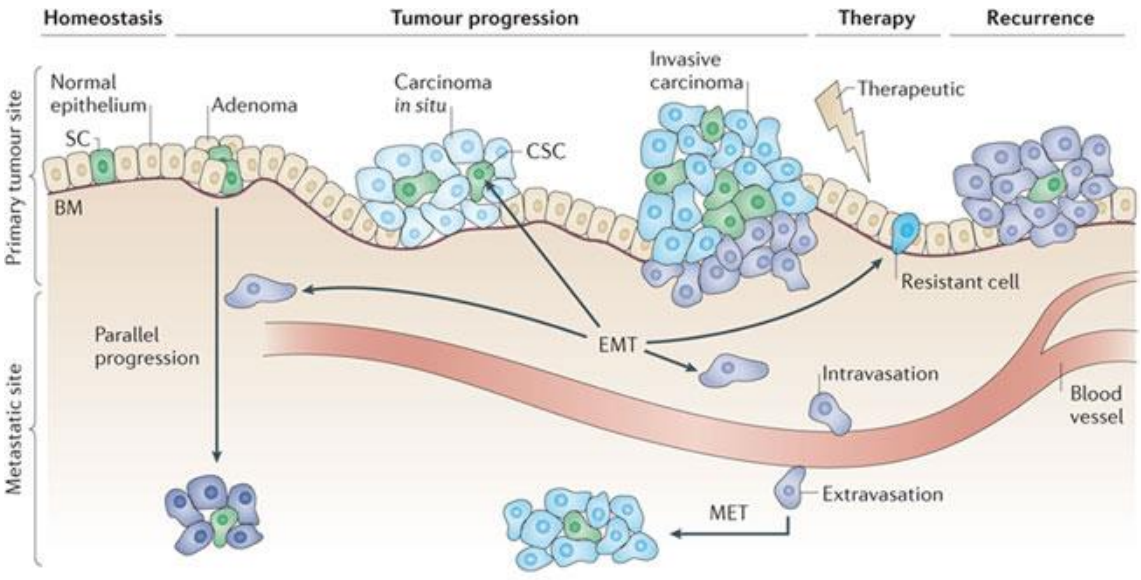

Figure 6. Role of EMT during cancer progression. During tumor invasion, cells that undergo an EMT experience a loss of cell-cell adhesion and polarity accompanied by cytoeskeleton rearrengements that allow them to become more invasive. These cells can break the basal lamina, invade the extracellular matrix and intravasate into the bloodstream. As a result of interactions with the new microenvironment,

they can reacquire the epithelial phenotype through a MET and proliferate in the metastatic site. In parallele, the EMT process may potentially, but not necessarily, confer dedifferentiated tumor cells with stem cell properties, which may further promote resistance during tumour therapy, leading to recurrence and a poor prognosis (195).

In tumour cells, epithelial to mesenchymal transition (EMT)-inducing transcription factors (EMT-TFs) may primarily redefine the epithelial status of the cell, potentially — but not necessarily — assigning stem cell (SC) characteristics to dedifferentiated tumour cells, or they may redefine resident genetically altered stem cells to be cancer stem cells (CSCs). The dissemination of tumour cells from the solid tumour and subsequent migration after breakdown of the basement membrane (BM) — the classical view of the role of EMT in cancer — can only be achieved when all component pathways of the network are activated and fully parallels the process that is seen in development: if the cancer cell has acquired the necessary genetic aberrations and receives the appropriate signals at the tumour–host interface, the cell is ready to move towards metastasis. At this point, the active contribution of the EMT-associated programme is probably to give survival signals and to maintain the mesenchymal status of the metastasizing cell. It is likely that EMT also has a role in parallel progression, in which tumour cells escape early and metastasis progresses in parallel to the primary tumour. EMT features may further promote resistance during tumour therapy, leading to recurrence and a poor prognosis. The degree of EMT during the different steps in cancer progression probably depends on the imbalance of several associated regulatory networks with activated oncogenic pathways. MET, mesenchymal to epithelial transition.

Furthermore, since EMT tends to suppress proliferation, cells that undergo EMT will tend to be more resistant to drugs targeting the proliferative machinery, as it is the case of many chemotherapeutic agents and radiation, leading to recurrence and poor outcome. Twist, Snail and other EMT-inducing factors have been associated to resistance highlighting the importance of combining EMT-inhibition with chemotherapy (196, 197).

The EMT program also confers immunosuppressive capacity. In breast cancer cells, EMT promotes escape from T-cell mediated lysis, in part by inducing autophagy (198, 199).

Finally, there is evidence suggesting that cells that undergo EMT acquire stem-like properties (30, 31, 200, 201) assigning stem cell characteristics to dedifferentiated tumor cells or facilitating the generation of CSC from genetically altered stem cells.

3.2 Metastasis: Dissecting the multi-step process

The end-point of cancer progression is metastasis, which is responsible for a 90% of the mortality associated with cancer. It is a complex multi-step process of biological events, collectively termed as the "invasion-metastasis cascade", whereby epithelial cells in primary tumors 1) locally invade the surrounding tissue, 2) intravasate the microvasculature of the lymph and blood systems, 3) survive circulation in the bloodstream, 4) arrest at distant organs, 5) extravasate into the parenchyma of distant tissues, 6) initially adapts to the foreign microenvironment of these tissues forming micrometastases and finally 7) reactivate proliferation generating macroscopic, clinically detectable tumors, the final step known as "metastatic colonization" (202).

3.2.1 Translocation from the primary tumor to the site of dissemination

Local invasion. In order for individual or small groups of cancer cells to break away from the primary tumor and initiate the metastatic process, these cells must acquire the ability to migrate and invade (Figure 7). Local invasion involves the entrance of cancer cells into the surrounding tumor-associated stroma and thereafter adjacent normal stroma by breaching the basement membrane via secretion of proteases (203). In turn, proteolytic disruption liberates other bioactive molecules that recruit stromal cells to favour invasion, leads to alteration in cell polarity, modulate migration, survival and angiogenesis (203). Dissolution of the basement membrane allows direct contact between cancer and stromal cells, which establish a double positive feedback loop by which carcinoma cells induce a reactive stroma and the latter enhances malignant traits and tumor progression (204).

Intravasation. Locally invasive carcinoma cells are able to intravasate lymphatic or blood vessels. The haematogenous system is the major way of dissemination but lymphatic spread is frequently observed in the clinic and it is an important prognostic indicator of tumor invasiveness and metastatic dissemination in several types of carcinomas (Figure 7) (127). Intravasation into the haematogenous circulation can be enhanced by cells of the stroma, such as perivascular macrophages (205, 206). However, the molecular mechanisms controlling intravasation have not been totally defined.

Survival into circulation. Once malignant cells have intravasated, they can disseminate all over the body through the arterial and venous circulation (Figure 7). These circulating tumor cells (CTCs) must survive several stresses, such as damage caused by the hemodynamic shear forces and immune-mediated killing. CTCs seem to evade both by displaying the tissue factor protein and/or L- and P-selectins, attracting aggregating platelets and travelling as microemboli. CTCs must also overcome anoikis induced by matrix detachment, which could be due to EMT-inducing factors such as the platelet-secreted TGF-β (207, 208). In fact, a recent study reported a detailed phenotypic analysis of CTCs from patients with metastatic breast cancer, revealing that a significant number of CTCs exhibited a partial or a full-blown EMT phenotype, supportive of an EMT-driven mechanism, though also epithelial CTCs were found (193). Contrary to what is expected, even CTCs with a mesenchymal phenotype were found to be in clusters (193). Blood samples from cancer patients can contain both single and clustered tumor cells but it has been reported that tumor cell clusters are more efficient at producing metastases (209). This behavior is explained in a recent paper that shows how clusters can transit through capillaries by reorganizing into single-file chains (210).

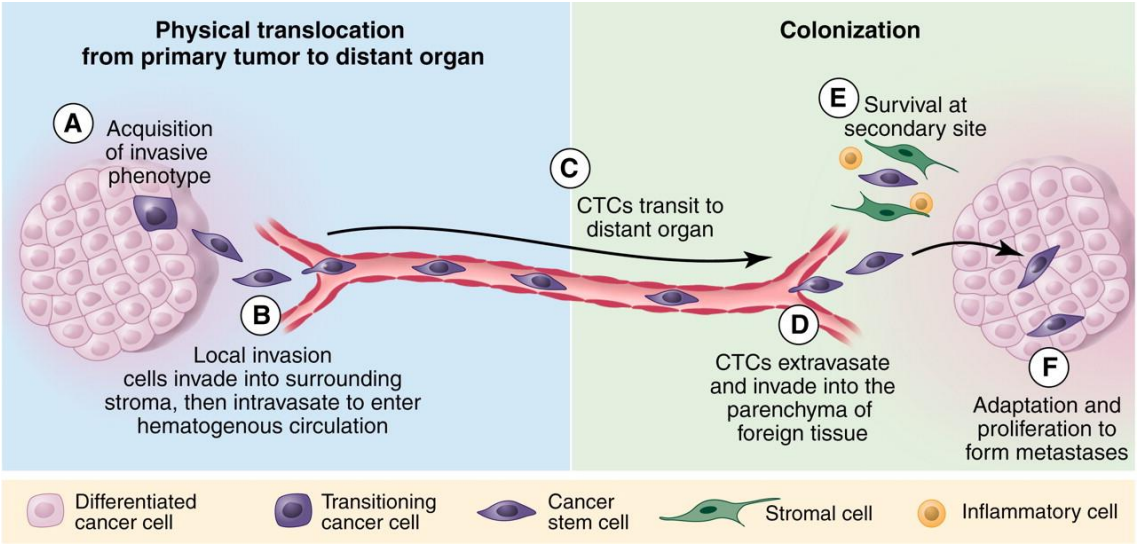

Figure 7. The metastatic process can be divided into two major phases that comprise a first phase translocation from the primary site to distant organs (A-D) that involve the acquisition of an invasive phenotype to penetrate into the surrounding stroma, until being able to intravasate into the bloodstream when arriving to blood vessels and finally extravasate into a distant organ. The second phase involves survival and colonization of the new microenvironment (E-F) (211).

Homing and extravasation. Circulating carcinoma cells have diameters of 20μm to 30μm and larger diameters if they travel as microemboli or clusters and so they can lodge into capillaries before extravasating. The first capillary bed that a CTC encounters is determined by the patterns of blood circulation in the body. However, each type of carcinoma displays a tissue tropism for certain organs, which is not only dependent on the passive process of circulation but also on an actively homing to specific organs (202).

The alternative active process involves expression of specific cell-surface receptors and ligands between cancer cells and cells form the target organ to establish adhesive interactions. For example, CTCs express $\alpha_3\beta_1$ integrin to bind laminin-5, a basement membrane component exposed in regions of the pulmonary vasculature, allowing its arrest in the lungs (212).

Different tissues oppose different barriers, which has a major impact on the fate of disseminated tumor cells (DTCs). Extravasation is not the reverse process of intravasation as the vasculature in the primary tumor is leaky due to neoangiogenesis while microvessels in the distant tissue have structural differences (213). For example, pulmonary microvessels and the blood-brain barrier are mostly impermeable in contrast to the sinusoid capillaries in liver and bone marrow, which are fenestrated. The presence of bone marrow DTCs is predictive of metastasis not only in bone but also in the liver, lungs and brain (214).

3.2.2 Colonization of the Disseminated cell to the Microenvironment at the Metastatic Site.

Metastatic colonization is the limiting step of metastasis. Both experimental and clinical observations demonstrate that it is a very inefficient process as most of the cells die upon infiltrating and only a minority of those that survive develop macrometastases. As an example, the number of CTCs in a patient surpass the number of macroscopic metastasis (Figure 8) (215).

The observation that disseminated cells find significant limitations to grow in distant organs and also that each type of carcinoma has predilection for specific distant organs was made by Stephen Paget almost 130 years ago, who postulated the “seed and soil hypothesis” (202), according to which, DTCs or “seeds” would only colonize organ microenvironments “soils” that were compatible with their growth. Every distant soil is deadly but some can be more hospitable than others. In fact, the most compatible soil is the tumor primary itself, which causes CTCs to reseed the primary tumor or distant organs that have been already colonized, in a process

known as tumor self-seeding (216). Although little is known about the precise mechanisms that leads to the seeding bottleneck, the lack of a supportive stroma (217) and the immune surveillance mechanisms in each organ are likely explanations. Cytotoxic T cells and natural killer cells are a strong antimetastatic defense. The brain has a very reactive stroma where astrocytes secrete the plasminogen activator (PA) to kill infiltrating cancer cells. To counteract this barrier, metastatic cells form breast and lung carcinomas secrete PA inhibitors (218).

Micrometastases formation. With the purpose of creating a more welcoming soil, primary tumors can send systemic signals to activate resident tissue fibroblasts and direct bone marrow-derived cells and macrophages, priming the microenvironment of distant organs for tumor cell accommodation by creating a **pre-metastatic niche** (Figure 8) (204, 219). According to a lung carcinoma model, primary tumors release systemic signals such as inflammatory cytokines, the growth factors VEGF and TGF-β, ECM remodeling enzymes including the lysyl oxidase LOX preconditioning the target tissue for infiltrating cells. Many of these are thought to be delivered in tumor-derived exosomes that specifically unload their cargo in specific tissues (220). These signals induce activation of resident fibroblasts, that upregulate fibronectin. This leads to mobilization of bone marrow-derived hematopoietic progenitor cells (HPCs), which interact with the deposited fibronectin by expressing integrin α4β1. HPCs further modify the microenvironment by secreting MMP-9, with the subsequent release of molecules retained in the ECM such as SDF-1, which chemoattracts CXCR4+ HPCs and metastatic cancer cells (221). It is not known the precise role of the premetastatic niche in cancer patients as metastasis can appear a long period after the removal of the primary tumor by entering in dormancy while in the experimental models, it promotes outgrowth of DTCs.

Occult micrometastases can persist either through the entry of cancer cells in a quiescence state or may alternatively be in a continuous proliferative state that is counterbalanced by a high apoptotic rate, which impairs growth of the micrometastases.

Entering and exiting dormancy. Dormancy is a defense mechanism that allow cancer cells to survive in environments they are not adapted to (Figure 8). Indeed, most DTCs in bone-marrow samples are found as quiescent single cells. To enhance their survival while in a dormant sate, they can establish adhesive and signaling interactions with the tumor microenvironment to induce Akt signaling (222). Activation of programs that suppress self-renewal may also favor exit

from the cell cycle as it has been described that paracrine BMP signaling promotes latency in breast cancer cells that have extravasated in the lungs by inhibiting their self-renewal ability while an antagonistic effect mediated by Coco promotes its reactivation, preventing its differentiation and promoting metastasis (223).

Microenvironmental cues are not only in the form of soluble ligands. The cytoskeletal architecture itself can promote reactivation of dormant cells. Increased matrix stiffness has been associated to proliferation in vitro (224).

In a similar way to what occurs with adult stem cells, it would be like that DTCs may be constantly entering and leaving the latency state and evolving during active periods to eventually awake their potential of initiating a metastasis in a distant organ.

Formation of overt metastasis. Despite the fact that overt metastasis eventually arise from disseminated cancer cells that have managed to survive in a different microenvironment, that does not mean that every DTC has potential as a metastasis-initiating cell (MIC) to reinitiate the growth of the tumor and form clinically significant metastatic colonies in distant organs (Figure 8).

Among the several hallmarks important for reinitiating tumor growth, exiting from dormancy is one of them. Due to technical limitations as it is following a single cell during years until it initiates secondary tumor growth, there is still limited information on the precise mechanism. Some have been mentioned above but different mechanisms are proposed depending on the tropism of metastatic cells (225). As in lung metastasis, production of BMP inhibitors by tumor cells is a mechanism for escaping dormancy, in a bone metastasis model, it is the elevated expression of VCAM-1 that promotes osteoclast activity and the progression to overt metastasis (226).

Most of the DTCs will either die by anoikis/apoptosis due to lack of survival signals or energy resources or will be killed by extrinsic apoptotic signals from the stroma (227). Therefore, resistance to induced cell death is another hallmark. Those mechanisms converging on activation of the AKT pathway will promote survival and inhibition of anoikis. For example, Src mediates AKT-dependent survival in response to CXCL12 and TRAIL, factors expressed in bone metastasis (222). Different organs have different metabolic requirements. In consequence, MICs must have a high metabolic plasticity to obtain energy from multiple sources (228, 229).

As already mentioned, it is believed that EMT allows the escape of cells from the primary tumor but that it is necessary a mesenchymal-to-epithelial transition (MET) for the outgrowth of tumor cells into macroscopic metastases (167). Importantly, EMT can induce stem cell-like properties and promote metastasis (30, 31, 167). It is hypothesized that metastasis-initiating cells may function like stem cells in their ability to initiate and propagate metastatic tumors. A recent study isolated metastatic cells with a stem-like gene expression signature that also displayed increased expression of EMT, pro-survival and dormancy-associated genes; supporting the notion that metastases are initiated by stem-like cells with EMT features (230). Nevertheless, in macrometastases, the EMT features were replaced by markers of proliferation and differentiation in line with the notion that MET is essential for the outgrowth. Furthermore, multiple studies have provided evidence that MET can also induce stem cell-like traits and metastasis colonization (172, 231-234).

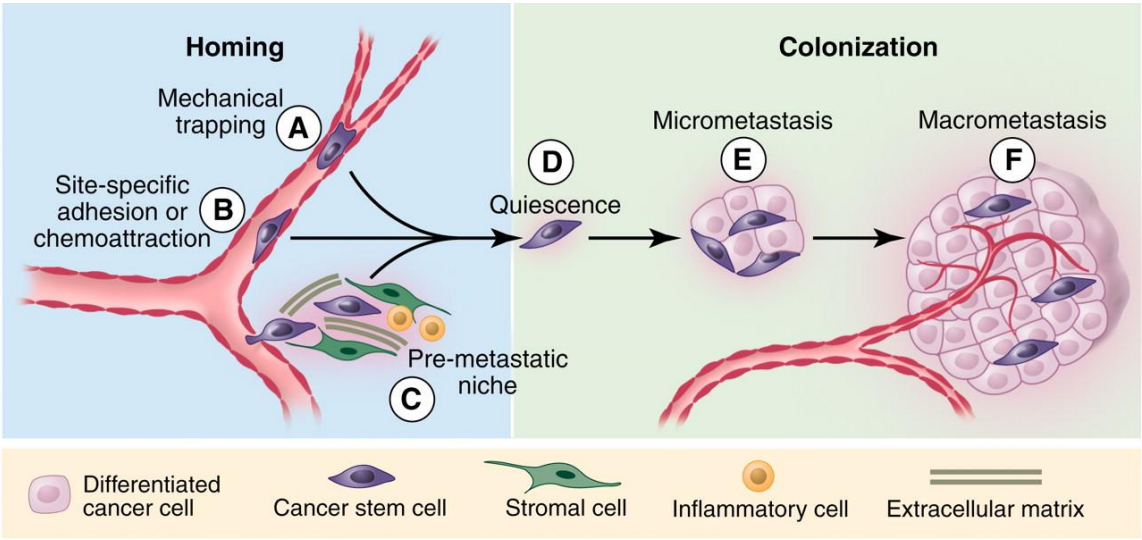

Figure 8. Colonization of a distant organ is also a complex process in which circulating tumor cells must first arrive at the distant organs either by a passive (A) or an active mechanism mediated by specific adhesion molecules that enable them to adhere to microvessels (B) or an active process of homing attracted by signals from a premetastatic niche (C). Once they have extravasated, they may enter in a dormancy state while they adapt to the new environment (D) until they start to proliferate forming micrometastasis (E) that are able to spawn into macrometastasis when they overcome the barriers of the immune system and proliferation exceeds apoptosis (F) (211).

4. MicroRNAs

MicroRNAs (miRNAs) are a class of endogenous, small noncoding RNAs that control target gene expression posttranscriptionally by binding to complementary mRNAs. The fact that each miRNA can target hundreds of mRNAs, explains their involvement in a wide variety of biological processes such as proliferation, cell signalling, differentiation, stress responses and DNA repair, cell adhesion and motility, inflammation, cel survival, senescence or apoptosis (235). Some of them act as master regulators of these functions. In consequence, their aberrant expression is associated with cancer, among other diseases, being involved from the initiation to the progression of cancer to metastatic spread.

4.1 Biogenesis of microRNAs

miRNAs are a class of small RNAs, the one that dominates in mostt somatic tissues. Their biogenesis is tightly regulated in a temporal and spatial manner, and their dysregulation is associated with many human diseases.

In humans, the majority of miRNAs are encoded by introns of coding or non-coding transcripts. Some of the miRNAs genes that are localized in the introns of protein-coding genes share the same promoter of the host gene. However, as they often have multiple transcription start sites, the promoters of intronic miRNAs are sometimes distinct from the promoters of their host genes. Other miRNAs are encoded by exonic regions. Adding further complexity, several miRNA loci are often found clustered constituting polycistronic transcription units that are transcribed independently or co-transcribed but can be separately regulated at the post-transcriptional level (236).

Following the canonical pathway, miRNA genes are transcribed by RNA polymerase II or RNA polymerase III into primary miRNA transcripts (pri-miRNAs) (237). Pri-miRNAs are a 1 to 3kb long primary transcripts organized into a stem-loop structure, in which one or several mature miRNAs are embedded. This pri-miRNAs present in the nucleus are recognized and processed by the Microprocessor complex, formed by the RNase III enzyme Drosha and its essential cofactor DGCR8 (DiGeorge critical region 8) (238, 239). Drosha cleaves the 5’ and 3’ arms of the pri-miRNA whereas DGCR8’s function is to directly interact and stabilize the pri-miRNA to determine the precise cleavage site, which is 11 bp away from the single-stranded RNA/double-stranded RNA

junction at the base of the hairpin stem. Cleavage of the pri-miRNA is thought to occur co-transcriptionally and before splicing of the protein-encoding or non-coding host pre-mRNA. Splicing is not inhibited by the cleavage, thereby ensuring both miRNA biogenesis and protein synthesis from a single primary transcript (240, 241). However, there are examples in which Drosha directly controls mRNA stability by cleaving miRNA-like hairpins on the exonic regions, such as on the DGCR8 mRNA (242, 243).

Drosha-mediated processing generates a 60-100 nts hairpin intermediate with a stem loop structure called the pre-miRNA. This precursor is then transported from the nucleus to the cytoplasm, where maturation can be completed. Translocation is mediated by the protein Exportin-5 (XPO5) in complex with Ran-GTP that also protects pre-miRNAs against nuclear digestion. A defined length of the double-stranded stem and the 3´overhangs are important for successful binding to Exportin-5, ensuring the export of only correctly processed pre-miRNAs (244, 245).

In the cytoplasm, the RNase Dicer in complex with the double-stranded RNA-binding protein TRBP cleaves the pre-miRNA hairpin liberating a 19 to 24 nt long miRNA duplex that is subsequently loaded onto an Argonaute (AGO) protein to form the effector complex called RNA-induced silencing complex (RISC). RISC assembly is carried out by the RISC-loading complex (RLC). RLC is a protein complex containing Dicer, a dsRNA-binding protein (TRBP or PACT), an AGO protein and potentially other proteins that are required for loading of small RNAs onto the AGO proteins (246, 247). In humans, by contrast with other species, all four AGO proteins (AGO1-4) can incorporate both miRNA duplexes, without selective loading of sets of miRNAs, and siRNAs with a preference for duplexes with central mismatches (positions 8-11) (248). Following the loading of the RNA duplex, the passenger strand is discarded and the guide strand (mature miRNA) remains in one of the AGO proteins to generate a mature RISC that mediates RNA-silencing. Unwinding of the miRNA duplexes is a most general process as most duplexes have central mismatches that prevent slicing but if the duplex is matched at the centre, AGO2 (the only one with slicing activity) can cleave the passenger strand. The miRNA strand with the less stable base pair at its 5’ end in the duplex and a U at nucleotide position 1 is selected as the guide or functional strand. The passenger strand is released and degraded. However, as strand selection in not completely strict, the strand that is not favoured can be selected. This passenger strand, termed miRNA*, is also active in silencing although less abundant and potently. The alternative strand selection or arm-switching may be explained by an alternative processing of Drosha, which changes the thermodynamic stability of the duplex ends.

Apart from the canonical miRNA biogenesis pathway, there are alternative pathways from which miRNAs can be generated. One of the first mechanisms described is Drosha and DGCR8 independent: miRNAs derive from introns that are released from the host transcript by splicing. After debranching, if the resulting intron has the appropriate size and mimics the structure of pre-miRNAs, it bypasses Drosha-mediated cleavage and is further processed in the cytoplasm by Dicer. These miRNAs are called mirtrons (249, 250). Some small nucleolar RNAs (snoRNAs), tRNAs and endogenous short hairpin RNAs (shRNAs) may also be cleaved to produce pre-miRNAs with independence of Drosha (251, 252). The miR-451 follows a different non-canonical pathway that is Dicer-independent. Its pre-miRNA is produced by Drosha, exported to the cytoplasm and loaded into AGO2. AGO2 cleaves the stem of the pre-miRNA generating AGO-cleaved-pre-mir-451 (ac-pre-mir-451), which is further trimmed to produce the mature miRNA (253, 254). The third pathway is Terminal uridylyl transferase (TUTase)-dependent. Drosha initiates microRNA maturation by cleaving a primary miRNA transcript and releasing a pre-miRNA with a 2 nt 3' overhang and Dicer recognizes the 2 nt 3' overhang structure to selectively process pre-miRNAs. Unlike prototypical or group I pre-miRNAs, some pre-miRNAs (group II) have a shorter 3' overhang that is not optimal for Dicer processing. Thus, they are extended by monouridylation mediated by terminal uridylyl transferases and subsequently processed by Dicer. The majority of let-7 and miR-105 follow this alternative pathway (255).

4.2 Mechanisms of action of microRNAs

4.2.1 Characteristics of miRNA target sites

Over half of the human transcriptome is predicted to be under miRNA regulation, as a single miRNA can regulate multiple mRNAs (256). The reason is that, despite its small size, the majority of miRNAs in animals bind to their target sites with partial complementarity, making the prediction of target sites a challenging issue. In general, binding is done through perfect base pairing between the seed region (nucleotides 2 to 8) of the miRNA and the 3'UTR of the target mRNA while the rest of the duplex contains mismatches and multiple nucleotide bulges. Sometimes, the absence of perfect seed pairing is compensated by the extensive pairing of the 3'end of the miRNA to the target sequence. Pairing of the miRNA middle sequences with their targets and regulation of mRNAs through coding regions instead of 3'UTRs, has also been

described (256, 257). It seems that, beyond pairing, there are additional factors that determine which genes are under the regulation of miRNAs.

4.2.2 Mechanisms of silencing by microRNAs

miRNAs can regulate target gene expression in several ways. A widespread mechanism is mRNA destabilization. Several studies demonstrate that miRNAs can act by first inhibiting translation in a deadenylation-independent manner, then subsequently causing deadenylation, decapping and destabilization of targeted mRNAs (258). However, the current evidence suggests that target mRNA degradation is the predominant mechanism of silencing by miRNAs, rather than mRNA destabilization (259). As a result of partial pairing between the miRNA complex and the 3'UTR target site, miRISC-associated GW182 proteins recruit the CCR4-NOT complex to the mRNA resulting in deadenylation and degradation (260, 261). microRNAs, in the miRISC complex, can also induce translational repression by blocking initiation (262) or inhibiting a step after translational initiation, such as promoting ribosome drop-off or stimulating proteolysis of the peptide (263). Perfect complementarity of a miRNA with its target site is a common mechanism in plants but is very rare in animals and it results in endonucleolytic cleavage of the mRNA by Argonaute and degradation.

4.3 microRNAs in Cancer

4.3.1 Key microRNAs involved in cancer

microRNA expression profiles comparing normal and tumor tissues have shown that most miRNAs are down-regulated in cancer (264). Given that this widespread underexpression is observed mainly in poorly differentiated tumors, some have proposed that many miRNAs may be involved in defining lineage-specific properties of differentiated cells. For example, there is a significant increase in miRNA levels upon inducing differentiation of the HL60 cancer cell line (264). An alternative explanation is that reduced levels of miRNAs may confer a selective advantage during tumorogenesis. In this regard, RNA-interference- based silencing of Dicer, DGCR8 and Drosha and consequent knockdown of mature miRNAs, promoted tumor formation and progression but only in already transformed cell lines (265). Given that Dicer haploinsufficiency is observed only in 27% of tumors, widespread downregulation cannot be explained only by defects in miRNA biogenesis (266).

However, the initial study by Lu and collaborators profiling multiple tumors, where half of the microRNAs were found down-regulated, only measured the expression levels of the 217 microRNAs known at the moment with flow cytometry (264). The first large microarray study on solid tumors showed a more complex role of microRNAs in cancer by identifying a specific miRNA expression signature of 36 overexpressed and 21 down-regulated miRNAs compared to normal tissue (267). miR-17-5p, miR-20a, miR-21, miR-92, miR-106a, and miR-155 were found to be strongly associated with tumorogenesis (267). This "miRNome" together with other genome-wide profiling studies, supports the notion that miRNA alterations in cancer consist of both downregulated and overexpressed miRNAs with putative tumor suppressive and oncogenic functions (268).

microRNAs with and oncogenic role, also termed "oncomiRs", are those that negatively regulate tumor-suppressors and its overexpression, amplification of loss of epigenetic silencing leads to tumor development. By contrast, deletion or epigenetic silencing of microRNAs that target oncogenes, considered tumor-suppressor miRNAs, will also have an oncogenic effect (269). Despite being classified often between these two categories, miRNAs can act as oncogenes or tumour suppressors depending on the cellular context in which they are expressed explained by the fact that they can have tens to hundreds of target mRNAs with opposing roles (270). A few classical well-characterized microRNAs involved in cancer are briefly described below.

miR-17-92. Generally classified as an oncogenic cluster, this cluster consists of six miRNAs (miR-17, miR-18a, miR-19a, miR-20a, miR-19-b-1, and miR-92a-1), some of which act as tumor suppressors (271). Several types of lymphoproliferative disorders (271) and solid tumors show amplification of the locus where it maps (272). It is also upregulated by direct binding and activation by MYC and E2Fs, the latter promoting proliferation (273). Activation by MYC induces suppression of chromatin regulators, the proapoptotic gene BIM or PTEN to maintain survival, autonomous proliferation and self-renewal (274).

miR-155. It has been mainly reported as an oncomiR in hematopoietic malignancies but also in breast, lung and pancreas (275). It targets significant tumor suppressor genes such as SOCSI in breast carcinomas (276). Targeting of several tumor suppressors (TP53INP1, CDC73, VHL, SEL1L) increases cell proliferation through the activation of the JAK2/STAT3, AKT, NF-kB, Wnt/βcatenin signaling pathway and activation of cell cycle, and sensitizes cancer cells to apoptosis induced by chemotherapeutic agents (277).

miR-21. In the microarray study by Volinia and colleagues, miR-21 was found to be the only miRNA overexpressed in six different solid tumors: breast, lung, stomach, pancreas, prostate and colon cancers (267). Its status as oncomiR was established after finding it overexpressed in most cancer types analysed (278-280). In breast cancer, it was progressively upregulated with increasing tumor stage (278). In colon cancer, increased levels positively correlated with cancer stage, lymph node presence and distant metastases and with poor response to 5-FU therapy and recurrence in end-stage patients (279). It was also upregulated in the most aggressive subtype of brain cancer, glioblastoma multiforme, where inhibition with anti-miR-21 would be an interesting approach (281). miR-21 is a negative regulator of PTEN, which leads to increased FAK, Akt and MMP expression, leading to increased cell migration and metastasis (282). Programmed cell death 4 (PDCD4), a tumor and metastasis suppressor downregulated in multiple cancer forms and upregulated during apoptosis, was identified as a miR-21 target in various cancers (283). Several other tumor suppressors are targeted by miR-21 (284).

let-7 family. The let-7 family of tumor suppressor miRNAs in humans has 12 homologs, encoded at multiple chromosomal sites, some of which are deleted in various human tumors. Increased let-7 can block tumour formation, progression and metastasis and can induce cell apoptosis through targeting downstream oncogenes, such as RAS, HMGA2 or LIN28 (285). In turn, members of the Lin-28 family, Lin-28A and Lin-28B, selectively inhibit the expression of let-7 through different mechanisms and increased abundance of Lin-28 in tumors is associated with poor prognosis in patients (286). Let-7 functions in differentiation and fate determination. It promotes differentiation of stem cells and can also regulate self-renewal, differentiation and polarity of CSCs (287).

miR -34 family. This family of miRNAs (miR-34a/b/c and miR449a/b/c) acts as tumor suppressors. They are induced p53 in response to DNA damage or oncogenic stress (288). It can repress in turn various inhibitors of p53 creating a feedback positive loop (289). miR-34 members regulates hundreds of genes, a majority involved in the cell signalling network and proliferation in response to growth factors. Therefore, it mitigates the proliferative and pro-survival effect by interfering with downstream pathways activated upon growth factor stimulation (290). It also inhibits EMT by targeting two EMT TFs, ZNF281 and SNAIL, the latter forming a double-negative feedback loop with miR-34 (291, 292).

4.3.2 microRNAs in Breast Cancer initiation and progression

4.3.2.1 Breast cancer signatures of microRNAs

In 2005, Ioro et al identified a set of miRNAs aberrantly expressed in breast cancer compared to normal tissue. The overall miRNA expression could clearly separate both groups, with the most significantly down-regulated miRNAs being miR-125b and miR-145 whereas miR-21 and miR-155 were up-regulated (293). In addition, they correlated miRNA expression with breast cancer biopathological features. miR-26, miR-30, miR-185, miR-191, miR-206, and miR-212 expression were found to be associated with ER status while that of let-7c, miR-26a, miR-29b, and miR-30 correlated with PR status. miR-206 was highly expressed in ER- cancers as it targets the ERα receptor, in addition to miR-221 and miR-222 which are also negative regulators (294, 295). Among others, miR-21, miR-181a and miR-30a–s correlated with tumor stage. Let-7 and miR-9-3 expression associated with lymph node metastasis while miR-9-3, miR-10b, miR-27a, miR29a, miR-123 and miR-205 with vascular invasion and let-7, miR-26, miR-30a-5p, miR-102, and miR-145 expression differed in tumors with high vs low proliferation index (293).

The first integrated analysis of miRNA expression, mRNA expression and genomic changes in human breast cancer was performed by Blenkiron and colleagues, which analyzed 93 primary human breast tumors, using a bead-based flow cytometric miRNA expression profiling method (296). This integration allowed classification of tumors into the intrinsic subtypes luminal A, luminal B, basal-like, HER2+ and normal-like based on the expression of a set of 31 miRNAs. Those miRNAs up-regulated in the basal-like and HER2+ subtypes positively associated with an ER- status and high-grade tumors and conversely (296).

A more recent study also integrating mRNA and miRNA expression profiling identified microRNAs implicated in the progression of patients with a 10-year follow-up (297). Four microRNAs (miR-135a, miR-128a, miR-767-3p, and miR-769-3p) defined a low-risk signature in ER+ tumors treated with tamoxifen that correlated with prolonged distant relapse-free survival (DRFS). In ER- tumors, a different low-risk signature appeared with expression of miR-30c, miR-150, miR-27b and miR-210, among others (297). miR-30c, which is regulated by GATA3, is an individual prognostic marker for breast tumor DRFS (298).

Except for a few microRNAs that will be described below, it is not known yet the role the majority of this microRNAs plays in the progression of breast cancer.

4.3.2.2 microRNAs, Breast Cancer Stem Cells and EMT

MicroRNAs are essential regulators of different aspects of tumor formation, including tumor initiation, growth, differentiation, and progression. Therefore, they are also potent regulators of the EMT process and CSCs.

In 2008, four different studies described the role of the miR-200s family in regulating the epithelial and mesenchymal phenotype of cancer cells (299-302) and since then, its role as a master regulator has been widely recognized. They were found to be downregulated in human BCSCs as well as in normal human and murine mammary stem/progenitor cells. Through BMI1 targeting, miR-200c inhibited the clonogenicity of BCSCs *in vitro* and tumorogenicity of BCSCs *in vivo* (303). Inhibition of miR-200b was also required for the formation of CSCs, mediated by expression of Suz12, which represses E-cadherin and is required to sustain CSC growth (303). Additionally, they were found consistently down-regulated in $CD44^{+}CD24^{-/low}$ putative BCSCs isolated from metastatic samples of breast cancer patients (304).

In synergy with the miR-200 family, miR-205 is also a negative regulator of EMT by targeting ZEB1 and ZEB2, thus reinforcing the epithelial phenotype (300). It has been found to have a low expression in metastatic breast cancer cell lines (305) and clinical samples (278). Furthermore, the expression of miR-205 was decreased in a metastatic miRNA signature in human breast cancers (306). It was shown to suppress invasion and metastasis of MDA-MB-231 cells in vivo by directly regulating ErbB3 and VEGF-A expression (307). It is enriched in normal mammary epithelial stem cells and its overexpression results in increased growth and expansion of the progenitor cell population, by targeting PTEN (308). miR-205 is an important determinant of mammary stem cell fate and its misregulation leads to tumorogenesis. A recent study reported that secretion of jagged1 by the tumor stroma repressed miR-205 expression inducing a CSC phenotype (309). Low expression of miR-205 resulted in disruption of epithelial polarity and aberrant mammary stem cell division, which in turn leads to an expansion of stem cell population and tumorigenesis (309).

The majority of studies exploring the role of miR-21 EMT/MET have been done in breast cancer. The use of an antagomir against miR-21 reversed EMT and the CSC phenotype in the metastatic MDA-MB-231 cells by down-regulation of its target PTEN (310). A previous study had demonstrated its role in promoting invasion and metastasis by inhibiting other tumor suppressors: TPM1, PDCD4 and maspin (282). On the contrary, re-expression of miR-21 in the

non-aggressive MCF7 cells promoted the acquisition of an EMT phenotype together with an increase in cell growth, migration and invasion, and also self-renewal and clonogenicity characteristics of CSCs (311).

Similarly to colon and lung cancers, a p53 loss-of-function or mutation promotes EMT in breast carcinoma cells by decreasing miR-34 levels, which are negative regulators of Snail1 (180). It has been recently shown that upon TGF-β-induced EMT, there is a first activation of the miR34-Snail1 negative feedback loop followed by activation of negative feedback loop between miR-200 and ZEB1 (312).

MiR-9 induces EMT in human epithelial cells and breast carcinoma cell lines by directly targeting E-cadherin (183). Downregulation of E-cadherin leads to release of β-catenin from the junctions and translocation into the nucleus, where it activates transcription of VEGF, increasing tumor angiogenesis. Overexpression of miR-9 leads to lose cell-cell contact, increased motility and migration in vitro and increased pulmonary metastases *in vivo*. Expression of miR-9 is activated by MYC and MYCN. In human cancers, miR-9 levels correlate with MYCN amplification, tumor grade and metastatic status (183). In fact, miR-9 has been found to be highly expressed in the more aggressive HER2+ and triple-negative subtypes and also in tumors showing a CSC phenotype, vimentin expression and loss of E-cadherin (313).

Let-7 was the first microRNA family to be identified in breast cancer stem cells. It was the most consistently down-regulated miRNA in tumor-initiating cells by serial passaging *in vivo* whereas its expression increased upon differentiation to non-tumorigenic cancer cells. Re-expression of let-7 reduced the stem cell-like properties of BCSCs, inhibited both tumor formation and its metastatic potential (287). Consequently, it is considered a potential biomarker of BCSCs and an ideal therapeutical candidate in anti-cancer therapy (314). Besides it role in BCSCs, let-7 but also miR-200's levels are downregulated in response to inflammation-induced EMT and this downregulation allows initiation and maintenance of the EMT phenotype (315).

Similar to the let-7 family, Yu and colleagues also demonstrated that miR-30e was downregulated in tumor initiating BCSCs from mammospheres and breast cancer patients (316). miR-30e expression impaired self-renewal of BCSCs through Ubc-9 and ITGB3 silencing and reduced tumorogenesis and metastatic growth (316). Consistent downregulation of the miR-30 family, especially miR-30, was also found in breast cancer cells grown in non-adherent conditions. They control expression of genes involved in apoptosis and proliferation BCSC (317).

miR-128 levels were also found to be reduced in BCSCs in primary breast tumors and in mammospheres. This reduction increased the protein levels of Bmi-1 and ATP-binding cassette sub-family C member 5 (ABCC5) triggering a pro-apoptotic and DNA-damaging effect when treated with a chemotherapeutical agent. Given the known intrinsic resistance of stem cell-like cells to chemotherapy and radiotherapy, miR-128 may have strong therapeutic potential. Furthermore, reduced expression of miR-128 in breast tumor tissues was associated with chemotherapeutic resistance and poor survival rates of breast cancer patients (318). Another study reported miR-128 to be the most frequently deleted miRNA in a series of primary breast tumors. Depletion of miR-128 induced stem cell-like features in immortalized non-tumorigenic mammary epithelial cells and was sufficient for oncogenic transformation. They also found that miR-128 was directly inhibited by SNAIL and repressed by TGF-β signaling, resulting in derepression of several direct targets of miR-128 involved in stem cell signaling (BMI1, CSF1, KLF4, LIN28A, NANOG and SNAIL) (319).

In addition to the above mentioned microRNAs, miR-34c levels are also reduced in BCSCs from two different breast cancer cell lines. Apparently, silencing of miR-34c was caused by hypermethylation of a single CpG site locate in a crucial regulatory element of the promoter. Silencing promoted self-renewal and EMT of breast TICs (320).

miR-7 was significantly downregulated in CSCs with high metastatic capacity to bone and brain and restoration of its expression abrogated this capacity by suppressing KLF4. KFL4 is one of the essential genes in reprogramming induced pluripotent stem cells that also prevents differentiation of embryonic stem cells. Moreover, an inverse correlation between miR-7 and KLF4 expression was found in brain metastatic samples (321).

On the contrary, the miR-181 family members are upregulated in tumor-initiating mammospheres compared to non-tumorigenic parental cells. TGF-β increases formation of mammospheres by upregulating miR-181 levels with the consequent downregulation of the tumor suppressor BRCA1, a direct target of miR-181. An inverse correlation between TGF-β and miR-181 with BRCA1 expression was observed in breast tumor samples (322).

Similarly, upregulated expression of miR-495 was found in two BCSCs subpopulations (PROCR+/ESA+ and CD44+/CD24-). The transcription factors E12 and E47 are responsible for miR-495 upregulation, which in turn contributes to down-regulation of E-cadherin and REDD1, finally resulting in maintaining a stem cell-like phenotype in breast cancer (323).

4.3.3 MetastamiRs

The termed metastamiR has been used to describe microRNAs that play roles in various steps of metastasis, some without necessarily needing to be involved in tumorogenesis. The majority have been identified in breast tumor cell lines.

4.3.3.1 Metastasis-suppressing microRNAs

A miRNA profiling on highly metastatic MDA-MB-231 derivatives to lung and bone revealed that expression of miR-335, miR-206 and miR-126 was specifically lost as cells developed metastatic potential (324). miR-126 suppressed overall tumor formation whereas miR-206 and miR-335 prevented cell invasion and the three of them suppressed metastasis *in vivo* (324). Low expression of these miRNAs correlated with a shorter time of metastatic relapse. In addition, low levels of miR-335 and miR-126 correlated with poor overall metastasis-free survival (324). Several studies in larger clinical cohorts of other cancer types have since validated their prognostic capacity (325-329). Transfection of cells with an antagomir targeting miR-335 showed a set of pro-metastatic genes regulated by miR-335 involved in control of extracellular matrix, cytoskeleton, signal transduction and cell migration. Suppression of invasion and lung colonization by miR-335 was mediated by specifically targeting SOX4 and Tenascin-C (324).

miR-126 is an example of microRNA that regulates the metastatic microenvironment. It inhibits the recruitment of endothelial cells by metastatic breast cancer cells (330). miR-126 is silenced in cancer cells, resulting in up-regulation of its target genes IGFBP2, MERTK and PITPNC1, which are pro-angiogenic. In consequence, IGFBP2 is secreted by cancer cells and promotes endothelial cell migration by activating the endothelial IGF1 receptor. PITPNC1 contributes by enhancing extracellular levels of IGFBP2 and the c-Mer tyrosine kinase receptor (MERKT) is proteolitically cleaved and secreted by cancer cells to sequester the endothelial cell migration inhibitor GAS6. The outcome of miR-126 silencing is metastatic endothelial recruitment, metastatic angiogenesis and metastatic colonization (330). Its tumor-suppressive function has been reported in other cancer types (331, 332).

Both miR-146a and miR-146b inhibit invasion and migration of metastatic breast cancer cells by negatively regulating IRAK1 and TRAF6, which positively regulate NF-kB activity (333). Breast

cancer metastasis suppressor 1 (BRMS1) exerts its anti-metastatic function by up-regulating miR-146 (334).

miR-31 is a pleiotropically acting metastasis-suppressor involved in multiple steps of the metastatic process. It inhibits breast cancer cell invasion and anoikis resistance through coordinate repression of integrin α5 (ITGA5), radixin (RDX) and RhoA (335-337). Besides inhibition of local invasion, it also prevents survival in the circulation through targeting these enhancers of anoikis resistance, intravasation and micrometastases formation. Finally, it also inhibits metastatic colonization through impairment of signal transduction pathways of relevance to cell survival and proliferation that are controlled by ITGA5 and RDX (338).

The miR-200 family has been considered a metastasis suppressor as it inhibits the early step of invasion. They do so by targeting proteins that regulate the reorganization of the actin cytoskeleton such as WAVE3, which promotes metastasis (339). In breast cancer, FHOD1 and PPM1F, direct regulators of the actin cytoskeleton that are associated with a mesenchymal phenotype, were validated as targets of miR-200c (340). Nevertheless, the classical suppressor role of miR-200s has been attributed to the fact that they maintain the epithelial phenotype by preventing EMT. They achieve so by directly targeting the master regulators of EMT, ZEB1 and ZEB2, which negatively regulate E-cadherin (300, 302, 341). The miR-200-Zeb1-E-cadherin axis was deregulated in metastases and a subset of primary tumors that presented a strong downregulation of the miR-200 family and ectopic expression of miR-200 blocked EMT and metastasis (342). Besides, part of the anti-metastatic activity of miR-200f can be explained through regulation of the Notch ligands Jagged1 and Jagged2 (343, 344). More recent studies have reported that miR-200 can inhibit cell migration in a ZEB1/2- independent manner (345) and by regulating multiple targets that act coordinately on Rho-ROCK signaling, invadopodia formation, MMP activity and focal adhesions to maintain cell morphology and prevent cell migration (346).

4.3.3.2 Metastasis-promoting microRNAs

One of the first microRNA drivers of metastasis identified was miR-10b (347). Ectopic expression of miR-10b increased invasiveness and was sufficient to promote metastasis formation *in vivo* while its action could be blocked by administration of an antagomir anti-miR-10b (348). Transcription of miR-10b is induced by Twist 1. Induction of miR-10b down-regulates the levels of HOXD10, which is a negative regulator of the expression of genes that are involved in cell

migration and extracellular matrix remodelling, such as RHOC (Figure 9a). As a consequence, it stimulates breast cancer cell migration and invasion (347). High miR-10b levels have been observed in the invasive front of xenografts but has no effect on the growth of the primary tumor. Additionally, a positive correlation between the levels of miR-10b and the presence of metastasis was found in the clinic (347).

At least in breast cancer, miR-373 and miR-520c confer metastatic potential to non-metastatic cancer cells. They were found to be highly expressed in cells with migratory capacity and their ability to induce migration and invasion was mediated by inhibition of CD44 (Figure 9b). miR-373 levels were upregulated and correlated inversely with CD44 expression in breast cancer metastasis samples (349).

Besides its role as an EMT inducer, miR-9 promotes migration, invasion and metastasis by targeting a breast cancer metastasis suppressor, the leukemia inhibitory factor receptor (LIFR) (Figure 9d). Downregulation of LIFR results in activation of the oncoprotein YAP. Loss of LIFR is associated with development of distant metastasis and lymph node metastasis in patients with invasive breast cancer (350).

The expression of the miR-103/107 has been also associated with metastasis and poor outcome in breast cancer patients. It was identified that miR-103/107 directly target and down-regulate Dicer to promote migration and invasion *in vitro* and fosters metastasis of non-metastatic cells *in vivo* (351). Besides, miR-103/107 induces EMT and is required for maintaining a mesenchymal gene expression through down-regulation of the miR-200 family as a consequence of the suppression of Dicer (351).

miR-21 is a potent oncomiR involved in multiple cell-signalling pathways that fosters invasion and metastasis through targeting PDCD4 and tropomyosin 1 (TPM1) (Figure 9e) (282, 352). It also drives invasion and metastasis in colorectal cancer cells by targeting PDCD4 and in other cancers (353-355).

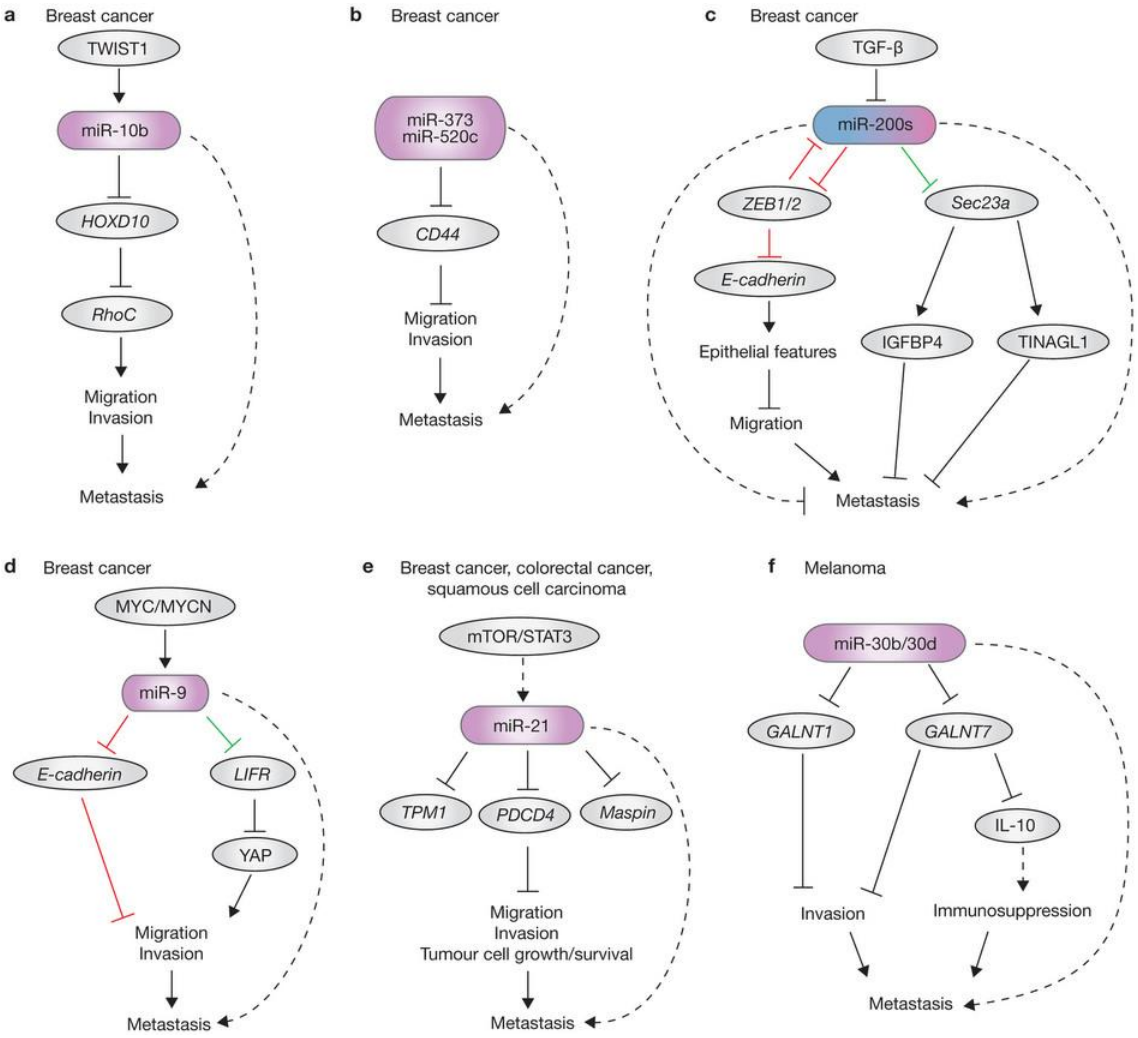

Figure 9. Regulatory networks in which microRNAs that promote metastasis are implicated.

While silencing of miR-200f facilitates early steps of metastasis, they might promote metastatic colonization by inducing a MET at distant organs. This was firstly observed by Dykxhoorn et al. (356). Subsequently, Korpal et al. stated that miR-200s promotes metastatic colonization in part by downregulating Sec23a, which mediates secretion of Tinagl1 and Igfbp4, two metastasis suppressors (Figure 9c) (341). More recently, it has been seen that the metastatic capacity can be transferred from metastatic to non-metastatic cancer cells via extracellular vesicles containing miR-200 microRNAs (357).

Other pro-metastatic microRNAs have been identified besides those in breast cancer. As an example, miR-30b/d enhance micrometastases formation by targeting the inhibitors of invasion GALNT1 and GALNT7 in melanoma cells (358) and miR-199a-3p, miR-199a-5p and miR-1908 are also drivers of melanoma metastasis (Figure 9e) (359).

Finally, microRNAs control metastasis by targeting multiple genes in a coordinate manner. The majority of the microRNAs described above exert a divergent mode of regulation, which is that a single microRNA could coordinately regulate the expression of multiple genes, all of which participate in a common metastatic phenotype with the advantage that the effect is more potent. However, others such as miR-199miR-199a-3p, miR-199a-5p and miR-1908, act in a convergent way to silence a single gene. Acting in cooperation ensures a stronger regulation on that gene (360).

Objectives

OBJECTIVE 1

- The identification of sets of microRNAs deregulated along metastatic progression of invasive ductal carcinomas in tissue and in the blood of metastatic breast cancer patients.

OBJECTIVE 2

- To explore the role of the miR200s family in metastatic progression in breast carcinomas by using a breast cellular model.

Material and Methods

miRNA microarray for FFPE breast samples

The cohort to study microRNA expression in breast cancer tissue consisted of 20 patients from each of the following samples: primary invasive ductal carcinoma (IDC) without lymph node involvement, for at least 5 years (PNM); lymph-node positive IDC (PM) and its paired lymph node metastasis (LN) and non-matching distant metastasis (DM). In addition, a pool of 10 normal breast samples from mammoplasties were studied (N). An immunohistochemical and histopathological examination was done to select a relatively homogeneous serie of medium/high-grade, estreogen receptor (ER) positive, progesterone receptor (PR) positive, HER2 negative breast ductal carcinomas from patients that had not received chemoradiotherapy before surgery. Formalin-fixed paraffin-embedded (FFPE) samples with >70% tumor epithelial enrichment were macrodissected to minimize stromal and lymphocytic presence. RNA extraction was performed with the miRNeasy FFPE kit (Qiagen). One hundred nanograms of total RNA was labeled using the Agilent miRNA Complete Labeling and Hybridization Kit (Agilent Technologies Incorporated, Santa Clara CA, USA) according to the manufacturer's instructions. The labeled RNA was hybridized to the Agilent Human miRNA Microarray (V12.0, Agilent) which contains probes for 866 mature human miRNAs and 89 human viral microRNAs. Arrays were scanned using an Agilent scanner and feature extracted using Agilent Feature Extraction Software, version 10.5.1.1. Quantile normalization applied to agilent microarrays was performed using the "normalize.quantiles" function from R package "affy" from the Bioconductor project (http://www.bioconductor.org). We use the method proposed by Pradervand to select invariant and highly expressed micrornas (361).

Real-time quantitative PCR (qPCR) of FFPE samples

Total RNA, including the microRNA fraction, was isolated from FFPE samples using the miRNeasy FFPE kit (Qiagen). RNA was retrotranscribed with the Universal cDNA Synthesis kit (Exiqon). Mature microRNAs were detected using the ExiLENT SYBR Green Master Mix (Exiqon) and the specific LNA^{TM} primers for the following microRNAs: hsa- miR-10b, hsa-miR-101, has-miR-148a, hsa-miR-181a, hsa-miR-181b, hsa-miR-200a, hsa-miR-200b, hsa-miR-429, hsa-miR210, hsa-miR-30a, hsa-miR-7. hsa-miR-16 and hsa-let7a were used as reference microRNAs. Real-time quantitative PCR assays were performed on a StepOnePlus Real-time PCR instrument (Life Technologies) and analyzed with the StepOne software (v2.3). Relative quantifications were

assessed by the ΔΔCp method. Normalized values were used in comparative analysis between categories of samples using either parametric (t-test) or non-parametric tests (Mann-Whitney)..

Real-time quantitative PCR (qPCR) of microRNA blood levels

The cohort used in this study consists of initially diagnosed breast cancer patients, lymph-node positive (N=26) and negative (N=34) that had not received any treatment prior to sample obtention. Additionally, distant metastatic patients (N=18) were included. As a control, samples from women with negative mammographies were also obtained (N=6). Collection of samples was done in collaboration with the Gynecology Department of Hospital Clínic and Hospital Clínic-IDIBAPS Biobank, a Generalitat de Catalunya authorized biobank registered at the Instituto de Salud Carlos III, and thus sample collection and processing fulfilled all ethical and legal requirements,

Total RNA was isolated from 2.5 mL blood contained in PAXgene Blood RNA tubes (Qiagen). Blood was processed using the Paxgene Blood miRNA kit (Qiagen). RNA was retrotranscribed with the Universal cDNA Synthesis kit (Exiqon) and microRNA levels were determined by real-time PCR as described above. miR-16 and miR-103a-3p were used as reference microRNAs.

Immunohistochemical analysis

For immunohistochemical analysis, tissue microarrays (TMA) were built from the paraffin-embedded samples , bearing 3-µm thick, 1-mm diameter triplicate cores for each of the samples from the miRNA microarray study. Orthotopic tumors and lung metastases were processed for the formation of FFPE blocks that were then cut into 2µm sections and processed for immunostaining as described below.

Samples were deparaffinized and rehydrated prior to antigen retrieval, followed by incubation with primary antibodies , incubation with polymer-peroxidase-conjugated secondary antibody and developed with diaminobenzidine. Slides were counterstained with haematoxylin, dehydrated and coverslipped. The pre-treatment process of deparaffinization, rehydration and epitope retrieval was performed in the PTLink module (Dako Omnis). For retrieval of antigen, the EnVision™ FLEX Target Retrieval Solution, Low pH (Dako Omnis) and High pH (Dako Omnis) were used. The Low pH solution is a pH 6.1 citrate buffer while the High pH solution is a pH 9

Tris/EDTA buffer. Retrieval was set at 97º for 20 minutes. The Cell Conditioning Solution (CC1) (Ventana) is a tris based buffer with a slightly basic pH; retrieval is set at 95º for 30 minutes. Immunohistochemistry reactions were carried out on Autostainer Link48 (Dako) when using the EnVision™ FLEX Target Retrieval Solution or on BenchMark (Ventana, Roche) when using the CC1 retrieval solution. The reaction specificity was ascertained by the absence of staining when using a non-specific isotype-matched primary antibody.

Nuclear immunostaining for ERα, PR and GATA3 was evaluated according to intensity (I; 0, 1, 2 and 3) and percentages of positive cells. For GATA3, the percentage of positive cells was classified within the following discrete intervals (P): 0, 1 (1 - 9%), 2 (10 - 49%), 3 (50 - 74%) and 4 (75 - 100%). The final staining histoscore (Hscore) was obtained by multiplying intensity and percentage interval values (Hscore = I x P), thus ranging from 0 to 12. Hscores ≥ 2 were considered as positive staining. Membrane staining for HER2 was assessed following recommendations of the American College of Pathologists (78) as 0, 1, 2 and 3. Only cases with HER2 score equal to 3 were considered HER2(+). Primary tumors were classified according to the expression of hormone receptors and HER2. Cases were considered “luminal” with positive ER and/or PR (≥ 10% positive cells) and negative HER2 (≤ 2). Membrane E-cadherin expression was scored according to the product between the intensity coefficient (0, negative; 1, low; 2, moderate; 3, strong) and the frequency of positivity coefficient (0, no stained cells; 1, 1–9%; 2, 10–49%; 3, 50–79%; 4, 80–100%) and categorised as follows: 0, negative score; 1+, score 1–4; 2+, score 5–8; 3+, score 9–12 (362).

Antibody	Clone	Source	Dilution	Manufacturer	Pretreatment
Cytokeratin	CAM 5.2	mouse	1:15	BD	CC1
Cytokeratin 18	H-80	rabbit	1:100	Santa Cruz	High pH
Cytokeratin 5/6	D5/16 B4	mouse	ready-to use	Dako	High pH
Cytokeratin 14	LLOO2	mouse	1:100	Cell Marque	High pH
E-cadherin	NCH-38	mouse	ready-to use	Dako	High pH
Estrogen Receptor α	EP1	rabbit	ready-to use	Dako	High pH
GATA-3	HG3-31	mouse	1:200	Santa Cruz	High pH
Ki-67	MIB-1	mouse	ready-to use	Dako	Low pH
p63	4A4	mouse	ready-to use	Roche	CC1
Actin (smooth muscle)	1A4	mouse	ready-to use	Dako	High pH
Sox2	D6D9	rabbit	1:30	Cell Signalling	Low pH
Vimentin	V9	mouse	ready-to use	Dako	Low pH

Table 1. List of antibodies used in immunohistochemistry

Ethics Statement

Patient selection and sample procurement complied with Spanish laws regarding data protection and written informed consent, which was obtained from all patients and stored at the Hospital Clinic Biobank, and were approved by the Hospital Clinic and IDIBAPS Ethics Committee and Review Board.

Cell culture

MCF10CA1a, MCF10CA1h, SCP2 and SCP6 cell lines were kindly provided by Dr. Yibin Kang (Department of Molecular Biology, Princeton University). MCF10CA1a and MCF10CA1h cells were maintained in DMEM/F12 supplemented with 5% horse serum, 10 μg/mL insulin (Sigma-Aldrich), 20 ng/mL epidermal growth factor (Life Technologies), 0.5 μg/mL hydrocortisone (Sigma-Aldrich), and 50 ng/mL cholera toxin (Sigma-Aldrich). PC-3M and PC-3S cells were maintained in RPMI 1640 medium supplemented with 10% fetal bovine serum, L-glutamine, non-essential aminoacids, sodium pyruvate, penicillin/streptomycin (PAA) and geneticin (Santa Cruz Biotechnologies, Santa Cruz, CA). The human breast carcinoma cell lines MDA-MB-468 and MCF7 , the MDA-MB-231 cell derivatives SCP2 and SCP6 and HEK293T cells were kept in DMEM culture medium (PAA) containing 10% fetal bovine serum and penicillin/streptomycin. Cells were grown at 37°C in a 5% CO2 atmosphere.

Generation of stable miR-200 overexpressing and shZEB2 knockdown lines

pMSCV-puro, pMSCV-hygro and the genomic fragments encoding cluster 1 (miR-200b/200a/429) and cluster 2 (miR-200c/141) cloned into pMSCV-puro and pMSCV-hygro retroviral vectors, respectively, were kindly provided by Dr. Yibin Kang. Each of these plasmids was co-transfected in the retrovirus packaging cell line PG13 with pVSVG and (Clontech, Mountain View, CA) using X-tremeGENE9 (Roche). Supernatants were collected for the following 48 h and filtered through 0.45 μm methylcellulose filters (Millipore). Virus-containing supernatants were used to infect the target cells in the presence of polybrene (8 μg/mL; Sigma-Aldrich), and selected with 1 μg/mL puromycin (Sigma-Aldrich) or 50 μg/mL hygromicinB for 7 days.

To overexpress hsa-miR-200b, the pCDH-CMV-MI0000342-EF1-copGFP lentiviral vector and its control were purchased from System Bioscience. Each plasmid was co-transfected in HEK293T cells with pVSVG and pCMVΔR8.91 (Clontech, Mountain View, CA) using X-tremeGENE9 (Roche). Supernatants were collected for the following 48 h and filtered through 0.45 μm methylcellulose filters (Millipore). Supernatants were used to infect the target cells in the presence of polybrene (8 μg/mL; Sigma-Aldrich). At 72 hours after infection, successful gene transfer was confirmed by visualization of GFP by fluorescence microscopy. These cells were enriched by fluorescence-activated cell sorting (FACS) (MoFlo, Beckman Coulter).

For the generation of stable ZEB2 knockdown cell lines, pLKO.1-Puro plasmids for control (shC002) and ZEB2-targeting shRNAs TRCN0000013528 (Sigma-Aldrich) were co-transfected in HEK293T cells with pVSVG and pCMVΔR8.91 (Clontech, Mountain View, CA) using X-tremeGENE9 (Roche). Supernatants were collected for the following 48 h and filtered through 0.45 µm methylcellulose filters (Millipore). Target cells were transduced with lentiviral particle-containing supernatans in the presence of polybrene (8 µg/mL; Sigma-Aldrich), and selected with 1 µg/mL puromycin (Sigma-Aldrich) for 7 days.

Mammosphere culture

Mammosphere culture was performed as originally described by Dontu et al (363). Single cells were plated in ultralow attachment plates (Corning) at a density of 1000 cells/mL in passages. Cells were grown in a serum-free DMEM:F12 medium supplemented with B27, 20 ng/mL EGF, 20 ng/mL bFGF (all from Life Technologies), 4 µg/mL heparin, 5 µg/mL insulin and 0.5 mg/mL hydrocortisone (all from Sigma-Aldrich) . For serial passaging, mammospheres were collected by gentle centrifugation (800 rpm) after 7–10 d and dissociated enzymatically (10 min in 0.05% trypsin, 0.53 mM EDTA-4Na; Invitrogen) and mechanically. The cells obtained from dissociation were sieved through a 70-µm sieve, analyzed microscopically for single-cellularity and replated. Colonies were counted following MTT assay in order to increase the contrast. After image acquisition, area was quantified by ImageJ.

Differentiating culture conditions

3D cultures were performed as previously described (364). Mammospheres were dissociated as described above. Cells were grown in DMEM:Ham's F-12 medium with 5% Matrigel Basement Membrane (Corning), 5% FBS, 10 µg/mL insulin, 1 µg/mL hydrocortisone, 10 µg/mL cholera toxin, 10 ng/mL EGF, and 1× Pen/Strep. Single cell suspensions were plated on Matrigel precoated plates at a density of 25000 cells/mL and allowed to form three-dimensional structures for 10-14 days.

Cell surface immunophenotyping

Cells were detached with 0.25% trypsin/ 0.1% EDTA, washed and incubated with primary antibodies CD44 (Alexa Fluor 647, anti-mouse/human, 1:4000 dilution, BioLegend) and CD24 (Alexa Fluor 488, anti-human, 1:20 dilution, BioLegend) in PBS/3% normal goat serum for 30 min in a shaker at 4 °C, washed and analyzed by flow cytometry on a Gallios Flow Cytometer instrument (Beckman Coulter).

ALDH activity

Aldehyde dehydrogenase was detected by Aldefluor kit (Stem Cell Technologies) used according to manufacturer's protocol and analyzed by flow cytometry on a Gallios Flow Cytometer instrument (Beckman Coulter).

RNA isolation, reverse transcription and real-time RT-PCR of cell culture samples.

For RNA extraction of adherent cells, cells were grown to 70-80% confluence and lysed directly on the plate with Qyazol lysis reagent. Mammospheres were collected by gentle centrifugation and resuspended with Qyazol. 3D structures were recovered from Matrigel using the non-enzymatic solution Matrisperse (Cultek), following manufacturer's instructions, and resuspended in Qyazol after gentle centrifugation. RNA was isolated with the miRNeasy Mini kit (Qiagen). cDNA was synthesized with the HighCapacity cDNA Reverse Transcription Kit (Applied Biosystems). Real-time quantitative PCR assays were performed on a LightCycler 480 instrument (Roche) and analyzed with the LightCycler 480 Software release 1.5.0. The Universal Probe Library system (UPL) (Roche) was used to quantify transcripts. Probes and sequences are shown in Table 2. RN18S1 amplification levels were used as an internal reference, and relative transcript quantification determined by the $\Delta\Delta$Cp method.

Gene	UPL probe	Forward 5'-3'	Reverse 5'-3'
CDH1	35	CCCGGGACAACGTTTATTAC	GCTGGCTCAAGTCAAAGTCC
EpCAM	3	CCATGTGCTGGTGTGTGAA	TGTGTTTTAGTTCAATGATGATCCA

ZEB1	3	GGGAGGAGCAGTGAAAGAGA	TTTCTTGCCCTTCCTTTCTG
ZEB2	68	AAGCCAGGGACAGATCAGC	GCCACACTCTGTGCATTTGA
KRT8	64	GATGAACCGGAACATCAGC	CATCCTTAATGGCCAGCTCT
KRT18	78	TGATGACACCAATATCACACGA	GGGCTTGTAGGCCTTTTACTTC
KRT5	1	GCAGATCAAGACCCTCAACAAT	CCACTTGGTGTCCAGAACCT
KRT14	18	CCTCCTCCCAGTTCTCCTCT	ATGACCTTGGTGCGGATTT
TP63	10	CGCCATGCCTGTCTACAA	TGACTAGGAGGGGCAATCTG
ACTA2	58	CTGTTCCAGCCATCCTTCAT	TCATGATGCTGTTGTAGGTGGT
SOX2	19	ATGGGTTCGGTGGTCAAGT	GGAGGAAGAGGTAACCACAGG
POU5F1	52	GTGCCTGCCCTTCTAGGAAT	GGCACAAACTCCAGGTTTTCT
KLF4	82	GCCGCTCCATTACCAAGA	TCTTCCCCTCTTTGGCTTG
MYC	34	CACCAGCAGCGACTCTGA	GATCCAGACTCTGACCTTTTGC
NANOG	69	ATGCCTCACACGGAGACTGT	AGGGCTGTCCTGAATAAGCA
BMI1	54	TGTAAAACGTGTATTGTTCGTTACC	CAATATCTTGGAGAGTTTTATCTGACC
FOXA1	1	AGGGCTGGATGGTTGTATTG	ACCGGGACGGAGGAGTAG
ESR1	24	TTACTGACCAACCTGGCAGA	ATCATGGAGGGTCAAATCCA
Elf5	87	AGTTGAGCAGAGCCCTGAGA	GCCTGGAGCAGATCATAGCTT
PTEN	48	GGGGAAGTAAGGACCAGAGAC	TCCAGATGATTCTTTAACAGGTAGC
SYNJ1	84	CTGCTTGGTGAAGATGCAGA	GAAGTTCCTCCACTTCAGCACTA
TSC1	65	CAACCAGAGCCAGGAATTACA	CAGCTCCGCAATCATGTTC
RN18S1	40	GGAGAGGGAGCCTGAGAAAC	TCGGGAGTGGGTAATTTGC

Table 2. List of primers used in this study for qPCR

Immunocytochemistry

Cells were seeded on sterile round glass coverslips, allowed to attach for 24 to 48 h, washed with PBS, fixed with cold 4% paraformaldehyde for 20 min, permeabilized with 1% Triton X-100 in PBS, blocked for 30 min with blocking buffer (3% BSA, 1% Triton X-100 in PBS), incubated with primary antibodies in blocking buffer (Table 3) for 2 h at room temperature, washed 3× with blocking buffer, and incubated for 30 min with Alexa Fluor 488-conjugated rabbit-anti-mouse antibodies (Life Technologies; 1:1,000 in blocking buffer), Alexa Fluor 568 goat anti-rabbit (Life Technologies; 1:1,000) and 4',6-Diamidino-2-phenylindole (DAPI) (Sigma; 0.1 µg/mL). After washes, samples were mounted on glass slides with Mowiol 18–88 (Sigma) and visualized under a Leica SP5 confocal microscope (Leica, Wetzlar, Germany).

Immunoblotting

Cell lysates were prepared in 80 mM Tris-HCl pH 6.8, 2% SDS buffer and sonicated (Branson digital sonicator) in two cycles of 20 s and 10% amplitude with a conic tip. Protein from the sonicated samples was quantified by the Lowry method (BioRad) and boiled in Laemmli buffer (60 mM Tris-Cl pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue). Samples (30 to 100 µg) were electrophoresed by SDS-PAGE and transferred to PVDF membranes (Immobilon-FL, Millipore). The efficiency of the electrophoretic transfer was monitored visually by complete transfer of pre-stained protein molecular weight standards (BioRad) electrophoresed in parallel with the samples, and by Ponceau Red protein staining of the membranes after transfer. Membranes were washed briefly with PBS and blocked for 1 h in an orbital shaker at room temperature with either Odyssey blocking buffer (Li-COR Biosciencies) or PBS/5% powder skim milk/0.05% Tween-20 (chemiluminescence blocking buffer) and incubated o/n at 4 ºC with primary antibodies diluted in blocking buffer (Table 3). After incubations, membranes were washed 3x with PBS (15 min) and incubated with secondary antibodies diluted in blocking buffer for 1 h. Membranes were washed 3x with PBS and reactions detected by fluorescence. Fluorescence detection was done with 1/10000 dilution in blocking buffer of fluorescent dye-conjugated secondary antibodies and the membranes were scanned with an Odyssey infrared imaging system (Li-COR Biosciences) and quantified with the software provided by the manufacturer. Sample loadings were normalized against actin or tubulin levels.

Protein	Clone	Source	Manufacturer	Dilution WB	Dilution ICC
E-cadherin	36	mouse	BD	1:8000	1:500
Fibronectin	Polyclonal	goat	Sigma-Aldrich	1:500	-
Vimentin	V9	mouse	LabVision	1:200	-
Akt	EPR16798	rabbit	AbCAM	1:10000	-
Akt (phospho S473)	EP2109Y	rabbit	AbCAM	1:5000	-
Tubulin	Tub 2.1	mouse	Sigma-Aldrich	1:2000	-

Table 3. List of antibodies used in immunocytochemistry and immunoblotting.

3'UTR luciferase reporter assays.

psiCHECK2-PTEN 3'UTR construct was obtained from Addgene (Plasmid 50936).. Due to size constraints, TSC1 3'UTR was cloned as three different fragments (Fragment #1:483bp; fragment #2: 941bp; fragment #3: 368bp) into the pCR® Blunt vector. TSC1 3'UTR fragments were subcloned into psiCHECK-2 (Promega) using XhoI and NotI restriction sites. The primers used are listed in table 4. Reporter assays were performed as follows: HEK293T were transduced with lentiviral particles carrying either pmiR-empty(∅) or pmiR-200b. Seventy-two hours post-infection, cells were seeded into 96-well plates and transfected with 50 ng of the indicated 3'UTR reporter vectors for an additional 24h (n=4 per condition). Luciferase activity was measured using the Dual-Glo Luciferase AssaySystem (Promega). Renilla luciferase activity was normalized to corresponding firefly luciferase activity and plotted as a percentage of the control.

3' UTR Reporter	Forward 5'-3'	Reverse 5'-3'
TSC1 (F1)	CTCGAGAAAGACTTGGGTGTGGAAGC	GCGGCCGCTGAACTTGCACTCAGACCCT
TSC1 (F2)	CTCGAGTCTGGGTGTGACTGATTCCC	GCGGCCGCACATTGGCCAAACAGCTGAG
TSC1 (F3)	CTCGAGCAGCGCCACATTATCCATCC	GCGGCCGCTCTTTCACTGATGGGACCCC

Table 4. List of primers used in the 3'UTR reporter assay.

In vivo tumorigenic and lung colonization assays

MCF10CA1h cells were transduced with pCMV-GFP/luc for the constitutively expression of the firefly luciferase gene and GFP. GFP+ cells were selected by fluorescence activated cell sorting (FACS). Luciferase expression was assessed by processing cell lysates with the Dual luciferase assay (Promega) and measuring them in an OrionII Microplate Luminometer (Berthold Detection Systems. These cells were then used for generating the MCf10ca1h-control, MCf10ca1h-miR200-C1 and C2 overexpressing cell lines. To assess the in vivo tumorigenic potential, female immunodeficient (SCID-NOD) mice aged 8-10 weeks were anesthetized with 100 mg/kg ketamine (Merial) and 3.3 mg/kg xilacine (Henry Schein) and injected with 50 µL of 2 x 10^6 cells suspended in 50%PBS/50% matrigel into the abdominal mammary fat pad. In vivo optical imaging of engrafted mice was performed after intraperitoneal injection with 150 mg/kg of D-luciferin (16.7 mg/mL in physiological serum) (Caliper Life Science). Tumor growth was monitored once a week in vivo at real time in anesthetized mice and after intraperitoneal injection of D-luciferin. Mice were then placed in the detection chamber of an ORCA-2BT imaging system (Hamamatsu Photonics) provided with a C4742–98-LWG-MOD camera and a 512 x 512 pixel, chargecoupled device (CCD) cooled at -80 ºC at a distance of 200 mm from the camera objective (HFP-Schneider Xenon 0.95/25 mm). Imaging was performed routinely 5 min after substrate injection. Two images were generated from each mouse, one using a light source inside the chamber to register the animal position and a second one, in total darkness, during a 5 min period to acquire photons from the light emitting cells. To increase detection sensitivity the readout noise of the recorded signal was reduced by adding the light events recorded by arrays of 8 x 8 adjacent pixels (binning 8 x 8) in the camera CCD. Mice were monitored during a 7week period; 8 control mice and 10 mice xenografted with MCF10ca1h-miR200-C1 cells. To evaluate the hystopathology, 2 mice were xenograted with MCF10ca1h-miR200-C2 cells. Quantification and analysis of photons recorded in the areas of interest from images were carried out using the Hokawo Imaging Software (Hamamatsu Photonics). To assess lung colonization, 150 µL of 500,000 cells were injected in the tail vein of immunodeficient (SCID-NOD) mice. 5 mice for each cell line were monitored during a 5 week period as described in the orthotopic implantation assay. The ortothopic xenograft tumors and lungs from the colonization assays were excised from the mice at necropsy, preserved in 10% formalin solution (Sigma) immediately after *ex vivo* imaging and subsequently processed for histological analysis.

Results

OBJECTIVE 1: Identification of sets of microRNAs deregulated along metastatic progression of invasive ductal carcinomas

A microRNA screening on invasive ductal carcinomas

Seeking to identify miRNAs with deregulated expression patterns along breast cancer metastatic progression, we proceeded to a miRNA microarray screening of FFPE samples from normal breast (N), ductal carcinomas *in situ* (DCIS), primary IDC without lymph node involvement (PNM), primary IDC with regional lymph node involvement (PM), lymph node metastases matched to the node-positive primary tumors (LNM) and distant metastases (DM). A pool of non-cancerous lymph nodes from the breast cancer patients was included as a negative control of lymphocytic contamination. As breast cancer is a heterogeneous disease, we only selected breast ductal carcinomas with a luminal molecular phenotype, based on positivity for ER and PR expression and negativity for the HER2 receptor by immunohistochemistry, which we further enriched by macrodissection of the epithelial component.

From the 866 human mature microRNAs analyzed, 39 miRNAs were significantly upregulated and 41 miRNAs (Fig. 1A, left) were downregulated in this breast cancer subtype (Fig. 1A, right).

This analysis revealed a set of 10 miRNAs with a general tendency to increase and a set of 9 miRNAs that tended to decrease along metastatic progression, in the transition from normal breast to primary tumors and distant metastases. We then reselected the most significant miRNAs, based on fold change and p-value, for further validation by qPCR on the same samples. miR-7, miR-210, miR-181a and miR-181b levels tend to increase and are significantly upregulated in distant metastasis relative to node-negative primary tumor (Fig. 1B, left). On the contrary, miR-10b and miR-101 levels are downregulated along progression; the lowest levels being in distant metastases (Fig. 1B, right). Also, miR-30a, miR-30a* and miR-148a show a general tendency to decrease along progression, with significant lower levels in distant metastases compared to non-metastatic primary tumors (Fig. 1B, right).

Focusing on the clinically relevant question of what is the difference, at a molecular level, between a primary tumor that metastasizes either to lymph nodes or distant organs and a non-metastatic tumor, we found that 17 miRNAs exhibited a differential expression ($p \leq 0.01$). 7 miRNAs had an increased expression in primary metastatic tumors while 10 microRNAs were

downregulated. qPCR and microarray data are corresponding: miR-7 is upregulated in primary metastatic tumors while miR-30a, miR-30a* and miR-148a levels decrease (Fig. 1B, right). However, it is the subset formed by miR-200b, miR-7, miR-210 and miR-148a that significantly discriminates node-negative from node-positive patients (Fig. 1C).

Circulating microRNAs in breast cancer patients

As circulating microRNAs have arisen as promising biomarkers for cancer detection, we profiled the selected microRNAs that we had found deregulated along metastatic progression in tumor tissue samples, in the blood from breast cancer patients. For this, we initiated a prospective study in which we collected total blood samples from newly diagnosed breast cancer patients that had not been subjected to any prior treatment, taken at the time-point of diagnosis. Patients were stratified according to their metastatic status in patients with a primary breast tumor and lymph node metastasis (PM) (N=29) or without lymph node metastasis (PNM) (N=40). Additionally, already diagnosed and treated patients suffering from distant metastasis were included (DM) (N=14). Six women with negative mammographies were included as normalization controls. The levels of miR-200b, miR-200c, miR-181a, miR-210, miR-101, miR-10b, miR-7, miR-30a and miR-148a were determined by qPCR (Fig. 2).

The analysis showed that miR-200b and miR-7 blood levels tend to be increasingly higher in patients with affected lymph nodes at diagnosis and distant metastatic patients and are statistically significant when compared to the levels in patients with node-negative tumors (Fig. 2, A-B). Considering its delta Ct as a measure of its expression, miR-7 levels in blood are relatively high while miR-200b exhibits the lowest blood levels. As in the tumoral tissue, miR-7 blood levels shows a tendency towards increased expression in node-positive compared to node-negative patients. On the contrary, miR-30a and miR-148a blood levels do not exhibit a significant decrease in PM patients (Fig. 2 E-F). Having seen such interesting results with miR-200b, we decided to analyze miR-200c as a representative member of the cluster 2 of the microRNA-200 family. However, no significant differences are appreciated despite its circulating levels are superior to that of miR-200b (Fig. 2H).

Among the microRNAs that we found deregulated in distant metastasis, miR-181a and miR-210 (Fig. 2 C-D), the expression of which is augmented in metastasis, exhibit significantly lower levels in the blood of distant metastatic patients. However, considering its expression compared to the rest of the analyzed microRNAs, they are found at considerably high levels in blood. Likewise,

miR-101 is also expressed at raised levels in blood and shows a tendency to be decreased in distant metastatic patient, exhibiting the same pattern as in tissue (Fig. 2G). The other microRNA downregulated along metastatic progression, miR-10b, was not expressed or not at sufficient detectable levels in blood.

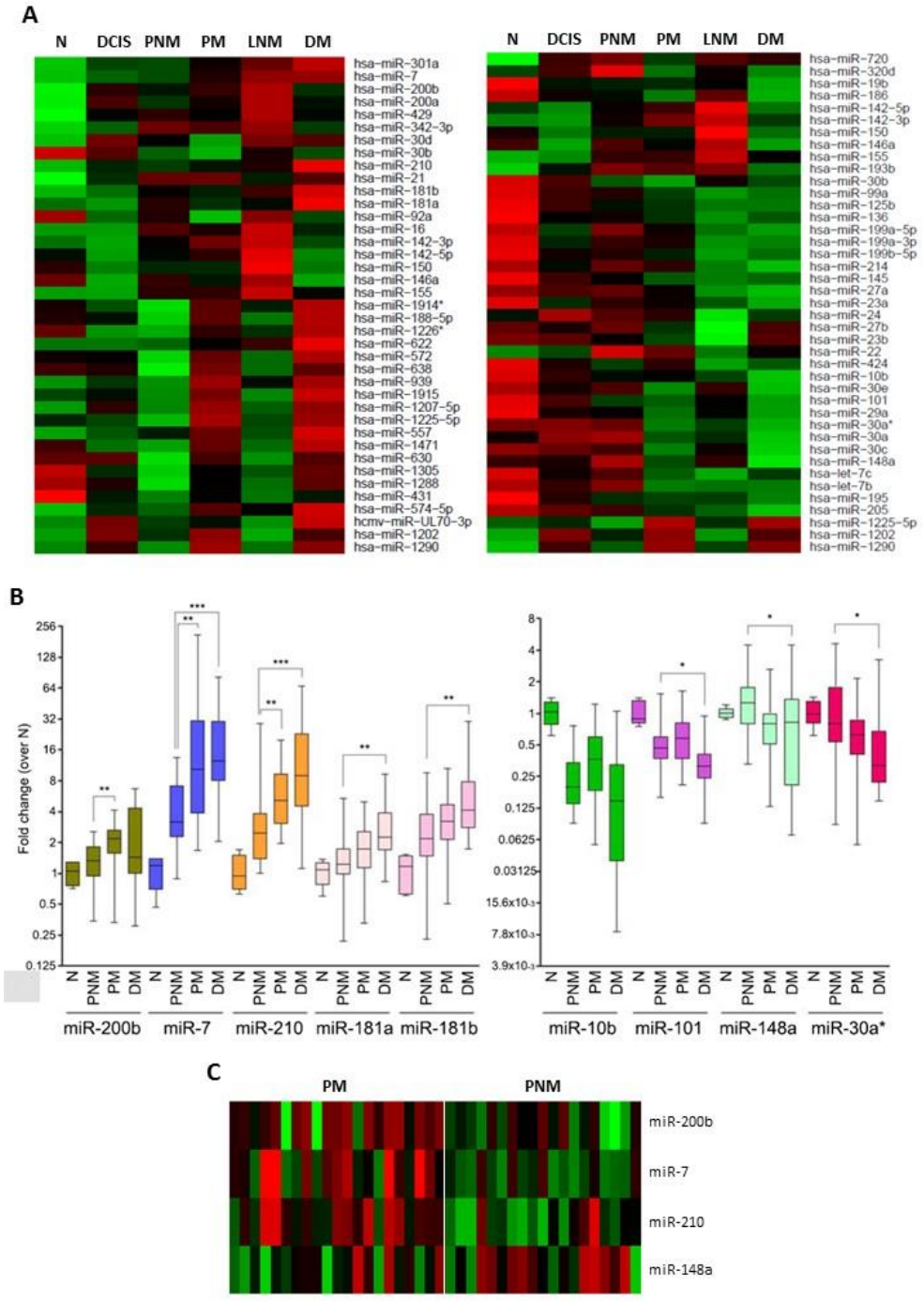

Figure 1. MicroRNAs dysregulated along malignant invasive ductal carcinoma (IDC) progression. (**A**) Microarray analysis of relative expression levels of miRNAs differentially expressed between samples of normal breast epithelium (N), ductal carcinoma in situ (DCIS), primary IDC with no regional lymph node involvement or distant metastasis (PNM), primary IDC with lymph node involvement (PM), matched lymph node metastases (LNM) or distant metastases (DM). Heatmaps were generated after probe normalization and selection of differentially expressed miRNAs upregulated (left) or downregulated (right) along progression (**B**) Quantification by qPCR of levels of selected miRNAs along metastatic progression of IDC. Reference probe-normalized values ($n^{-\Delta Ct}$) are shown relative to the median of values for normal breast epithelial tissues (N). (C) Heatmap generated with qPCR data of the differentially expressed miRNAs that discriminate PNM from PM tumors.

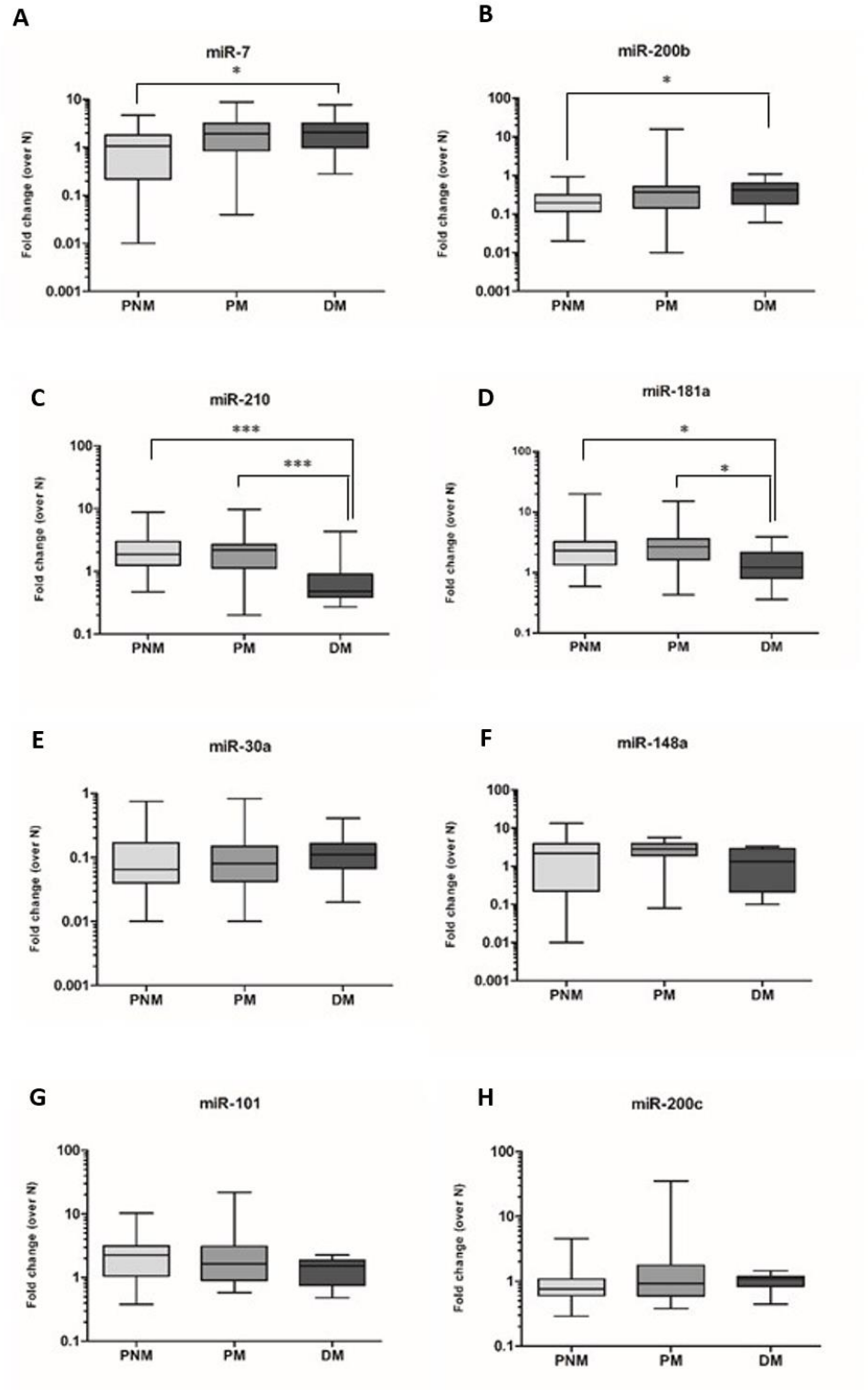

Figure 2. Circulating levels in blood of a set of significant microRNAs (A-H) deregulated along progression in the microarray study . Reference probe-normalized values (- ΔCt) from each group of patients are shown relative to the median of values for blood from a control group of women with no history of breast cancer. Data is represented in a logarithmic scale.

OBJECTIVE 2: Role of the miR200s family in metastatic progression

Expression of miR-200f and E-cadherin in lymph node metastases

In the absence of distant metastases, lymph node status is one of the most important prognostic factors considered by clinicians to guide treatment decisions. The presence of tumoral cells in the lymph nodes reflects the ability of the tumor to metastasize and to colonize distant organs with greater probability than lymph node negative tumors. We therefore, were interested in the microRNA expression profile of lymph node metastasis and its corresponding primary tumors. Besides enrichment of the epithelial component of the primary tumors, we also procured the selection of the carcinoma cells that had metastasized to lymph nodes by macrodissection.

The results from microarrays show a cluster of 11 microRNAs markedly upregulated in lymph node metastasis (Fig. 3A). Among them, we were particularly interested in the miR-200 family members miR-200a, miR-200b and miR-429, known EMT inhibitors mainly described at the moment as metastasis-inhibiting microRNAs. We further determined the miR-200 cluster 1 expression by qPCR and observed that 40% of the lymph node metastasis exhibit over a 1.5 fold change expression for miR-200a and 45% in the case of miR-200b and miR-429, over their matching primary tumors. The remaining lymph node metastasis, 45% for miR-200a and miR-429 and 50% for miR-200b, increase discreetly or maintain the same expression as the primary tumor (Figure 3B). Among the three family members, miR-200a and miR-200b show the highest expression in both PM and LNM having a ΔCt mean of 1.66 and 1.56, each microRNA respectively, and highest correlation in their levels (r^2=0.96). miR-429 is the most different member in terms of tissue expression, with a 4.12 ΔCt mean, and correlation with miR-200a and miR-200b (r^2=0.73 and 0.67, respectively) (Fig. 3B).

As the miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2, we performed immunohistochemistry on a cohort of 49 primary tumors and its paired lymph node metastasis, to assess the E-cadherin expression (Fig. 3C). Considering only the samples in which miR-200 levels were measured, 80% of LNM showed an increase and 15% maintained the E-cadherin expression compared to the primary tumors they come from. When considering also the rest of samples, 90% of LNM either maintain (51%) or have increased expression (39%)(Fig. 3C).

As expected, miR-200 levels and E-cadherin expression are mainly corresponding. The majority of metastatic cells from ductal carcinomas with a luminal phenotype express miR-200 and/or E-

cadherin at considerable levels. Therefore, both are implicated at least in the colonization stage of the metastatic process.

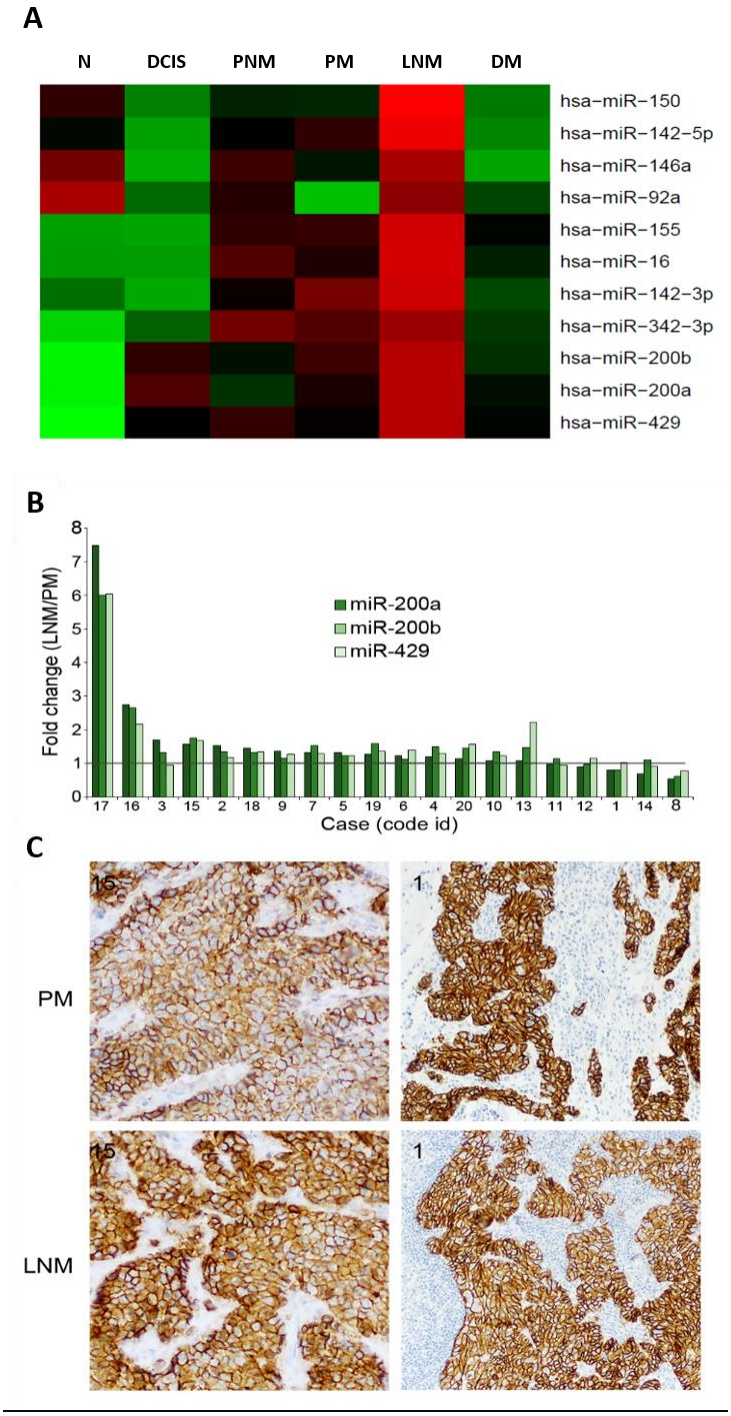

Figure 3. Involvement of miR-200f and the epithelial phenotype in lymph node metastases. (A) Heatmap generated from the microarray data showing the microRNAs with highest expression in lymph node metastases (LNM). (B) miR-200's are frequently expressed at higher levels in lymph node metastases than matched primary tumors. Shown are ratios (fold-change; - ΔΔCt) of qPCR values (ΔCt) for the indicated miRs in lymph node metastases (LNM) and their matched primary tumor (PM). The cut-off threshold was set at a fold change of 1.5 (C) Immunohistochemical staining of E-cadherin in paired LNM and PM.

Differential microRNA expression levels in breast cancer cell lines

The above results suggest that metastatic breast cancer cells display both an enhanced epithelial gene program and increased miR-200 levels. Many studies have found an association of low miR200 levels with invasive cancers and poor prognosis and show that constitutive expression of miR-200 family members can suppress invasiveness and metastasis in mouse xenotransplanation models (302, 365). Recent studies, however, have challenged these conclusions and rather suggest a pro-metastatic function for miR-200's (341, 356, 357). To determine the expression profiles of miR200 in cellular models displaying differential metastatic potential, we quantified miR-200b levels in cell lines derived from breast cancers corresponding to distinct molecular subtypes and also from non-cancerous MCF10 series breast cells (Fig. 4A). Remarkably, the highly metastatic MDA-MB-486 cells, derived from basal-like breast cancers with a strong mesenchymal-like phenotype (366) expressed significantly higher levels of miR-200b than the luminal-type MCF7 cells, which are poorly metastatic and display a prominent epithelial phenotype (Fig. 4A). Moreover, we found that two clonal lines derived from MD-MBA-231 basal-like cancer cells, SCP2 and SCP6 (367), presented a striking differential expression of miR-200b, with the highly metastatic SCP2 cells expressing 60-fold more miR-200b than the poorly metastatic SCP6 cells (Fig. 4A). Likewise, the aggressive PC-3M cells, derived from the PC-3 prostate cancer cell line, expressed more than 10-fold higher levels of miR-200b than the non-metastatic PC-3S subpopulation (368) (Fig. 4A). Finally, two subpopulations derived from the non-cancerous MCF10 cell line (369, 370), the moderately metastatic MCF10C1a and the non-metastatic MCF10CA1h cell lines, also presented a remarkable differential expression profile for miR200, with MCF10CA1a cells expressing miR200b levels up to 3 orders of magnitude higher than MCF10CA1h cells (Fig. 4A). These observations suggest an association of high miR-200 levels with metastatic potential and conform to our observations of a tendency of metastatic breast cancer samples to express higher levels of miR-200's than primary tumors.

The only other microRNAs showing a similar expression profile consistent with that of the metastatic samples, in which they are upregulated, are miR-181a and miR-7 in the metastatic SCP2 and MDA-MB-468, also miR-181b in the SCP2 cells and miR-210 in the metastatic MCF10CA1a. However, miR-200b is the only microRNA upregulated in all of the metastatic cell line counterparts.

Overexpression of the miR-200 microRNA family in the MCF10CA1h cell line

Given that the MCF10CA1 cells are fully malignant derived clones with a common genetic background, we selected the mesenchymal-like, weakly metastatic MCF10CA1h cell clone to characterize the phenotype upon miR-200's overexpression. We stably overexpressed the miR-200 cluster 1 (miR-200 C1) and cluster 2 (mir-200 C2) by using a retroviral vector containing the genomic fragment encoding miR-200a/miR-200b/miR-429 and miR200c/miR-141, respectively. A lentiviral vector containing miR-200b was also used to generate MCF10CA1h miR-200b stable cell lines that co-express GFP.

MCF10CA1h cells exhibit a stellate fibroblast-like shape (Fig. 4D). Upon miR-200 overexpression, they acquire a more epithelial-like morphology presenting a wider cytoplasm, retraction of filopodia and grouping in clusters; showing an intermediate morphology between MCF10CA1h cells and the epithelial-like MCF10CA1a cells (Fig. 4D). No evident changes in morphology are observed in cells transduced with retroviral particles carrying the corresponding empty expression vector (Fig. 4D). RT-qPCR of EMT markers clearly support the induction of an epithelial gene program with an upregulation of the epithelial genes E-cadherin and EpCAM (Fig. 4B,C). However, mesenchymal markers vimentin or fibronectin are not always downregulated, suggesting that there exists a varying portion of the population that does not undergo a complete mesenchymal to epithelial transition; rather it achieves an intermediate state (Fig. 4B,C).

We next determined whether miR-200 overexpression affected the ability of MCF10CA1h cells to proliferate in suspension, as nonadherent mammospheres. This *in vitro* assay is an indicator of tumor cell self-renewal and is a good predictor of the tumorigenic potential of cells when xenografted in immunodeficient mice (363). As the sphere formation assay allows for serial propagation of cells in an undifferentiated state, we assessed the self-renewal ability of MCF10CA1h miR-200 C1 and C2 lines by serial cultivation of the mammospheres in a second generation. The results show that both miR-200 cluster 1 and cluster 2 endow MCF10CA1h cells with increased self-renewal ability and this capacity is sustained with serial passaging (Fig. 5A).

To further confirm the above results regarding miR-200's overexpression, the aldehyde dehydrogenase (ALDH) enzymatic activity was measured. High levels of ALDH expression has been reported for normal and cancer precursor cells of various lineages, such as primitive human hematopoietic progenitor cells, cancer stem cells, primary human mammary epithelial cells, human breast cancer cell lines (371). While less than 5% of either of the different control

populations have high ALDH activity, a fraction that comprises a 30% to 50% of miR-200 cluster 1 or cluster 2 expressing cells, display ALDH activity (Fig. 5B). These features are consistent with an induction by miR-200f of either a stem cell or a progenitor state in mammary cell lineage specification (24, 27, 29, 372)

We also assessed the expression of the cell surface markers CD44 and CD24 by flow cytometry, two cell-surface markers whose expression in the $CD44^{high}/CD24^{low}$ configuration is associated with both human breast CSCs and normal mammary epithelial stem cells (31, 125, 373). Practically all of the MCF10CA1h control cells display a $CD44^{high}$ $CD24^{low}$ phenotype and upon miR-200 cluster 1 overexpression, almost the entire population acquires a $CD44^{high}/CD24^{high}$ expression pattern (Fig. 5C). Overexpression of miR-200 cluster 2 also induces a shift to the $CD44^{high}$ $CD24^{high}$ phenotype but roughly in 30% of the population (Fig. 5C).

In summary, overexpression of either miR-200 cluster 1 or cluster 2 engages the acquisition of epithelial characteristics in the mesenchymal-like MCF10CA1h cells coupled to the ability to self-renew in non-adherent conditions and the expression of markers associated to progenitor or stem-like features.

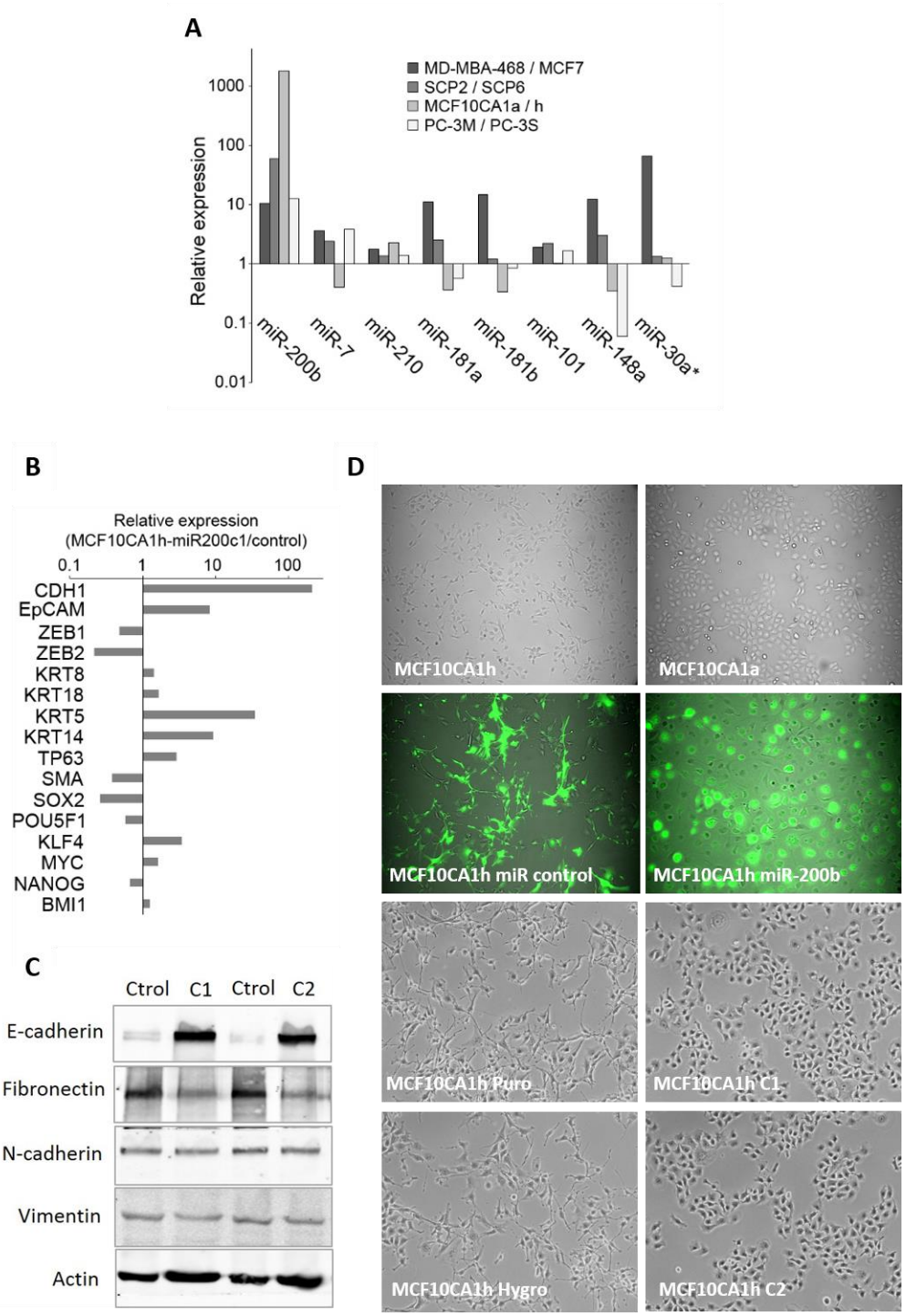

Figure 4. (A) qPCR expression analysis for the indicated miRs comparing metastatic *vs* non/weakly metastatic cell lines. Shown are ratios (fold-change; - ΔΔCt) of qPCR values (ΔCt) for the indicated miRs represented in a logarithmic scale. (B) qPCR expression analysis of selected mRNAs expressed in MCF10CA1h miR-200 C1 cells relative to MCF10CA1h control cells . Shown are ratios (fold-change; - ΔΔCt) of qPCR values (ΔCt) for the indicated mRNAs represented in a logarithmic scale. (C) Western blotting of total extracts of cells probed with anti-E-cadherin, anti-fibronectin, anti-N-cadherin and anti-vimentin, normalized to actin. (D) Bright field images of MCF10CA1h, MCF10CA1a, MCF10CA1h control and miR-200 C1 and C2 overexpressing cells. Fluorescence images showing GFP expression as an indicative of the expression of a control miR (second raw, left) or miR-200b (second raw, right) in MCF10CA1h cells.

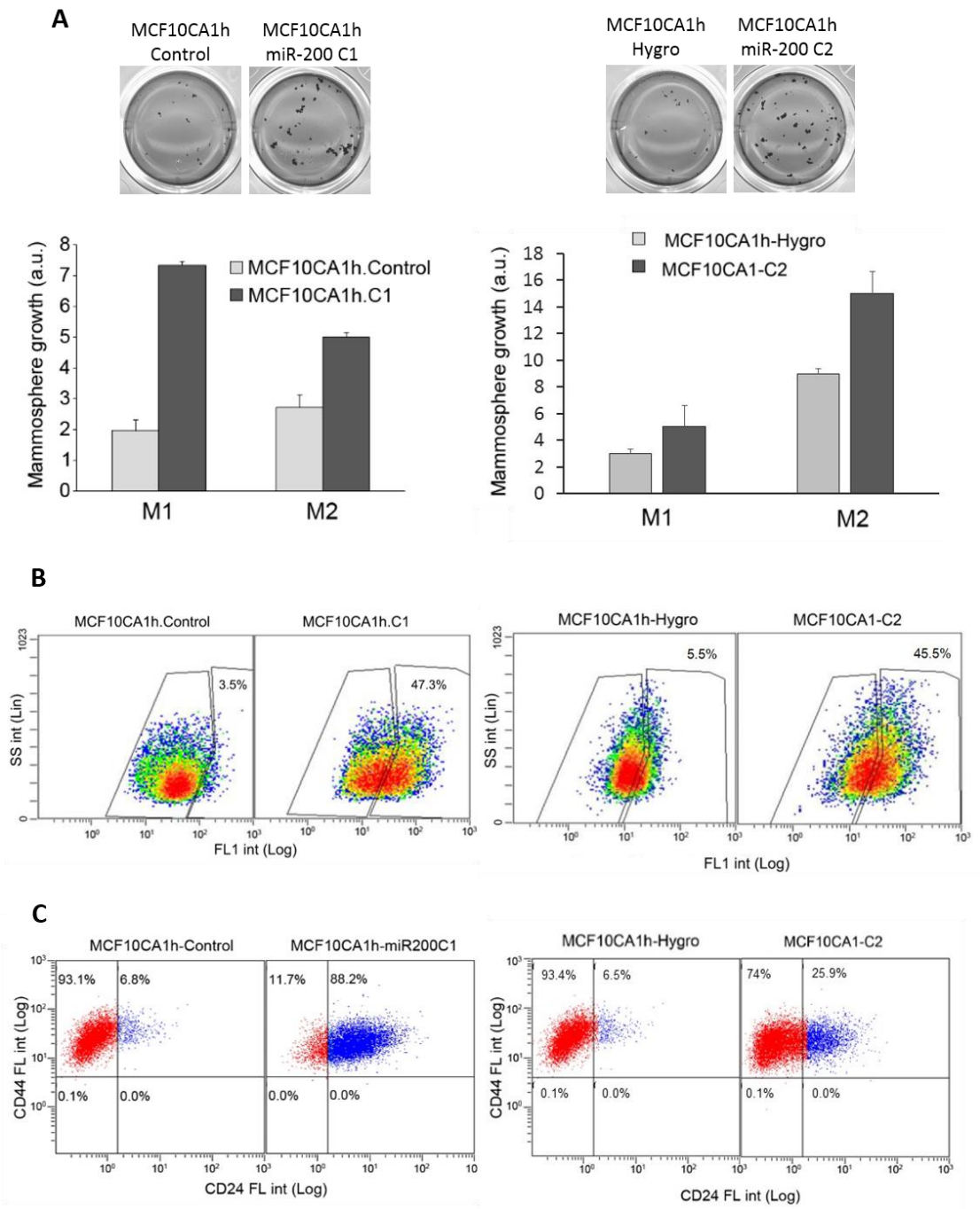

Figure 5. Characterization of stem cell properties in the MCF10CA1h model upon miR-200 expression. (A) miR-200 C1 and C2 potentiate the mammosphere growth of MCF10CA1h cells. Cells (6 x 10^3/well) were seeded in low-attachment plates and allowed to grow for 7 days (primary mammospheres, M1), disgregated and reseeded for secondary mammosphere formation (M2). Quantification was done after MTT staining and image analysis (n = 3). Shown are representative bright field images of M1 mammospheres (top) and graphic representing the quantification (bottom). (B) miR-200 C1 and C2 strongly induce aldehyde dehydrogenase activity in MCF10CA1h cells. ALDH activity was determined with the Aldefluor assay and flow cytometry. Windows for baseline fluorescence levels were established for cells incubated with the Aldefluor reagent and the ALDH inhibitor diethylaminobenzaldehyde (DEAB). (C)) miR-200 C1 and C2 upregulate cell-surface CD24 in MCF10CA1h cells. Cells in suspension were incubated with Alexa Fluor 647-labeled anti-CD44 and Alexa Fluor-488-labeled anti-CD24 and analyzed by flow cytometry.

Activation of the PI3K-Akt signaling pathway and knockdown of Zeb2 partially phenocopy the stem-like characteristics attributed to miR-200 expression

Taking into account that miR-200 overexpression promotes the acquisition of stem-like traits in MCF10CA1h cells, we next wanted to investigate which was the target whose direct mRNA repression and silencing by the miR-200 family could explain the acquirement of this phenotype.

With this purpose, we searched several microRNA target prediction databases (microRNA.org, TargetScan, PicTar) to obtain a list of possible candidates that directly interacted with the cluster 1 and cluster 2 members. Among the most probable targets, the ones with the highest score, we also searched the published literature to evaluate their general role in cancer progression and metastasis and more specifically if they were involved in the epithelial-to-mesenchymal transition or cancer stem cell generation. This allowed us to scale down to a few candidates: ZEB1, ZEB2, QKI, FBXW7, RECK, SYNJ1, TSC1 and PTEN.

As an initial exploratory phase, we first determined their mRNA levels by qPCR (Fig. 6A). Upon miR-200 cluster 1 expression, transcript levels for all these genes were downregulated in MCF10CA1h cells, indicating that they were directly or indirectly regulated by miR-200a, miR-200b and/or miR-429. ZEB2 and PTEN were markedly downregulated (Fig. 6A). Since ZEB2 has long been validated as a direct target, we were interested in wether PTEN and TSC1, both part of the PI3K/Akt signaling pathway, were subjected to direct regulation by miR-200.

In order to functionally validate the interaction between miR-200 and its predicted targets, 3'-UTR luciferase reporter assays were performed (Fig. 6B). The 3'UTR fragment of PTEN and TSC1 were cloned into the psiCHECK™-2 Vector in a 3' location to the *Renilla* luciferase gene under control of the SV40 promoter. Since miR-200a and miR-200b are part of functionally different clusters, we chose them as representative candidates in case they had different target affinity. To this end, HEK293T cells stably expressing miR-200b and HEK293T containing miR-200a mimics were cotransfected with the vector containing the 3'UTR region of the candidate target gene. If miR-200a or miR-200b binds to the target mRNA and initiates the RNAi process, the fused *Renilla* luciferase: 3'UTR region of the interest mRNA is cleaved and subsequently degraded, decreasing the *Renilla* luciferase signal.

PTEN did not decreased luciferase activity, despite having one binding site for miR-200a/141 and miR-200b/c/429 (Fig. 6B). Therefore, the PTEN transcript is not directly regulated by the miR-200 family. In contrast, the second 3'URT fragment of TSC1 (TSC1 3'UTR F2), the only one containing two binding sites for miR-200b/c/429, displayed a 20% reduction of luciferase activity

upon miR-200b expression but no significant change when miR-200a wasexpressed (Fig. 6B). Conversely, the 3'UTR TSC1 fragments 1 (F1) and 3 (F3), each one containing one binding site for miR-200a/141, decreased luciferase activity by 60% and 20%, respectively, when miR-200a but not miR-200b was expressed (Fig. 6B).

To assess the involvement of the PI3K/Akt pathway in the the gain in mammosphere formation conferred by miR-200s, both control cells and cells overexpressing miR-200 C1 were treated with the PI3K inhibitor, LY294002 or the mTOR inhibitor, rapamycin and grown in non-adherent conditions . Both of them were able to block the increase in mammosphere growth promoted by miR-200 (Fig. 6C). Furthermore, expression of either cluster 1 or cluster 2 induces activation and phophorylation of Akt suggesting the implication of this signaling pathway in mediating, at least partially, the effect of miR-200 in self-renewal ability (Fig. 6D).

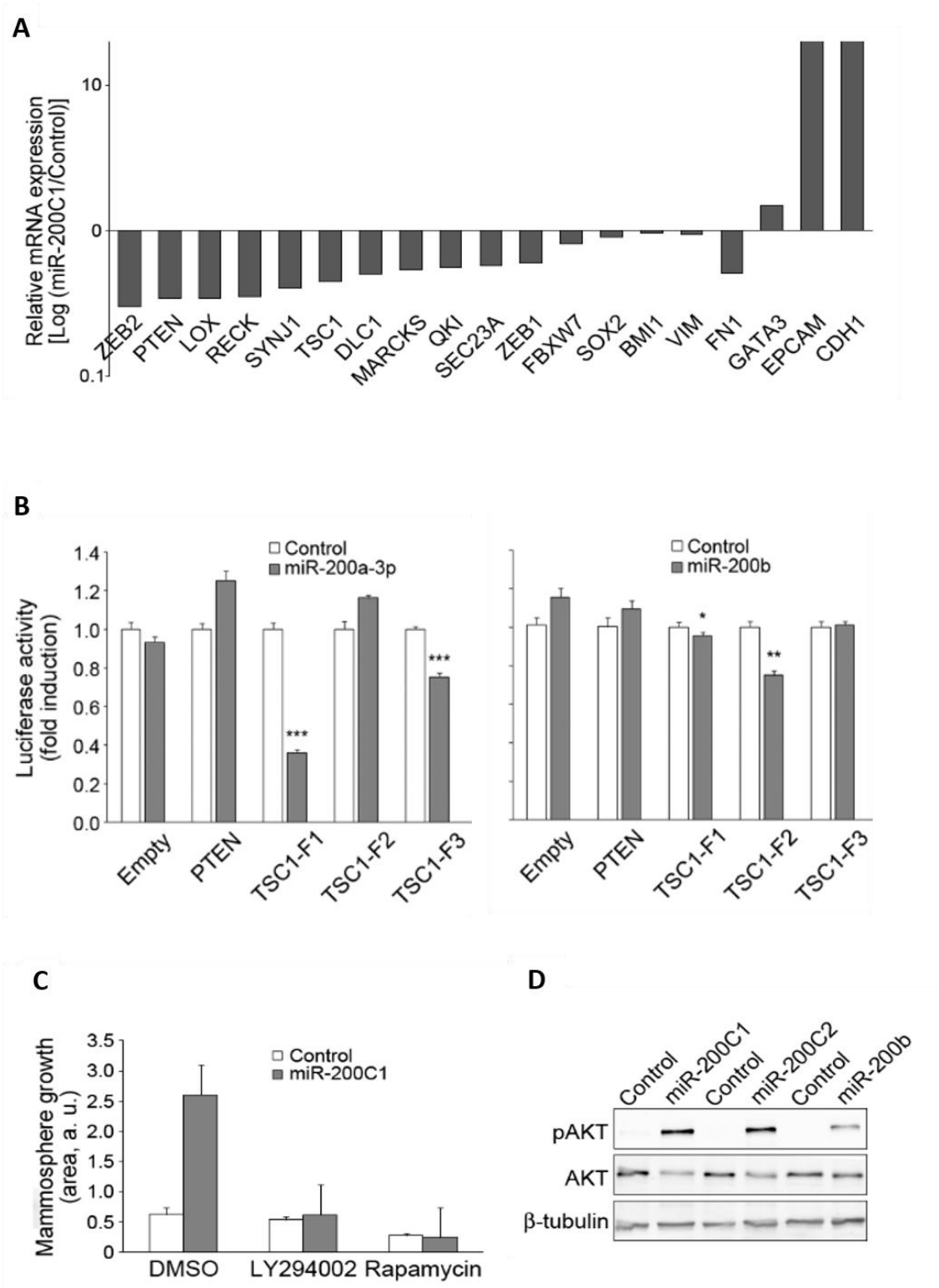

Figure 6. Activation of the PI3K-Akt signaling pathway by expression of miR-200 in MCF10CA1h cells. (A) qPCR expression analysis of known and predicted miR-200 mRNA targets. (B) Luciferase reporter assays in HEK293T cells cotransfected with miR-200a or miR-200b and the 3'UTR fragments of PTEN or TSC1. (C) Mammosphere growth assay upon treatment of MCF10CA1h control and miR-200 C1 cells with the PI3K signaling pathway inhibitors LY294002 and rapamycin. (D) Western blotting of total extracts of cells serum-starved for 24 h, probed with anti-pAkt (Ser473), total Akt or β-tubulin.

We also investigated the role of ZEB2 by silencing its expression in MCF10CA1h cells. Knockdown of ZEB2 in MCF10CA1h cells resulted in a significant increase in mammospheres formation (Fig. 7A) accompanied by a remarkable induction of ALDH activity, equivalent to those achieved through expression of miR-200's (Fig. 7B). However, MCF10CA1h shZEB2 cells are mainly $CD44^{high}$ $CD24^{low}$ (Fig. 7C). Protein levels of the mesenchymal markers vimentin and fibronectin 1 did not vary and neither did E-cadherin, which we expected to be increased (Fig. 8A). Confocal imaging of E-cadherin confirms that ZEB2 silenced and control cells lack E-cadherin membrane staining while it is clearly induced in MCF10CA1h miR-200 C2 cells (Fig. 8B).

As has been noted, the acquisition of mammary luminal progenitor properties induced by miR-200's in MCF10CA1h cells is in part driven or reinforced by downregulation of ZEB2 and activation of the PI3K/Akt signaling pathway.

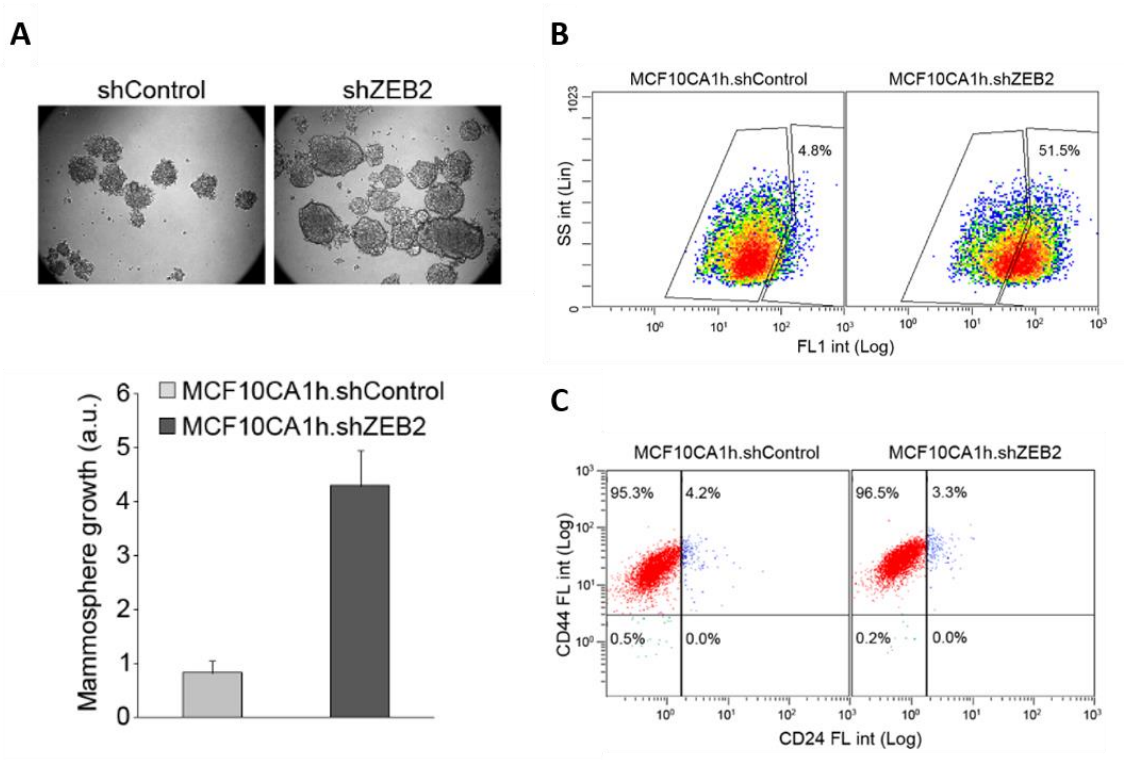

Figure 7. Characterization of stem-like traits in knockdown MCF10CA1h cells of ZEB2. (A) Knockdown of ZEB2 potentiates the mammosphere growth of MCF10CA1h cells . Bright field images of mammospheres (top) and graphic showing primary mammospheres quantified as in figure 5A (bottom). (B) Knockdown of ZEB2 induces ALDH activity in MCF10CA1h cells, determined as in figure 5B. (C) Knockdown of ZEB2 fails to upregulate cell-surface CD24 in MCF10CA1h cells, determined as in figure 5C.

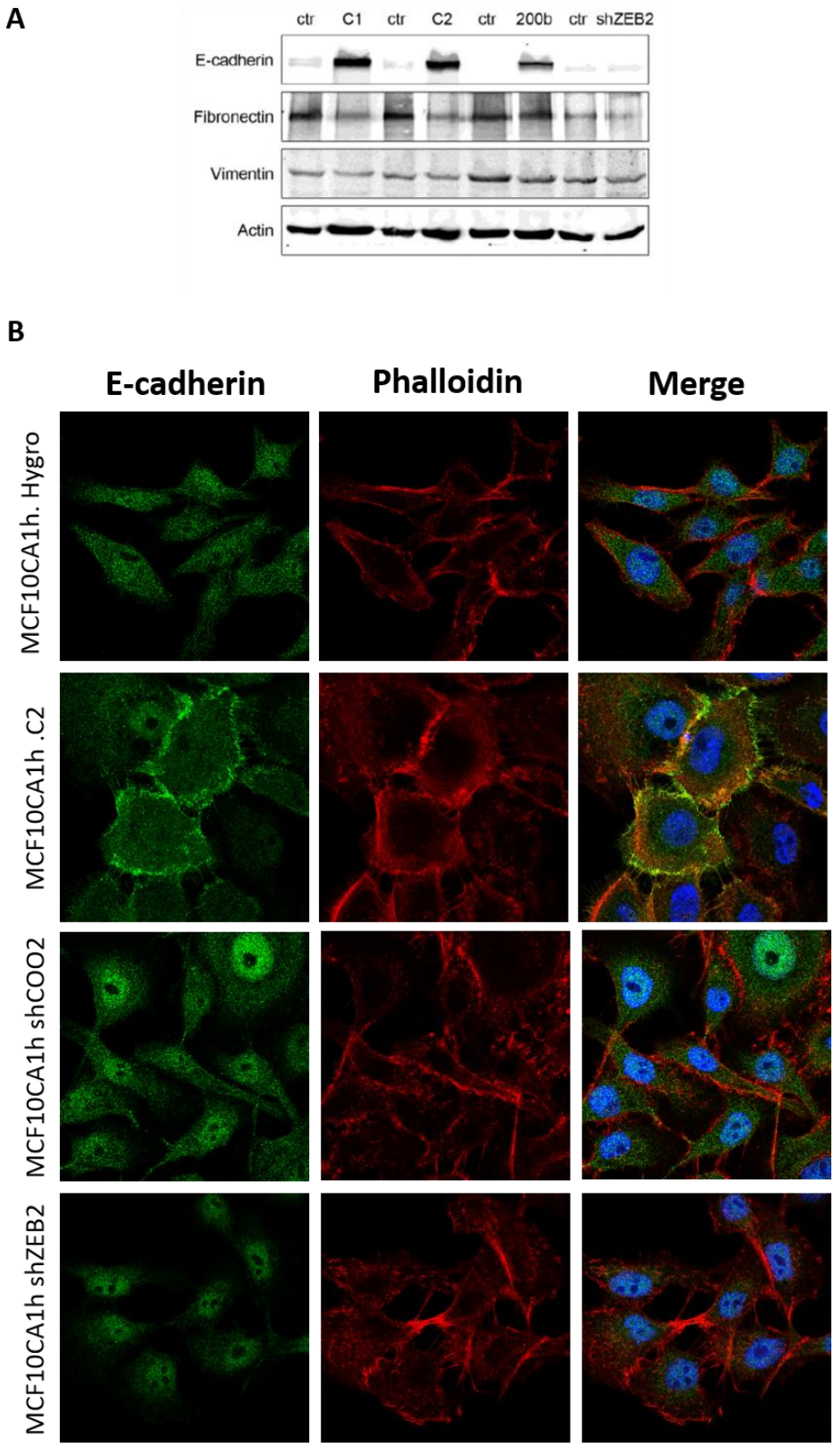

Figure 8. Knockdown of ZEB2 fails to upregulate E-cadherin. (A) Western blotting of total extracts of cells probed with anti-E-cadherin, anti-fibronectin and anti-vimentin, normalized to β-tubulin. (B) Confocal microscopy images showing expression of E-cadherin, phalloidin and both by immunofluoescence in MCF10CA1h miR-200 C1, C2 and their respective controls.

Differential growth and lung colonization by miR-200 C1 *in vivo*

As previously mentioned, the *in vitro* spheroid growth assay is a surrogate test for assessing the self-renewal potential of tumor cells and also a good predictor of the tumorigenic and/or metastatic potentials of these cells in immunodeficient mice. Considering that the miR-200 family considerably increases the formation of mammospheres *in vitro*, we wanted to test the *in vivo* tumorogenicity of miR-200-expressing MCF10CA1h cells.

Control and miR-200 cluster 1 overexpressing cells were modified in order to stably integrate in their genomes the firefly luciferase gene under the transcriptional control of a constitutive promoter and selected on the basis of GFP expression. Subsequently, these cells were implanted in cleared mammary fat pads of NOD-SCID mice and tumor growth was monitored at real time by bioluminescence (Fig. 9A). Both control and miR-200-expressing MCF10CA1h cells grew at robust rates for the first 7 days after implantation, after which the growth of control cells stagnated while that of MCF10CA1h-miR200 cells continued at an exponential rate (Fig 9A). MCF10CA1h-miR200 cells also showed a significantly higher proliferation index than control cells, as determined by Ki67 staining of these tumors (Fig. 14). Neither control nor miR200-expressing MCF10CA1h cells produced detectable metastatic growth outside of their sites of implantation for the duration of local growth monitoring. After removal of the implanted tumors, cells were additionally followed up for two months to explore the possibility that some cells had escaped from the orthotopic site, entered the systemic circulation and were able to colonize distant organs. However, neither control nor miR-200-expressing cells were detected into circulation or distant organs during this period.

Alternative models to characterize differences in the metastatic capacity of cells are experimental metastases models, in which cells are directly injected into the systemic circulation. We injected control and MCF10CA1h-miR200C1 cells expressing the firefly luciferase gene that could be monitored by bioluminescence (Fig. 9B). After 42 days, MCF10CA1h-miR200 cells resulted in tumor colonization of the lungs at significantly higher rates than control cells (Fig. 9B). Lungs were then removed, fixed, embedded in paraffin and sectioned for histopathological analysis. In concordance with the bioluminiscence measure, control mice showed a significant lower mean of metastases per case and a smaller size, whilst metastatic lungs from MCF10CA1h-miR200C1 injected mice, displayed a significantly higher number of metastases per case ranging from 1 to 3 mm with variable size (Fig. 9C).

These experimental *in vivo* assays demonstrate that the miR-200 cluster 1 enhances the tumorigenic and metastatic growth of MCF10CA1h cells. However, although miR-200-expressing cells possess an increased ability to establish and grow at a distant organ, they do not possess the ability to escape from the local tumor and enter the circulation, or at least not during the period we monitored. It can be inferred from the results that enhanced metastatic colonization and growth can be uncoupled from local escape functions of tumor cells.

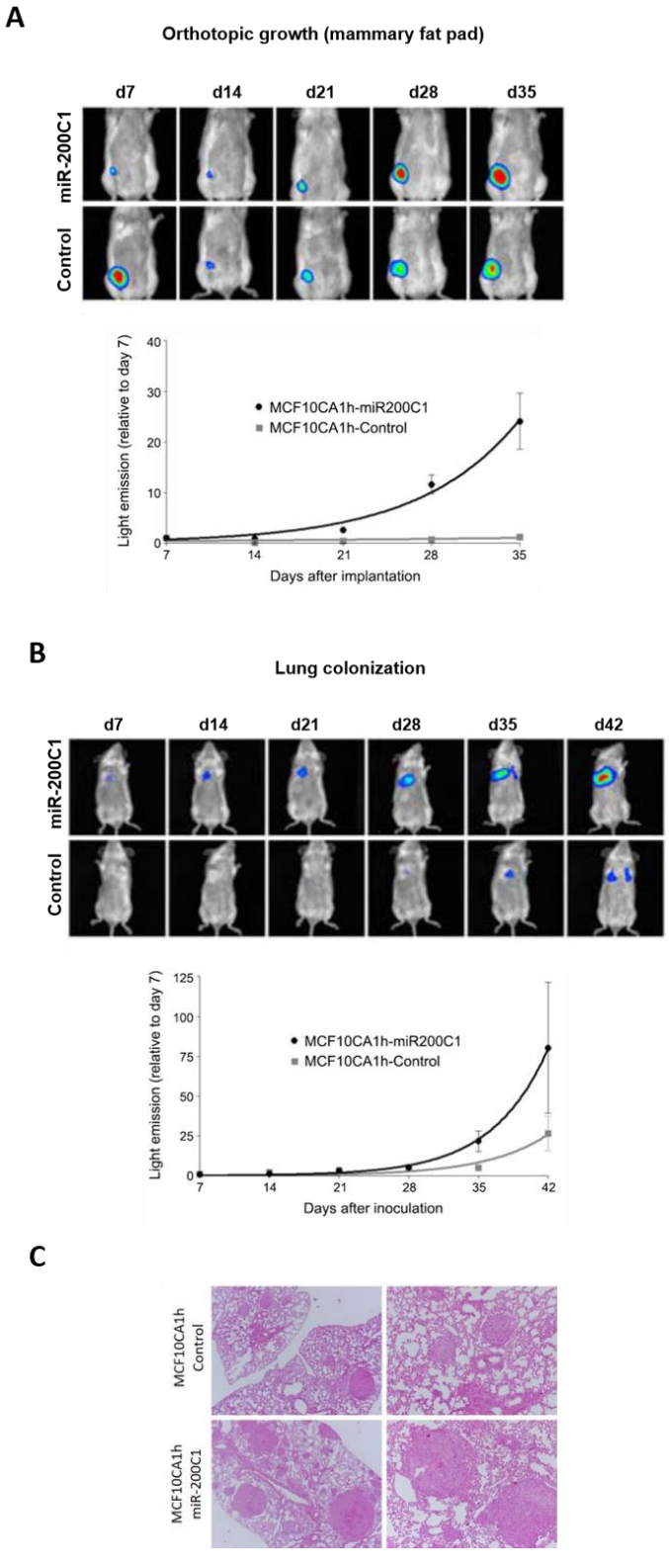

Figure 9. miR-200 promotes tumor growth and metastatic colonization of lungs. (A) Control and miR-200 C1-expressing cells were transduced for stably express firefly luciferase gene and injected in cleared mammary fat pads of NOD-SCID mice. Monitoring of growing tumors was done after injection of luciferin at indicated days and photon counts were quantified. (B) Same procedure described as in (A) but cells were injected intravenously. (C) Hematoxilyn and eosin staining of MCF10CA1h control and miR-200 C1-derived metastases from the experiment performed in (B).

Characterization of cell lineage markers in differentiating conditions

Regarding the fact that MCF10CA1h miR-200 cells possess stem-like or luminal progenitor features, we cultured them in conditions mimicking the extracellular matrix. In order to do so, we selected those cells with higher self-renewal and stem-like potential by firstly growing cells in anchorage-independent conditions until mammospheres were formed. These mammospheres were then disaggregated and grown in 5% Matrigel 3-D lattices for 10 to 14 days (Fig. 10).

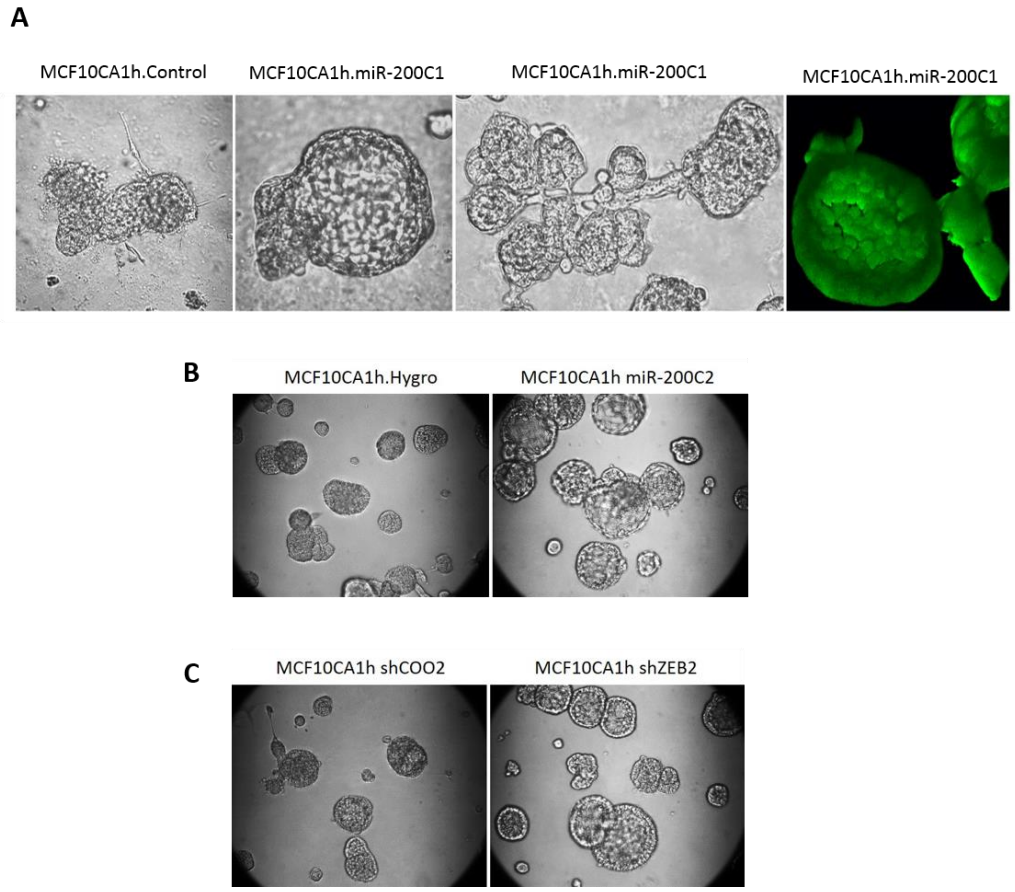

Figure 10. miR-200f promotes the morphogenesis and differentiation of MCF10CA1h cells, partially mediated by ZEB2 silencing. (a) miR-200f promotes the formation of pseudoalveolar (5-d culture) and complex tubuloalveolar (10-d culture) structures by MCF10CA1h cells grown in 5% Matrigel. The pseudoalveolar structures are hollow cavities formed by a single-cell layer as shown in 3-D reconstructions for GFP-expressing MCF10CA1.200b cells (right panel). (B) Culture of MCF10CA1h control and miR-200 C2-expressing cells in Matrigel for 10 days. (C) Culture of control and knockdown of ZEB2 cells in Matrigel for 10 days.

MCF10CA1h cells expressing the miR-200 family, formed highly organized branched tubular structures with multiple terminal hollow spheres morphologically similar to alveoli (Fig. 10A). These structures were reminiscent of the complex structures that explanted normal mammary epithelial progenitor cells form *in vitro* (374). In contrast, control MCF10CA1h cells only formed amorphous or spherical structures with no evidence of tubulogenesis, branching or alveologenesis under these conditions (Fig. 10A). ZEB2 knocked down cells were also able to form the spherical structures, though not the tubular ones (Fig. 10C). In order to visualize what the spatial position of cells expressing miR-200 was, confocal microscopy of MCF10CA1h-miR200b cells grown in Matrigel was performed. As they co-express green fluorescent protein, we were able to see that these cells contributed equally to ductal structures and to alveoli and that the latter were hollow spherical structures lined by a single cell layer (Fig. 10A). Control cells also co-expressing green fluorescent protein, formed the same amorphous mass of cells as control cells with the puromicin or hygromicin empty vector (Fig. 10A-C).

We have previously observed that miR-200 promotes the acquisition by MCF10CA1h cells of an epithelial phenotype with self-renewing stem-like ability. In addition, these results suggest that, when grown in an appropriate environment, they further acquire luminal traits that allow them to organize into alveoli-like structures with a lumen.

In view of the results, we studied the expression of several markers indicative of a luminal or basal phenotype when cells are grown in the three-dimensional Matrigel culture that stimulates their differentiation (Fig. 11, 12 and 13).

As expected, both miR-200 cluster 1 and cluster 2 extremely induced the expression of E-cadherin at levels comparable to those seen in invasive ductal carcinomas (Fig. 11, 12). This was accompanied by a discreet increase of luminal keratins in cells overexpressing miR-200 cluster 1 but curiously, not in those expressing miR-200 cluster 2 (Fig. 11 and 12). Strikingly, both basal markers cytokeratin 5 and p63 showed an abrupt induction. While p63 exhibited a diffused expression, KRT5 was expressed at varying intensities, being intensely expressed even in cells with a luminal position (Fig. 11, 12). The other marker characteristic of a basal phenotype, keratin 14, was also induced in both clusters, separately (Fig. 13). Consistent with the induction of an epithelial phenotype, vimentin expression was reduced to approximately 60% of cells overexpressing C1 or C2 at a moderate intensity compared to their respective controls, in which all cells expressed vimentin at the maximum intensity (Fig. 13).

In spite of the morphological maturation in 3-D Matrigel cultures of tubuloalveolar structures with near-physiological appearance, these cells failed to express the luminal differentiation markers GATA3 (only discreetly in miR-200 C2 cells) or estrogen receptor (ESR1) nor they did express smooth muscle myosin heavy chain (SMMHC), a marker of mature myoepithelial cell differentiation (Fig. 13). Expression of SOX2 was also negative both in control and cells overexpressing miR-200f, suggesting that it did not had any significant role in the acquisiton of stem-like traits (Fig. 13).

Therefore, expression of miR-200 confers MCF10CA1h cells with the ability to form complex tubuloalveolar structures upon culture in 3-D Matrigel lattices, but falls short of engaging luminal or myoepithelial terminal differentiation gene programs in these cells.

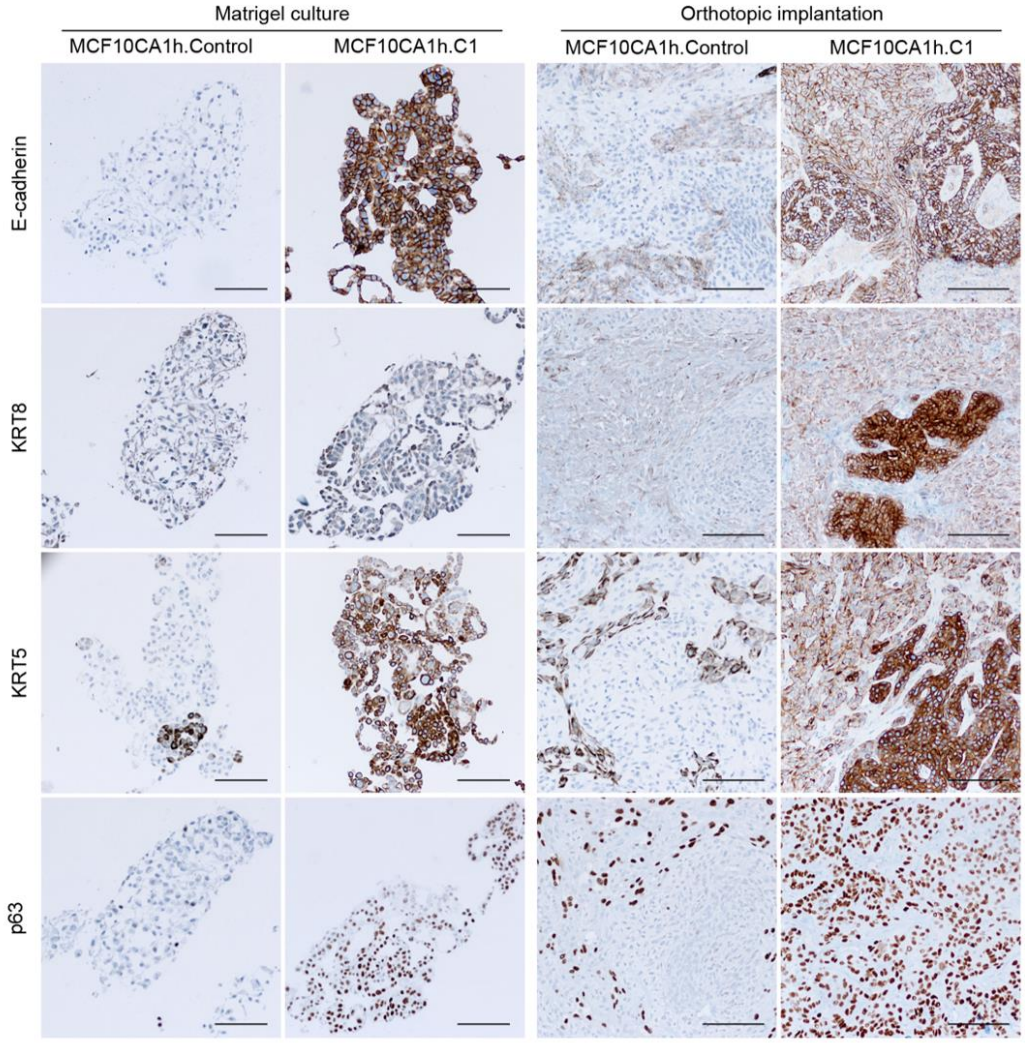

Figure 11. miR-200 C1 induce an epithelial program (E-cadherin) and the concomitant upregulation of luminal (KRT8) and basal (KRT5) keratins in vitro (5% Matrigel 3D culture) and in vivo (orthotopic implantation in cleared mammary fat pads of NOD-SCID mice). Cells- and tumors were formalin fixed and paraffin embedded and processed for immunostaining and diaminobenzidine-based detection.

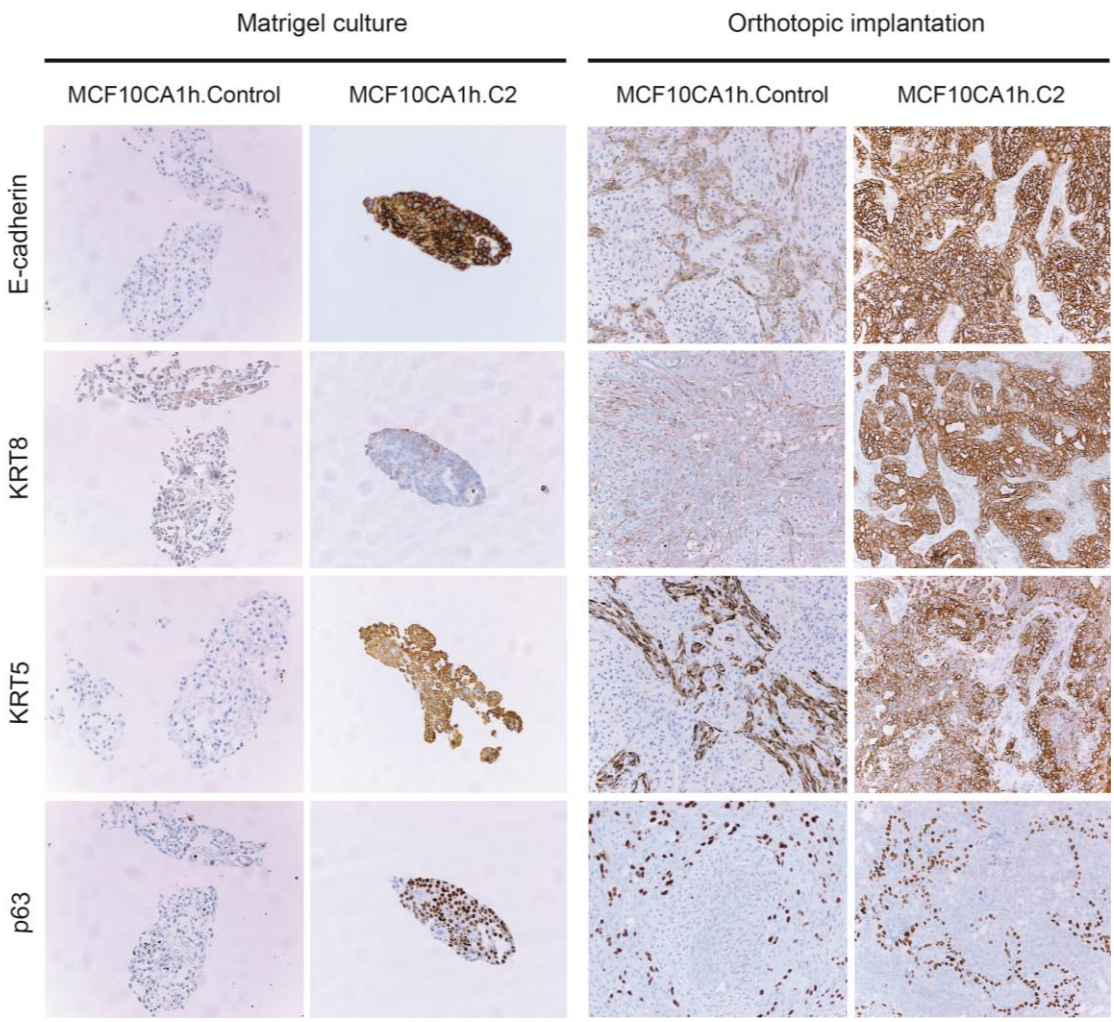

Figure 12. miR-200 C2 induce an epithelial program (E-cadherin) and the concomitant upregulation of luminal (KRT8) and basal (KRT5) keratins in vitro (5% Matrigel 3D culture) and in vivo (orthotopic implantation in cleared mammary fat pads of NOD-SCID mice). Cells- and tumors were formalin fixed and paraffin embedded and processed for immunostaining and diaminobenzidine-based detection.

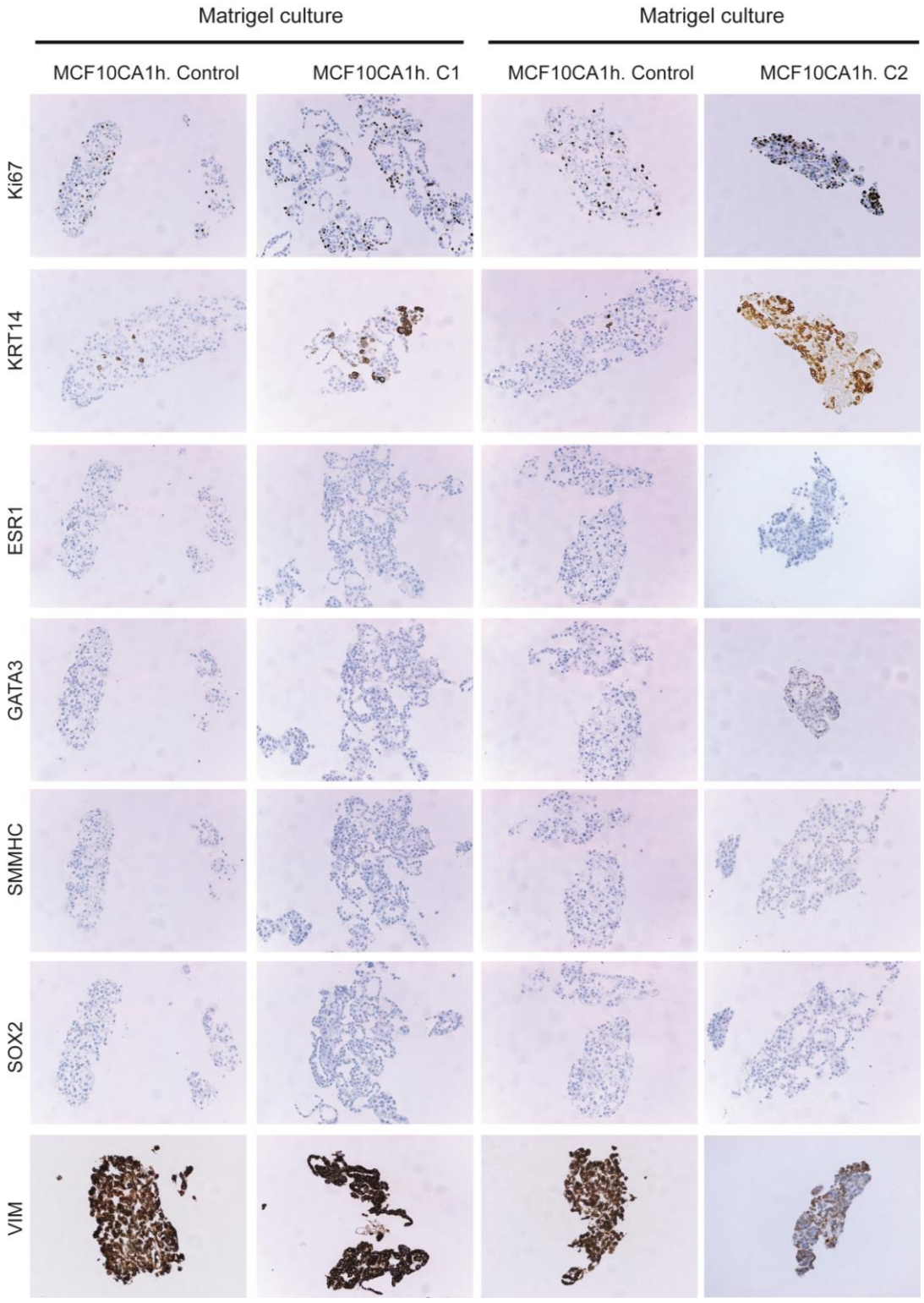

Figure 13. miR-200f does not induce other markers of a luminal (GATA3, ESR1) or myoepithelial terminal differentiation (SMMHC) in differentiating culture conditions. Cells were formalin fixed and paraffin embedded and processed for immunostaining and diaminobenzidine-based detection.

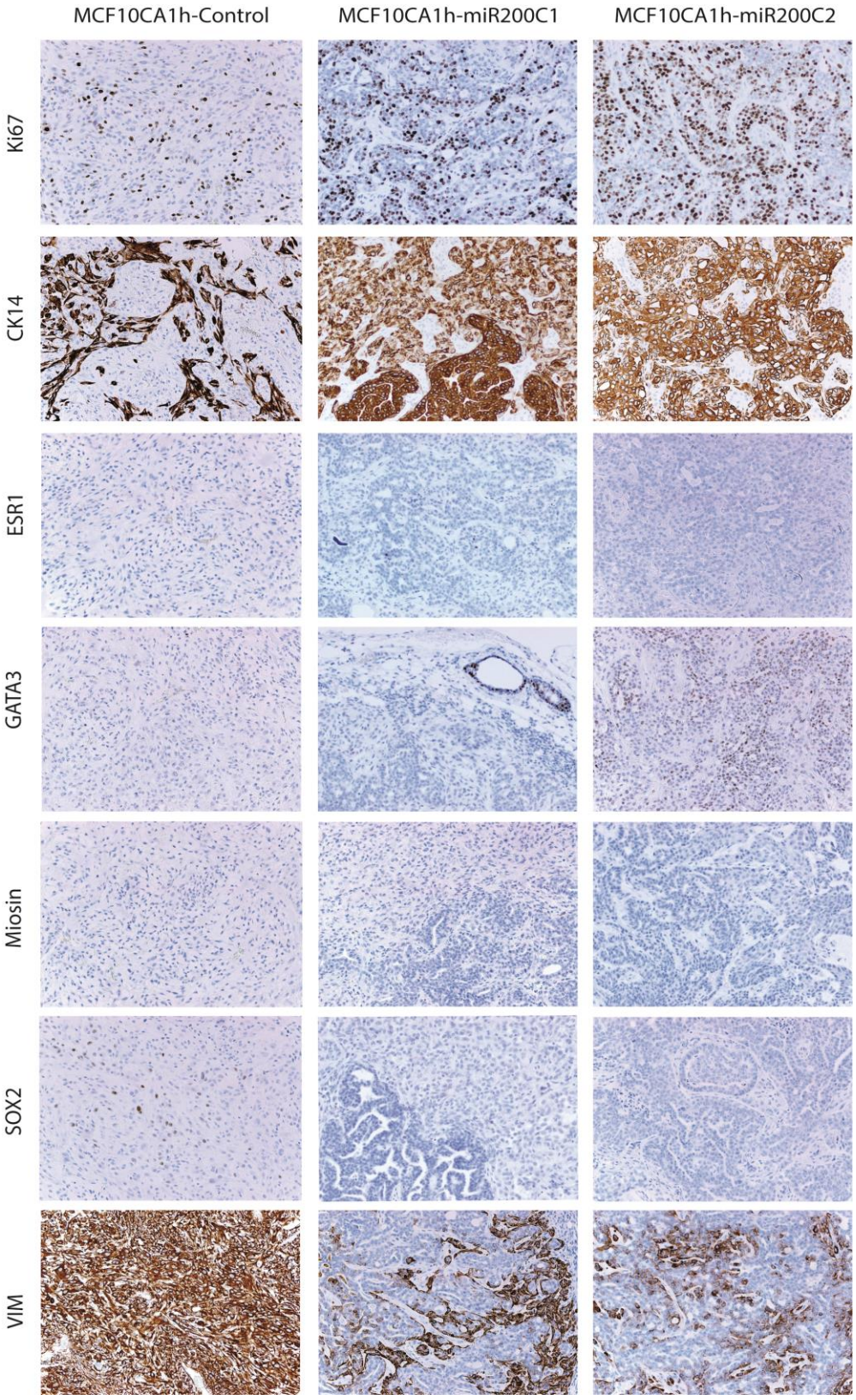

Figure 14. miR-200f does not induce other markers of a luminal (GATA3, ESR1) or myoepithelial terminal differentiation (SMMHC) in orthotopic tumors. Orthotopic tumors were formalin fixed and paraffin embedded and processed for immunostaining and diaminobenzidine-based detection

Characterization of the phenotype of miR-200 cells in orthotopic tumors

Tumors generated in control mice showed morphologically heterogeneous areas which included predominantly spindle-shaped and mesenchymal-like elements in some parts together with areas with more epithelioid appearance (Fig. 11, 12 and 14). However, they failed to form glandular structures. In contrast, this double component was more evident in tumours formed in mice by MCF10CA1h-miR200 cells, which contained larger areas of epithelioid appearance, displayed morphological differentiation into gland-like structures adjacent to spindle-cells sarcoma-like mesenchymal-like areas and areas of morphological transition (Fig. 11, 12 and 14).

We assessed the expression of the same phenotypical markers evaluated in the 3D culture. Both MCF10CA1h-miR200 cluster 1 and cluster 2 expressing cells induced a strong E-cadherin expression, which was more intense in the differentiated glandular areas of MCF10CA1h miR-200 C1 tumors (Fig. 11, 12). This difference was more evident when staining for the luminal cytokeratin 8, which shows the same expression pattern as E-cadherin. miR-200 cluster 1-expressing cells showed higher cytokeratin 8 expression intensity in accordance with the more glandular differentiated areas than cluster 2-expressing cells, which displayed a more homogeneous expression of both E-cadherin and cytokeratin 8 (Fig. 11, 12). Tumours generated by MCF10CA1h-miR200 cells also showed more intense and diffuse expression the basal cytokeratins 5 and 14 in both the mesenchymal and epithelial-like areas when compared to control tumours, and such expressions were more evident in the differentiated epithelioid cells (Fig. 11, 12 and 14). p63 expression was also higher in miR-200 expressing tumors, where positive cells are periglandular, surrounding CK8 positive gland-like structures while in the non-glandular area, where epithelial cells were less differentiated, it was diffusedly expressed (Fig. 11, 12). Despite the expression of p63 and CK8 markers was mutually exclusive, some MCF10CA1h-miR200 tumor areas did coexpress CK5 and p63 as in control tumors.

Nonetheless, miR-200-expressing tumours did not show a more intense staining of the luminal differentiation markers GATA3 or estrogen receptor (Fig. 14). They also lacked expression of progesterone receptor and HER2 that together with the estrogen receptor negativity allowed to molecularly define both MCF10CA1h control and miR-200-derived tumours as triple-negative tumours. As in the matrigel cultures, neither control nor MCF10CA1h-miR200 tumours expressed heavy-chain smooth muscle myosin (Fig. 14). In contrast, MCF10CA1h control tumours displayed a strong expression of vimentin, comparable to that of mesenchymal cells of the tumor stroma, while its expression was drastically reduced in the glandular component of miR-200 expressing tumors (Fig. 14).

In summary, MCF10CA1h control cells form triple-negative tumors upon implantation. The miR-200 microRNA family is able to induce a double component formed by a mesenchymal-like area that express vimentin at a higher intensity than E-cadherin accompanied by a weak cytokeratin 8 expression and moderate expression of basal markers. The glandular component, on the other hand, shows strong E-cadherin expression while a complete lack of expression of the mesenchymal marker vimentin. These differentiated cells coexpress luminal (CK8) and basal (CK5, CK14) markers , both at higher intensity than in the mesenchymal component and control tumors. However, similar to what it is observed *in vitro*, they do not exhibit luminal or myoepithelial terminal differentiation gene programs.

Characterization of the phenotype of miR-200 cells in lung metastases

Besides assessing the expression of breast luminal and basal markers in orthotopic tumors, we also evaluated its immunohistochemical expression in metastatic lungs from the tumor colonization assay.

Whilst metastases from control mice are formed by well-defined nodiums with predominantly spindle-shaped cells, metastases from MCF10CA1h-miR200C1 injected mice have tumor nests with neoplastic cells that have in general a more epithelioid morphology (Fig. 15). They also show several morphologies including round and differentiated cells growing in sheets, small foci of squamous-like elements with white neophiosilic citoplasm and areas with glandular differentiation although they exhibit less luminal differentiation than their orthotopic counterparts (Fig. 15).

Comparable to what was observed in orthotopic tumors, 90% of the metastatic MCF10CA1h-miR200C1 cells exhibited intense staining for E-cadherin, cytokeratins 8 and 18, p63 and cytokeratin 5, whereas only 10-30% of metastatic control cells did express these markers (Fig. 15). Double staining for p63 and CK8/18 revealed that there were two types of MCF10CA1h-miR200 cluster 1 metastases: some with almost complete p63 expression and focal expression of CK8/18 and others with excluding populations of p63 and CK8/18 positive cells (Fig. 15). As CK8/18 expression in control metastatic cells is restricted, in general, there is not coexpression (Fig. 15).

It seems that metastases recapitulate the morphology and expression of luminal and basal markers in orthotopic tumours. Cells expressing miR-200 with high expression levels of epithelial, luminal and basal markers show a more efficient colonization ability which allows the establishment of a higher number and increased size of metastases. Nevertheless, miR-200 cluster 1 cells do not exhibit the same extent of differentiation in a distant organ, remarking the important role that the microenvironment stimuli play in cancer progression.

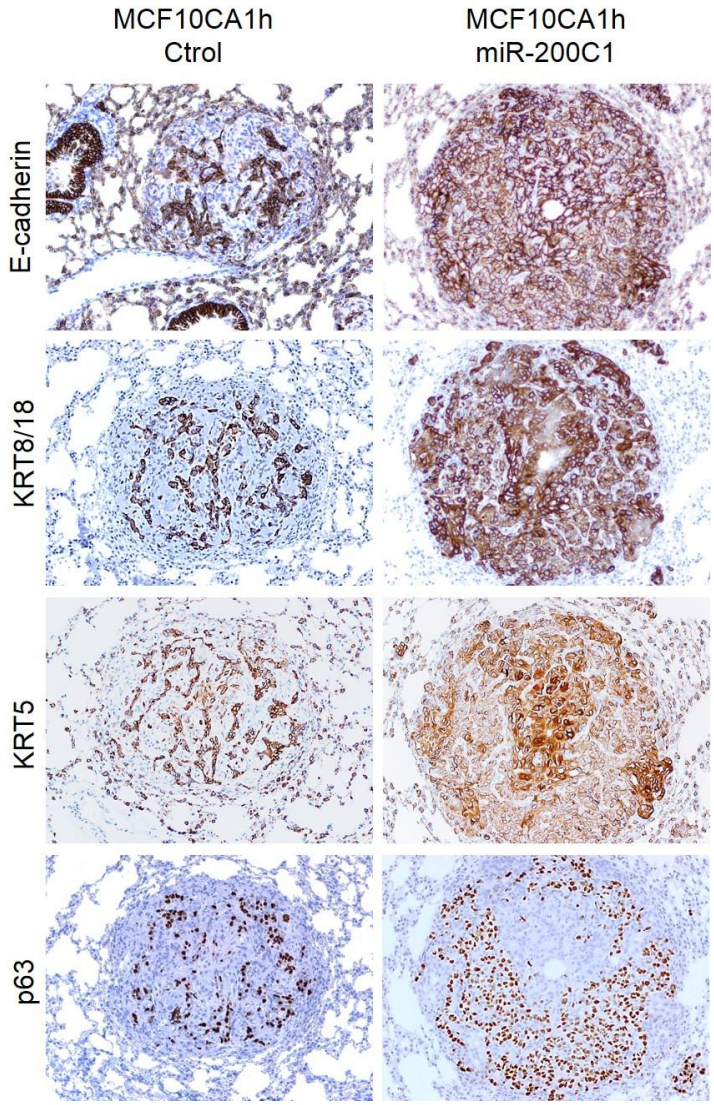

Figure 15. Metastases from MCF10CA1h control or miR-200 C1 cells recapitulate the expression of cell lineage and EMT markers observed in orhtotopic tumors. Lung metastases were formalin fixed and paraffin embedded and processed for immunostaining and diaminobenzidine-based detection.

The epithelial components of metaplastic cancers of the breast express high levels of miR-200 and marker profiles suggestive of luminal progenitor cells

As mentioned, the tumors grown in mice by MCF10CA1h-miR200 cells presented a complex morphological appearance which consisted of a double component with areas with epithelial traits that can form glandular structures, adjacent to areas of spindle cells with sarcomatous appearance, joined by areas of morphological transition, suggestive of the coexistence of cells displaying heterogeneous phenotypes (Fig. 11, 12 and 14). These morphological features are reminiscent of metaplastic carcinomas of the breast (MCB) with spindle cell component (375). Thus, MCB tumors harbor glandular and non-glandular (metaplastic) tumor components, the latter having been ascribed a mesenchymal-like phenotype presumably resulting from epithelial-mesenchymal transition affecting epithelial tumor cells (92, 376-378). Given these morphological similarities, we hypothesized that the epithelial/glandular components of human MCB would express higher levels of miR-200 and co-express luminal and basal markers. As this class of tumors encompasses several histological subtypes with likely underlying genetic and epigenetic differences (377), we restricted this analysis to the carcinosarcoma subtype, with well-delimited double epithelial and mesenchymal components as determined by mutually exclusive expression of cytokeratins and vimentin, respectively.

Having this in mind, we macrodissected the epithelial and mesenchymal components in 4 cases and separately quantified the expression of miR-200a, miR-200b, miR-200c, miR-429 and miR-141 by qPCR. The epithelial components in 3 out of the 4 tumors analyzed expressed significantly higher levels of at least one of the four miR-200 microRNAs than the mesenchymal components (Fig. 16A).

As expected, E-cadherin was consistently and diffusely expressed in the epithelial components of all cases except one, in which staining was limited to 10% of cells (Fig. 16B). Similarly, the luminal cytokeratin CK8 was expressed in the epithelial component of all five carcinosarcoma MCB cases analyzed. Remarkably, the same cellular areas co-expressed the basal cytokeratin CK5 in all but one case (Fig. 16B). The breast luminal progenitor marker ALDH and the basal marker p63 were also expressed in all cases, both in the mesenchymal and the epithelial components, with variable intensities in each morphological component (Fig. 16B). The luminal differentiation markers estrogen receptor and GATA3 were undetectable in these samples, as was HER2.

Taken together, these observations suggest that the epithelial components of MCB with double mesenchymal and epithelioid components (carcinosarcomas) express markers of breast luminal progenitor cells, including ALDH and double luminal-basal cytokeratins, coincident with high levels of miR-200s. Because metastatic tumors associated with MCB of the carcinosarcoma type invariably display epithelial morphologies and in light of the higher metastatic potential conferred to breast epithelial cells by an epithelial program engaged by miR200, we suggest that the highly malignant and metastatic behavior of double-component MCB may reside primarily in their epithelial components.

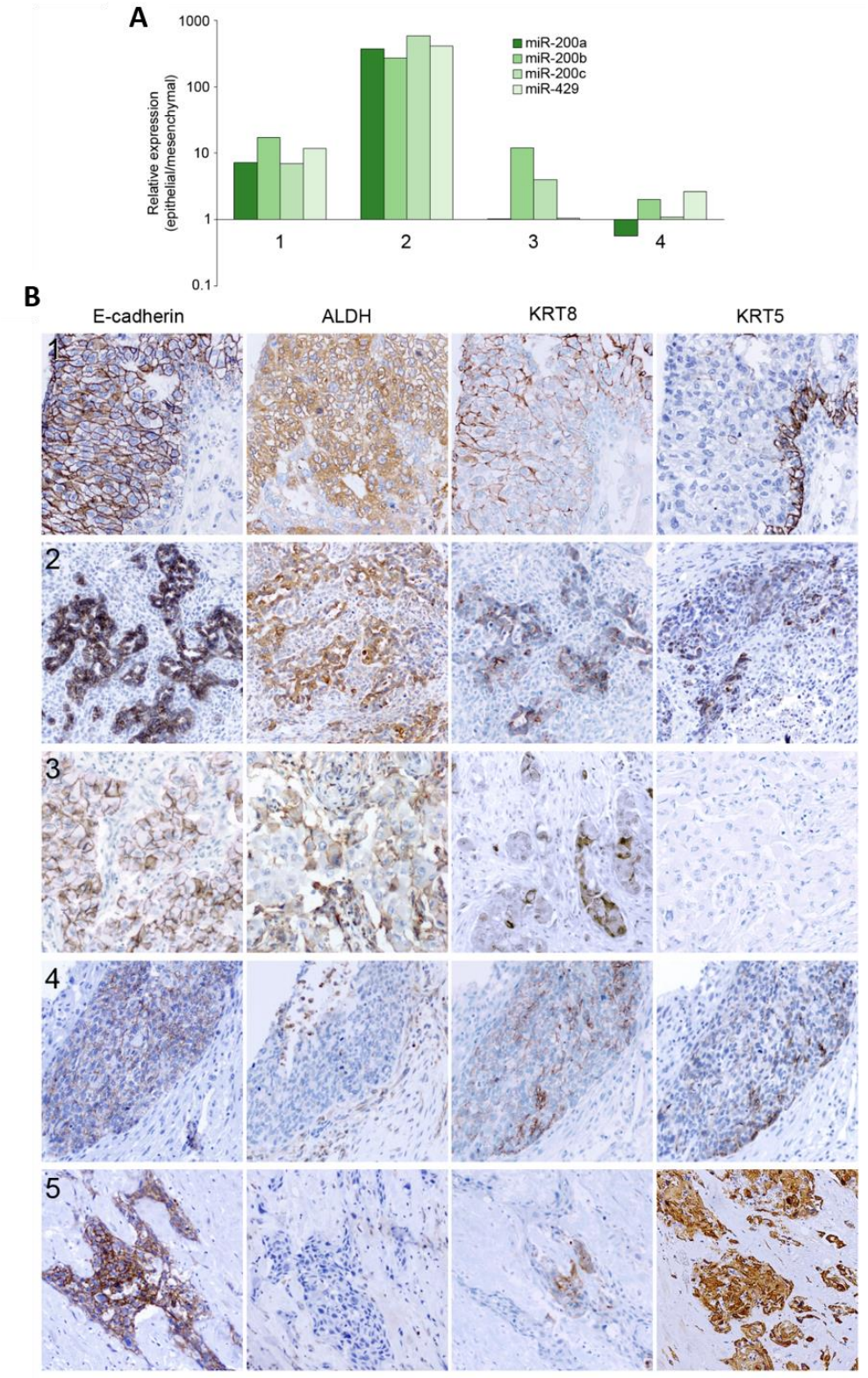

Figure 16. The epithelial components of metaplastic breast cancer can express high levels of miR-200s and markers of luminal progenitor cells. (A) Expression levels of miR-200a, miR-200b, miR-200c and miR-429 in enriched epithelial components relative to mesenchymal components of MBC samples, determined by qPCR. (B) The epithelial components of MBC samples can co-express E-cadherin, KRT8, KRT5 and ALDH, determined by immunohistochemical analysis. Samples labeled 1-4 correspond to the samples with the same labels in (A).

Discussion

In this study, we have identified sets of miRNAs that are dysregulated along the metastatic progression of IDC, all of which have been previously reported in similar associations, including miRNAs significantly associated with metastatic progression in breast cancer and other tumor types (360). Of a selection of miRNAs identified as differentially expressed in metastatic vs. non-metastatic IDC tissues, two, namely miR-200b and miR-7, were found to circulate at higher levels in blood from patients with metastasis than patients with no overt metastasis at the time of diagnosis, in concordance with the trend found in tissue samples along metastatic progression. Focusing on the biological significance of the observed higher levels of expression of miR-200 family member in association with metastatic progression, we have discovered that these microRNAs induce in MCF10CA1h cells a phenotype consistent with a switch to EpCAM+, ALDH+ luminal progenitor cells, which may provide a mechanistic explanation for the association of high levels of miR-200's with metastasis and a tendency of primary tumors to metastasize. What follows is a more detailed discussion of the most relevant observations in this study.

MicroRNAs differentially expressed in node-positive vs. node-negative samples

Four microRNA were significantly deregulated between non-metastatic and lymph node positive patients: miR-7, miR-30a, miR-148a and let-7 members.

Besides the well known tumor suppressor role of the let-7 family, miR-30a and miR-148a have been mainly reported as tumor suppressors and all of them were donwregulated in non-metastatic tumors. A downregulated expression of miR-30a in other breast cancer specimes has been reported. A role as an inhibitor of breast cancer cell proliferation and migration has been shown by targeting eye absent 2 (Eya2) (379) ubiquitin Protein ligase E3C (UBE3C) or vimentin (VIM) (380) It may also play an important role as an inhibitor of metastasis, as it does by targeting Slug and promoting MET, which could explain our observations (381). In fact, a negative correlation between miR-30a levels and the extent of lymph node and lung metastasis was found in patients (382) consistent with the downregulation we observed in lymph node positive compared to lymph node negative patients. It is also an inhibitor of invasiveness and metastasis in other cancers (383-386).

A downregulated expression of mir-148a has also been found in various cancers including colorectal (387), pancreatic (388) or esophageal (389). Moreover, circulating miR-148 levels were downregulated in non-small cell lung cancer patients linked to tumor size and lymph node

metastasis (390) and this association was also observed in colorectal cancers (391, 392). An inverse correlation with tumor grade and also with lymph node metastasis was reported in breast cancer (393).

Ioro and colleagues found an inverse association of the tumor suppressor let-7 with lymph node metastasis in a breast cancer study (293). Besides let-7 downregulation, they also reported downregulated levels of miR-125, miR-145 and miR-10b and upregulation of miR-21 and miR-210, which are among the top deregulated microRNAs in our study as it is seen in figure 1.

On the other hand, miR-7 levels were higher in distant metastatic than in node-positive primary tumors (PM) and higher levels in PM than in node-negative primary tumors and circulating miR-7 levels followed the same trend, being statistically significant in tumors and patients. Based on its major target mRNAs coding for components of the EGFR, Ras and MAPK signaling pathways, miR-7 has been assigned a tumor suppressor funcion . However, the opposite effect has also been reported, depending on the tumor type . Interestingly, miR-7 has a similar behaviour regarding EMT as miR-200, which shows the same trend in blood and tumors as miR-7. miR-7 indirectly upregulates E-cadherin by targeting insulin-like growth factor 1 receptor (55), focal adhesion kinase (59) or SETDB1 (60), resulting in reduced epithelial-to-mesenchymal transition and enhanced tumor keratinization (58, 61), a poor prognostic factor in several types of tumors (62), along with enhanced proliferation, anchorage-independent growth and tumorogenesis in immunodeficient mice (58). The variable outcomes of miR-7 expression may be related to its availability and functionality, which is regulated by circular RNAs that act as “sponges” or inhibitors (63-65).

microRNAs in distant metastases

We furher observed an increasing tendency of miR-210, miR-181a/b from non-cancerous to distant metastatic samples while decreasing levels of miR-10b and miR-101. In consonance with our observations, miR-101 has been described as an inhibitor of migration and invasion in bladder (394) and hepatocarcinoma cancer cells (395), the loci of which is lost in both localized and metastatatic prostate cancer (396). Though we did not obtain significant differences in its blood levels in metastatic *vs* non-metastatic patients, a downregulation of miR-101 levels in plasma may be a predictor of worse disease-free survival in hepatocellular carcinoma patients (397). Similarly, an inverse correlation with the clinical stage, lymph node metastasis and prognosis has been described in breast cancer patients (398). On the contrary, expression of

miR-10b correlates with high-grade and metastatic behaviour in various cancers, including breast cancer (306, 347, 348). It is higly expressed in metastatic breast cancer cell lines and promotes invasion and metastasis *in vivo* of otherwise non-metastatic cells (347) Besides, its overexpression induces self-renewal and expression of stemness and EMT markers, such as Twist (347, 399). However, these studies find higher levels of miR-10b in basal-like cells with a triple negative profile than in cells with a luminal phenotype that express E-cadherin, such as MCF-7. No studies with a positive association between miR-10b and metastatic progression in breast tumors with a molecular luminal profile have been published so far, which may explain why we observe a downregulation. In fact, only Ioro and colleagues reported a downregulation of miR-10b in primary tumors compared to normal breast samples (293). In addition, we were not able to detect circulating miR-10b in blood.

On the other hand, several reports show that elevated miR-181 family members are associated with worse phenotype and prognosis in breast cancer (400-402). miR-181 has been established as a metastamiR, regulated by TGF-β, that promotes metastasis by inducing EMT, migration and invasion of breast cancer cells (400-402). It also favors the expansion of breast cancer stem-like cells by targeting ataxia telangiectasia mutated (ATM) kinase (400-402) and elevated levels have also been observed in EpCAM-positive hepatic cancer stem cells (403). Moreover, increased miR-181 expression has been seen in MCF-7 cells resistant to tamoxifen, used in the treatment of ER+ breast cancers. A tumor-promoting role for miR-181 has also been observed in other cancers (404-407) as well as for miR-210. Several meta-analysis including studies on different carcinomas sistematically report high miR-210 expression to be a predictor for poor overall survival (OS) and also metastasis free survival or distant relapse free survival (MFS/DRFS), and disease specific survival (DSS); especially in breast cancer (408-410). It is not only associated with a worse outcome in the most aggressive phenotypes but also in ER+ breast cancer tumors (411, 412). It has been consistenly reported that expression of miR-210 is induced under hypoxia (413-415), which is an adverse prognostic factor in many tumor types (416). The established link between miR-210 and poor outcome is not suprising considering that it promotes angiogenesis by modulating VEGF levels through direct targeting of ephrin-A3 (EFNA3) (417-419)and also fosters invasion and metastasis (420-424). If in our study, induction of miR-210 reflects an hypoxic environment that promotes progression of tumor through advanced stages, it is not surprising to find the highest expression in distant metastases, which have been subjected to conventional treatments, since hypoxic cells are notorious for their enhanced resistance to chemoterapy and radiotherapy (416). Furthermore, miR-210 has been reported to have an antiapoptotic and cytoprotective effect on cancer cells (378, 425-427). Several studies analyzing

microRNAs in serum or plasma in various cancers have reported elevated miR-210 levels in metastatic patients or as a prognostic factor associated with worse outcome (428-430). In sharp contrast, we have observed that circulating levels of miR-210, as well as miR-181 levels, are significantly downregulated in the blood from metastatic breast cancer patients compared to non-metastatic or node-positive metastatic patients. However, if we do not consider its relative expression but rather we take into account the -ΔCt parameter, both miR-181 and miR-210 are among the highest expressed microRNAs in blood. Moreover, these studies tend to compare the expression of miR-210 to a healthy control group, thus obtaining high circulating levels in metastatic patients. One possible explanation for observing the lowest expression in our group of metastatic breast cancer (MBC) patients, may be that these patients have undergone surgical resection of the primary tumor, which could be the main source of delivery of microRNAs. Another possibility could be that the levels of miR-210 are reflecting a good response for chemotherapy in the majority of MBC patients, as it has been reported in other studies (428, 431). In a clinical trial with breast cancer patients under trastuzumab treatment, plasma levels of miR-210 were higher in patients before than after surgery and in those with lymph node metastatis. Also, circulating levels were higher in patients who had residual disease compared to those who achieved a complete reponse (431).

miR-200 and breast cancer progression

In our exploration, a set of 9 microRNAs were overexpressed in lymph node metastasis relative to their matched primary tumors along with 8 other microRNAs. Some of these miRNAs are known to be involved in immune system regulation: miR-150 (432), miR-146a (433), miR-142-3p and miR-142-5p (434).

Another relevant microRNA family from the microarray profiling was the miR-200 family members, specifically the miR-200b/a/429 cluster. More than 95% of lymph node metastases either maintained or had increased levels of miR-200 cluster 1 members along with E-cadherin expression, supporting observations from other groups (341). Though distant metastases showed greater variability regarding miR-200 levels, there was a non-significant trend towards increased expression compared to primary tumors. Moreover, approximately 70% of samples showed a moderate to intense E-cadherin staining. Similarly, other laboratories have reported increased miR-200b expression in distant metastases and stronger or equal E-cadherin staining intensity compared to its corresponding primary tumors (435). Therefore, at least in ductal

carcinomas with a luminal phenotype, both miR-200 and E-cadherin are expressed at considerable levels and may be required for metastatic outgrowth. Korpal and colleagues also reported higher expression in lymph node metastases and showed that colonization is strongly enhanced by re-expression of miR-200s and subsequent epithelial differentiation (341). As we have observed, in some cases, metastases form epithelial cancers can show a greater degree of cellular differentiation than the corresponding primary tumor.

In addition, miR-200b, which we selected as the representative candidate of the cluster, significantly discriminated primary metastatic from non-metastatic tumors along with miR-7 and miR-210, for which we have argued to be promoters of tumor progression and poor prognosis factors. Moreover, circulating levels of miR-200b but not miR-200c tended to increase from non-metastatic to lymph node metastatic patients and were significantly upregulated in distant metastases. Interestingly, one laboratory has reported that miR-200a, miR-200b, miR-200c and miR-141 showed increased levels in the plasma of metastatic breast cancer patients and were useful predictors of overall survival and progression-free survival combined with four other microRNAs (436). Furthermore, this combination of miRNAs had increased levels in MBC patients with circulating tumor cells and miR-200b alone was sufficient to discriminate CTC-positive fom CTC-negative metastatic patients (Madhavan). Given that the percentage of CTCs in blood is low (437), it may explain the relatively low expression of miR-200b that we are able to detect, possibly reflecting the epithelial nature of CTCs. More recently, the same group has published a prognostic panel of sixteen microRNAs that includes the five members of the miR-200f, associated with overall survival. Again, miR-200a, miR-200b and miR-200c were significantly elevated in patients who later developed metastasis within 2 years; therefore, they are early indicators of metastasis (436) Elevated miR-200 levels have also been associated with prognosis of prostate cancer (438) and with the development of distant metastases in colorectal cancer patients (439).

In additional support of a prometastatic function of the miR-200 family, we have found miR-200b to be the most dysregulated microRNA at significantly higher levels in several cell lines exhibiting metastatic potential compared to its non/weakly metastatic counterparts. Given the pronounced difference in miR-200b expression levels and the fact that the malignant MCF10CA1 derivatives have a common genetic origin and that the epithelial MCF10CA1a cells are the ones displaying the ability to metastasize (370), we selected this model to further explore the relationship between miR-200s, the epithelial-mesenchymal transition and the metastatic process. As mentioned in the introduction, it has been reported that enforced expression of EMT

transcription factors in breast cells confers them stem-like features and favors their migratory behaviour at the invasive front and the initiation of metastasis (31). However, it has also been proposed that metastatic stem cells must reacquire an epithelial phenotype to expand into metastatic outgrowth (172, 232).

Luminal progenitor phenotype?

Since MCF10CA1h cells clearly gained epithelial characteristics upon overexpression of miR-200's, we perfomed mammospherse assays to test the capacity of the cells to form spherical colonies *in vitro* under anchorage-independent conditions, which is a feature of normal and malignant stem cells (363). Additionally, we performed two other approaches classically used to identify and isolate breast cancer stem cells: determination of the expression of the cell surface markers CD24 and CD44 and expression of the enzyme ALDH by means of the ALDEFLUOR assay. Al-Hajj *et al* were the first ones to report that human breast cancer cells from patients with tumor-initiating ability when implanted in NOD/SCID mice had a $CD44^{pos}$ $CD24^{neg}$ cell surface profile (125). Later on, it was reported that normal and malignant breast stem cells that express ALDH are also tumorigenic (136). Afterwards, CD44, CD24 and ALDH expression has been detected in CSCs from lung, colon, ovary, pancreas or head and neck carcinomas (440-444).

We have observed that, while less than 20% of MCF10CA1h cells or control cells expressed ALDH, overexpression of miR-200 C1 or C2 induced ALDH expression in about 50% of the population. Nevertheless, ALDH expression did not correlate with the classical profile of stem cell, as ~90% of MCF10CA1h control cells exhibited a $CD44^{pos}$ $CD24^{neg}$ immunophenotype. The discrepancy may reflect that this CSC markers identify distinct types of CSC. Immunofluorescence of these CSC markers in primary human breast tumors revealed that they do not need to necessarily overlap: CD44+ CD24- cells were located at the tumor invasive edge in contact with the stroma while ALDH+ cells located more centrally (445, 446). Nevertheless, both populations share the expression of genes with that of cells forming tumorspheres (446). Since only a relatively low number of MCF10CA1h cells are able to form mammospheres, it could be attributed to the ALDH+ fraction or to cells expressing both CD44+CD24- and ALDH+, which have been found to be the cells displaying the greatest tumor-initiating ability (136). Curiously, overexpression of miR-200 induced a shift towards a CD44+CD24+ immunophenotype in approximately 90% of cells. Considering the elevated percentage of ALDH positive cells, it is plausible to think that a greater fraction of the population overexpressing miR-200 expresses ALDH, CD44 and CD24. It

has been reported that independently of the CD24 status, CD44neg cells lack self-renewal potential (137) and since both MCF10CA1h and MCF10CA1h-miR-200 express CD44, at least part of the increase in the tumorigenic potential observed in the formation of mammospheres may be attributed to the expression alone or in combination with ALDH. Regarding CD24 positivity, it has been frequently associated with a terminally differentiated, luminal phenotype (373, 447, 448).

Importantly, analyses on mammary epithelial subsets demonstrated that in human breast tissue, ALDH activity is a feature of committed luminal progenitor cells whereas the stem cell populations showed extremely low activity (13, 18). Additionaly, the fact that only MCF10CA1h cells overexpressing miR-200 were able to form imperfect tubuloalveolar structures rather similar to that of the mammary gland, though monolayered, encouraged us to think that miR-200 is inducing a luminal progenitor phenotype on the basal/mesenchymal MCF10CA1h cells. Similarly, others have described the existence of two CSC populations in the MCF10CA1a cell line: the non-invasive, epithelial-like CD44posCD24pos and the invasive, mesenchymal CD44posCD24neg (137). Specifically, these authors demonstrated that cell lines with a high percentage of CD24pos cells expressed luminal keratins while cell lines with a high percentage of CD24neg cells expressed basal keratins. Furthermore, an interconversion between CD44posCD24pos and CD44posCD24neg was observed, demonstrating the plasticity between phenotypes (137, 446).

Epithelial vs. mesenchymal gene programs as proliferative and metastasis drivers: plasticity

There is a strong evidence that BCSC exist in alternative epithelial-like and mesenchymal-like states, between which they transition in response to the tumor microenvironment (446). Our results are in agreement with the existence of these two states, altough we observe a mesenchymal state with less tumorigenic potential. Although we did not specifically design assays to study interconversion between both phenotypes, we observed indeed that xenografts from MCF10CA1h miR-200 cells have an epithelial differentiated glandular component in a morphological continuum from a less differentiated component. A possible explanation for this morphological heterogeneity in vivo would be that there are at least two subopulations of MCF10CA1h cells transduced with miR-200s, one that express high levels of miR-200s and one that fails to express these microRNAs, and that the latter might have expanded *in vivo*. However, this is an unlikely explanation because these cells still express E-cadherin, although at lower

levels. We favor an alternative explanation, namely that MCF10CA1h miR-200 cells, that have experienced a MET relative to their parental cells, exhibit plasticity in their expression of epithelial or mesenchymal phenotypes, giving rise to a continuum of intermediate states and morphologies depending on external stimuli.

Considering the induction of an epithelial phenotype upon expression of miR-200, it is not surprising that MCF10CA1h-miR-200 cells grew at higher rates when implanted in NOD/SCID mice (449, 450). Likewise, they did generate an increased number and size of metastasis in the lungs, which correlates with the fact that despite being more differentiated they display stem-like features and greater tumorogenic potential than control cells. Besides, a positive correlation between the expression levels of CD24 and E-cadherin has been described in other breast cancer cell lines (451), as we have observed. In addition, evidence from others indicate that independently of the type of primary tumor, CD24 is upregulated in metastases from different locations (447) and CD24 positive tumor cells are enriched in stem/progenitor features displaying a strong tumorigenic and metastatic capacity than CD24 negative cells (452). Moreover, expression of CD24 has been associated with tumor progression and metastatic behavior in other contexts (453-456).Though we did not analyzed CD24 epxression, lung colonies formed by MCF10CA1h miR-200C1 cells (an experimental surrogate of metastatic tumors) intensely stained for E-cadherin and the luminal cytokeratins 8 and 18, besides the basal markers, indicative of a higher degree of differentiation of metastases relative to control cells..

Counter to the above evidences, a more broadly extended view is that miR-200 prevents metastasis through the reinforcement of an epithelial gene program in detriment of EMT and invasive properties of cells (302, 457). These views are at strong variance with more recent evidences demonstrating that expression of miR-200 is required for metastatic colonization in several models (341, 356, 357). Moreover, it has been demonstrated that metastatic colonization is favoured by a shift to a MET coupled to a CSC state as part of a process of phenotypic plasticity. It is important to note that undifferentiated metastases from carcinomas can exist that do not need a redifferentiation or MET to colonize a secondary site. That is the case for triple-negative breast cancers, characterized by low expression of miR-200, an EMT and stem-cell like phenotype (163). Treatment with several rounds of chemotherapy can also induce a switc from a differentiated to a less differentiated tumor that has recurred and it is now highly metastatic (154).

Possible mechanisms underlying miR-200-driven prometastatic functions

As previously described, a crucial pathway regulating EMT is controlled by the double negative feedback loop formed by the ZEB1/2 transcription factors, targeted by the miR-200 family of microRNAs, which are in turn transcriptionally repressed by ZEB1/2 (299, 300, 302, 343) (15, 26, 39, 50). The ZEB2 protein is responsible for repressing the key epithelial gene E-cadherin (458) and ZEB2 but not ZEB1 was downregulated in our MCF10CA1h cell model upon miR-200 expression. In our experiments, knockdown of ZEB2 in MCF10CA1h cells phenocopied two key parameters promoted by the expression of miR-200's, mammosphere growth and ALDH activity, suggesting that the capacity to self-renew when grown in non-adherent conditions is marked by the expression of ALDH. Despite the fact that we observed a morphological change towards an epithelial phenotype, it was not accompanied by the expression of E-cadherin nor by a substantial shift towards a CD24 positive expression, which is reasonable as we have already referred to their correlation (451). These features may be important for the spatial organization of cells into more complex tubular structures, considering that ZEB2 silenced cells are only able to form spherical structures. We cannot exclude a compensatory mechanism on the loss of ZEB2 exerted by ZEB1, for which a double knockdown of ZEB factors should be accomplished to at least induce a partial MET. Others have reported that knockdown of ZEB2 was sufficient to mimic the effect of miR-200s in promoting a MET and reprogrammation of fibroblasts to induced pluripotent stem cells (iPSC) (459). This same article reports that Oct4/Sox2 directly activates the miR-200 clusters by binding to their promoter regions, which induces a MET and generation of iPSC. Indeed, we studied SOX2 expression as a possible mediator of the self-renewal properties acquired upon overexpression of miR-200. However, we observed a downregulation at the transcriptional level, suggesting that SOX2 does not play a significant role.

Bipotent luminal/basal states: plasticity in cancer cells?

In breast cancers and breast cancer cell lines, CSCs are suggestive of a rudimentary lineage differentiation hierarchy and thus we determined the expression levels of luminal (CK8, CK18) and basal (CK5, CK14, p63) markers as well as characteristic EMT markers. We observed that whether miR-200 expressing cells grow as mammospheres or in a three-dimensional culture in matrigel, there is a substance induction of E-cadherin accompanied by basal markers and only in matrigel, cells are also expressing luminal keratins. This induction is further exacerbated in

xenografts from MCF10CA1h miR-200 cells, which show a very intense expression of luminal and basal markers accompanied by E-cadherin expressionand lack of expression of vimentin. Although it would be desirable to perform double immunofluorescence assays in oder to demonstrate coexpression of luminal and basal markers in the same cells, we could clearly observe that CK5 is expressed in epithelial cells in very differentiated areas that also express luminal cytokeratins. Again, this suggests the existence of mesenchymal/basal-like CSCs and epithelial/luminal-like CSCs that can interconvert between both states (446) and what we observe is a variety of cells with an intermediate phenotype, possibly retaining bipotent features. Our results accomodate well to the model of the "stemness window" in which a bipotent or hybrid EMT/MET state with stem-like features exists in a small fraction of the tumor population, that is bidirectional and displays a gradient of partial states toward either extreme (460, 461). Despite exhibiting luminal differentiation, both control and miR-200-expressing tumors pertain to the triple negative subtype. As an evidence of the cellular heterogeneity of breast tumors and specially this subtype, others have found luminally differentiated cells in primary basal tumor and cell lines and *vice versa* (462). In addition, these luminal-like cells, which differentiated from a hierarchy of basal origin, showed even more tumor-iniatining capacity than their basal-like counterparts (462). Likewise in our model, a MET induced by miR-200 in basal-like MCF10CA1h cells, generates epithelial-like cells with luminal progenitor features that maintain the basal phenotype and display greater tumorigenic and metastatic behaviour.

Metaplastic breast cancer: which component (epithelial, mesenchymal) is the more aggressive one?

Our analysis of metaplastic breast cancer samples indicates that tumor cells expressing epithelial programs and miR-200 also co-express luminal and basal keratins and of ALDH. Co-expression of luminal and basal keratins in MBC has been reported previously and interpreted as a reflection of the cells of origin of these tumor cells being bipotent stem cells or myoepithelial cells (463-466). Although such interpretations may apply to some subtypes of MBC, they may not suit the carcinosarcomatous cases that we have studied here, since they fail to consider that normal breast epithelial stem cells express low levels of ALDH (18) and that myoepithelial cells do not express luminal keratins (29). In view of the evidences presented for the MCF10CA1h cell model, our finding of co-expression of luminal and basal keratins and ALDH in the epithelial components of 3 out of 5 MBC samples that express high levels of miR-200's suggests that the epithelial neoplastic components of MBC may have originated in cells expressing an early

progenitor luminal phenotype retaining bipotent features. This interpretation approximates our proposed cells of origin of MBC to those proposed for most other types of breast cancer, for which evidences suggest that they originate from luminal progenitor cells (29, 467), with a minority of tumor types, including claudin-low and some metaplastic tumors (377), possibly originating from less differentiated cells that might correspond to mammary stem cells (29).

In line with other authors (92, 300, 376, 378), we also postulate that the epithelial components of MBC may harbor sufficient phenotypic plasticity to give rise to mesenchymal (and other metaplastic) neoplastic components through the engagement of epithelial-mesenchymal transition in response to epigenetic cues or genetic mutations. In addition, based on our observations of enhanced aggressiveness upon induction of an epithelial gene program through transduction of miR-200 in MCF10CA1h cells, we propose that the more intrinsically aggressive components of MBC are the epithelial components, rather than the mesenchymal components. Indeed, metastatic samples from MBC tend to display hallmarks of epithelial differentiation (468, 469) and subtypes with a higher representation of epithelial components, such as spindle cell or carcinomatous MBC, have been reported as associated with higher rates of distant metastasis and worse outcomes (470-473) than those with a greater representation of mesenchymal components, such as fibromatosis-like or matrix-producing MBC (471, 474). Nevertheless, we do not rule out that the presence of a mesenchymal component might contribute to the overall aggressiveness of these tumors (92, 468, 471, 475) through co-operative interactions that foster the local escape of tumor cells with strong epithelial phenotypes and thus with low intrinsic invasiveness, as shown in several experimental models (232, 476-478).

Conclusions

- We have identified sets of microRNAs that are dysregulated along the progression of invasive ductal carcinoma (IDC) from primary node-negative to primary node-positive tumors, lymph node metastasis and distant metastasis. Of particular significance and interest is the upregulation in this progression of miR-181a, miR-181b, miR-210, miR-7 and miR-200s.

- Blood levels of miR-7 and miR-200s tend to be higher in patients with metastatic breast cancer at the time of diagnosis, thus constituting potential early markers of metastatic dissemination in IDC.

- Using the mesenchymal-like MCF10CA1h cell line model, we have found that expression of microRNAs of the miR-200 family induces a strong epithelial gene program that is accompanied by a strong increase of the self-renewal capacity *in vitro*, a shift towards a more committed $CD44^{high}$ $CD24^{high}$ population, increased ALDH activity, ability to further differentiate under appropiate stimuli *in vitro* and undergo glandular differentiation *in vivo*, accompanied by a strong upregulation of luminal and basal markers. These features suggest that miR-200s promote luminal progenitor properties that, coupled to an epithelial phenotype, are associated with an increased metastatic colonization ability *in vivo*.

- The acquisition of mammary luminal progenitor properties induced by miR-200s is in part driven or reinforced by activation of the PI3K-Akt signaling pathway and downregulation of ZEB2. However, the engagement of a full epithelial gene program to counter the mesenchymal state of MCF10CA1h cells requires additional activities mediated by miR-200s.

- The epithelial components of metaplastic breast carcinomas express markers of breast luminal progenitors, including ALDH and double luminal-basal cytokeratins, coincident with high levels of miR-200s, suggesting that this component may be responsible for the aggressive behavior of these rare tumors.

Bibliography

1. Guinebretiere JM, Menet E, Tardivon A, Cherel P, Vanel D. Normal and pathological breast, the histological basis. European journal of radiology. 2005;54(1):6-14.
2. Gudjonsson T, Adriance MC, Sternlicht MD, Petersen OW, Bissell MJ. Myoepithelial cells: their origin and function in breast morphogenesis and neoplasia. Journal of mammary gland biology and neoplasia. 2005;10(3):261-72.
3. Polyak K, Kalluri R. The role of the microenvironment in mammary gland development and cancer. Cold Spring Harbor perspectives in biology. 2010;2(11):a003244.
4. Hansen RK, Bissell MJ. Tissue architecture and breast cancer: the role of extracellular matrix and steroid hormones. Endocrine-related cancer. 2000;7(2):95-113.
5. Jacobs L, Finlayson CA. Breast cancer. Philadelphia: Elsevier/Saunders; 2011. 1 recurs electrònic (xiv, 388 p.) p.
6. Watson CJ. Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. Breast cancer research : BCR. 2006;8(2):203.
7. Deome KB, Faulkin LJ, Jr., Bern HA, Blair PB. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer research. 1959;19(5):515-20.
8. Hoshino K, Gardner WU. Transplantability and life span of mammary gland during serial transplantation in mice. Nature. 1967;213(5072):193-4.
9. Smith GH. Experimental mammary epithelial morphogenesis in an in vivo model: evidence for distinct cellular progenitors of the ductal and lobular phenotype. Breast cancer research and treatment. 1996;39(1):21-31.
10. Kordon EC, Smith GH. An entire functional mammary gland may comprise the progeny from a single cell. Development. 1998;125(10):1921-30.
11. Stingl J, Eirew P, Ricketson I, Shackleton M, Vaillant F, Choi D, et al. Purification and unique properties of mammary epithelial stem cells. Nature. 2006;439(7079):993-7.
12. Shackleton M, Vaillant F, Simpson KJ, Stingl J, Smyth GK, Asselin-Labat ML, et al. Generation of a functional mammary gland from a single stem cell. Nature. 2006;439(7072):84-8.
13. Shehata M, Teschendorff A, Sharp G, Novcic N, Russell IA, Avril S, et al. Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. Breast cancer research : BCR. 2012;14(5):R134.
14. Smalley MJ, Kendrick H, Sheridan JM, Regan JL, Prater MD, Lindeman GJ, et al. Isolation of mouse mammary epithelial subpopulations: a comparison of leading methods. Journal of mammary gland biology and neoplasia. 2012;17(2):91-7.
15. Stingl J, Eaves CJ, Zandieh I, Emerman JT. Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. Breast cancer research and treatment. 2001;67(2):93-109.
16. Stingl J, Raouf A, Emerman JT, Eaves CJ. Epithelial progenitors in the normal human mammary gland. Journal of mammary gland biology and neoplasia. 2005;10(1):49-59.
17. Stingl J, Eaves CJ, Kuusk U, Emerman JT. Phenotypic and functional characterization in vitro of a multipotent epithelial cell present in the normal adult human breast. Differentiation; research in biological diversity. 1998;63(4):201-13.
18. Eirew P, Stingl J, Raouf A, Turashvili G, Aparicio S, Emerman JT, et al. A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. Nature medicine. 2008;14(12):1384-9.
19. Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nature medicine. 2009;15(8):907-13.
20. Villadsen R, Fridriksdottir AJ, Ronnov-Jessen L, Gudjonsson T, Rank F, LaBarge MA, et al. Evidence for a stem cell hierarchy in the adult human breast. The Journal of cell biology. 2007;177(1):87-101.
21. Van Keymeulen A, Rocha AS, Ousset M, Beck B, Bouvencourt G, Rock J, et al. Distinct stem cells contribute to mammary gland development and maintenance. Nature. 2011;479(7372):189-93.
22. Prater MD, Petit V, Alasdair Russell I, Giraddi RR, Shehata M, Menon S, et al. Mammary stem cells have myoepithelial cell properties. Nature cell biology. 2014;16(10):942-50, 1-7.
23. van Amerongen R, Bowman AN, Nusse R. Developmental stage and time dictate the fate of Wnt/beta-catenin-responsive stem cells in the mammary gland. Cell stem cell. 2012;11(3):387-400.

24. Rios AC, Fu NY, Lindeman GJ, Visvader JE. In situ identification of bipotent stem cells in the mammary gland. Nature. 2014;506(7488):322-7.
25. Wagner KU, Boulanger CA, Henry MD, Sgagias M, Hennighausen L, Smith GH. An adjunct mammary epithelial cell population in parous females: its role in functional adaptation and tissue renewal. Development. 2002;129(6):1377-86.
26. Chang TH, Kunasegaran K, Tarulli GA, De Silva D, Voorhoeve PM, Pietersen AM. New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. Breast cancer research : BCR. 2014;16(1):R1.
27. Fu N, Lindeman GJ, Visvader JE. The mammary stem cell hierarchy. Current topics in developmental biology. 2014;107:133-60.
28. Asselin-Labat ML, Vaillant F, Sheridan JM, Pal B, Wu D, Simpson ER, et al. Control of mammary stem cell function by steroid hormone signalling. Nature. 2010;465(7299):798-802.
29. Visvader JE, Stingl J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. Genes & development. 2014;28(11):1143-58.
30. Guo W, Keckesova Z, Donaher JL, Shibue T, Tischler V, Reinhardt F, et al. Slug and Sox9 cooperatively determine the mammary stem cell state. Cell. 2012;148(5):1015-28.
31. Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell. 2008;133(4):704-15.
32. Nassour M, Idoux-Gillet Y, Selmi A, Come C, Faraldo ML, Deugnier MA, et al. Slug controls stem/progenitor cell growth dynamics during mammary gland morphogenesis. PloS one. 2012;7(12):e53498.
33. Zeng YA, Nusse R. Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. Cell stem cell. 2010;6(6):568-77.
34. Badders NM, Goel S, Clark RJ, Klos KS, Kim S, Bafico A, et al. The Wnt receptor, Lrp5, is expressed by mouse mammary stem cells and is required to maintain the basal lineage. PloS one. 2009;4(8):e6594.
35. Lindvall C, Zylstra CR, Evans N, West RA, Dykema K, Furge KA, et al. The Wnt co-receptor Lrp6 is required for normal mouse mammary gland development. PloS one. 2009;4(6):e5813.
36. Scheel C, Eaton EN, Li SH, Chaffer CL, Reinhardt F, Kah KJ, et al. Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. Cell. 2011;145(6):926-40.
37. Gu B, Watanabe K, Sun P, Fallahi M, Dai X. Chromatin effector Pygo2 mediates Wnt-notch crosstalk to suppress luminal/alveolar potential of mammary stem and basal cells. Cell stem cell. 2013;13(1):48-61.
38. Liu S, Dontu G, Mantle ID, Patel S, Ahn NS, Jackson KW, et al. Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malignant human mammary stem cells. Cancer research. 2006;66(12):6063-71.
39. Pietersen AM, Evers B, Prasad AA, Tanger E, Cornelissen-Steijger P, Jonkers J, et al. Bmi1 regulates stem cells and proliferation and differentiation of committed cells in mammary epithelium. Current biology : CB. 2008;18(14):1094-9.
40. Pal B, Bouras T, Shi W, Vaillant F, Sheridan JM, Fu N, et al. Global changes in the mammary epigenome are induced by hormonal cues and coordinated by Ezh2. Cell reports. 2013;3(2):411-26.
41. Bouras T, Pal B, Vaillant F, Harburg G, Asselin-Labat ML, Oakes SR, et al. Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. Cell stem cell. 2008;3(4):429-41.
42. Cicalese A, Bonizzi G, Pasi CE, Faretta M, Ronzoni S, Giulini B, et al. The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. Cell. 2009;138(6):1083-95.
43. Choi YS, Chakrabarti R, Escamilla-Hernandez R, Sinha S. Elf5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: failure of Stat5 activation and functional differentiation in the absence of Elf5. Developmental biology. 2009;329(2):227-41.
44. Zhou J, Chehab R, Tkalcevic J, Naylor MJ, Harris J, Wilson TJ, et al. Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. The EMBO journal. 2005;24(3):635-44.
45. Oakes SR, Naylor MJ, Asselin-Labat ML, Blazek KD, Gardiner-Garden M, Hilton HN, et al. The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes & development. 2008;22(5):581-6.
46. Chakrabarti R, Wei Y, Romano RA, DeCoste C, Kang Y, Sinha S. Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. Stem cells. 2012;30(7):1496-508.

47. Chakrabarti R, Hwang J, Andres Blanco M, Wei Y, Lukacisin M, Romano RA, et al. Elf5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. Nature cell biology. 2012;14(11):1212-22.
48. Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell. 2006;127(5):1041-55.
49. Asselin-Labat ML, Sutherland KD, Barker H, Thomas R, Shackleton M, Forrest NC, et al. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nature cell biology. 2007;9(2):201-9.
50. Asselin-Labat ML, Sutherland KD, Vaillant F, Gyorki DE, Wu D, Holroyd S, et al. Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. Molecular and cellular biology. 2011;31(22):4609-22.
51. Rodilla V, Dasti A, Huyghe M, Lafkas D, Laurent C, Reyal F, et al. Luminal progenitors restrict their lineage potential during mammary gland development. PLoS biology. 2015;13(2):e1002069.
52. Bhat V, Sun YJ, Weger S, Raouf A. Notch-Induced Expression of FZD7 Requires Noncanonical NOTCH3 Signaling in Human Breast Epithelial Cells. Stem cells and development. 2016;25(7):522-9.
53. Raouf A, Zhao Y, To K, Stingl J, Delaney A, Barbara M, et al. Transcriptome analysis of the normal human mammary cell commitment and differentiation process. Cell stem cell. 2008;3(1):109-18.
54. Lafkas D, Rodilla V, Huyghe M, Mourao L, Kiaris H, Fre S. Notch3 marks clonogenic mammary luminal progenitor cells in vivo. The Journal of cell biology. 2013;203(1):47-56.
55. Sale S, Lafkas D, Artavanis-Tsakonas S. Notch2 genetic fate mapping reveals two previously unrecognized mammary epithelial lineages. Nature cell biology. 2013;15(5):451-60.
56. Michalak EM, Nacerddine K, Pietersen A, Beuger V, Pawlitzky I, Cornelissen-Steijger P, et al. Polycomb group gene Ezh2 regulates mammary gland morphogenesis and maintains the luminal progenitor pool. Stem cells. 2013;31(9):1910-20.
57. Hennighausen L, Robinson GW. Information networks in the mammary gland. Nature reviews Molecular cell biology. 2005;6(9):715-25.
58. Yamaji D, Na R, Feuermann Y, Pechhold S, Chen W, Robinson GW, et al. Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. Genes & development. 2009;23(20):2382-7.
59. Cui Y, Riedlinger G, Miyoshi K, Tang W, Li C, Deng CX, et al. Inactivation of Stat5 in mouse mammary epithelium during pregnancy reveals distinct functions in cell proliferation, survival, and differentiation. Molecular and cellular biology. 2004;24(18):8037-47.
60. Anderson E, Clarke RB. Steroid receptors and cell cycle in normal mammary epithelium. Journal of mammary gland biology and neoplasia. 2004;9(1):3-13.
61. Brisken C, Park S, Vass T, Lydon JP, O'Malley BW, Weinberg RA. A paracrine role for the epithelial progesterone receptor in mammary gland development. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(9):5076-81.
62. Mallepell S, Krust A, Chambon P, Brisken C. Paracrine signaling through the epithelial estrogen receptor alpha is required for proliferation and morphogenesis in the mammary gland. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(7):2196-201.
63. Sleeman KE, Kendrick H, Robertson D, Isacke CM, Ashworth A, Smalley MJ. Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. The Journal of cell biology. 2007;176(1):19-26.
64. Joshi PA, Jackson HW, Beristain AG, Di Grappa MA, Mote PA, Clarke CL, et al. Progesterone induces adult mammary stem cell expansion. Nature. 2010;465(7299):803-7.
65. Lee HJ, Gallego-Ortega D, Ledger A, Schramek D, Joshi P, Szwarc MM, et al. Progesterone drives mammary secretory differentiation via RankL-mediated induction of Elf5 in luminal progenitor cells. Development. 2013;140(7):1397-401.
66. Forster N, Saladi SV, van Bragt M, Sfondouris ME, Jones FE, Li Z, et al. Basal cell signaling by p63 controls luminal progenitor function and lactation via NRG1. Developmental cell. 2014;28(2):147-60.
67. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, Marchio C, Reis-Filho JS. Breast cancer precursors revisited: molecular features and progression pathways. Histopathology. 2010;57(2):171-92.
68. Stingl J, Caldas C. Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. Nature reviews Cancer. 2007;7(10):791-9.

69. Allred DC, Medina D. The relevance of mouse models to understanding the development and progression of human breast cancer. Journal of mammary gland biology and neoplasia. 2008;13(3):279-88.
70. Li CI, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. British journal of cancer. 2005;93(9):1046-52.
71. Diab SG, Clark GM, Osborne CK, Libby A, Allred DC, Elledge RM. Tumor characteristics and clinical outcome of tubular and mucinous breast carcinomas. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1999;17(5):1442-8.
72. Glazebrook KN, Reynolds C, Smith RL, Gimenez EI, Boughey JC. Adenoid cystic carcinoma of the breast. AJR American journal of roentgenology. 2010;194(5):1391-6.
73. Wang XX, Jiang YZ, Liu XY, Li JJ, Song CG, Shao ZM. Difference in characteristics and outcomes between medullary breast carcinoma and invasive ductal carcinoma: a population based study from SEER 18 database. Oncotarget. 2016;7(16):22665-73.
74. Hsiao YH, Tsai HD, Chou MC, Man YG. The myoepithelial cell layer may serve as a potential trigger factor for different outcomes of stage-matched invasive lobular and ductal breast cancers. International journal of biological sciences. 2011;7(2):147-53.
75. Simpson PT, Reis-Filho JS, Gale T, Lakhani SR. Molecular evolution of breast cancer. The Journal of pathology. 2005;205(2):248-54.
76. Farabegoli F, Champeme MH, Bieche I, Santini D, Ceccarelli C, Derenzini M, et al. Genetic pathways in the evolution of breast ductal carcinoma in situ. The Journal of pathology. 2002;196(3):280-6.
77. Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. Breast cancer research : BCR. 2010;12(4):207.
78. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2013;31(31):3997-4013.
79. Bertos NR, Park M. Breast cancer - one term, many entities? The Journal of clinical investigation. 2011;121(10):3789-96.
80. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. The New England journal of medicine. 2001;344(11):783-92.
81. Smith I, Procter M, Gelber RD, Guillaume S, Feyereislova A, Dowsett M, et al. 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. Lancet. 2007;369(9555):29-36.
82. Bartlett JM, Brookes CL, Robson T, van de Velde CJ, Billingham LJ, Campbell FM, et al. Estrogen receptor and progesterone receptor as predictive biomarkers of response to endocrine therapy: a prospectively powered pathology study in the Tamoxifen and Exemestane Adjuvant Multinational trial. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29(12):1531-8.
83. Stotter A, Walker R. Tumour markers predictive of successful treatment of breast cancer with primary endocrine therapy in patients over 70 years old: a prospective study. Critical reviews in oncology/hematology. 2010;75(3):249-56.
84. Junttila TT, Akita RW, Parsons K, Fields C, Lewis Phillips GD, Friedman LS, et al. Ligand-independent HER2/HER3/PI3K complex is disrupted by trastuzumab and is effectively inhibited by the PI3K inhibitor GDC-0941. Cancer cell. 2009;15(5):429-40.
85. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. The Journal of clinical investigation. 2011;121(7):2750-67.
86. Weigelt B, Peterse JL, van 't Veer LJ. Breast cancer metastasis: markers and models. Nature reviews Cancer. 2005;5(8):591-602.
87. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.

88. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America. 2001;98(19):10869-74.
89. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. Genome biology. 2007;8(8):R157.
90. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. Molecular oncology. 2011;5(1):5-23.
91. Bruna A, Greenwood W, Le Quesne J, Teschendorff A, Miranda-Saavedra D, Rueda OM, et al. TGFbeta induces the formation of tumour-initiating cells in claudinlow breast cancer. Nature communications. 2012;3:1055.
92. Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, et al. Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. Cancer research. 2009;69(10):4116-24.
93. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27(8):1160-7.
94. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC genomics. 2006;7:96.
95. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(16):9212-7.
96. Perreard L, Fan C, Quackenbush JF, Mullins M, Gauthier NP, Nelson E, et al. Classification and risk stratification of invasive breast carcinomas using a real-time quantitative RT-PCR assay. Breast cancer research : BCR. 2006;8(2):R23.
97. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(14):8418-23.
98. Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. Journal of the National Cancer Institute. 2009;101(10):736-50.
99. Millikan RC, Newman B, Tse CK, Moorman PG, Conway K, Dressler LG, et al. Epidemiology of basal-like breast cancer. Breast cancer research and treatment. 2008;109(1):123-39.
100. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. Clinical cancer research : an official journal of the American Association for Cancer Research. 2007;13(15 Pt 1):4429-34.
101. Brenton JD, Carey LA, Ahmed AA, Caldas C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2005;23(29):7350-60.
102. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2010;16(21):5222-32.
103. Johnston S, Pippen J, Jr., Pivot X, Lichinitser M, Sadeghi S, Dieras V, et al. Lapatinib combined with letrozole versus letrozole and placebo as first-line therapy for postmenopausal hormone receptor-positive metastatic breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2009;27(33):5538-46.
104. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clinical cancer research : an official journal of the American Association for Cancer Research. 2005;11(16):5678-85.
105. Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(8):1275-81.
106. Paik S, Kim C, Wolmark N. HER2 status and benefit from adjuvant trastuzumab in breast cancer. The New England journal of medicine. 2008;358(13):1409-11.

107. Rottenberg S, Jaspers JE, Kersbergen A, van der Burg E, Nygren AO, Zander SA, et al. High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(44):17079-84.
108. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.
109. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.
110. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. Cancer cell. 2012;21(3):309-22.
111. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. Nature. 2013;501(7467):355-64.
112. Axelrod R, Axelrod DE, Pienta KJ. Evolution of cooperation among tumor cells. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(36):13474-9.
113. Anderson K, Lutz C, van Delft FW, Bateman CM, Guo Y, Colman SM, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature. 2011;469(7330):356-61.
114. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481(7381):306-13.
115. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, et al. The life history of 21 breast cancers. Cell. 2012;149(5):994-1007.
116. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. The New England journal of medicine. 2012;366(10):883-92.
117. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature. 2013;501(7467):346-54.
118. Nowell PC. The clonal evolution of tumor cell populations. Science. 1976;194(4260):23-8.
119. Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. Nature reviews Cancer. 2008;8(10):755-68.
120. Wang JC, Dick JE. Cancer stem cells: lessons from leukemia. Trends in cell biology. 2005;15(9):494-501.
121. Wang Y, Krivtsov AV, Sinha AU, North TE, Goessling W, Feng Z, et al. The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. Science. 2010;327(5973):1650-3.
122. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. Nature medicine. 1997;3(7):730-7.
123. Ricci-Vitiani L, Lombardi DG, Pilozzi E, Biffoni M, Todaro M, Peschle C, et al. Identification and expansion of human colon-cancer-initiating cells. Nature. 2007;445(7123):111-5.
124. Singh SK, Hawkins C, Clarke ID, Squire JA, Bayani J, Hide T, et al. Identification of human brain tumour initiating cells. Nature. 2004;432(7015):396-401.
125. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(7):3983-8.
126. Quintana E, Shackleton M, Foster HR, Fullen DR, Sabel MS, Johnson TM, et al. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. Cancer cell. 2010;18(5):510-23.
127. Gupta GP, Massague J. Cancer metastasis: building a framework. Cell. 2006;127(4):679-95.
128. Campbell LL, Polyak K. Breast tumor heterogeneity: cancer stem cells or clonal evolution? Cell cycle. 2007;6(19):2332-8.
129. Roesch A, Fukunaga-Kalabis M, Schmidt EC, Zabierowski SE, Brafford PA, Vultur A, et al. A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. Cell. 2010;141(4):583-94.
130. Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C, et al. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell. 2011;146(4):633-44.
131. Chaffer CL, Brueckmann I, Scheel C, Kaestli AJ, Wiggins PA, Rodrigues LO, et al. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(19):7950-5.
132. Pardal R, Clarke MF, Morrison SJ. Applying the principles of stem-cell biology to cancer. Nature reviews Cancer. 2003;3(12):895-902.

133. Merlos-Suarez A, Barriga FM, Jung P, Iglesias M, Cespedes MV, Rossell D, et al. The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. Cell stem cell. 2011;8(5):511-24.
134. Krivtsov AV, Twomey D, Feng Z, Stubbs MC, Wang Y, Faber J, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. Nature. 2006;442(7104):818-22.
135. Clarke MF, Dick JE, Dirks PB, Eaves CJ, Jamieson CH, Jones DL, et al. Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. Cancer research. 2006;66(19):9339-44.
136. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell stem cell. 2007;1(5):555-67.
137. Meyer MJ, Fleming JM, Lin AF, Hussnain SA, Ginsburg E, Vonderhaar BK. CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor-negative breast cancer. Cancer research. 2010;70(11):4624-33.
138. Gibbs KD, Jr., Jager A, Crespo O, Goltsev Y, Trejo A, Richard CE, et al. Decoupling of tumor-initiating activity from stable immunophenotype in HoxA9-Meis1-driven AML. Cell stem cell. 2012;10(2):210-7.
139. Goardon N, Marchi E, Atzberger A, Quek L, Schuh A, Soneji S, et al. Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer cell. 2011;19(1):138-52.
140. Boiko AD, Razorenova OV, van de Rijn M, Swetter SM, Johnson DL, Ly DP, et al. Human melanoma-initiating cells express neural crest nerve growth factor receptor CD271. Nature. 2010;466(7302):133-7.
141. Quintana E, Shackleton M, Sabel MS, Fullen DR, Johnson TM, Morrison SJ. Efficient tumour formation by single human melanoma cells. Nature. 2008;456(7222):593-8.
142. Ishizawa K, Rasheed ZA, Karisch R, Wang Q, Kowalski J, Susky E, et al. Tumor-initiating cells are rare in many human tumors. Cell stem cell. 2010;7(3):279-82.
143. Takebe N, Miele L, Harris PJ, Jeong W, Bando H, Kahn M, et al. Targeting Notch, Hedgehog, and Wnt pathways in cancer stem cells: clinical update. Nature reviews Clinical oncology. 2015;12(8):445-64.
144. Wong DJ, Liu H, Ridky TW, Cassarino D, Segal E, Chang HY. Module map of stem cell genes guides creation of epithelial cancer stem cells. Cell stem cell. 2008;2(4):333-44.
145. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76.
146. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nature genetics. 2008;40(5):499-507.
147. Oskarsson T, Batlle E, Massague J. Metastatic stem cells: sources, niches, and vital pathways. Cell stem cell. 2014;14(3):306-21.
148. Ye J, Wu D, Wu P, Chen Z, Huang J. The cancer stem cell niche: cross talk between cancer stem cells and their microenvironment. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine. 2014;35(5):3945-51.
149. Cabarcas SM, Mathews LA, Farrar WL. The cancer stem cell niche--there goes the neighborhood? International journal of cancer. 2011;129(10):2315-27.
150. Chanmee T, Ontong P, Konno K, Itano N. Tumor-associated macrophages as major players in the tumor microenvironment. Cancers. 2014;6(3):1670-90.
151. Fessler E, Dijkgraaf FE, De Sousa EMF, Medema JP. Cancer stem cell dynamics in tumor progression and metastasis: is the microenvironment to blame? Cancer letters. 2013;341(1):97-104.
152. Sneddon JB, Werb Z. Location, location, location: the cancer stem cell niche. Cell stem cell. 2007;1(6):607-11.
153. Li X, Lewis MT, Huang J, Gutierrez C, Osborne CK, Wu MF, et al. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. Journal of the National Cancer Institute. 2008;100(9):672-9.
154. Creighton CJ, Li X, Landis M, Dixon JM, Neumeister VM, Sjolund A, et al. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(33):13820-5.
155. Domingo-Domenech J, Vidal SJ, Rodriguez-Bravo V, Castillo-Martin M, Quinn SA, Rodriguez-Barrueco R, et al. Suppression of acquired docetaxel resistance in prostate cancer through depletion of notch- and hedgehog-dependent tumor-initiating cells. Cancer cell. 2012;22(3):373-88.

156. Yang W, Yan HX, Chen L, Liu Q, He YQ, Yu LX, et al. Wnt/beta-catenin signaling contributes to activation of normal and tumorigenic liver progenitor cells. Cancer research. 2008;68(11):4287-95.
157. Fletcher JI, Haber M, Henderson MJ, Norris MD. ABC transporters in cancer: more than just drug efflux pumps. Nature reviews Cancer. 2010;10(2):147-56.
158. Bao S, Wu Q, McLendon RE, Hao Y, Shi Q, Hjelmeland AB, et al. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature. 2006;444(7120):756-60.
159. Gieni RS, Ismail IH, Campbell S, Hendzel MJ. Polycomb group proteins in the DNA damage response: a link between radiation resistance and "stemness". Cell cycle. 2011;10(6):883-94.
160. Calabrese C, Poppleton H, Kocak M, Hogg TL, Fuller C, Hamner B, et al. A perivascular niche for brain tumor stem cells. Cancer cell. 2007;11(1):69-82.
161. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. The Journal of clinical investigation. 2009;119(6):1420-8.
162. Hay ED. An overview of epithelio-mesenchymal transformation. Acta anatomica. 1995;154(1):8-20.
163. Sarrio D, Rodriguez-Pinilla SM, Hardisson D, Cano A, Moreno-Bueno G, Palacios J. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. Cancer research. 2008;68(4):989-97.
164. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, et al. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(35):15449-54.
165. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nature reviews Cancer. 2007;7(6):415-28.
166. Nieto MA. Epithelial plasticity: a common theme in embryonic and cancer cells. Science. 2013;342(6159):1234850.
167. Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. Cell. 2009;139(5):871-90.
168. Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nature reviews Cancer. 2009;9(4):265-73.
169. De Craene B, Gilbert B, Stove C, Bruyneel E, van Roy F, Berx G. The transcription factor snail induces tumor cell invasion through modulation of the epithelial cell differentiation program. Cancer research. 2005;65(14):6237-44.
170. Vandewalle C, Comijn J, De Craene B, Vermassen P, Bruyneel E, Andersen H, et al. SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. Nucleic acids research. 2005;33(20):6566-78.
171. Yang J, Mani SA, Donaher JL, Ramaswamy S, Itzykson RA, Come C, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell. 2004;117(7):927-39.
172. Ocana OH, Corcoles R, Fabra A, Moreno-Bueno G, Acloque H, Vega S, et al. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. Cancer cell. 2012;22(6):709-24.
173. Graff JR, Gabrielson E, Fujii H, Baylin SB, Herman JG. Methylation patterns of the E-cadherin 5' CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. The Journal of biological chemistry. 2000;275(4):2727-32.
174. Lombaerts M, van Wezel T, Philippo K, Dierssen JW, Zimmerman RM, Oosting J, et al. E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. British journal of cancer. 2006;94(5):661-71.
175. Herranz N, Pasini D, Diaz VM, Franci C, Gutierrez A, Dave N, et al. Polycomb complex 2 is required for E-cadherin repression by the Snail1 transcription factor. Molecular and cellular biology. 2008;28(15):4772-81.
176. Peinado H, Ballestar E, Esteller M, Cano A. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. Molecular and cellular biology. 2004;24(1):306-19.
177. Sanchez-Tillo E, Lazaro A, Torrent R, Cuatrecasas M, Vaquero EC, Castells A, et al. ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. Oncogene. 2010;29(24):3490-500.

178. McDonald OG, Wu H, Timp W, Doi A, Feinberg AP. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. Nature structural & molecular biology. 2011;18(8):867-74.
179. Brabletz S, Brabletz T. The ZEB/miR-200 feedback loop--a motor of cellular plasticity in development and cancer? EMBO reports. 2010;11(9):670-7.
180. Kim NH, Kim HS, Li XY, Lee I, Choi HS, Kang SE, et al. A p53/miRNA-34 axis regulates Snail1-dependent cancer cell epithelial-mesenchymal transition. The Journal of cell biology. 2011;195(3):417-33.
181. Iliopoulos D, Lindahl-Allen M, Polytarchou C, Hirsch HA, Tsichlis PN, Struhl K. Loss of miR-200 inhibition of Suz12 leads to polycomb-mediated repression required for the formation and maintenance of cancer stem cells. Molecular cell. 2010;39(5):761-72.
182. Tellez CS, Juri DE, Do K, Bernauer AM, Thomas CL, Damiani LA, et al. EMT and stem cell-like properties associated with miR-205 and miR-200 epigenetic silencing are early manifestations during carcinogen-induced transformation of human lung epithelial cells. Cancer research. 2011;71(8):3087-97.
183. Ma L, Young J, Prabhala H, Pan E, Mestdagh P, Muth D, et al. miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. Nature cell biology. 2010;12(3):247-56.
184. Yanagisawa M, Huveldt D, Kreinest P, Lohse CM, Cheville JC, Parker AS, et al. A p120 catenin isoform switch affects Rho activity, induces tumor cell invasion, and predicts metastatic disease. The Journal of biological chemistry. 2008;283(26):18344-54.
185. Horiguchi K, Sakamoto K, Koinuma D, Semba K, Inoue A, Inoue S, et al. TGF-beta drives epithelial-mesenchymal transition through deltaEF1-mediated downregulation of ESRP. Oncogene. 2012;31(26):3190-201.
186. Reinke LM, Xu Y, Cheng C. Snail represses the splicing regulator epithelial splicing regulatory protein 1 to promote epithelial-mesenchymal transition. The Journal of biological chemistry. 2012;287(43):36435-42.
187. Acevedo VD, Gangula RD, Freeman KW, Li R, Zhang Y, Wang F, et al. Inducible FGFR-1 activation leads to irreversible prostate adenocarcinoma and an epithelial-to-mesenchymal transition. Cancer cell. 2007;12(6):559-71.
188. Brown RL, Reinke LM, Damerow MS, Perez D, Chodosh LA, Yang J, et al. CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. The Journal of clinical investigation. 2011;121(3):1064-74.
189. Thiery JP. Epithelial-mesenchymal transitions in tumour progression. Nature reviews Cancer. 2002;2(6):442-54.
190. Blick T, Widodo E, Hugo H, Waltham M, Lenburg ME, Neve RM, et al. Epithelial mesenchymal transition traits in human breast cancer cell lines. Clinical & experimental metastasis. 2008;25(6):629-42.
191. Choi Y, Lee HJ, Jang MH, Gwak JM, Lee KS, Kim EJ, et al. Epithelial-mesenchymal transition increases during the progression of in situ to invasive basal-like breast cancer. Human pathology. 2013;44(11):2581-9.
192. Tarin D, Thompson EW, Newgreen DF. The fallacy of epithelial mesenchymal transition in neoplasia. Cancer research. 2005;65(14):5996-6000; discussion -1.
193. Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT, et al. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. Science. 2013;339(6119):580-4.
194. Giampieri S, Manning C, Hooper S, Jones L, Hill CS, Sahai E. Localized and reversible TGFbeta signalling switches breast cancer cells from cohesive to single cell motility. Nature cell biology. 2009;11(11):1287-96.
195. De Craene B, Berx G. Regulatory networks defining EMT during cancer initiation and progression. Nature reviews Cancer. 2013;13(2):97-110.
196. Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong ST, et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. Nature. 2015;527(7579):472-6.
197. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature. 2015;527(7579):525-30.
198. Akalay I, Janji B, Hasmim M, Noman MZ, Andre F, De Cremoux P, et al. Epithelial-to-mesenchymal transition and autophagy induction in breast carcinoma promote escape from T-cell-mediated lysis. Cancer research. 2013;73(8):2418-27.
199. Akalay I, Tan TZ, Kumar P, Janji B, Mami-Chouaib F, Charpy C, et al. Targeting WNT1-inducible signaling pathway protein 2 alters human breast cancer cell susceptibility to specific lysis through regulation of KLF-4 and miR-7 expression. Oncogene. 2015;34(17):2261-71.

200. Morel AP, Lievre M, Thomas C, Hinkal G, Ansieau S, Puisieux A. Generation of breast cancer stem cells through epithelial-mesenchymal transition. PloS one. 2008;3(8):e2888.
201. Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, et al. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. Nature cell biology. 2009;11(12):1487-95.
202. Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nature reviews Cancer. 2003;3(6):453-8.
203. Bissell MJ, Hines WC. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. Nature medicine. 2011;17(3):320-9.
204. McAllister SS, Weinberg RA. The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. Nature cell biology. 2014;16(8):717-27.
205. Harney AS, Arwert EN, Entenberg D, Wang Y, Guo P, Qian BZ, et al. Real-Time Imaging Reveals Local, Transient Vascular Permeability, and Tumor Cell Intravasation Stimulated by TIE2hi Macrophage-Derived VEGFA. Cancer discovery. 2015;5(9):932-43.
206. Wyckoff JB, Wang Y, Lin EY, Li JF, Goswami S, Stanley ER, et al. Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. Cancer research. 2007;67(6):2649-56.
207. Kim YN, Koo KH, Sung JY, Yun UJ, Kim H. Anoikis resistance: an essential prerequisite for tumor metastasis. International journal of cell biology. 2012;2012:306879.
208. Labelle M, Begum S, Hynes RO. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. Cancer cell. 2011;20(5):576-90.
209. Talmadge JE, Fidler IJ. AACR centennial series: the biology of cancer metastasis: historical perspective. Cancer research. 2010;70(14):5649-69.
210. Au SH, Storey BD, Moore JC, Tang Q, Chen YL, Javaid S, et al. Clusters of circulating tumor cells traverse capillary-sized vessels. Proceedings of the National Academy of Sciences of the United States of America. 2016;113(18):4947-52.
211. Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. Science. 2011;331(6024):1559-64.
212. Wang H, Fu W, Im JH, Zhou Z, Santoro SA, Iyer V, et al. Tumor cell alpha3beta1 integrin and vascular laminin-5 mediate pulmonary arrest and metastasis. The Journal of cell biology. 2004;164(6):935-41.
213. Carmeliet P, Jain RK. Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases. Nature reviews Drug discovery. 2011;10(6):417-27.
214. Braun S, Vogl FD, Naume B, Janni W, Osborne MP, Coombes RC, et al. A pooled analysis of bone marrow micrometastasis in breast cancer. The New England journal of medicine. 2005;353(8):793-802.
215. Nagrath S, Sequist LV, Maheswaran S, Bell DW, Irimia D, Ulkus L, et al. Isolation of rare circulating tumour cells in cancer patients by microchip technology. Nature. 2007;450(7173):1235-9.
216. Kim MY, Oskarsson T, Acharyya S, Nguyen DX, Zhang XH, Norton L, et al. Tumor self-seeding by circulating cancer cells. Cell. 2009;139(7):1315-26.
217. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organ-specific colonization. Nature reviews Cancer. 2009;9(4):274-84.
218. Valiente M, Obenauf AC, Jin X, Chen Q, Zhang XH, Lee DJ, et al. Serpins promote cancer cell survival and vascular co-option in brain metastasis. Cell. 2014;156(5):1002-16.
219. Kaplan RN, Riba RD, Zacharoulis S, Bramley AH, Vincent L, Costa C, et al. VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. Nature. 2005;438(7069):820-7.
220. Hoshino A, Costa-Silva B, Shen TL, Rodrigues G, Hashimoto A, Tesic Mark M, et al. Tumour exosome integrins determine organotropic metastasis. Nature. 2015;527(7578):329-35.
221. Psaila B, Lyden D. The metastatic niche: adapting the foreign soil. Nature reviews Cancer. 2009;9(4):285-93.
222. Zhang XH, Wang Q, Gerald W, Hudis CA, Norton L, Smid M, et al. Latent bone metastasis in breast cancer tied to Src-dependent survival signals. Cancer cell. 2009;16(1):67-78.
223. Gao H, Chakraborty G, Lee-Lim AP, Mo Q, Decker M, Vonica A, et al. The BMP inhibitor Coco reactivates breast cancer cells at lung metastatic sites. Cell. 2012;150(4):764-79.
224. Schrader J, Gordon-Walker TT, Aucott RL, van Deemter M, Quaas A, Walsh S, et al. Matrix stiffness modulates proliferation, chemotherapeutic response, and dormancy in hepatocellular carcinoma cells. Hepatology. 2011;53(4):1192-205.

225. Sosa MS, Bragado P, Aguirre-Ghiso JA. Mechanisms of disseminated cancer cell dormancy: an awakening field. Nature reviews Cancer. 2014;14(9):611-22.
226. Lu X, Mu E, Wei Y, Riethdorf S, Yang Q, Yuan M, et al. VCAM-1 promotes osteolytic expansion of indolent bone micrometastasis of breast cancer by engaging alpha4beta1-positive osteoclast progenitors. Cancer cell. 2011;20(6):701-14.
227. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nature medicine. 2013;19(11):1423-37.
228. Aguilar E, Marin de Mas I, Zodda E, Marin S, Morrish F, Selivanov V, et al. Metabolic Reprogramming and Dependencies Associated with Epithelial Cancer Stem Cells Independent of the Epithelial-Mesenchymal Transition Program. Stem cells. 2016;34(5):1163-76.
229. Chen EI, Hewel J, Krueger JS, Tiraby C, Weber MR, Kralli A, et al. Adaptation of energy metabolism in breast cancer brain metastases. Cancer research. 2007;67(4):1472-86.
230. Lawson DA, Bhakta NR, Kessenbrock K, Prummel KD, Yu Y, Takai K, et al. Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature. 2015;526(7571):131-5.
231. Beck B, Lapouge G, Rorive S, Drogat B, Desaedelaere K, Delafaille S, et al. Different levels of Twist1 regulate skin tumor initiation, stemness, and progression. Cell stem cell. 2015;16(1):67-79.
232. Celia-Terrassa T, Meca-Cortes O, Mateo F, Martinez de Paz A, Rubio N, Arnal-Estape A, et al. Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. The Journal of clinical investigation. 2012;122(5):1849-68.
233. Schmidt JM, Panzilius E, Bartsch HS, Irmler M, Beckers J, Kari V, et al. Stem-cell-like properties and epithelial plasticity arise as stable traits after transient Twist1 activation. Cell reports. 2015;10(2):131-9.
234. Stankic M, Pavlovic S, Chin Y, Brogi E, Padua D, Norton L, et al. TGF-beta-Id1 signaling opposes Twist1 and promotes metastatic colonization via a mesenchymal-to-epithelial transition. Cell reports. 2013;5(5):1228-42.
235. Hata A, Lieberman J. Dysregulation of microRNA biogenesis and gene silencing in cancer. Science signaling. 2015;8(368):re3.
236. Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. The EMBO journal. 2002;21(17):4663-70.
237. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. Nature structural & molecular biology. 2006;13(12):1097-101.
238. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. Genes & development. 2004;18(24):3016-27.
239. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, et al. The Microprocessor complex mediates the genesis of microRNAs. Nature. 2004;432(7014):235-40.
240. Kim YK, Kim VN. Processing of intronic microRNAs. The EMBO journal. 2007;26(3):775-83.
241. Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. Primary microRNA transcripts are processed co-transcriptionally. Nature structural & molecular biology. 2008;15(9):902-9.
242. Han J, Pedersen JS, Kwon SC, Belair CD, Kim YK, Yeom KH, et al. Posttranscriptional crossregulation between Drosha and DGCR8. Cell. 2009;136(1):75-84.
243. Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nature cell biology. 2009;11(3):228-34.
244. Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. Rna. 2004;10(2):185-91.
245. Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear export of microRNA precursors. Science. 2004;303(5654):95-8.
246. MacRae IJ, Ma E, Zhou M, Robinson CV, Doudna JA. In vitro reconstitution of the human RISC-loading complex. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(2):512-7.
247. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nature reviews Molecular cell biology. 2009;10(2):126-39.
248. Dueck A, Ziegler C, Eichner A, Berezikov E, Meister G. microRNAs associated with the different human Argonaute proteins. Nucleic acids research. 2012;40(19):9850-62.
249. Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell. 2007;130(1):89-100.

250. Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. Nature. 2007;448(7149):83-6.
251. Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. Genes & development. 2008;22(20):2773-85.
252. Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, et al. A human snoRNA with microRNA-like functions. Molecular cell. 2008;32(4):519-28.
253. Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature. 2010;465(7298):584-9.
254. Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, et al. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. Science. 2010;328(5986):1694-8.
255. Heo I, Ha M, Lim J, Yoon MJ, Park JE, Kwon SC, et al. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell. 2012;151(3):521-32.
256. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nature reviews Genetics. 2010;11(9):597-610.
257. Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP. Expanding the microRNA targeting code: functional sites with centered pairing. Molecular cell. 2010;38(6):789-802.
258. Fabian MR, Sonenberg N. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. Nature structural & molecular biology. 2012;19(6):586-93.
259. Huntzinger E, Izaurralde E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nature reviews Genetics. 2011;12(2):99-110.
260. Chekulaeva M, Mathys H, Zipprich JT, Attig J, Colic M, Parker R, et al. miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. Nature structural & molecular biology. 2011;18(11):1218-26.
261. Fabian MR, Cieplak MK, Frank F, Morita M, Green J, Srikumar T, et al. miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. Nature structural & molecular biology. 2011;18(11):1211-7.
262. Humphreys DT, Westman BJ, Martin DI, Preiss T. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(47):16961-6.
263. Petersen CP, Bordeleau ME, Pelletier J, Sharp PA. Short RNAs repress translation after initiation in mammalian cells. Molecular cell. 2006;21(4):533-42.
264. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. Nature. 2005;435(7043):834-8.
265. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. Nature genetics. 2007;39(5):673-7.
266. Kumar MS, Pester RE, Chen CY, Lane K, Chin C, Lu J, et al. Dicer1 functions as a haploinsufficient tumor suppressor. Genes & development. 2009;23(23):2700-4.
267. Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(7):2257-61.
268. Di Leva G, Croce CM. Roles of small RNAs in tumor formation. Trends in molecular medicine. 2010;16(6):257-67.
269. Samantarrai D, Dash S, Chhetri B, Mallick B. Genomic and epigenomic cross-talks in the regulatory landscape of miRNAs in breast cancer. Molecular cancer research : MCR. 2013;11(4):315-28.
270. Svoronos AA, Engelman DM, Slack FJ. OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. Cancer research. 2016;76(13):3666-70.
271. Dal Bo M, Bomben R, Hernandez L, Gattei V. The MYC/miR-17-92 axis in lymphoproliferative disorders: A common pathway with therapeutic potential. Oncotarget. 2015;6(23):19381-92.
272. Mendell JT. miRiad roles for the miR-17-92 cluster in development and disease. Cell. 2008;133(2):217-22.
273. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature. 2005;435(7043):839-43.
274. Li Y, Choi PS, Casey SC, Dill DL, Felsher DW. MYC through miR-17-92 suppresses specific target genes to maintain survival, autonomous proliferation, and a neoplastic state. Cancer cell. 2014;26(2):262-72.

275. Medina PP, Slack FJ. microRNAs and cancer: an overview. Cell cycle. 2008;7(16):2485-92.
276. Jiang S, Zhang HW, Lu MH, He XH, Li Y, Gu H, et al. MicroRNA-155 functions as an OncomiR in breast cancer by targeting the suppressor of cytokine signaling 1 gene. Cancer research. 2010;70(8):3119-27.
277. Chen Z, Ma T, Huang C, Hu T, Li J. The pivotal role of microRNA-155 in the control of cancer. Journal of cellular physiology. 2014;229(5):545-50.
278. Iorio MV, Visone R, Di Leva G, Donati V, Petrocca F, Casalini P, et al. MicroRNA signatures in human ovarian cancer. Cancer research. 2007;67(18):8699-707.
279. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, et al. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. Jama. 2008;299(4):425-36.
280. Si ML, Zhu S, Wu H, Lu Z, Wu F, Mo YY. miR-21-mediated tumor growth. Oncogene. 2007;26(19):2799-803.
281. Corsten MF, Miranda R, Kasmieh R, Krichevsky AM, Weissleder R, Shah K. MicroRNA-21 knockdown disrupts glioma growth in vivo and displays synergistic cytotoxicity with neural precursor cell delivered S-TRAIL in human gliomas. Cancer research. 2007;67(19):8994-9000.
282. Zhu S, Wu H, Wu F, Nie D, Sheng S, Mo YY. MicroRNA-21 targets tumor suppressor genes in invasion and metastasis. Cell research. 2008;18(3):350-9.
283. Asangani IA, Rasheed SA, Nikolova DA, Leupold JH, Colburn NH, Post S, et al. MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. Oncogene. 2008;27(15):2128-36.
284. Jazbutyte V, Thum T. MicroRNA-21: from cancer to cardiovascular disease. Current drug targets. 2010;11(8):926-35.
285. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, et al. RAS is regulated by the let-7 microRNA family. Cell. 2005;120(5):635-47.
286. Viswanathan SR, Powers JT, Einhorn W, Hoshida Y, Ng TL, Toffanin S, et al. Lin28 promotes transformation and is associated with advanced human malignancies. Nature genetics. 2009;41(7):843-8.
287. Yu F, Yao H, Zhu P, Zhang X, Pan Q, Gong C, et al. let-7 regulates self renewal and tumorigenicity of breast cancer cells. Cell. 2007;131(6):1109-23.
288. Chang TC, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KH, et al. Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. Molecular cell. 2007;26(5):745-52.
289. Okada N, Lin CP, Ribeiro MC, Biton A, Lai G, He X, et al. A positive feedback between p53 and miR-34 miRNAs mediates tumor suppression. Genes & development. 2014;28(5):438-50.
290. Lal A, Thomas MP, Altschuler G, Navarro F, O'Day E, Li XL, et al. Capture of microRNA-bound mRNAs identifies the tumor suppressor miR-34a as a regulator of growth factor signaling. PLoS genetics. 2011;7(11):e1002363.
291. Hahn S, Jackstadt R, Siemens H, Hunten S, Hermeking H. SNAIL and miR-34a feed-forward regulation of ZNF281/ZBP99 promotes epithelial-mesenchymal transition. The EMBO journal. 2013;32(23):3079-95.
292. Siemens H, Jackstadt R, Hunten S, Kaller M, Menssen A, Gotz U, et al. miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. Cell cycle. 2011;10(24):4256-71.
293. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, et al. MicroRNA gene expression deregulation in human breast cancer. Cancer research. 2005;65(16):7065-70.
294. Adams BD, Furneaux H, White BA. The micro-ribonucleic acid (miRNA) miR-206 targets the human estrogen receptor-alpha (ERalpha) and represses ERalpha messenger RNA and protein expression in breast cancer cell lines. Molecular endocrinology. 2007;21(5):1132-47.
295. Zhao JJ, Lin J, Yang H, Kong W, He L, Ma X, et al. MicroRNA-221/222 negatively regulates estrogen receptor alpha and is associated with tamoxifen resistance in breast cancer. The Journal of biological chemistry. 2008;283(45):31079-86.
296. Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin SF, Dunning MJ, et al. MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. Genome biology. 2007;8(10):R214.

297. Buffa FM, Camps C, Winchester L, Snell CE, Gee HE, Sheldon H, et al. microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. Cancer research. 2011;71(17):5635-45.
298. Bockhorn J, Dalton R, Nwachukwu C, Huang S, Prat A, Yee K, et al. MicroRNA-30c inhibits human breast tumour chemotherapy resistance by regulating TWF1 and IL-11. Nature communications. 2013;4:1393.
299. Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, Spaderna S, et al. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. EMBO reports. 2008;9(6):582-9.
300. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nature cell biology. 2008;10(5):593-601.
301. Korpal M, Lee ES, Hu G, Kang Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. The Journal of biological chemistry. 2008;283(22):14910-4.
302. Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes & development. 2008;22(7):894-907.
303. Shimono Y, Zabala M, Cho RW, Lobo N, Dalerba P, Qian D, et al. Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. Cell. 2009;138(3):592-603.
304. Lim YY, Wright JA, Attema JL, Gregory PA, Bert AG, Smith E, et al. Epigenetic modulation of the miR-200 family is associated with transition to a breast cancer stem-cell-like state. Journal of cell science. 2013;126(Pt 10):2256-66.
305. Sempere LF, Christensen M, Silahtaroglu A, Bak M, Heath CV, Schwartz G, et al. Altered MicroRNA expression confined to specific epithelial cell subpopulations in breast cancer. Cancer research. 2007;67(24):11612-20.
306. Baffa R, Fassan M, Volinia S, O'Hara B, Liu CG, Palazzo JP, et al. MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. The Journal of pathology. 2009;219(2):214-21.
307. Wu H, Zhu S, Mo YY. Suppression of cell growth and invasion by miR-205 in breast cancer. Cell research. 2009;19(4):439-48.
308. Greene SB, Gunaratne PH, Hammond SM, Rosen JM. A putative role for microRNA-205 in mammary epithelial cell progenitors. Journal of cell science. 2010;123(Pt 4):606-18.
309. Chao CH, Chang CC, Wu MJ, Ko HW, Wang D, Hung MC, et al. MicroRNA-205 signaling regulates mammary stem cell fate and tumorigenesis. The Journal of clinical investigation. 2014;124(7):3093-106.
310. Han M, Liu M, Wang Y, Chen X, Xu J, Sun Y, et al. Antagonism of miR-21 reverses epithelial-mesenchymal transition and cancer stem cell phenotype through AKT/ERK1/2 inactivation by targeting PTEN. PloS one. 2012;7(6):e39520.
311. Han M, Liu M, Wang Y, Mo Z, Bi X, Liu Z, et al. Re-expression of miR-21 contributes to migration and invasion by inducing epithelial-mesenchymal transition consistent with cancer stem cell characteristics in MCF-7 cells. Molecular and cellular biochemistry. 2012;363(1-2):427-36.
312. Zhang J, Tian XJ, Zhang H, Teng Y, Li R, Bai F, et al. TGF-beta-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. Science signaling. 2014;7(345):ra91.
313. Gwak JM, Kim HJ, Kim EJ, Chung YR, Yun S, Seo AN, et al. MicroRNA-9 is associated with epithelial-mesenchymal transition, breast cancer stem cell phenotype, and tumor progression in breast cancer. Breast cancer research and treatment. 2014;147(1):39-49.
314. Barh D, Malhotra R, Ravi B, Sindhurani P. MicroRNA let-7: an emerging next-generation cancer therapeutic. Current oncology. 2010;17(1):70-80.
315. Guo L, Chen C, Shi M, Wang F, Chen X, Diao D, et al. Stat3-coordinated Lin-28-let-7-HMGA2 and miR-200-ZEB1 circuits initiate and maintain oncostatin M-driven epithelial-mesenchymal transition. Oncogene. 2013;32(45):5272-82.
316. Yu F, Deng H, Yao H, Liu Q, Su F, Song E. Mir-30 reduction maintains self-renewal and inhibits apoptosis in breast tumor-initiating cells. Oncogene. 2010;29(29):4194-204.
317. Ouzounova M, Vuong T, Ancey PB, Ferrand M, Durand G, Le-Calvez Kelm F, et al. MicroRNA miR-30 family regulates non-attachment growth of breast cancer cells. BMC genomics. 2013;14:139.

318. Zhu Y, Yu F, Jiao Y, Feng J, Tang W, Yao H, et al. Reduced miR-128 in breast tumor-initiating cells induces chemotherapeutic resistance via Bmi-1 and ABCC5. Clinical cancer research : an official journal of the American Association for Cancer Research. 2011;17(22):7105-15.
319. Qian P, Banerjee A, Wu ZS, Zhang X, Wang H, Pandey V, et al. Loss of SNAIL regulated miR-128-2 on chromosome 3p22.3 targets multiple stem cell factors to promote transformation of mammary epithelial cells. Cancer research. 2012;72(22):6036-50.
320. Yu F, Jiao Y, Zhu Y, Wang Y, Zhu J, Cui X, et al. MicroRNA 34c gene down-regulation via DNA methylation promotes self-renewal and epithelial-mesenchymal transition in breast tumor-initiating cells. The Journal of biological chemistry. 2012;287(1):465-73.
321. Okuda H, Xing F, Pandey PR, Sharma S, Watabe M, Pai SK, et al. miR-7 suppresses brain metastasis of breast cancer stem-like cells by modulating KLF4. Cancer research. 2013;73(4):1434-44.
322. Wang Y, Yu Y, Tsuyada A, Ren X, Wu X, Stubblefield K, et al. Transforming growth factor-beta regulates the sphere-initiating stem cell-like feature in breast cancer through miRNA-181 and ATM. Oncogene. 2011;30(12):1470-80.
323. Hwang-Verslues WW, Chang PH, Wei PC, Yang CY, Huang CK, Kuo WH, et al. miR-495 is upregulated by E12/E47 in breast cancer stem cells, and promotes oncogenesis and hypoxia resistance via downregulation of E-cadherin and REDD1. Oncogene. 2011;30(21):2463-74.
324. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, et al. Endogenous human microRNAs that suppress breast cancer metastasis. Nature. 2008;451(7175):147-52.
325. Dohi O, Yasui K, Gen Y, Takada H, Endo M, Tsuji K, et al. Epigenetic silencing of miR-335 and its host gene MEST in hepatocellular carcinoma. International journal of oncology. 2013;42(2):411-8.
326. Donnem T, Lonvik K, Eklo K, Berg T, Sorbye SW, Al-Shibli K, et al. Independent and tissue-specific prognostic impact of miR-126 in nonsmall cell lung cancer: coexpression with vascular endothelial growth factor-A predicts poor survival. Cancer. 2011;117(14):3193-200.
327. Png KJ, Yoshida M, Zhang XH, Shu W, Lee H, Rimner A, et al. MicroRNA-335 inhibits tumor reinitiation and is silenced through genetic and epigenetic mechanisms in human breast cancer. Genes & development. 2011;25(3):226-31.
328. White NM, Bao TT, Grigull J, Youssef YM, Girgis A, Diamandis M, et al. miRNA profiling for clear cell renal cell carcinoma: biomarker discovery and identification of potential controls and consequences of miRNA dysregulation. The Journal of urology. 2011;186(3):1077-83.
329. Xu Y, Zhao F, Wang Z, Song Y, Luo Y, Zhang X, et al. MicroRNA-335 acts as a metastasis suppressor in gastric cancer by targeting Bcl-w and specificity protein 1. Oncogene. 2012;31(11):1398-407.
330. Png KJ, Halberg N, Yoshida M, Tavazoie SF. A microRNA regulon that mediates endothelial recruitment and metastasis by cancer cells. Nature. 2012;481(7380):190-4.
331. Hamada S, Satoh K, Fujibuchi W, Hirota M, Kanno A, Unno J, et al. MiR-126 acts as a tumor suppressor in pancreatic cancer cells via the regulation of ADAM9. Molecular cancer research : MCR. 2012;10(1):3-10.
332. Taulli R, Bersani F, Foglizzo V, Linari A, Vigna E, Ladanyi M, et al. The muscle-specific microRNA miR-206 blocks human rhabdomyosarcoma growth in xenotransplanted mice by promoting myogenic differentiation. The Journal of clinical investigation. 2009;119(8):2366-78.
333. Bhaumik D, Scott GK, Schokrpur S, Patil CK, Campisi J, Benz CC. Expression of microRNA-146 suppresses NF-kappaB activity with reduction of metastatic potential in breast cancer cells. Oncogene. 2008;27(42):5643-7.
334. Hurst DR, Edmonds MD, Scott GK, Benz CC, Vaidya KS, Welch DR. Breast cancer metastasis suppressor 1 up-regulates miR-146, which suppresses breast cancer metastasis. Cancer research. 2009;69(4):1279-83.
335. Valastyan S, Benaich N, Chang A, Reinhardt F, Weinberg RA. Concomitant suppression of three target genes can explain the impact of a microRNA on metastasis. Genes & development. 2009;23(22):2592-7.
336. Valastyan S, Chang A, Benaich N, Reinhardt F, Weinberg RA. Activation of miR-31 function in already-established metastases elicits metastatic regression. Genes & development. 2011;25(6):646-59.
337. Valastyan S, Reinhardt F, Benaich N, Calogrias D, Szasz AM, Wang ZC, et al. A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. Cell. 2009;137(6):1032-46.
338. Valastyan S. Roles of microRNAs and other non-coding RNAs in breast cancer metastasis. Journal of mammary gland biology and neoplasia. 2012;17(1):23-32.

339. Sossey-Alaoui K, Bialkowska K, Plow EF. The miR200 family of microRNAs regulates WAVE3-dependent cancer cell invasion. The Journal of biological chemistry. 2009;284(48):33019-29.
340. Jurmeister S, Baumann M, Balwierz A, Keklikoglou I, Ward A, Uhlmann S, et al. MicroRNA-200c represses migration and invasion of breast cancer cells by targeting actin-regulatory proteins FHOD1 and PPM1F. Molecular and cellular biology. 2012;32(3):633-51.
341. Korpal M, Ell BJ, Buffa FM, Ibrahim T, Blanco MA, Celia-Terrassa T, et al. Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. Nature medicine. 2011;17(9):1101-8.
342. Olson P, Lu J, Zhang H, Shai A, Chun MG, Wang Y, et al. MicroRNA dynamics in the stages of tumorigenesis correlate with hallmark capabilities of cancer. Genes & development. 2009;23(18):2152-65.
343. Brabletz S, Bajdak K, Meidhof S, Burk U, Niedermann G, Firat E, et al. The ZEB1/miR-200 feedback loop controls Notch signalling in cancer cells. The EMBO journal. 2011;30(4):770-82.
344. Vallejo DM, Caparros E, Dominguez M. Targeting Notch signalling by the conserved miR-8/200 microRNA family in development and cancer cells. The EMBO journal. 2011;30(4):756-69.
345. . !!! INVALID CITATION !!!
346. Bracken CP, Li X, Wright JA, Lawrence DM, Pillman KA, Salmanidis M, et al. Genome-wide identification of miR-200 targets reveals a regulatory network controlling cell invasion. The EMBO journal. 2014;33(18):2040-56.
347. Ma L, Teruya-Feldstein J, Weinberg RA. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature. 2007;449(7163):682-8.
348. Ma L, Reinhardt F, Pan E, Soutschek J, Bhat B, Marcusson EG, et al. Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. Nature biotechnology. 2010;28(4):341-7.
349. Huang Q, Gumireddy K, Schrier M, le Sage C, Nagel R, Nair S, et al. The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. Nature cell biology. 2008;10(2):202-10.
350. Chen D, Sun Y, Wei Y, Zhang P, Rezaeian AH, Teruya-Feldstein J, et al. LIFR is a breast cancer metastasis suppressor upstream of the Hippo-YAP pathway and a prognostic marker. Nature medicine. 2012;18(10):1511-7.
351. Martello G, Rosato A, Ferrari F, Manfrin A, Cordenonsi M, Dupont S, et al. A MicroRNA targeting dicer for metastasis control. Cell. 2010;141(7):1195-207.
352. Zhu S, Si ML, Wu H, Mo YY. MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1). The Journal of biological chemistry. 2007;282(19):14328-36.
353. Bornachea O, Santos M, Martinez-Cruz AB, Garcia-Escudero R, Duenas M, Costa C, et al. EMT and induction of miR-21 mediate metastasis development in Trp53-deficient tumours. Scientific reports. 2012;2:434.
354. Hiyoshi Y, Kamohara H, Karashima R, Sato N, Imamura Y, Nagai Y, et al. MicroRNA-21 regulates the proliferation and invasion in esophageal squamous cell carcinoma. Clinical cancer research : an official journal of the American Association for Cancer Research. 2009;15(6):1915-22.
355. Chen Y, Liu W, Chao T, Zhang Y, Yan X, Gong Y, et al. MicroRNA-21 down-regulates the expression of tumor suppressor PDCD4 in human glioblastoma cell T98G. Cancer letters. 2008;272(2):197-205.
356. Dykxhoorn DM, Wu Y, Xie H, Yu F, Lal A, Petrocca F, et al. miR-200 enhances mouse breast cancer cell colonization to form distant metastases. PloS one. 2009;4(9):e7181.
357. Le MT, Hamar P, Guo C, Basar E, Perdigao-Henriques R, Balaj L, et al. miR-200-containing extracellular vesicles promote breast cancer cell metastasis. The Journal of clinical investigation. 2014;124(12):5109-28.
358. Gaziel-Sovran A, Segura MF, Di Micco R, Collins MK, Hanniford D, Vega-Saenz de Miera E, et al. miR-30b/30d regulation of GalNAc transferases enhances invasion and immunosuppression during metastasis. Cancer cell. 2011;20(1):104-18.
359. Pencheva N, Tran H, Buss C, Huh D, Drobnjak M, Busam K, et al. Convergent multi-miRNA targeting of ApoE drives LRP1/LRP8-dependent melanoma metastasis and angiogenesis. Cell. 2012;151(5):1068-82.
360. Pencheva N, Tavazoie SF. Control of metastatic progression by microRNA regulatory networks. Nature cell biology. 2013;15(6):546-54.
361. Pradervand S, Weber J, Thomas J, Bueno M, Wirapati P, Lefort K, et al. Impact of normalization on miRNA microarray expression profiling. Rna. 2009;15(3):493-501.

362. Querzoli P, Coradini D, Pedriali M, Boracchi P, Ambrogi F, Raimondi E, et al. An immunohistochemically positive E-cadherin status is not always predictive for a good prognosis in human breast cancer. British journal of cancer. 2010;103(12):1835-9.
363. Dontu G, Abdallah WM, Foley JM, Jackson KW, Clarke MF, Kawamura MJ, et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes & development. 2003;17(10):1253-70.
364. Lee GY, Kenny PA, Lee EH, Bissell MJ. Three-dimensional culture models of normal and malignant breast epithelial cells. Nature methods. 2007;4(4):359-65.
365. Li X, Roslan S, Johnstone CN, Wright JA, Bracken CP, Anderson M, et al. MiR-200 can repress breast cancer metastasis through ZEB1-independent but moesin-dependent pathways. Oncogene. 2014;33(31):4077-88.
366. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, et al. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer cell. 2006;10(6):515-27.
367. Minn AJ, Kang Y, Serganova I, Gupta GP, Giri DD, Doubrovin M, et al. Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. The Journal of clinical investigation. 2005;115(1):44-55.
368. Celia-Terrassa T, Meca-Cortes O, Mateo F, de Paz AM, Rubio N, Arnal-Estape A, et al. Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. The Journal of clinical investigation. 2012;122(5):1849-68.
369. Dawson PJ, Wolman SR, Tait L, Heppner GH, Miller FR. MCF10AT: a model for the evolution of cancer from proliferative breast disease. Am J Pathol. 1996;148(1):313-9.
370. Santner SJ, Dawson PJ, Tait L, Soule HD, Eliason J, Mohamed AN, et al. Malignant MCF10CA1 cell lines derived from premalignant human breast epithelial MCF10AT cells. Breast cancer research and treatment. 2001;65(2):101-10.
371. Ma I, Allan AL. The role of human aldehyde dehydrogenase in normal and cancer stem cells. Stem cell reviews. 2011;7(2):292-306.
372. Eirew P, Kannan N, Knapp DJ, Vaillant F, Emerman JT, Lindeman GJ, et al. Aldehyde dehydrogenase activity is a biomarker of primitive normal human mammary luminal cells. Stem Cells. 2012;30(2):344-8.
373. Sleeman KE, Kendrick H, Ashworth A, Isacke CM, Smalley MJ. CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. Breast cancer research : BCR. 2006;8(1):R7.
374. Vaillant F, Lindeman GJ, Visvader JE. Jekyll or Hyde: does Matrigel provide a more or less physiological environment in mammary repopulating assays? Breast cancer research : BCR. 2011;13(3):108.
375. Tan PH, Ellis IO. Myoepithelial and epithelial-myoepithelial, mesenchymal and fibroepithelial breast lesions: updates from the WHO Classification of Tumours of the Breast 2012. J Clin Pathol. 2013;66(6):465-70.
376. Zhang Y, Toy KA, Kleer CG. Metaplastic breast carcinomas are enriched in markers of tumor-initiating cells and epithelial to mesenchymal transition. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc. 2012;25(2):178-84.
377. Cooper CL, Karim RZ, Selinger C, Carmalt H, Lee CS, O'Toole SA. Molecular alterations in metaplastic breast carcinoma. Journal of clinical pathology. 2013;66(6):522-8.
378. Castilla MA, Diaz-Martin J, Sarrio D, Romero-Perez L, Lopez-Garcia MA, Vieites B, et al. MicroRNA-200 family modulation in distinct breast cancer phenotypes. PloS one. 2012;7(10):e47709.
379. Fu J, Xu X, Kang L, Zhou L, Wang S, Lu J, et al. miR-30a suppresses breast cancer cell proliferation and migration by targeting Eya2. Biochemical and biophysical research communications. 2014;445(2):314-9.
380. Xiong J, Wei B, Ye Q, Liu W. MiR-30a-5p/UBE3C axis regulates breast cancer cell proliferation and migration. Biochemical and biophysical research communications. 2016.
381. Chang CW, Yu JC, Hsieh YH, Yao CC, Chao JI, Chen PM, et al. MicroRNA-30a increases tight junction protein expression to suppress the epithelial-mesenchymal transition and metastasis by targeting Slug in breast cancer. Oncotarget. 2016;7(13):16462-78.
382. Zhang N, Wang X, Huo Q, Sun M, Cai C, Liu Z, et al. MicroRNA-30a suppresses breast tumor growth and metastasis by targeting metadherin. Oncogene. 2014;33(24):3119-28.

383. Kumarswamy R, Mudduluru G, Ceppi P, Muppala S, Kozlowski M, Niklinski J, et al. MicroRNA-30a inhibits epithelial-to-mesenchymal transition by targeting Snai1 and is downregulated in non-small cell lung cancer. International journal of cancer. 2012;130(9):2044-53.
384. Qi F, He T, Jia L, Song N, Guo L, Ma X, et al. The miR-30 Family Inhibits Pulmonary Vascular Hyperpermeability in the Premetastatic Phase by Direct Targeting of Skp2. Clinical cancer research : an official journal of the American Association for Cancer Research. 2015;21(13):3071-80.
385. Wang W, Lin H, Zhou L, Zhu Q, Gao S, Xie H, et al. MicroRNA-30a-3p inhibits tumor proliferation, invasiveness and metastasis and is downregulated in hepatocellular carcinoma. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology. 2014;40(11):1586-94.
386. Zhong M, Bian Z, Wu Z. miR-30a suppresses cell migration and invasion through downregulation of PIK3CD in colorectal carcinoma. Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology. 2013;31(2-3):209-18.
387. Takahashi M, Cuatrecasas M, Balaguer F, Hur K, Toiyama Y, Castells A, et al. The clinical significance of MiR-148a as a predictive biomarker in patients with advanced colorectal cancer. PloS one. 2012;7(10):e46684.
388. Hanoun N, Delpu Y, Suriawinata AA, Bournet B, Bureau C, Selves J, et al. The silencing of microRNA 148a production by DNA hypermethylation is an early event in pancreatic carcinogenesis. Clinical chemistry. 2010;56(7):1107-18.
389. Wijnhoven BP, Hussey DJ, Watson DI, Tsykin A, Smith CM, Michael MZ, et al. MicroRNA profiling of Barrett's oesophagus and oesophageal adenocarcinoma. The British journal of surgery. 2010;97(6):853-61.
390. Li L, Chen YY, Li SQ, Huang C, Qin YZ. Expression of miR-148/152 family as potential biomarkers in non-small-cell lung cancer. Medical science monitor : international medical journal of experimental and clinical research. 2015;21:1155-61.
391. Chen Y, Song Y, Wang Z, Yue Z, Xu H, Xing C, et al. Altered expression of MiR-148a and MiR-152 in gastrointestinal cancers and its clinical significance. Journal of gastrointestinal surgery : official journal of the Society for Surgery of the Alimentary Tract. 2010;14(7):1170-9.
392. Sun J, Song Y, Wang Z, Wang G, Gao P, Chen X, et al. Clinical significance of promoter region hypermethylation of microRNA-148a in gastrointestinal cancers. OncoTargets and therapy. 2014;7:853-63.
393. Xu Q, Jiang Y, Yin Y, Li Q, He J, Jing Y, et al. A regulatory circuit of miR-148a/152 and DNMT1 in modulating cell transformation and tumor angiogenesis through IGF-IR and IRS1. Journal of molecular cell biology. 2013;5(1):3-13.
394. Lei Y, Li B, Tong S, Qi L, Hu X, Cui Y, et al. miR-101 suppresses vascular endothelial growth factor C that inhibits migration and invasion and enhances cisplatin chemosensitivity of bladder cancer cells. PloS one. 2015;10(2):e0117809.
395. Liu Z, Wang J, Mao Y, Zou B, Fan X. MicroRNA-101 suppresses migration and invasion via targeting vascular endothelial growth factor-C in hepatocellular carcinoma cells. Oncology letters. 2016;11(1):433-8.
396. Varambally S, Cao Q, Mani RS, Shankar S, Wang X, Ateeq B, et al. Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer. Science. 2008;322(5908):1695-9.
397. Zheng F, Liao YJ, Cai MY, Liu TH, Chen SP, Wu PH, et al. Systemic delivery of microRNA-101 potently inhibits hepatocellular carcinoma in vivo by repressing multiple targets. PLoS genetics. 2015;11(2):e1004873.
398. Li JT, Jia LT, Liu NN, Zhu XS, Liu QQ, Wang XL, et al. MiRNA-101 inhibits breast cancer growth and metastasis by targeting CX chemokine receptor 7. Oncotarget. 2015;6(31):30818-30.
399. Bahena-Ocampo I, Espinosa M, Ceballos-Cancino G, Lizarraga F, Campos-Arroyo D, Schwarz A, et al. miR-10b expression in breast cancer stem cells supports self-renewal through negative PTEN regulation and sustained AKT activation. EMBO reports. 2016;17(7):1081.
400. Bisso A, Faleschini M, Zampa F, Capaci V, De Santa J, Santarpia L, et al. Oncogenic miR-181a/b affect the DNA damage response in aggressive breast cancer. Cell cycle. 2013;12(11):1679-87.
401. Neel JC, Lebrun JJ. Activin and TGFbeta regulate expression of the microRNA-181 family to promote cell migration and invasion in breast cancer cells. Cellular signalling. 2013;25(7):1556-66.

402. Taylor MA, Sossey-Alaoui K, Thompson CL, Danielpour D, Schiemann WP. TGF-beta upregulates miR-181a expression to promote breast cancer metastasis. The Journal of clinical investigation. 2013;123(1):150-63.
403. Ji J, Yamashita T, Budhu A, Forgues M, Jia HL, Li C, et al. Identification of microRNA-181 by genome-wide screening as a critical player in EpCAM-positive hepatic cancer stem cells. Hepatology. 2009;50(2):472-80.
404. Pichiorri F, Suh SS, Ladetto M, Kuehl M, Palumbo T, Drandi D, et al. MicroRNAs regulate critical genes associated with multiple myeloma pathogenesis. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(35):12885-90.
405. Pichler M, Winter E, Ress AL, Bauernhofer T, Gerger A, Kiesslich T, et al. miR-181a is associated with poor clinical outcome in patients with colorectal cancer treated with EGFR inhibitor. Journal of clinical pathology. 2014;67(3):198-203.
406. Wang B, Hsu SH, Majumder S, Kutay H, Huang W, Jacob ST, et al. TGFbeta-mediated upregulation of hepatic miR-181b promotes hepatocarcinogenesis by targeting TIMP3. Oncogene. 2010;29(12):1787-97.
407. Yang CC, Hung PS, Wang PW, Liu CJ, Chu TH, Cheng HW, et al. miR-181 as a putative biomarker for lymph-node metastasis of oral squamous cell carcinoma. Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology. 2011;40(5):397-404.
408. Tang Y, Zhou X, Ji J, Chen L, Cao J, Luo J, et al. High expression levels of miR-21 and miR-210 predict unfavorable survival in breast cancer: a systemic review and meta-analysis. The International journal of biological markers. 2015;30(4):e347-58.
409. Wang J, Zhao J, Shi M, Ding Y, Sun H, Yuan F, et al. Elevated expression of miR-210 predicts poor survival of cancer patients: a systematic review and meta-analysis. PloS one. 2014;9(2):e89223.
410. Xie X, Wu W, Liang L, Han S, Chen T, Pan S, et al. Prognostic role of microRNA-210 in various carcinomas: a meta-analysis. International journal of clinical and experimental medicine. 2015;8(9):15283-9.
411. Foekens JA, Sieuwerts AM, Smid M, Look MP, de Weerd V, Boersma AW, et al. Four miRNAs associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(35):13021-6.
412. Rothe F, Ignatiadis M, Chaboteaux C, Haibe-Kains B, Kheddoumi N, Majjaj S, et al. Global microRNA expression profiling identifies MiR-210 associated with tumor proliferation, invasion and poor clinical outcome in breast cancer. PloS one. 2011;6(6):e20980.
413. Chan SY, Loscalzo J. MicroRNA-210: a unique and pleiotropic hypoxamir. Cell cycle. 2010;9(6):1072-83.
414. Devlin C, Greco S, Martelli F, Ivan M. miR-210: More than a silent player in hypoxia. IUBMB life. 2011;63(2):94-100.
415. Ivan M, Huang X. miR-210: fine-tuning the hypoxic response. Advances in experimental medicine and biology. 2014;772:205-27.
416. Vaupel P, Mayer A. Hypoxia in cancer: significance and impact on clinical outcome. Cancer metastasis reviews. 2007;26(2):225-39.
417. Fasanaro P, D'Alessandra Y, Di Stefano V, Melchionna R, Romani S, Pompilio G, et al. MicroRNA-210 modulates endothelial cell response to hypoxia and inhibits the receptor tyrosine kinase ligand Ephrin-A3. The Journal of biological chemistry. 2008;283(23):15878-83.
418. Hua Z, Lv Q, Ye W, Wong CK, Cai G, Gu D, et al. MiRNA-directed regulation of VEGF and other angiogenic factors under hypoxia. PloS one. 2006;1:e116.
419. Tadokoro H, Umezu T, Ohyashiki K, Hirano T, Ohyashiki JH. Exosomes derived from hypoxic leukemia cells enhance tube formation in endothelial cells. The Journal of biological chemistry. 2013;288(48):34343-51.
420. Cui H, Grosso S, Schelter F, Mari B, Kruger A. On the Pro-Metastatic Stress Response to Cancer Therapies: Evidence for a Positive Co-Operation between TIMP-1, HIF-1alpha, and miR-210. Frontiers in pharmacology. 2012;3:134.
421. Qin Q, Furong W, Baosheng L. Multiple functions of hypoxia-regulated miR-210 in cancer. Journal of experimental & clinical cancer research : CR. 2014;33:50.
422. Qu A, Du L, Yang Y, Liu H, Li J, Wang L, et al. Hypoxia-inducible MiR-210 is an independent prognostic factor and contributes to metastasis in colorectal cancer. PloS one. 2014;9(3):e90952.

423. Ying Q, Liang L, Guo W, Zha R, Tian Q, Huang S, et al. Hypoxia-inducible microRNA-210 augments the metastatic potential of tumor cells by targeting vacuole membrane protein 1 in hepatocellular carcinoma. Hepatology. 2011;54(6):2064-75.
424. Zhang Y, Yang P, Wang XF. Microenvironmental regulation of cancer metastasis by miRNAs. Trends in cell biology. 2014;24(3):153-60.
425. Grosso S, Doyen J, Parks SK, Bertero T, Paye A, Cardinaud B, et al. MiR-210 promotes a hypoxic phenotype and increases radioresistance in human lung cancer cell lines. Cell death & disease. 2013;4:e544.
426. Yang W, Sun T, Cao J, Liu F, Tian Y, Zhu W. Downregulation of miR-210 expression inhibits proliferation, induces apoptosis and enhances radiosensitivity in hypoxic human hepatoma cells in vitro. Experimental cell research. 2012;318(8):944-54.
427. Yang W, Wei J, Guo T, Shen Y, Liu F. Knockdown of miR-210 decreases hypoxic glioma stem cells stemness and radioresistance. Experimental cell research. 2014;326(1):22-35.
428. Chen J, Wang W, Zhang Y, Chen Y, Hu T. Predicting distant metastasis and chemoresistance using plasma miRNAs. Medical oncology. 2014;31(1):799.
429. Madhavan D, Zucknick M, Wallwiener M, Cuk K, Modugno C, Scharpff M, et al. Circulating miRNAs as surrogate markers for circulating tumor cells and prognostic markers in metastatic breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2012;18(21):5972-82.
430. Shen J, Todd NW, Zhang H, Yu L, Lingxiao X, Mei Y, et al. Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. Laboratory investigation; a journal of technical methods and pathology. 2011;91(4):579-87.
431. Jung EJ, Santarpia L, Kim J, Esteva FJ, Moretti E, Buzdar AU, et al. Plasma microRNA 210 levels correlate with sensitivity to trastuzumab and tumor presence in breast cancer patients. Cancer. 2012;118(10):2603-14.
432. Bezman NA, Chakraborty T, Bender T, Lanier LL. miR-150 regulates the development of NK and iNKT cells. The Journal of experimental medicine. 2011;208(13):2717-31.
433. Xu WD, Lu MM, Pan HF, Ye DQ. Association of MicroRNA-146a with autoimmune diseases. Inflammation. 2012;35(4):1525-9.
434. Andreopoulos B, Anastassiou D. Integrated Analysis Reveals hsa-miR-142 as a Representative of a Lymphocyte-Specific Gene Expression and Methylation Signature. Cancer informatics. 2012;11:61-75.
435. Gravgaard KH, Lyng MB, Laenkholm AV, Sokilde R, Nielsen BS, Litman T, et al. The miRNA-200 family and miRNA-9 exhibit differential expression in primary versus corresponding metastatic tissue in breast cancer. Breast cancer research and treatment. 2012;134(1):207-17.
436. Madhavan D, Peng C, Wallwiener M, Zucknick M, Nees J, Schott S, et al. Circulating miRNAs with prognostic value in metastatic breast cancer and for early detection of metastasis. Carcinogenesis. 2016;37(5):461-70.
437. Riethdorf S, Fritsche H, Muller V, Rau T, Schindlbeck C, Rack B, et al. Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: a validation study of the CellSearch system. Clinical cancer research : an official journal of the American Association for Cancer Research. 2007;13(3):920-8.
438. Brase JC, Johannes M, Schlomm T, Falth M, Haese A, Steuber T, et al. Circulating miRNAs are correlated with tumor progression in prostate cancer. International journal of cancer. 2011;128(3):608-16.
439. Toiyama Y, Hur K, Tanaka K, Inoue Y, Kusunoki M, Boland CR, et al. Serum miR-200c is a novel prognostic and metastasis-predictive biomarker in patients with colorectal cancer. Annals of surgery. 2014;259(4):735-43.
440. Eramo A, Lotti F, Sette G, Pilozzi E, Biffoni M, Di Virgilio A, et al. Identification and expansion of the tumorigenic lung cancer stem cell population. Cell death and differentiation. 2008;15(3):504-14.
441. Huang EH, Hynes MJ, Zhang T, Ginestier C, Dontu G, Appelman H, et al. Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis. Cancer research. 2009;69(8):3382-9.
442. Kryczek I, Liu S, Roh M, Vatan L, Szeliga W, Wei S, et al. Expression of aldehyde dehydrogenase and CD133 defines ovarian cancer stem cells. International journal of cancer. 2012;130(1):29-39.
443. Li C, Heidt DG, Dalerba P, Burant CF, Zhang L, Adsay V, et al. Identification of pancreatic cancer stem cells. Cancer research. 2007;67(3):1030-7.

444. Prince ME, Sivanandan R, Kaczorowski A, Wolf GT, Kaplan MJ, Dalerba P, et al. Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(3):973-8.
445. Conley SJ, Gheordunescu E, Kakarala P, Newman B, Korkaya H, Heath AN, et al. Antiangiogenic agents increase breast cancer stem cells via the generation of tumor hypoxia. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(8):2784-9.
446. Liu S, Cong Y, Wang D, Sun Y, Deng L, Liu Y, et al. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. Stem cell reports. 2014;2(1):78-91.
447. Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, et al. Molecular definition of breast tumor heterogeneity. Cancer cell. 2007;11(3):259-73.
448. Villadsen R. In search of a stem cell hierarchy in the human breast and its relevance to breast cancer evolution. APMIS : acta pathologica, microbiologica, et immunologica Scandinavica. 2005;113(11-12):903-21.
449. Balic M, Lin H, Young L, Hawes D, Giuliano A, McNamara G, et al. Most early disseminated cancer cells detected in bone marrow of breast cancer patients have a putative breast cancer stem cell phenotype. Clinical cancer research : an official journal of the American Association for Cancer Research. 2006;12(19):5615-21.
450. Muller V, Stahmann N, Riethdorf S, Rau T, Zabel T, Goetz A, et al. Circulating tumor cells in breast cancer: correlation to bone marrow micrometastases, heterogeneous response to systemic therapy and low proliferative activity. Clinical cancer research : an official journal of the American Association for Cancer Research. 2005;11(10):3678-85.
451. Blick T, Hugo H, Widodo E, Waltham M, Pinto C, Mani SA, et al. Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44(hi/)CD24 (lo/-) stem cell phenotype in human breast cancer. Journal of mammary gland biology and neoplasia. 2010;15(2):235-52.
452. Rostoker R, Abelson S, Genkin I, Ben-Shmuel S, Sachidanandam R, Scheinman EJ, et al. CD24(+) cells fuel rapid tumor growth and display high metastatic capacity. Breast cancer research : BCR. 2015;17:78.
453. Baumann P, Cremers N, Kroese F, Orend G, Chiquet-Ehrismann R, Uede T, et al. CD24 expression causes the acquisition of multiple cellular properties associated with tumor growth and metastasis. Cancer research. 2005;65(23):10783-93.
454. Bircan S, Kapucuoglu N, Baspinar S, Inan G, Candir O. CD24 expression in ductal carcinoma in situ and invasive ductal carcinoma of breast: an immunohistochemistry-based pilot study. Pathology, research and practice. 2006;202(8):569-76.
455. Kristiansen G, Winzer KJ, Mayordomo E, Bellach J, Schluns K, Denkert C, et al. CD24 expression is a new prognostic marker in breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2003;9(13):4906-13.
456. Mylona E, Giannopoulou I, Fasomytakis E, Nomikos A, Magkou C, Bakarakos P, et al. The clinicopathologic and prognostic significance of CD44+/CD24(-/low) and CD44-/CD24+ tumor cells in invasive breast carcinomas. Human pathology. 2008;39(7):1096-102.
457. Lamouille S, Subramanyam D, Blelloch R, Derynck R. Regulation of epithelial-mesenchymal and mesenchymal-epithelial transitions by microRNAs. Current opinion in cell biology. 2013;25(2):200-7.
458. Comijn J, Berx G, Vermassen P, Verschueren K, van Grunsven L, Bruyneel E, et al. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. Molecular cell. 2001;7(6):1267-78.
459. Wang G, Guo X, Hong W, Liu Q, Wei T, Lu C, et al. Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(8):2858-63.
460. Jolly MK, Boareto M, Huang B, Jia D, Lu M, Ben-Jacob E, et al. Implications of the Hybrid Epithelial/Mesenchymal Phenotype in Metastasis. Frontiers in oncology. 2015;5:155.
461. Jolly MK, Jia D, Boareto M, Mani SA, Pienta KJ, Ben-Jacob E, et al. Coupling the modules of EMT and stemness: A tunable 'stemness window' model. Oncotarget. 2015;6(28):25161-74.
462. Kim J, Villadsen R, Sorlie T, Fogh L, Gronlund SZ, Fridriksdottir AJ, et al. Tumor initiating but differentiated luminal-like breast cancer cells are highly invasive in the absence of basal-like activity. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(16):6124-9.

463. Cimino-Mathews A, Subhawong AP, Elwood H, Warzecha HN, Sharma R, Park BH, et al. Neural crest transcription factor Sox10 is preferentially expressed in triple-negative and metaplastic breast carcinomas. Human pathology. 2013;44(6):959-65.
464. Keller PJ, Arendt LM, Skibinski A, Logvinenko T, Klebba I, Dong S, et al. Defining the cellular precursors to human breast cancer. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(8):2772-7.
465. Reis-Filho JS, Milanezi F, Paredes J, Silva P, Pereira EM, Maeda SA, et al. Novel and classic myoepithelial/stem cell markers in metaplastic carcinomas of the breast. Applied immunohistochemistry & molecular morphology : AIMM / official publication of the Society for Applied Immunohistochemistry. 2003;11(1):1-8.
466. Weigelt B, Kreike B, Reis-Filho JS. Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis. Breast cancer research and treatment. 2009;117(2):273-80.
467. Ibarra I, Erlich Y, Muthuswamy SK, Sachidanandam R, Hannon GJ. A role for microRNAs in maintenance of mouse mammary epithelial progenitor cells. Genes & development. 2007;21(24):3238-43.
468. Rakha EA, Tan PH, Varga Z, Tse GM, Shaaban AM, Climent F, et al. Prognostic factors in metaplastic carcinoma of the breast: a multi-institutional study. British journal of cancer. 2015;112(2):283-9.
469. van Deurzen CH, Lee AH, Gill MS, Menke-Pluijmers MB, Jager A, Ellis IO, et al. Metaplastic breast carcinoma: tumour histogenesis or dedifferentiation? The Journal of pathology. 2011;224(4):434-7.
470. Lee H, Jung SY, Ro JY, Kwon Y, Sohn JH, Park IH, et al. Metaplastic breast cancer: clinicopathological features and its prognosis. Journal of clinical pathology. 2012;65(5):441-6.
471. McKinnon E, Xiao P. Metaplastic carcinoma of the breast. Archives of pathology & laboratory medicine. 2015;139(6):819-22.
472. Wargotz ES, Norris HJ. Metaplastic carcinomas of the breast. I. Matrix-producing carcinoma. Human pathology. 1989;20(7):628-35.
473. Wargotz ES, Norris HJ. Metaplastic carcinomas of the breast. IV. Squamous cell carcinoma of ductal origin. Cancer. 1990;65(2):272-6.
474. Wargotz ES, Deos PH, Norris HJ. Metaplastic carcinomas of the breast. II. Spindle cell carcinoma. Human pathology. 1989;20(8):732-40.
475. Zhang Y, Lv F, Yang Y, Qian X, Lang R, Fan Y, et al. Clinicopathological Features and Prognosis of Metaplastic Breast Carcinoma: Experience of a Major Chinese Cancer Center. PloS one. 2015;10(6):e0131409.
476. Chapman A, Fernandez del Ama L, Ferguson J, Kamarashev J, Wellbrock C, Hurlstone A. Heterogeneous tumor subpopulations cooperate to drive invasion. Cell reports. 2014;8(3):688-95.
477. Cleary AS, Leonard TL, Gestl SA, Gunther EJ. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. Nature. 2014;508(7494):113-7.
478. Tsuji T, Ibaragi S, Hu GF. Epithelial-mesenchymal transition and cell cooperativity in metastasis. Cancer research. 2009;69(18):7135-9.

Resumen de la tesis doctoral

INTRODUCCIÓN

1. El cáncer de mama

El cáncer de mama es un grupo muy heterogéneo de neoplasias tanto a nivel histológico como de evolución clínica. A nivel histológico se han llegado a categorizar en 18 subtipos distintos, que en términos generales puede ser dividido en carcinoma *in situ* e invasivo/infiltrante, que a su vez pueden ser ductales o lobulillares según su localización (67). Es frecuente encontrar la lesión ductal in situ (CDIS) adyacente al componente infiltrante, lo cual llevó a hipotetizar que era un estadio previo a la fase invasiva pero que no tiene por qué ser necesariamente así (75). Con frecuencia, la progresión no es lineal sino que es un proceso complejo que evoluciona a través de distintas vías (75, 76). En cuanto a los carcinomas infiltrantes, el carcinoma ductal invasivo (CDI) es el más frecuente y que por ello hemos seleccionado para nuestro estudio, abarcando un 70-80% de los casos. En cambio, el carcinoma lobulillar infiltrante (CLI) comprende un 10% y el resto pertenecen al subtipo denominado "tipos especiales" (70). El tipo histológico está asociado al pronóstico teniendo el CLI, por ejemplo, mejor pronóstico que el CDI (74).

A nivel clínico, existen varios marcadores que se determinan por immunohistoquímica para definir subtipos con distinto pronóstico e identificar tumores que puedan beneficiarse de tratamientos dirigidos. Los más comunes son el receptor de estrógenos (RE), de progesterona (RP) y el receptor 2 del factor de crecimiento epidérmico humano (HER2). La medida de expresión se da combinando el porcentaje de células tumorales que expresan la proteína y la intensidad de la tinción para determinar la positividad o negatividad. Cuando los resultados de HER2 son ambiguos por immunohistoquímica, se determina si hay amplificación del gen por hibridación *in situ* fluorescente (FISH) (78). A cada caso se le asigna una de las siguientes categorías en función del resultados: RE+ (RE+/HER2-), HER2+ (RE-/HER2+), triple negativo (TN; RE-/PR-/HER2-) y triple positivo (RE+/RP+/HER2+) (79). Des del punto de vista de pronóstico, los RE+ son los que tienen mejor evolución. Los HER2+ se asociaban a un mal pronóstico pero gracias al uso de tratamiento anti-HER2, ahora exhiben un mejor pronóstico (80, 81). Los tumores con peor pronósitco son los triple negativos mientras que los triple positivos tienen un pronóstico intermedio entre casos RE+ y HER2+ (79). El patólogo combina la información de marcadores con el subtipo histológico y el grado de diferenciación del tumor, que es un indicador potente del pronóstico basado en la combinación del grado nuclear, formación de túbulos y la tasa mitótica (77).

De cara a diagnóstico clínico también se utiliza el sistema de clasificación TNM, que permite comunicar información clínica sin ambigüedad. Éste tiene en cuenta la extensión del tumor primario (T), si hay afectación ganglionar (N) y la presencia de metástasis a distancia (M). Existen cinco estadios en los que el 0-II se considera una fase temprana en la que puede haber como máximo invasión a ganglios linfáticos. Cuando éstos se ven extensamente afectados, se pasa a estadio III y en el estadio IV ya ha metastatizado a distancia (86).

Con la llegada de las nuevas tecnologías que permitían hacer perfiles de expresión génica a gran escala, la clasificación histológica fue redefinida por las clasificaciones moleculares que recapitulan en parte las clases definidas por immunohistoquímica. La primera clasificación molecular con mayor impacto en cáncer de mama fue propuesta por Perou y colaboradores (87, 88, 97), en la que definen 4 subtipos intrínsecos: *Luminal-like, Basal-like, Normal-like* y *HER-2-enriched*. Estudios más recientes identificaron otro subtipo intrínseco denominado *Claudin-low* o *mesenchymal-like* (90). La heterogeneidad en el cáncer de mama hace difícil asociar los parámetros clinicopatológicos con los datos moleculares. Los clásicos marcadores RE, RP y HER2 usados por los patólogos no recapitulan totalmente los subtipos intrínsecos (93).

Los luminales son el subtipo más común y posteriormente se subdividieron en *Luminal A* y *Luminal B.* Ambos son además RE+, como en nuestro estudio, y expresan las citoqueratinas luminales 8 y 18. Los luminales A se asocian a menor proliferación (94, 96) y mejor pronóstico (88, 94, 97) que los luminales B, los cuales tienen menor expresión de receptores hormonales y pueden ser HER2+ (98). El resto de subtipos contiene casos predominantemente RE- (87). El subtipo molecular HER2 solapa en general con los tumores definidos clínicamente como HER2+ y son de alto grado (88). El *normal-like* recuerda al tejido epitelial normal y podría contener casos en los que las muestras contienen mucho tejido no tumoral (88, 93, 94). Los tumores *basal-like* solapan sustancialmente con los triples negativos. Suelen tener una edad más temprana de aparición y se asocian a alto grado. Tienen unas características histológicas definidas, expresan las citoqueratinas 5/6 de la capa basal epitelial y el receptor del factor de crecimiento epidérmico (EGFR) (99, 100). El *claudin-low* también es un triple negativo y se caracteriza por una alta expresión de genes mesenquimales como ZEB1 y ZEB2, así como una baja expresión de los niveles de claudinas 3/4/7 y E-cadherina (90). Parece estar enriquecido en células con propiedades biológicas asociadas a células madre mamarias y potencial para iniciar un tumor (19,85,91,92). Los tumores metaplásicos parecen pertenecer a este subtipo, el cual tiene una evolución poco favorable (92). De hecho, los *luminal A* son los que exhiben mejor pronóstico

mientras que no hay diferencias significativas de supervivencia entre el resto de subtipos con mal pronóstico como son los *luminal B, HER2-enriched, basal-like* o *claudin-low* (89).

2. MicroRNAs en Cáncer de mama

2.1 Mecanismo de acción y desregulación de los microRNAs

Los microRNAs (miRNAs) son una clase de pequeños RNAs endógenos no codificantes de unos 19-24 nucleótidos, que controlan a nivel postranscripcional la expresión génica mediante la unión complementaria a sus mRNAs diana. Cada microRNA puede controlar la expresión de cientos de mRNAs y a su vez, cada mRNA puede estar controlado por varios microRNAs distintos. Esto explica que estén implicados en una gran variedad de procesos biológicos desde la proliferación, señalización celular, diferenciación, reparación del DNA, adhesión celular y motilidad, inflamación, supervivencia o apotosis (235). En consecuencia, la alteración en su expresión está asociada a cáncer y otras enfermedades, desde el inicio a la progresión hacia la diseminación metastática.

Se ha predecido que más de la mitad del transcriptoma humano es regulado por miRNAs (256). En general, la unión se produce mediante complementariedad perfecta entre la región "seed", que comprende desde los nucleótidos 2 a 8, del miRNA y el 3'UTR del mRNA diana mientras que el resto del duplex contiene múltiples sitios que no se alinean (256,257). Existen varios mecanismos por los que pueden regular la expresión; uno muy extendido es la desestabilización del mRNA (258). El complejo entre el microRNA y las proteínas del complejo de silenciamiento inducido por ARN (RISC), puede reprimir la traducción bloqueando el inicio o un paso después, bien impidiendo la unión de ribosomas o estimulando la proteólisis del péptido (262, 263). Sin embargo, la evidencia actual sugiere que el mecanismo predominante de silenciamiento es la degradación, más que la desestabilización (259).

Además de la especificidad de la secuencia diana, hay otros factores que influencian la capacidad de un miRNA para reconocer y regular mRNAs. Uno de ellos es el contexto en el que se encuentra la secuencia diana. También la acción de RNA binding proteins (RBPs), como HuR, que pueden interferir con el complejo miRISC antagonizando su función. Además, el complejo miRISC puede asociarse a varios cofactores que promuevan o impidan su función, como los miembros de la familia TRIM-NHL como LIN41. Finalmente, modificaciones postranscripcionales en las proteínas Argonaute del complejo miRISC pueden estabilizar la unión al mRNA o lo contrario (259).

Varios factores contribuyen a la alteración de la expresión de microRNAs en cáncer. Por un lado las alteraciones genómicas pueden resultar en deleción, amplificación o translocación de los genes que codifican para miRNAs. De hecho, el 50% de éstos se localizan en regiones susceptibles de estar alteradas en cáncer (267). Por otro lado, una regulación epigenética anormal de los genes de miRNAs, ya sea por hipermetilación anormal de miRNAs que actúan como supresores tumorales, la hipometilación global del DNA o la modificación post-transcripcional de histonas (275). Muchos oncongenes o supresores tumorales son factores de transcripción y pueden activar la transcripción de miRNAs. Por ejemplo, la pérdidad de p53 provoca una disminución de los niveles de miR-34 (288). Los factores implicados en la biogénesis de microRNAs también pueden actuar como oncogenes o supresores tumorales, como es el caso de Dicer1 y Drosha, cuya baja expresión correlaciona con un peor pronóstico en múltiples cánceres (235). Por otro lado, mientras que mutaciones en la secuencia madura del miRNA son infrecuentes, las variantes genéticas en las formas precursoras de miRNAs son más frecuentes en las células cancerígenas pero no es probable que tengan una relevancia fisiológica mientras que las variaciones en la secuencia diana del mRNA a la que se une el miRNA sí tienen un impacto. Se han encontrado polimorfismos de nucleótido único (SNP) en el 3'UTR del sitio de unión en tumores.

2.2 microRNAs, células madre tumorales de mama y transición epitelio-mesénquima

Los microRNAs regulan distintos aspectos de la formación de un tumor des de su iniciación, proliferación, diferenciación y progresión. Todo ello es consecuencia de su potente implicación como reguladores de la transición epitelio-mesénquima así como de la formación de células madre cancerígenas.

Una familia crucial que regula ambos procesos es la de miR-200s. Los miembros de la familia miR-200 son inductores de la transición mesénquima-epitelio (TME) e importantes encargados de mantener el estado epitelial de las células (299-302). Se han encontrado infraexpresados en células madre tumorales de mama en humanos así como en células madre y progenitoras mamarias. miR-200c inhibe la clonogenicidad *in vitro* y tumorogenicidad *in vivo* mediante inhibición de Bmi1 (303). Por otro lado, la inhibición de miR-200b también es necesarias para la formación de células madre tumorales, lo cual permite la expresión de Suz12, el cual reprime E-cadherina (303). Además, se han encontrado infraexpresados de forma consistente en células

$CD44^{+}CD24^{-}$, definidas como céluas madre tumorales mamarias, aisladas de muestras metastásicas de pacientes con cáncer de mama (304).

En conjunto con la familia miR-200, miR-205 también es un regulador negativo de la transición epitelio-mesénquima (TEM) mediante inhibición de ZEB1 y ZEB2, reforzando en consecuencia el fenotipo epitelial (300). Se ha visto que tiene una baja expresión en líneas celulares metastáticas (305) y muestras clínicas de cáncer de mama (278). Además, se ha encontrado una expresión reducida de miR-205 en una signatura metastásica de cáncer de mama (306). Suprime invasión y metástasis a través de ErbB3 y VEGF-A (307). Al contrario, se ha observado un enriquecimiento en céulas madre epiteliales mamarias normales y que su sobreexpresión promueve la expansión de la población progenitora mediante supresión de PTEN (308). miR-205 es un determinante importante de la diferenciación de células madre mamarias y por ello su expresión anormal promueve la tumorogénesis. Se vio que una baja expresión de miR-205 altera la polaridad epitelial y la división de células madre, provocando la expansión de esta población y resultando en tumorogénesis (309).

miR-21 se encuentra sobreexpresado en la mayoría de tumores y es considerado un oncomiR (278-280). Promueve la invasión y metástasis inhibiendo los supresores tumorales TPM1, PDCD4 y maspin (282). Se ha demostrado que es un inductor de transición epitelio-mesénquima y de la capacidad de autorrenovación y clonogenicidad características del fenotipo de células madre tumoral (310).

La pérdida de p53 promueve también la TEM en células de carcinoma de mama mediante la disminución de los niveles de miR-34, que es un regulador negativo de Snail1 (180). De hecho, la activación de este bucle negativo va seguida de la activación del eje miR-200-ZEB1, reforzando la TEM inducida por TGFb (312).

miR-9 también induce una TEM mediante inbición directa de E-cadherina, lo que promueve la liberación de β-catenina y activación de VEGF, favoreciendo la angiogénesis (183). Su sobreexpresión aumenta la migración *in vitro* y la formación de metástasis pulmonares *in vivo* (183). Se ha encontrado una expresión elevada en los subtipos más agresivos, HER2+ y triple negativos, así como en tumores con un fenotipo de células madre tumoral (313).

Por el contrario, se ha visto que la familia de microRNAs let-7 se encuentra consistentemente infraexpresada en células tumorogénicas. Promueve la diferenciación y reduce las propiedades de célula madre en células madre tumorales de mama, inhibiendo la formación tumoral y capacidad metastática (316). De hecho, también está infraexpresado en nuestro estudio.

De forma similar, miR-30e también está infraexpresado en células madre tumorales de mama tanto en mamoesferas como en pacientes, impidiendo la autorrenovación inhibiendo Ubc-9 e ITGB3, lo cual reduce la tumorogénesis y metástasis (317). Otros microRNAs como miR-128 y miR-34c también se han encontrado infraexpresados en células madre tumorales de mama en tumores primarios, líneas celulares o mamoesferas (318, 319).

También alterado en nuestro estudio, miR-7 está infraexpresado en células madre tumorales con gran capacidad metastática, lo que se explica al ser un regulador negativo de KLF4, el cual es esencial en reprogramación de células pluripotentes inducidas. Además se observó una correlación inversa entre ambos en metástasis cerebrales (321).

En cambio, los miembros de la familia miR-181, que nosotros observamos sobreexpresados en metástasis a distancia, se encuentran elevados en mamoesferas con capacidad iniciadora de tumores (322).

2.3. MetastamiRs

Este término ha sido utilizado para describir aquellos microRNAs que tienen una implicación en una o varias etapas de la metástasis, aunque no necesariamente tienen que estar implicados en la tumorogénesis.

microRNAs supresores de metástasis

La expresión de miR-335, miR-206 y miR-126 se pierde específicamente en derivados de MDA-MB-231, conforme adquieren capacidad metastática (324). miR-126 suprime la formación del tumor mientras que miR-206 y miR-335 impiden la invasión; todos suprimen la formación de metástasis *in vivo*. Además, niveles bajos de estos tres microRNAs correlacionan con un periodo más corto de recidiva metastática (324). miR-335 ejerce su acción supresora regulando un conjunto de genes pro-metastáticos implicados en el control de la matriz extracelular, citoesqueleto, transducción de señal y migración (324). En cambio, miR-126 regula el microambiente metastático impidiendo el reclutamiento de células endoteliales por las células metastáticas (330).

Otro microRNA clave es miR-31, que tiene un papel en múltiples etapas del proceso metastático. Por un lado, inhibe la invasión y resistencia a anoikis mediante la represión coordinada de ITGA5, RDX y RhoA (335-337). Además, impide la supervivencia en la circulación mediante represión de

promotores de la resistencia a anoikis, la entrada en vasos sanguíneos y la formación de micrometástasis. Por último, también impide la colonización metastática (338).

También deregulados en nuestro estudio, tanto miR-146a como miR-146b inhiben migración, invasión y metástasis mediante el silenciamiento de IRAK1 y TRAF6, los cuales regulan positivamente la actividad NF-kB (334).

Finalmente, como parte central de nuestro estudio, la familia miR-200 ha sido considerada una potente supresora de metástasis al inhibir las fases iniciales de invasión tumoral. El mecanismo por el que se expica es la inhibición de proteínas que regulan la reorganización del citoesqueleto de actina, como WAVE3, las cuales promueven metástasis (339). En cáncer de mama, FHOD1 y PPM1F, reguladores directos del citoesqueleto de actina asociados al fenotipo mesenquimal, fueron validados como dianas de miR-200c (340). Sin embargo, el clásico rol de supresor tumoral le ha sido atribuido al ser el principal represor de la transición epitelio-mesénquima en favor del mantenimiento del fenotipo epitelial. El mecanismo a través del que lo consigue es mediante inhibición de reguladores master que inducen una TEM, como ZEB1 y ZEB2, que regulan negativamente la expresión de E-cadherina (300, 302, 341). La vía miR-200-ZEB1-E-cadherina se ha encontrado alterada en metástasis y en un conjunto de tumores primarios que presentaban una marcada infraexpresión de la familia miR-200 y se ha observado que la expresión ectópica de miR-200 bloque la TEM y metástasis (342). Parte de su actividad anti-metastática se explica por la regulación negativa de los ligandos de Notch Jagged1 y Jagged2 (343,344). Estudios más recientes han econtrado que miR-200 puede inhibir la migración de forma independiente de ZEB1/2 (345) y regulando múltiples dianas que actúan coordinadamente en la señalización Rho-ROCK, la formación de invadopodios, la actividad de metaloproteinasas de la matriz y adhesiones focales para mantener la morfología celular y prevenir la migración (346).

microRNAs promotores de metástasis

Uno de los primeros microRNAs identificados fue miR-10b (347) , que en nuestro estudio aparece infraexpresado en metástasis a distancia. Su expresión incrementaba la invasividad y era suficiente para promover la formación de metástasis *in vivo* mientras que este efecto podía ser bloqueado con el uso de un antagomir anti-miR-10b (348). Twist 1 induce la transcripción de miR-10b, el cual disminuye los niveles de HOXD10, que es un regulador negativo de la expresión de genes implicados en migración y remodelamiento de la matriz extracelular. En consecuencia, estimula la migración e invasión de células tumorales mamarias y de hecho, se han observado

niveles elevados de miR-10b en el frente invasivo (347). Además, se observó una correlación positiva entre niveles elevados y la presencia de metástasis en la clínica (347).

En cáncer de mama, miR-373 y miR-520c confieren potencial metastático a células no metastáticas. Se expresan en células con gran capacidad para migrar e invadir; habilidad mediada por la regulación negativa de CD44. Los niveles de miR-373 y CD44 correlacionan inversamente en muestras metastásicas (349).

Además de ser un inductor de TEM, miR-9 también promueve la migración, invasión y metástasis a través de la regulación de el receptor del factor inhibidor de leucemia (LIFR), que es un supresor de metástasis. La pérdida de LIFR se asocia al desarrollo de metástasis regionales y a distancia en pacientes con cáncer de mama invasivo (350).

La expresió de miR-103/107 también se asocia a la aparición de metástasis y una evolución no favorable en pacientes con cáncer de mama. Se identificó que su diana es *Dicer1*, la infraexpresión del cual, promueve la migración e invasión *in vitro* y la metástasis de células no metastáticas *in vivo* (351). Además, miR-103/107 inducen una transición epitelio-mesénquima y es requerido para el mantenimiento del fenotipo mesenquimal mediante supresión de la expresión de miR-200, lo cual es consecuencia de la disminución de los niveles de *Dicer1* (351).

miR-21 es un potente oncomiR implicado en multiples vías de señalización que promueve la invasión y metástasis disminuyendo los niveles de PDCD4 y tropomisoina 1 (TPM1) (282, 355).

En cuanto a la familia miR-200, ésta tiene un papel dual. Por un lado impide la fase inicial de metástasis previendo la invasión. Por otro lado, promueve la fase de colonización a distancia al inducir la transicion mesénquima-epitelio (356). Se ha descrito un mecanismo que promueve la colonización, alternativo a la reversión de la TEM, que consiste en la infraexpresión de Sec23a, el cual estimula la secreción de dos supresores de metástasis (Tinagl1 e Igfbp4), lo cual resulta en la reducción de su secreción (341). Recientemente se ha visto que la capacidad metástática puede ser transferida desde células metastáticas a células no metastáticas a través de vesículas extracelulares que contienen microRNAs miR-200s (357).

Los microRNAs controlan el proceso de metástasis regulando múltiples dianas de una forma coordinada. La mayoría de los descritos actúan de forma divergente, en la que un único microRNA puede coordinar la expresión de múltiples genes, los cuales participan en un fenotipo metastático común con la ventaja de que el efecto es más potente. Sin embargo, algunos microRNAs actúan de forma convergente para silenciar un único gen, lo cual asegura una regulación mucho más fuerte de éste (360).

RESULTADOS Y DISCUSIÓN

OBJETIVO 1: Identificación de microRNAs alterados a lo largo de la progresión del carcinoma ductal invasivo.

Con el objetivo de estudiar la implicación de los microRNAs en la progresión metastática de carcinomas ductales invasivos, el subtipo más frecuente de cáncer de mama, seleccionamos una serie de muestras fijadas en formol y parafinadas (FFPE) de carcinoma ductal *in situ* (DCIS), tumores primarios no metastáticos (PNM), tumores primarios metastáticos (PM) y sus correspondientes metástasis ganglionares (LNM) así como metástasis a distancia (DM) incluyendo además un *pool* de muestras de mama normal procedentes de reducciones mamarias (N). Se incluyó también un *pool* de ganglios linfáticos libres de carcinoma como control negativo de contaminación linfocítica para las metástasis ganglionares. Dada la heterogeneidad no solamente morfológica sino molecular del cáncer de mama, seleccionamos carcinomas ductales con un fenotipo molecular luminal, basándonos en la expresión positiva de los receptores hormonales de estrógenos (ER) y progesterona (PR) así como su negatividad para el receptor HER2 evaluados por immunohistoquímica. Para ser todavía más selectivos, enriquecimos la muestra por macrodisección del componente carcinomatoso. Posteriormente, se procedió a la hibridación de estas muestras a un microarray de 866 microRNAs humanos (Agilent, versión 12.0), de los cuales 39 microRNAs aparecían sobreexpresados y 41 microRNAs estaban infraexpresados con respecto a las muestras no tumorales.

Análisis posteriores mostraron un conjunto de 10 microRNAs con una tendencia general a aumentar y 9 microRNAs con una tendencia a disminuir a lo largo de la progresión metastática, en la transición desde la mama normal a tumor invasivo y metástasis regional o a distancia. Considerando aquellos microRNAs diferencialmente expresados entre tumores primarios con capacidad o no de metastatizar, miR-7 tenía mayor expresión en los metastáticos mientras que miR-148a, miR-30a, miR-30a* y let-7b se expresaban más en los no metastáticos. En cambio, la expresión de miR-181a, miR-181b y mir-210 aumentaba progresivamente, siendo máxima en metástasis a distancia mientras que miR-101 y miR-10b mostraban una disminución progresiva. La validación por PCR cuantitativa (qPCR) de los microRNAs más estadísticamente significativos, basándonos en el *fold-change* y el *p-value*, confirmó estas tendencias observadas por microarrays. Los niveles de miR-7, miR-210, miR-181a y miR-181b tendían a aumentar y estaban significativamente sobreexpresados en metástasis a distancia *vs* tumores primarios sin

capacidad de metastatizar. Al contrario, los niveles de miR-10b y miR-101 disminuían, siendo mínimos en metástasis a distancia. miR-30a, miR-30a* y miR-148a también mostraban una tendencia general a disminuir a lo largo de la progresión estando significativamente menos expresados en metástasis a distancia respecto de tumores no metastáticos. Ninguno de estos microRNAs discriminaba significativamente entre tumores primarios con capacidad de metástasis, la mayoría regional, y las metástasis a distancia a pesar de que éstas provenían de otros tumores primarios distintos a los analizados. Por qPCR, el conjunto formado por miR-7, miR-210, miR-200b y miR-148a discriminaban significativamente tumores metastáticos de no metastáticos siendo miR-148a el único infraexpresado en primarios metastáticos.

Además del estudio con muestras tumorales parafinadas de pacientes, también iniciamos un estudio prospectivo en el que recogimos sangre total de pacientes recién diagnosticadas que por tanto, no habían sido sometidas a ningún tratamiento previo. Las pacientes fueron clasificadas según su condición metastática en pacientes con tumor primario y afectación ganglionar (PM, N=29) o sin metástasis ganglionares (PNM, N=40). Adicionalmente, se incluyeron pacientes ya diagnosticadas y tratadas que presentaban metástasis a distancia (DM, N=14). Determinamos los niveles por qPCR de aquellos microRNAs encontrados significativamente alterados en tejido en el estudio de expresión de microarrays.

Tanto miR-200b como miR-7 mostraron una tendencia a incrementar en pacientes con metástasis ganglionares y a distancia que es estadísticamente significativa con respecto a pacientes no metastáticas. Por el contrario, ni miR-148a ni miR-30a mostraron niveles significativamente inferiores en sangre en pacientes con metástatis ganglionares.

Contrariamente a la tendencia observada en tejido, los niveles de miR-181a y miR-210 en sangre son elevados pero su expresión en pacientes con metástasis a distancia es signficativamente inferior a la del resto. De igual manera, los niveles circulantes de miR-101 son elevados y exhiben una tendencia a ser inferiores en pacientes con metástasis a distancia aunque de forma no significativa. Los niveles de miR-10b no eran suficientes como para poder ser detectado.

Los microRNAs que nosotros encontramos deregulados en carcinomas ductales ya se han econtrado asociados significativamente a la progresión metastática en cáncer de mama y otros tipos de tumores (360). miR-30 se ha encontrado igualmente infraexpresado en tumores de mama, en los que tiene un papel como inhibidor de proliferación y migración (380, 381). Dado que el factor de transcripción Slug es una de sus dianas, también podría tener un papel como inhibidor de metástasis en mama promoviendo la transición epithelio-mesénquima (MET) (381),

lo cual concuerda con nuestras observaciones. Su capacidad para inhibir invasión y metástasis se ha encontrado en otros tumores (383-386). De igual manera, también se ha observado una infraexpresión de mir-148a en varios carcinomas (387-389) además de una correlación inversa entre el grado tumoral y la capacidad de metastatizar a ganglios linfáticos en cáncer de mama (393). También let-7, cuyo papel como supresor tumoral ha sido extensamente descrito, se encontró asociado inversamente a la presencia de metástasis ganglionares en un estudio de cáncer de mama (290), en el cual miR-210 figuraba entre los principales microRNAs sobreexpresados. Por otro lado, a miR-7 se le ha atribuido tanto una función de supresor tumoral como de promotor, dependiendo del tipo tumoral (357, 358). Es interesante destacar que miR-7, que exhibe el mismo comportamiento que miR-200b en sangre y muestras tumorales, también tiene un comportamiento similar en cuanto a transición epitelio-mesénquima (EMT) se refiere. miR-7 promueve indirectamente la expresión de E-cadherina mediante inhibición de IGF1R (355 tim), FAK (359) o SETDB1 (360), lo cual resulta en una disminución de la transición EMT y un aumento de la queratinización del tumor (358, 361) que se asocia a un peor pronóstico en varios tipos de tumores (362).

En consonancia con nuestros resultados, en los que miR-101 tiene a disminuir en tejido o sangre de pacientes con metástasis a distancia, ha sido descrito com un inhibidor de migración e invasión en otros tumores (394, 395). Además, su expresión correlaciona inversamente con el estadio clínico, la presencia de metástasis ganglionares y el pronóstico en pacientes con cáncer de mama además de otros tipos (398). En contra de los que observamos, la expresión de miR-10b se correlaciona con un alto grado y capacidad metastática en varios cánceres, incluido en cáncer de mama (306, 347, 348). Tiene una expresión elevada en líneas celulares de mama metastáticas y promueve invasión y metástasis *in vivo* en células no metastáticas (347). Además, su sobreexpresión induce la autorrenovación y expresión de marcadores de EMT y células madre, como Twist (347, 399). Sin embargo, estos estudios observan niveles más elevados de miR-10b en céulas de tipo basal con un perfil molecular de triple negativo que en células con un fenotipo luminal que expresan E-cadherina, como el de los tumores que nosotros hemos analizado. Hasta ahora no se han publicado estudios en los que haya una asociación positiva entre niveles de miR-10b y progresión metastática en tumores de mama con un perfil molecular luminal, lo cual explicaría por qué nosotros encontramos una baja expresión. De hecho, Ioro y colaboradores encontraron menor expresión de miR-10b en tumores primarios de mama comparado con tejido no tumoral (293).

Respecto a los miembros de la familia miR-181, los cuales observamos asociados de forma positiva a progresión metastática en tejido, también se ha encontrado una correlación entre altos niveles y peor pronóstico en cáncer de mama (400-402). miR-181 se ha establecido como metastamiR, regulado por TGFb, que promueve metástasis induciendo una EMT, la migración e invasión de células de cáncer de mama (400-402). Además, favorece la expansión del tipo célula madre en éstas mediante inhibición de la kinasa ATM (400-402), lo cual podría explicar que observemos niveles más elevados en metástasis a distancia, donde se podrían haber seleccionado este tipo de células. Asimismo, células MCF-7 resistentes a tamoxifeno tienen una mayor expresión que aquellas no tratadas (322). Esto también concordaría con nuestras observaciones, ya que por un lado las metástasis de nuestra serie provienen de tumores ER+ que han sido tratados con tamoxifeno y pueden estar enriquecidas en células resistentes que se han seleccionado tras permanecer en un estado de quiescencia y han conseguido proliferar cuando el microambiente lo ha permitido generando macrometástasis. El papel de miR-181 como promotor de progresión tumoral se ha visto en varios tipos tumorales (404-407) así como también el de miR-210. Varios metaanálisis que incluyen estudios en distintos tipos de carcinoma han publicado sistemáticamente que miR-210 es un predictor de baja supervivencia total, supervivencia libre de metástasis y supervivencia específica de enfermedad; especialmente en cáncer de mama (408-410). No solamente se asocia a peor pronóstico en los fenotipos más agresivos sino también en tumores de mama ER+ (411, 412). Se ha establecido de forma consistente que la expresión de miR-210 es inducida por hipoxia (413-415), la cual es un factor pronóstico adverso en muchos tumores (416). La relación establecida entre miR-210 y peor pronóstico no es sorprendente considerando que promueve angiogénesis modulando los niveles de VEGF (417-419) y también favorece la invasión y metástasis (420-424). Si en nuestro estudio la inducción de miR-210 refleja un ambiente hipóxico que promueve la progresión del tumor hasta estadios más avanzados, no es sorprendente encontrar que muestre una expresión elevada en metástasis a distancia, que en general aparecen posteriormente en pacientes ya sometidas a tratamiento convencionales, ya que las células hipóxicas se caracterizan por su resistencia a quimioterapia y radioterapia (416). Además, miR-210 tiene un efecto antiapoptótico y citoprotector en células cancerígenas (378, 425-427). Por tanto, tanto miR-181 como miR-210 podrían estar reflejando una mayor capacidad de resistencia a tratamiento en células metastáticas.

Varios estudios que analizan microRNAs en suero o plasma han observado niveles de miR-210 superiores en pacientes metastáticas o como factor pronóstico asociado a una peor evolución de la enfermedad (428-430). Contrariamente, hemos detectado que los niveles circulantes de

miR-210 y miR-181 son significativamente inferiores en la sangre de pacientes con metástasis a distancia. Sin embargo, si consideramos los niveles de expresión absolutos en vez de relativos, ambos microRNAs son los que mayor expresión presentan de entre los analizados. Además, la mayoría de estudios comparan el grupo de pacientes con cáncer *vs* un control libre de enfermedad, por lo que no es de extrañar que la expresión sea superior. Una posible explicación para que observemos niveles inferiores de miR-210 y miR-181 en la sangre de pacientes metastáticas podría ser que a estas pacientes se las intervino y se hizo una resección del tumor primario, que es la principal fuente de secreción de estos microRNAs. Otra posible explicación sería que los niveles de miR-210 sean el reflejo de una mejor respuesta a la quimioterapia en la mayoría de estas pacientes, tal y como se ha publicado en otros estudios (428-431). En un estudio clínico en pacientes tratadas con trastuzumab se observó que los niveles circulantes era más elevados en pacientes con enfermedad residual que en aquellas que alcanzaban una respuesta completa (431).

OBJETIVO 2: Implicación de la familia miR-200 en la progresión metastática

Otra familia relevante que aparecía en el estudio de perfiles de expresión por microarrays es la familia de miR-200. Concretamente, el cluster formado por los miembros miR-200a, miR-200b y miR-429 estaba fuertemente expresado en metástasis ganglionares con respecto a sus tumores primarios pareados, resultados que fueron validados por qPCR. En ausencia de metástasis a distancia, la afectación ganglionar es uno de los factores pronósticos más importantes considerados por los clínicos para guiar la decisión sobre el tratamiento más óptimo. La presencia de céluas tumorales en los ganglios linfáticos refleja la capacidad de un tumor para metastatizar con mayor probabilidad que un tumor negativo para ganglios linfáticos. Otros 8 microRNAs aparecían sobreexpresados además de la familia miR-200. Algunos de éstos tienen funciones conocidas de regulación del sistema immune: miR-150 (432), miR-146a (433), miR-142-3p y miR-142-5p (434).

Dado que miR-200 determina el fenotipo epitelial de las células promoviendo la expresión de E-cadherina a través de la inhibición de los factores de transcripción promotores de EMT, ZEB1 y ZEB2, estudiamos a nivel immunohistoquímico la expresión de E-cadherina en una serie más amplia, en la que se mantenía o aumentaba en metástasis ganglionares respecto de sus primarios pareados. Korpal y sus colaboradores también habían observado una mayor expresión de miR-200 en metástasis ganglionares (341). A pesar de que los niveles de miR-200 en

metástasis a distancia mostraban una gran variabilidad, había una tendencia a ser más elevados en la progresión desde muestras no tumorales a tumores primarios y a metástasis a distancia. Nosotros ya observamos una intensa expresión de E-cadherina en las metástasis incluidas en nuestro estudio. De igual manera, otros han encontrado una expresión aumentada de miR-200b en metástasis a distancia así como una igual o superior intensidad en la tinción de E-cadherina comparado con sus respectivos tumores primarios (435). Además de observar una mayor expresión en metástasis ganglionares, Korpal y sus colaboradores también demostraron que la colonización es fuertemente promovida por la re-expresión de la familia miR-200 y la consecuente diferenciación epitelial (341).

Adicionalmente, miR-200b, el cual seleccionamos como candidato representativo del cluster miR-200b/a/429, discriminaba significativamente tumores primarios con afectación ganglionar de aquello no metastáticos. Al igual que miR-7 y miR-210, ambos promotores de progresión tumoral y factores de peor pronóstico, miR-200b también presentaba una mayor expresión en tumores con capacidad de metastatizar regionalmente. Dado que miR7 también promueve indirectamente la expresión de E-cadherina, podría actuar en sinergia para reforzar la acción de miR-200 manteniendo un fenotipo epitelial, el cual se asocial a una mayor capacidad metastática. Además, también observamos que los niveles circulantes de miR-200b en sangre aumentaban en pacientes con afectación ganglionar comparado con pacientes no metástaticas en un diagnóstico inicial y que estaba sobreexpresado en pacientes en estadio avanzado con metástasis a distancia. Asimismo, Madhavan y colaboradores encontraron niveles más elevados en el plasma de pacientes metastáticas de cáncer de mama de miR-200a, miR-200b, miR-200c y miR-141, que en conjunto con otros 4 microRNAs eran capaces de predecir supervivencia total y libre de enfermedad (436). Además, miR-200b era suficiente para discriminar pacientes metastáticas con células tumorales circulantes (CTCs) de aquellas que no (436). Dado que el porcentaje de CTCs en sangre es bajo (437), esto podría explicar los niveles relativamente bajos de miR-200b que somos capaces de detectar, reflejando posiblemente la naturaleza epitelial de las CTCs. Este mismo grupo encontró que la familia miR-200 son indicadores temprano de metástasis, ya que encontraron niveles elevados en pacientes que posteriormente desarrollaron metástasis en 2 años (436).

En favor de una función prometastática de la familia miR-200s, encontramos que miR-200b era el microRNA que presentaba una expresión diferencial más notable en varias líneas celulares con distinto potencial metastático, siendo su expresión considerablemente más elevada en líneas altamente metastáticas en comparación con aquellas con baja o ninguna capacidad para

metastatizar. En general todas las líneas con capacidad metastática eran además epiteliales, a excepción de las MDA-MB-468 que a pesar de tener un fenotpo mesenquimal y derivar de tumores de tipo basal, presentaban una mayor expresión de miR-200b que las luminales MCF7 (366). Esto indica que otros factores que pueden estar regulados por miR-200 y que son independientes de la necesidad de un fenotipo epitelial, pueden determinar la capacidad metastática.

Dada la diferencia abismal de expresión de miR-200b en el model celular MCF10CA1, seleccionamos éste para caracterizar el fenotipo derivado de la expresión de la familia miR-200 en la línea de tipo mesenquimal MCF10CA1h, que tiene una capacidad de metastatizar limitada comparada con la línea MCF10CA1a, de tipo epitelial y altamente metastática. Disponíamos de vectores clonados con el fragmento genómico del cluster 1 (miR-200b/a/429) y del cluster 2 (miR-200c/141) y sus respectivos controles. También un vector comercial para expresar únicamente miR-200b, que se coexpresaba con GFP para poder monitorizar visualmente su expresión y el vector control que no codificaba para ningún microRNA existente y que también coexpresaba GFP. La sobreexpresión tanto del cluster 1 (miR-200 C1), del cluster 2 (miR-200 C2) como únicamente de miR-200b, induce un claro cambio morfológico hacia un fenotipo epitelial. Las células expresan E-cadherina y EpCAM mientras que la expresión de marcadores de EMT, tales como vimentina, fibronectina o N-cadherina, disminuye. Sin embargo, mientras que siempre hay una inducción de marcadores epiteliales, la represión de marcadores de un fenotipo mesenquimal no siempre es tan abrupta, sugiriendo que existe una parte variable de la población que no experimenta una transición mesénquima-epitelio (MET) completa sino que adquieren un fenotipo intermedio.

Existen dos tipos de corrientes en cuanto a la asociación del proceso de transición epitelio-mesénquimal y adquisición de características de células madre. Se ha demostrado que la expresión de factores de transcripción EMT en células mamarias les confiere un fenotipo de célula madre y que favorece la migración de éstas en el frente invasivo de los tumores favoreciendo las etapas iniciales de metástasis (31). Sin embargo, también se ha propuesto que las células metastáticas con características de célula madre deben readquirir un fenotipo epitelial para poder colonizar otros órganos (172, 232). Dado que las células MCF10CA1h adquieren un fenotipo epitelial al expresar miR-200s, realizamos ensayos de crecimiento de mamoesferas para testar su capacidad de autorrenovación (363). El crecimiento de células en estas condiciones independientes de anclaje es característico de células normales o malignas con característica tipo célula madre (363). Observamos una clara ganacia en la formación de

mamoesferas al expresar tanto miR-200 C1, miR-200 C2 como miR-200b comparado con sus respectivos controles. Adicionalmente, determinamos la expresión de los marcadores de superficie CD24 y CD44 así como el nivel de actividad enzimática de la aldheido deshidrogenasa (ALDH) mediante el ensayo ALDEFLUOR, ambos usados para aislar células madre tumorales de mama. Al-Hajj y colaboradores fueron los primeros en detectar células tumorales de pacientes con cáncer de mama con capacidad para iniciar tumores tras ser implantadas en ratones immunodeficientes NOD/SCID y éstas exhibían un perfil de marcador de superficie positivo para CD44 y negativo para CD24 ($CD44^{pos}$ $CD24^{neg}$) (125). Más tarde se demostró que tanto las células madre mamarias normales como malignas que expresaban ALDH también eran tumorogénicas (136). Posteriormente, CD44, CD24 y la expresión de ALDH se han detectado en células madre tumorales de pulmón, colon, ovario, pancreas y carcinomas de cuello y cabeza (440-444). Nosotros observamos que mientras que menos del 20% de células MCF10CA1h o control expresaban ALDH, la sobreexpresión de miR-200 C1 o C2 inducía la expresión de ALDH en aproximadamente un 50% de la población. Sin embargo, la expresión de ALDH no correlacionaba con el perfil clásico de célula con característica de célula madre, ya que el 90% de las MCF10CA1h control mostraban el immunofenotipo $CD44^{pos}$ $CD24^{neg}$ pero tenían poca actividad ALDH. Esta discrepancia reflejaría que estos marcadores identifican distintos tipos de de célula madre tumoral. Immunofluorescencias en tumores primarios humanos de mama revelan que no tienen necesariamente por qué solaparse: las células $CD44^{pos}$ $CD24^{neg}$ se localizaron en el frente invasivo del tumor en contacto con el estroma mientras que las positivas para ALDH tenían una localización más central (445, 446). Aún así, ambas poblaciones comparten la expresión de genes con células que son capaces de formar tumoroesferas (446). Dado que solamente un número relativamente bajo de células MCF10CA1h son capaces de formar mamoesferas, esto se podría atribuir a la fracción ALDH+ o a células que expresaran $CD44^{pos}$ $CD24^{neg}$ y ALDH+, que son las células que se ha visto que muestran la mayor capacidad para iniciar tumores (136).

Curiosamente, la sobreexpresión de miR-200 indujo un desplazamiento hacia un fenotipo $CD44^{pos}$ $CD24^{pos}$ en aproximadamente el 90% de la población. Considerando que prácticamente la mitad de la población tiene actividad ALDH, es plausible pensar que una fracción considerable de la población que sobreexpresa miR-200, esté expresando también ALDH, CD44 y CD24. Se ha encontrado que independientemente del estado de CD24, las células negativas para CD44 no tienen potencial de autorrenovación (137). Dado que tanto MCF10CA1h control como MCF10CA1h miR-200 C1 o C2 expresan CD44, como mínimo parte de este potencial, observado en el ensayo de mamoesferas, se podría atribuir únicamente a la expresión de ALDH. En cuanto a la positividad para CD24, se ha asociado frecuentemente a un fenotipo luminal con una

diferenciación terminal (373, 447, 448). Además, el análisis de subconjuntos epiteliales mamarios en tejido humano demostró que la actividad ALDH es una característica propia de células progenitoras comprometidas hacia una diferenciación luminal mientras que las poblaciones de células madre mostraban una extremada baja actividad ALDH (13, 18). Esto nos induce a pensar que la sobreexpresión de miR-200 en las células MCF10CA1h, más indiferenciadas, las conduce hacia un estadio de célula progenitora con capacidad de diferenciación luminal. Por este motivo, las sembramos en un cultivo tridimensional con Matrigel, que imita la matriz extracelular al contener hormonas, factores de crecimiento y señalización. Mientras los controles formaban estructuras amorfas, las células MCF10CA1h miR-200s se organizaban en estrutcturas tubulares que conectaban esferas con una luz central que recordaban las estructuras tubuloalveolares del tejido mamario normal, aunque estructuradas en una monocapa en vez de bicapa. Esto confirma nuestra teoría de que estas células adquieren características de progenitoras luminales. De hecho, en la línea MCF10CA1a se han descrito dos poblaciones de célula madre tumoral: una población no invasiva de tipo epitelial positiva para CD44 y CD24, similar a las MCF10CA1h miR-200, y otra invasiva de tipo mesenquimal que únicamente expresa CD44, igual que las MCF10CA1h (137). Estos autores demostraron que las células con un porcentaje elevado de células negativas para CD24 expresaban queratinas basales. Además, observaron una interconversión entre ambas poblaciones de células madre, demostrando la plasticidad entre fenotipos (137, 446). El laboratorio de Max Wicha también mostró evidencias de la existencia de células madre cancerígenas de mama en dos estados alternativos, unas de tipo epitelial y otras de tipo mesenquimal, entre los cuales pueden transitar en respuesta a señales del microambiente tumoral (446). Nuestras observaciones apoyan la existencia de estos dos estados aunque observamos un estado mesenquimal con menor potencial tumorogénico.

Teniendo en cuenta la inducción de un fenotipo epitelial tras la expresión de miR-200s, no es sorprendente que las células MCF10CA1h miR-200 C1 crecieran a un ritmo superior a las control al ser implantadas en ratones NOD/SCID (449, 450). Los tumores derivados de células que sobreexpresan miR-200 tenían un componente epitelial que mostraba una diferenciación glandular en un continuo desde un componente menos diferenciado que también expresaba E-cadherina aunque a una intensidad menor. Esto sugiere la existencia de un continuo de estados intermedios de células que tras la inducción de una transición mesénquima-epitelio mediada por miR-200, han adquirido características epiteliales pero que son plásticas y aún mantienen el programa mesenquimal, pudiendo hacer la transición reversa.

En el ensayo de colonización pulmonar por inyección intravenosa también generaron un mayor número y tamaño de metástasis pulmonares, lo cual correlaciona con el hecho de que a pesar de ser más diferenciadas muestran características de células madre y mayor potencial tumorogénico que las control. Como en nuestro caso, se ha observado una correlación positiva entre la expresión de CD24 y E-cadherina en otras líneas celulares de mama (451). Además, se ha visto un incrementeo notable de células que expresan CD24 en metástasis de distintas localizaciones (447) y que precisamente células positivas para CD24 están enriquecidas con caracteríticas de células madre o progenitoras y tienen un gran potencial tumorogénico y metastático (452). A pesar de que no analizamos la expresión de CD24, las metástasis derivadas de MCF10CA1h miR-200 C1 tienen una expresión intensa de E-cadherina y citoqueratinas luminales (CK8, CK18) además de marcadores basales (p63, CK5), mostrando un mayor grado de diferenciación respecto de las control.

En contraposición a las evidencias anteriores, la visión más extendida es la de que miR-200 previen la metástasis al inducir un fenotipo epitelial en detrimento de la TEM y de la capacidad invasiva de las células (302, 457). Sin embargo, nuestros resultados apoyan publicaciones más recientes en las que la expresión de miR-200 es necesaria para la colonización metastática y el que sea favorecida por una transición mesénquima-epitelio ligada a la adquisición de características de célula madre (341, 356, 357). Es importante resaltar que existen metástasis indiferenciadas derivadas de carcinomas que no necesitan que haya necesariamente una rediferenciación para colonizar otros órganos. Éste sería el caso de los cánceres de mama triple negativos, caracterizados por una baja expresión de miR-200, una transición epitelio-mesénquima y un fenotipo de célula madre (163).

Como ya se ha descrito previamente, una de las vías de señalización cruciales que regulan la transición epitelio-mesénquima es la controlada por los factores de transcripción ZEB1/2, los cuales son dianas de la familia miR-200, que a su vez son reprimidos transcripcionalmente por ZEB1/2 (299, 300, 302, 343). ZEB2 es responsable de reprimir E-cadherina (458) y es ZEB2, no ZEB1, el que aparece infraexpresado en nuestro modelo tras la expresión de miR-200. En nuestros experimentos, el silenciamiento de ZEB2 fenocopia tanto el mayor crecimiento de mamoesferas como actividad de ALDH promovida por miR-200, lo cual sugiere que la mayor capacidad de autorrenovación cuando crecen en condiciones de no adherencia está marcada por la expresión de ALDH. A pesar de que observamos un cambio morfológico hacia un fenotipo epitelial, ello no va acompañado de la expresión de E-cadherina ni de un desplazamiento sustancial hacia una expresión positiva de CD24, lo cual es razonable ya que como hemos

referido, correlacionan positivamente (451). Estas características parecen ser importantes para que haya una organización espacial formando estructuras tubulares complejas, ya que las células silenciadas con ZEB2 solamente son capaces de formar estructuras esféricas. Sin embargo, no podemos excluir que ZEB1 se exprese para compensar la pérdida de ZEB2 y que ello impida la expresión de E-cadherina, por lo que sería interesante conseguir un doble *knockdown* y la inducción, al menos parcial, de una transición mesénquima-epitelio. Se ha visto que el silenciamiento de ZEB2 es suficiente para reproducir el efecto de miR-200 en la reprogramación de fibroblastos a células madre pluripotentes inducidas (iPSC), lo cual va acaompañado de una transición mesénquima-epitelio (459). Este mismo artículo muestras que el dímero Oct4/Sox2 activa directamente los clusters miR-200 mediante unión directa a sus regiones promotoras, lo cuan induce la transición y reprogramación a células iPS. Nosotros estudiamos la expresión de SOX2 como posible mediador en la adquisición de la capacidad de autorrenovación en células que sobreexpresan miR-200. Sin embargo, observamos una infraexpresión de SOX2 a nivel transcripcional, lo cual sugiere que no tiene un papel significativo.

Tanto en tumores primarios como líneas tumorales de mama, la presencia de células madre tumorales sugiere la existencia de un linaje de diferenciación jerárquico rudimentario y por ello determinamos la expresión de marcadores luminales y basales así como marcadores característicos de la transición epitelio-mesénquima. Observamos que tanto si las células que sobreexpresan miR-200 crecen como mamoesferas o en un cultivo de diferenciación en matrigel, hay un incremento considerable de la expresión de E-cadherina que se acompaña de una fuerte expresión de marcadores basales y solamente en matrigel, también expresan queratinas luminales. Esta inducción es aún más exagerada en los tumores ortotópicos de células MCF10CA1h miR-200, que muestran una expresión muy intensa de marcadores basales y luminales acompañada de la expresión de E-cadherina y la pérdidad de expresión de vimentina. Aunque sería necesario hacer dobles immunofluorescencias que demostrasen la coexpresión de marcadores luminales y basales en la misma célula, como mínimo la citoqueratina 5 se expresa en células epiteliales en áreas muy diferenciadas que también expresan citoqueratinas luminales. De nuevo, esto sugiere la existencia de una población de células madre tumorales de tipo mesenquimal/basal y otra de tipo epitelial/luminal que puede fluctuar entre sí (446). Lo que nosotros observamos podría ser una variedad de células con un fenotipo intermedio, no solamente basales epiteliales sino también probablemente células entre un fenotipo basal y luminal que coexpresan ambos marcadores además de tener un fenotipo epitelial, que podrían retener una capacidad bipotencial. Nuestros resultados se ajustan al modelo de “window of stemness” según el cual existe una pequeña fracción de la

población en un estado bipotente o híbrido EMT/MET con características de células madre, que muestra un gradiente de estados intermedios entre cada extremo (460, 461). A pesar de que las MCF10CA1h miR-200 exhiben una diferenciación luminal, tanto éstas como el control forman tumores pertenecientes al subtipo triple negativo. Como evidencia de la heterogeneidad celular de los tumores mamarios y en especial de este subtipo, se han encontrado células con diferenciación luminal en tumores primarios y líneas celulares basales y *viceversa* (462). Además, estas células de tipo luminal, que provienen de una jerarquía de diferenciación con un origen basal, muestran una capacidad para iniciar tumores superior a la de sus equivalente basales (462). De igual manera, en nuestro modelo una transicición mesénquima-epitelio inducido por miR-200 en células MCF10CA1h de tipo basal, genera células de tipo epitelial con características de progenitor luminal que mantienen un fenotipo basal y que muestran una mayor capacidad tumorogénica y metastática.

Los tumores ortotópicos de células MCF10CA1h miR-200 presentan un doble componente con áreas epiteliales que pueden formar estructuras glandulares adyacentes a áreas de células alargadas de apariencia sarcomatosa, unidas por áreas morfológicas de transición, lo cual es sugestivo de la coexistencia de células que muestran fenotipos heterogéneos. Estas características morfológicas recuerdan el fenotipo de los carcinomas metaplásicos de mama. Este subtipo morfológico tiene un componente glandular y otro no glandular (metaplásico). A éste último se le atribuye un fenotipo mesenquimal resultante de una transición epitelio-mesénquima que afecta a células tumorales epiteliales (92, 376-378). Dadas las similitudes morfológicas, hipotetizamos que el componente epitelial/glandular de tumores metaplásicos humanos expresaría niveles más elevados de miR-200 que el componente no glandular y que co-expresaría marcadores luminales y basales así como ALDH. Restringimos el análisis a los carcinosarcomas con un doble componente epitelial y mesenquimal muy bien delimitado, en los que se confirmaron nuestras hipótesis. La coexpresión de queratinas luminales y basales en tumores metaplásicos ya se había observado con anterioridad y por ello se había interpretado que el origen de estos tumores eran células madre bipotentes o mioepiteliales (463-466). Aunque estas interpretaciones son válidas para otros subtipos de metaplásicos, no serían aplicables a los carcinosarcomas que nosotros estudiamos, ya que no tienen en consideración que las células madre epiteliales de mama normal tienen baja expresión de ALDH (18) y que las mioepiteliales no expresan queratinas luminales (29). En base a nuestro modelo, en el que encontramos una coexpresión de queratinas luminales, basales y ALDH en 3 casos de los 5 analizados que además expresan altos niveles de miR-200, esto sugiere que el componente epitelial neoplásico se podría haber originado en células que tienen un fenotipo de progenitor

luminal que retienen características bipotentes. Nuestra hipótesis se aproxima a la posible célula de origen que se propone en otros tipos tumorales de mama en los que las evidencias sugieren que se originan a partir de progenitores luminales (29, 467), con una minorIa de tipos, incluyendo los tumores con baja expresión de claudina y algunos metaplásicos (377), que podrían originarse a partir de células menos diferenciadas que podrían corresponder a células madre (29).

De acuerdo con otros autores (92, 300, 376, 378), proponemos que el componente epitelial de los tumores metaplásicos podría tener suficiente plasticidad fenotípica para dar lugar al componente neoplásico mesenquimal mediante una transición epitelio-mesénquima. Basándonos en nuestros resultados, según los cuales la inducción de un programa epitelial mediante miR-200 confiere agresividad a las células MCF10CA1h, proponemos que el componente más agresivo de los metaplásicos es el componente epitelial en vez del componente mesenquimal. De hecho, las muestras metastásicas de metaplásicos tienden a mostrar características propias de la diferenciación epitelial (468, 469) y los subtipos con mayor representación del componente epitelial, como los carcinomas de células fusiformes y carcinosarcomas, se asocian a mayores tasas de metástasis a distancia y peor evolución (470-473) que aquellos con mayor representación del componente mesenquimal, como los productores de matriz (471, 474). Sin embargo, no excluimos que la presencia del componente mesenquimal pueda contribuir a la agresividad del tumor (92, 468, 471, 475) cooperando para favorecer la invasión local de células epiteliales que intrínsecamente tienen poca capacidad invasiva, como se ha demostrado en varios modelos experimentales (232, 476-478).

CONCLUSIONES

- Hemos identificado un conjunto de microRNAs cuya expresión está alterada a lo largo de la progresión de carcinomas ductales invasivos (CDI) desde tumores primarios no metastáticos a tumores primarios con afectación ganglionar, sus correspondientes metástasis ganglionares y metástasis a distancia. En particular, son interesantes los microRNAs miR-181a, miR-181b, miR-210, miR-7 y miR-200s, cuyos niveles son elevados en la progresión metastática.

- Los niveles circulantes en sangre de miR-7 y miR-200s tienden a ser elevados en pacientes con cáncer de mama metastático en el momento inicial del diagnóstico, siendo indicativo de su potencial como marcadores tempranos de la diseminación metastática de CDI.

- Utilizando el modelo celular MCF10CA1h de tipo mesenquimal, hemos encontrado que la expresión de la familia miR-200 induce un programa epitelial génico muy marcado acompañado de un fuerte incremento de la capacidad de autorrenovación *in vitro*, un desplazamiento hacia una población $CD44^{pos}CD24^{pos}$ más diferenciada, un aumento de la actividad ALDH, la capacidad para diferenciarse aún más *in vitro* bajo los estímulos adecuados y adquirir una diferenciación glandular *in vivo*, todo ello acompañado de una intensa sobreexpresión de marcadores basales y luminales. Estas características sugieren que miR-200s induce propiedades de céluas progenitoras luminales, las cuales unidas a un fenotipo epitelial, se asocian a una mayor capacidad de colonización *in vivo*.

- La adquisición de propiedades de progenitoras luminales inducida por miR-200s es en parte reforzada por la activación de la ruta de señalización PI3K-Akt y la disminución de la expresión de ZEB2. Sin embargo, la adquisición de un programa génico epitelial completo en contraposición al fenotipo mesenquimal de las MCF10CA1h requiere actividades adicionales mediadas por miR-200s.

- El componente epitelial de carcinomas metaplásicos expresa marcadores de progenitoras luminales mamarias, incluyendo ALDH así como citoqueratinas luminales y basales, coincidiendo con unos niveles más elevados de miR-200s. Esto sugiere que el componente epitelial podría ser el responsable del comportamiento agresivo de estos raros tumores, con independencia de que el componente mesenquimal pueda contribuir.